A comprehensive review of reviews of school-based interventions to improve sexual-health.

Health Psychology Review

Sarah Denford[1]

Charles Abraham[1]

Rona Campbell[2]

Heide Busse[2]

[1] Psychology Applied to Health group, Institute of Health Research, University of Exeter Medical School, Exeter, UK

[2] School of Social and Community Medicine, University of Bristol, Bristol, UK

Corresponding author:

Charles Abraham

2.06 College house,

St Lukes Campus,

University of Exeter.

EX1 2LU

C.Abraham@exeter.ac.uk

Acknowledgements

This study was funded as part of National Institute of Health Research's School for Public Health Research (NIHR SPHR) project with additional support from the NIHR Collaboration for Leadership in Applied Health Research and Care of the South West Peninsula (PenCLAHRC). The views expressed are those of the author(s) and not necessarily those of

the NHS, the NIHR or the Department of Health. Author order represents extent of contribution.

Funding statement

The School for Public Health Research and the Collaboration for Leadership in Applied Health Research and Care of the South West Peninsula is funded by the National Institute for Health Research (NIHR). SPHR is a partnership between the Universities of Sheffield, Bristol, Cambridge, Exeter, UCL; The London School for Hygiene and Tropical Medicine; the LiLaC collaboration between the Universities of Liverpool and Lancaster and Fuse; The Centre for Translational Research in Public Health, a collaboration between Newcastle, Durham, Northumbria, Sunderland and Teesside Universities.

Abstract

Objective: To systematically review systematic reviews of school-based sexual-health and relationship Education (SHRE) programmes and, thereby, identify interventions and intervention components that promote reductions in risky sexual behaviour among young people.

Methods: Electronic bibliographies were searched systematically to identify systematic reviews of school-based interventions targeting sexual-health. Results were summarised using a narrative synthesis.

Results: Thirty seven systematic reviews (summarizing 224 primary RCTs) met our inclusion and quality assessment criteria. In general, these reviews analysed distinct sets of primary studies, and no comprehensive review of available primary studies was identified. Interventions were categorized into 5 types that segment this review literature. Unfortunately, many reviews reported weak and inconsistent evidence of behaviour change. Nonetheless, integration of review findings generated a list of 32 design, content, and implementation characteristics that may enhance effectiveness of school-based, sexual-health interventions. Abstinence-only interventions were found to be ineffective in promoting positive changes in sexual behaviour. By contrast, comprehensive interventions, those specifically targeting HIV prevention, and school-based clinics were found to be effective in improving knowledge and changing attitudes, behaviours and health-relevant outcomes.

Conclusions: School-based interventions targeting risky sexual behaviour can be effective. Particular design, content and implementation characteristics appear to be associated with greater effectiveness. We recommend consideration of these characteristics by designers of school-based sexual-health interventions.

Key words

Systematic review; behaviour change; sexual-health; school-based

Background

Young people are at risk of unwanted pregnancies and sexually transmitted infections (STIs). For example, a UK survey of more than 15,000 men and women aged 16-74 found that 31% of heterosexual men and 29% of women had sexual intercourse before they were 16 (Mercer et al., 2013). Consequently, in the UK in 2010, there were 6,674 under-16 conceptions and 34,633 under-18 conceptions; over half of which ended in abortion (The Office for National Statistics, 2012). Similarly, in the United States of America (USA), more than 400,000 teenage women aged 15–19 years give birth (Hamilton, Martin, & Ventura, 2014). Many of these are unplanned; up to 77% according to a UK survey of adolescents (Mosher, Jones and Abma, 2012). Such pregnancies have substantial social and financial costs (Hoffman, 2011) and children born to teenagers have below average educational, behavioural and health outcomes throughout their lives (Hoffman & Maynard., 2008).

Public Health England (2012) reported that the majority of UK STIs diagnoses are among those aged 15-24 years. For example, UK genital Chlamydia diagnoses are highest amongst under 25s with this group accounting for 64% of all new cases. Many STIs, such as Chlamydia, may not be diagnosed in young people and for young women; in particular, this may have serious consequences for future health and fertility. Similarly in the USA nearly half of the 19 million new STI cases each year are among young people aged 15–24 years (Weinstock, Berman, & Cates, Jr., 2004).

Numerous trials and systematic reviews have been conducted to assess the effectiveness of SHRE (e.g.,Kirby et al; Jones, Bates, Downing, Sumnal, & Bellis, 2009). This literature is diverse and reviews to date have specific, non-overlapping aims and inclusion criteria, and use varying quality-assessment methods. Consequently, there is little overlap in the primary studies included in available reviews, and no one review is representative of the literature as a whole.

Consider, for example, two large inclusive reviews of SHRE. A review by Jones et al. (2009a) included 65 studies and grouped into five potentially-overlapping groups, namely, UK-based (n=13), abstinence-only (n=10), abstinence-plus (n=24), HIV and sexual risk reduction (n=11), and others (n=7). Two years earlier, Kirby et al. (2007) had reviewed 56 studies classified as abstinence-only (n=8) and comprehensive (n=48). Yet only 17 studies included in the review by Kirby were also included in the review by Jones (21%). Moreover, two years earlier a Cochrane systematic review had examined the effectiveness of 13 abstinence interventions (Underhill et al, 2007) but only two of these were included in the 10 categorised "abstinence-only" interventions included by Jones et al. (2009a). The lack of overlap in such reviews does not imply inadequate search strategies or reviewer error. Rather, differences in the aims and inclusion criteria of reviews lead to differences in sampling of available primary studies.

We are aware of 12 reviews of reviews of the literature (Cheesbrough 2002; Downing, Jones, Cook & Bellis 2006; Fullerton 2004; Fullerton 1997; Jackson, Haw & Frank 2010; Jepson, Harris, Platt and Tannahill 2010; Jones, Bates, Downing, Sumnall, & Bellis 2009; Lister-Sharp, Chapman, Stewart-Brown, Sowden 1999; Nation, Crusto, Wandersman, Kumpfer, Seybolt, Morrissey-Kane, et al 2003; Peters, Kok, Ten Dam, Buijs & Paulussen 2009; Poobalan, Pitchforth, Imamura, Tucker, Philip, Spratt, et al 2009; Swann, Bowe, K., McCormick, G., & Kosmin 2003). Again these reviews of reviews have quite different, independent aims such that, on average, reviews (across all 12 reviews of reviews) are included in just two reviews of reviews (see online supplement 1).

We aimed to synthesise and simplify this complex literature by undertaking an inclusive, systematic review of reviews that would provide an overview of what is known about effectiveness of SHRE.

Aims

1. To systematically review existing reviews of SHRE for young people and to summarise evidence relating to effectiveness.

2. To identify interventions that are effective in reducing risky sexual behaviour in young people, and highlight key characteristics of those interventions.

Methods

The review was conducted and reported in accordance with a modified version of the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009).

Search strategy

To identify reviews, we electronically searched the Cochrane Database of Systematic Reviews Studies (CDSR), EMBASE, MEDLINE, MEDLINE-in-process, Web of Science, and PsycINFO using the search strategy presented in Supplement 2. Reference lists of included reviews were hand searched, and key authors and experts in the field were contacted.

Inclusion Criteria

Publication date

Prior to the 1980's SHRE was found mainly in lessons on reproduction in science classes (Reiss, 2015). The HIV/AIDS epidemic in the mid-1980s prompted radical changes designed to help children and young people protect themselves. These changes combined with developments in technology, social media, health services and family dynamics means that early school sexual education is unlikely to be relevant to young people in the twenty first century (Irvine, 2002). We decided, therefore, that interventions evaluated before 1990 were less likely to provide insights into the development of effective interventions and limited our search to the period 1990-2016.

Study Design

We considered systematic reviews and meta-analyses of randomised controlled trials (RCTs) cluster randomised trials (C-RCTs) and studies using a quasi-experimental design (including

non-randomised trials (NRCTs), controlled before and after studies or controlled interrupted time series). We define systematic reviews as those in which primary studies are identified using a systematic search strategy. Findings from narrative reviews and review of reviews were not be analysed.

Intervention participants

Reviews of interventions targeting children and young people (aged 4-18 years) in full-time education were eligible for inclusion. This included primary schools, secondary schools, senior colleges and further education colleges, but not tertiary educational establishments such as universities. Reviews in which studies of interventions targeted adults were eligible if the majority of the primary studies were targeting people between the ages of 4-19 years.

Intervention and control groups

Reviews of the effectiveness of any of the following school-based interventions compared to control conditions were eligible: sexual-health interventions; school-linked sexual-health services; or interventions to combat multiple risk behaviour patterns.

We defined school-based interventions as those in which the primary focus of the intervention was the school setting. Such interventions can be delivered outside the school premises, and may be delivered by providers external to the school. Eligible intervention recipients were teachers, school students, parents or a combination of the three. All control conditions comprised those receiving usual care or a modified (including simplified) version of the intervention. Reviews of interventions that were based in multiple settings, or targeted multiple health-related issues were only considered for inclusion if the majority of primary studies were linked to the school, and targeted sexual-health.

Reviews were excluded if their primary focus was sexual-health screening, sexual assault or abuse, or prevention of rape. We also excluded reviews of studies targeting children with developmental disorders.

Outcome Measures

We could have investigated the effectiveness of interventions in relation to health and social outcomes related to sexual-health. However, only a minority of studies reported such outcomes, thus we extracted data on any of the following primary outcomes where reported.

Primary outcomes

(1) Sexual behaviour (2) Health and social outcomes related to sexual-health.

Secondary outcomes

(1) Knowledge and understanding of sexual-health and relationship issues; (2) Personal and social skills (3) Attitudes and values

Risk of bias and assessment of study quality

Review quality was assessed by the lead author using the AMSTAR (Shea et al., 2007). This is an 11 item tool assessing the quality of the review's design, its search strategy, inclusion and exclusion criteria, quality assessment of included studies, methods used to combine the findings, likelihood of publication bias and statements of conflict of interest. The maximum quality score is 11.

Data extraction

Candidate reviews included large numbers of studies (mainly trials) with varying aims. Therefore, wherever possible, we extracted data only on studies relevant to the aims of the current review of reviews. Data extracted included; (i) the aim of the review (ii) the inclusion and exclusion criteria (iii) the number of primary studies included and the number of primary studies that met the inclusion criteria for the current review and (iv) a summary of the findings. This included summarising how effective interventions were, and the types or categories of interventions that were found to be effective as well as features or characteristics typical of effective interventions.

Reliability

Titles and abstracts of identified articles were screened by the first author, and a random 10% checked by the fourth author. Potentially relevant articles were retrieved in full and assessed against the inclusion criteria by the first author with a random 10% checked by the fourth author. Data extraction and quality assessment were conducted by the first author, with a random 10% checked by the forth author.

Cohen's Kappa has been traditionally used to assess inter-rater reliability but because Kappa displays prevalence bias (Byrt, Bishop and Carlin, 1993), alternatives have been developed. Gwet (2002) tested a number of such reliability indices and concluded the "AC1" statistic had optimal output characteristics, particularly when the frequencies of occurrence are small. Consequently, we used this statistic to check inter-rater reliability.

Data synthesis

Given the considerable variation in research aims, inclusion criteria and data presented in included reviews, we could not conduct a meta-analysis. Findings of each review were tabulated along with an assessment of the quality of the evidence provided and a narrative synthesis (Popay et al., 2006) was employed to present the data in accordance with the aims described above. Whilst it was not possible to combine findings from the eleven reviews that conducted meta-analyses the results of these studies are highlighted.

Results

The initial search retrieved 10,699 articles, of which 3,536 were duplicates. This resulted in 7163 articles, of which 303 were retrieved in full (see Figure 1). A total of 56 systematic reviews (Supplement 3, 4 and 5), 31 narrative reviews (lacking systematic searches), and 12 review of reviews were identified. Narrative reviews and reviews of reviews were not analysed in detail, but we have provided a brief overview of the identified reviews of reviews as a brief supplementary report (Supplement 1).

Reliability

Agreement was good for study selection based on titles and abstract (Gwet's AC1 of .89) and full text (Gwet's AC1 of .91). Disagreements were resolved via discussion, and approaches to study selection were modified accordingly. Agreement between reviewers for data extraction was 100% (Gwet's AC1 of 1.0).

Quality Assessment

According to the AMSTAR quality assessment tool's developers, reviews that score between 0 and 4 should be considered low quality. Those scoring between 5 and 8 should be considered of moderate quality, and those scoring between 9 and 11 high quality. Judged by these criteria, 19 of the 56 systematic reviews are of low quality, 26 of moderate quality, and 11 of high quality (Supplement 4). Scores ranged from 1-11 (mean score 5.9). Only one review scored 11 out of 11 (Oringanje et al., 2009). Overall, the majority of reviews performed a comprehensive literature search and provided detail on the characteristics of included studies; however, it was consistently unclear if duplicate study selection and data extraction had occurred, whether publication bias was assessed, whether an *a-priori* design was used, and whether or not there were any conflict of interests. Poor quality reviews may not provide reliable evidence so those scoring less than 5 should be regarded sceptically. We therefore only included reviews scoring 5 and above.

Reviews excluded from this review of reviews are presented in Supplement 6.

Meta-analyses

Thirteen reviews (Albarracin, 2005; Chin, 2012; DiCenso, 2002; Fonner, 2014; Franklin, 1997; Guyatt, 2000; Harden, 2009; Huedo-Medina, 2010; Johnson, 2003; Kim, 2008; Michielsen 2010; Scher, 2006 and Shepherd, 2010) performed meta-analysis of the data.

Five categories of intervention included in systematic reviews

The 37 retained systematic reviews included between 3 and 354 primary studies; of which 224 were relevant to our research aims. This included RCTs, cluster RCTs, studies using a quasi-experimental design, and controlled before and after studies.

Interventions evaluated in these 224 studies could be categorised into 5 types: (1) sexual abstinence-only programmes; (2) comprehensive programmes; (3) pregnancy-prevention programmes; (4) HIV-prevention programmes; and (5) school-based or school-linked clinics. These categories were clearly defined such that ever intervention could be allocated to just one category. Effectiveness within each category was then investigated in in the short term (less than 6 months), medium term (6-12 months) and longer term (more than 12 months) drawing on evidence form the 37 included reviews.

1. *Abstinence-only interventions*

Eight systematic reviews (Blank, Baxter, Payne, Guillaume, & Pilgrim, 2010; Chin et al., 2012; DiCenso, Guyatt, Willan, & Griffith, 2002; Jones et al, 2009a; Kirby, 2007; Scher, Maynard, & Stagner, 2006; Underhill, Operario, & Montgomery, 2007; Wight & Fullerton, 2013) considered the effectiveness of abstinence-only programmes. These reviews included experimental and quasi-experimental studies of variable quality. Common methodological weaknesses included; small sample sizes, reliance on self-reported outcomes, short follow-up periods, and inappropriately outcomes assessed. Overall, these reviews consistently found that interventions promoting abstinence-only can be effective in improving knowledge about how abstinence can protect against STIs, about STIs and about the risks and consequences of unprotected sex and pregnancy (DiCenso et al., 2002; Jones et al., 2009a; Jones et al., 2009b; Kirby, 2007; Scher et al., 2006; Underhill et al., 2007), but are *not* effective in changing behaviour (DiCenso et al., 1999; DiCenso et al., 2002; Jones et al, 2009a; Kirby et al., 2007; Scher et al., 2006; Underhill et al., 2007). Furthermore, there is tentative evidence that such

programmes may increase sexual activity, STIs and pregnancy (DiCenso et al., 2002; Scher et al., 2006; Underhill et al., 2007).

One review of moderate quality examined the effectiveness of interventions aiming to reduce pregnancies and STIs amongst adolescents in the US (Kirby, 2007) included studies of 56 curriculum-based interventions (of which 59% are set in schools) of which 14 were abstinence-only programmes. There was no evidence to support the effectiveness of any abstinence-only intervention in relation to behaviour change. Similar findings emerge from studies of school-based group interventions (Chin et al., 2012), and school-based interventions in which parents are involved (Wight et al., 2013).

A large high-quality review undertaken to support the development of the English NICE guidance included 75 trials of sexual-health and relationship education for adolescents aged 11-19. Ten of these were abstinence-only interventions, all employed in the USA. These were generally effective in improving knowledge and attitudes towards abstinence, but had no positive effect sexual behaviour. Whilst definitions of knowledge varied between studies (Underhill et al., 2009); improvements were seen in all areas. Interventions targeted knowledge about STIs (including AIDS and HIV); contraception and emergency contraception; methods to prevent pregnancy and STIs (including abstinence), risks and consequences of unprotected sex, and sexual health. However, changes in knowledge did not translate into actions. Indeed, participants in one intervention reported greater lifetime sexual experience than controls but this difference was not replicated. A review of interventions targeting younger children (aged 5-11) found one abstinence-only intervention reduced sexual behaviour. This programme targeted participants' knowledge relating to abstinence, self-esteem, and decision making skills; however, due to the average sample age, numbers of participants reporting frequency of sexual behaviour in this population is likely to very small. Therefore, any significant findings must be interpreted with caution (Jones et al., 2009b).

A review published in the Cochrane Library included 13 abstinence-only trials (7 school-based) and concluded that such interventions are not effective in reducing the incidence of unprotected sex, frequency of sex, number of sexual partners, sexual initiation, or increasing condom use (Underhill et al., 2007). One trial found that participants who had taken part in abstinence-only interventions were more likely to report sexually transmitted infections, pregnancy, and increased frequency of sex. However, the reviewers noted that the high attrition in this study means these findings should be interpreted with caution.

In a review and meta-analysis of interventions designed to reduce pregnancy (DiCenso et al., 2002), four abstinence-only programmes and were associated with an increase in number of pregnancies among partners of young male participants (OR 1.54; CI 1.03 to 2.29). Furthermore, a high-quality review and meta-analysis of pregnancy-prevention interventions for adolescents (Scher et al., 2006) reported a pooled analysis of three abstinence-only trials that showed a significantly increased pregnancy incidence among intervention participants (Scher et al., 2006).

2. Comprehensive interventions

Thirteen reviews considered the effectiveness of comprehensive programmes (Akers, Holland & Bost, 2010, Chin et al., 2012; Fonner, 2014; Jackson, 2010; Jones, 2009a; Kim et al 2008; Kirby et al., 2005; Kirby et al., 2007; Robin, et al., 2004; Shepherd et al., 2010; Tolli, 2012; Underhill et al, 2008; Yamada et al., 1999). These programmes *"aim to prevent, stop, or decrease sexual activity, but also promote condom use and other safer-sex strategies as alternatives for sexually active participants"* (Underhill et al., 2008). Reviews focused on studies conducted in the USA (Kirby, 2007); North America (Underhill et al., 2008); or worldwide (Jones et al, 2009a; Kirby et al, 2005), and there were considerable variation in the interventions in terms of populations targeted, theoretical underpinning, delivery provider,

exposure, content, and emphasis on abstinence/ delay of sexual debut (DiCenso et al., 1999; Jones et al., 2009a; Kirby et al., 2005; Kirby, 2007; Underhill et al, 2008).

Overall, these interventions led to improvements in knowledge, attitudes and skills (Jones et al., 2009a; Kim, 2008; Kirby et al., 2005; Kirby, 2007; Underhill et al., 2008). Reviewer's extracted data on changes in a range of skills including: life skills, social skills, and skills relating to sexual risk prevention; however, these were often grouped together, and it was rarely clear which specific skills were targeted. Knowledge included information about STIs (including HIV and AIDs); contraception and emergency contraception; risks and consequences of unprotected sex, STIs and pregnancy; and sexual knowledge. A range of behavioural outcomes were also assessed, largely using self-reported measures. Whilst positive changes in reported behaviour were observed in some studies, findings were not consistent enough to draw firm conclusions (Jones et al., 2009a; Kim, 2008; Kirby et al., 2005; Kirby, 2007; Underhill et al., 2008; Yamada et al., 1999). Indeed, some studies found improvements while others reported negative or null effects for the same outcome. Health-related outcomes were rarely reported, and when they were, few positive changes were observed (Jones et al., 2009a; Underhill et al., 2008; Kirby et al., 2007; Kirby et al., 2005; DiCenso et al., 1999). One review presented evidence that, in some instances, comprehensive programmes may increase sexual intercourse (Kirby et al., 2005) but it is not clear from this review whether or not increased intercourse was protected.

A high-quality Health Technology Assessment (HTA) review explored the effectiveness of 12 RCTs of behavioural interventions for reducing sexually transmitted diseases conducted in the USA, Africa or Europe (Shepherd et al., 2010). Overall, programmes led to changes in knowledge, attitudes, skills, self-efficacy, and behavioural intentions; however few effects on behaviour were found. Seven of the twelve RCTs reported that the intervention had a significant effect on at least one behavioural outcome; however, in

three cases, this was only for a subgroup of participants. Furthermore, other behavioural outcomes were not significantly changed as a result of the interventions. The remaining five RCTs did not report that the intervention had any statistically significant behavioural effects. Meta-analysis revealed that, collectively, the interventions did not result in a significant increase in debut of sexual activity by young people or to an increase in condom use (Shepherd et al., 2010).

One systematic review (Fonner et al, 2014) included 64 studies of SHRE in low to middle income countries. The majority of studies included in the review were described by Fonner 2014 as comprehensive (n=55); however, a further 9 studies were described as abstinence or abstinence plus. Of the 64 studies, 33 provided data suitable for meta-analysis. Results of the meta-analysis showed that studies had a positive overall impact on self-reported: HIV knowledge (Hedges g = 0.63, 95% Confidence Interval (CI): 0.49–0.78, p,0.001), condom use (Odds Ratio = 1.34, 95% CI: 1.18–1.52, p,0.001), and number of sexual partners (Odds Ratio = 0.75, 95% CI:0.67–0.84, p,0.001).

One systematic review considered the effectiveness of programmes that target sexual-health and alcohol misuse simultaneously (Jackson, Haw & Frank., 2010). Randomised trials of these interventions found some positive effects on substance use and sexual risk behaviour. Jackson concluded that the most promising interventions were those that target multiple risk and protective factors (i.e., those that target the school, family, and community).

Jones et al, (2009a) included 18 studies (ten RCTs and eight NRCTs) of interventions categorised as “comprehensive”; and 12 studies (of seven) interventions categorised as “sexual risk reduction programmes in the UK”. Interventions were generally effective in increasing knowledge (as described above). However, effects on behavioural intentions, attitudes towards sexual behaviour, and self-efficacy were inconsistent and often different for young men and women. Although interventions led to short- to medium-term improvements

in skills relating to sexual risk prevention, the impact of programmes on self-reported behaviours such as sexual debut and number and frequency of intercourse was variable. One of five programmes reporting on pregnancy had a positive effect on this outcome.

Underhill et al., (2008) reviewed 39 comprehensive interventions (10 school-based), all set in North America. Twenty four of 39 trials resulted in self-reported change on at least one behavioural outcome. However, there was no evidence of effectiveness among comprehensive programmes in relation to self-reported STIs, and only limited evidence of a protective effect on self-reported pregnancy (Underhill et al., 2008).

Kirby (2007) reviewed 42 comprehensive programmes in the US. Two thirds of these were found to have positive effects on at least one self-reported behavioural outcome measure (including delay in debut of sexual intercourse, reduced frequency of sex, reduced number of sexual partners, increased condom use, increased contraceptive use, and reduced risky sexual behaviour). Nearly two-fifths of the programmes had positive effects on two or more of these behaviours. There was no evidence that comprehensive programmes reduced self-reported STIs or pregnancy. Furthermore, despite concerns, there was no evidence that the inclusion of a contraceptive component increased sexual activity.

A review of 83 comprehensive interventions found similar effects (Kirby et al., 2005). Of the 83 studies 18 were described as being conducted in developing countries and 54 of the programmes were implemented in schools. The authors found evidence for the effectiveness of school-based interventions on knowledge, attitudes and intentions, with the majority of programmes having a significant positive impact on these outcomes. Just under half of the programmes had a positive impact on at least one behavioural outcome. Very few studies measured health outcomes. Three of 13 studies reporting on pregnancy incidence reduction found a positive impact, and two of 10 studies reporting on STI reduction found a positive impact. Moreover, three programmes implemented in developed countries increased

frequency of intercourse (but again it is not clear if this was protected or unprotected intercourse), one US intervention increased pregnancy rates, and, worryingly, two (of 10) studies showed an increase in STI rates. Overall then, these interventions tend not to affect health outcomes with very small number of studies reporting both positive and negative health effects.

In a review of 20 trials of comprehensive interventions, DiCenso et al., (1999) found limited evidence of effectiveness in relation to delaying intercourse debut, improving contraceptive use, and reducing pregnancies. However, all trials included in this review were methodologically weak. This limits confidence in the findings of both the primary studies and subsequent systematic reviews. Moreover, a poor description limits our capacity to identify and describe effective intervention components. This does not mean that current approaches are not effective, but it highlights the need for high quality trials. However, the two strongest trials had no positive effects on any behavioural or health outcomes.

One review and meta-analysis reports evidence to suggest that group-based comprehensive programmes may reduce pregnancy, with some studies showing a statistically significant effect (Chin et al., 2012) but there is no evidence to support the effectiveness of peer delivery of such interventions (Tolli, 2012; Kim et al, 2008). A review and meta-analysis (Kim et al 2008) including thirteen studies of peer-led interventions aiming to promote sexual health, found inconsistent evidence for the effectiveness of such trials on behaviour. Meta-analysis of seven trials found no effects of peer-led interventions on condom use. Although one study reported a reduced risk of chlamydia, another reported no impact on incidence of STIs (Kim et al, 2008). Two reviews suggest that interventions involving parents can influence communication (Akers et al., 2011), but there is limited evidence for the effectiveness on behaviour change (Wight & Fullerton, 2012).

A review of moderate quality assessed the impact of 24 studies of SHRE in the US (Robin et al., 2004). Studies were included if they were of high quality, evaluated using quasi-experimental or experimental methods, and published in the 1990s. Of the 24 included, half had positive effects, five had no effect, and three had negative effects on at least one self-reported behavioural outcome. Two studies had both positive and negative effects. Studies were most likely to demonstrate a positive increase in condom use (8 of 12 studies had a positive impact), and least likely to show an impact on debut of sexual intercourse (4 of 11 studies had a positive impact).

3. Pregnancy programmes

Nine reviews examined effectiveness of interventions to prevent pregnancy (Blank et al., 2010; DiCenso et al., 2002; Franklin et al., 1997; Guyatt et al., 2000; Harden, Brunton, Fletcher, & Oakley, 2009; Lopez, Tolley, Grimes, & Chen-Mok, 2009; Lopez, Otterness, Chen, Steiner, & Galo, 2013; Oringanje et al., 2009; Scher et al., 2006). These interventions used various strategies, specifically to prevent unwanted pregnancy. Again, programmes were diverse in terms of content, providers, sexual-health topics covered, length and intensity, theoretical basis, and origin of interventions. Studies were variable in terms of the comparator used, and characteristics of participants.

Overall, pregnancy programmes appear to be effective in improving knowledge about STIs, contraception (including female contraception), reproductive health and sexual risks (DiCenso et al 2009a) but all reviews state that the impact of these interventions on, attitudes, behaviours and skills is variable; with some studies demonstrating improvements, and others showing no change (DiCenso et al.,2009a; Franklin, 1997; Guyatt et al., 2000; Lopez et al., 2009; Lopez et al., 2013; Oringanje et al., 2009). There is also little evidence to support the use of interventions focusing exclusively on promoting condom use for dual protection (Lopez et al., 2009; Lopez et al., 2013). However, those using infant simulators may be

effective in changing attitudes (Blank et al., 2010). “Multifaceted” pregnancy prevention programmes (DiCenso, et al., 2002; Oringanje et al., 2009; Scher et al., 2006) and programmes targeting social disadvantage (Harden et al., 2009) may lead to reductions in pregnancy.

In a review of programmes aiming to prevent adolescent pregnancy, Franklin et al, (1997) provide little evidence for the effectiveness of programmes for reducing sexual activity; with effect sizes of 17 studies ranging from -0.4 to 0.5. Whilst there is better evidence of the impact of programmes on contraceptive use and pregnancy rates, the effectiveness of school based programmes may be exaggerated by the fact that community-based programmes were also included. In a series of subgroup analyses, community based programmes were found to be more effective than school-based programmes – although school based programmes were still effective.

One high-quality Cochrane review of 41 RCTs of worldwide school-based pregnancy-prevention trials found that interventions may lead to small but significant effects on self-reported unintended pregnancy, while the evidence for the effectiveness of trials on other behavioural outcomes is unclear (Oringanje et al., 2009). Programmes did not appear to be effective in reducing STIs or increasing contraceptive use. However, two Cochrane reviews of behavioural interventions for improving condom use for dual protection (Lopez et al., 2009; Lopez et al., 2013) suggest the opposite, that is, that these interventions do not change pregnancy incidence, but may increase condom use and decrease STI rates.

One review of moderate quality found little evidence to support the effectiveness of pregnancy prevention interventions (DiCenso et al., 2002). Of the 22 trials , 8 were considered to be of high quality and 11 were school-based. More than half of the participants were African American or Hispanic. Meta-analysis found that interventions did not have a significant impact on any behavioural outcome, or on pregnancy rates. However, subgroup

analysis revealed significantly fewer pregnancies in young women who received a "multifaceted programme" (Odds Ratio 0.41; CI 0.20 to 0.83). Similarly, a high-quality Campbell Collaboration review by Scher et al., (2006) also found no evidence for the effectiveness of pregnancy prevention interventions in relation to self-reported sexual behaviour or pregnancy risk. However, the 18 school-based sex education programmes with a contraceptive focus led to a reduction in sexual debut and the likelihood of becoming (or getting someone) pregnant. Seven intensive, multi-component youth development interventions involving higher-risk adolescents showed an overall statistically significant effect pregnancy-risk behaviours and pregnancy rates (Scher et al., 2006). However, it is not clear what the author means by "high risk" adolescents.

More optimistically, in a review and meta-analysis of 10 controlled studies of interventions targeting social disadvantage, Harden et al., (2009) presents evidence from the highest quality trials showing a statistically significant reduction in teenage pregnancy. Programmes included in the review aimed to improve young people's life opportunities and financial circumstances; for example, through educational, social or income support, providing guidance, and raising aspirations through career development and work experience.

4. HIV-prevention interventions

Nine reviews explored the effectiveness of interventions targeting HIV prevention (Albarracin et al., 2005; Gallant et al., 2004; Huedo Medina et al., 2010; Johnson, Carey, March, Levin, Scott-Sheldon, 2003; Jones et al., 2009a; Magnussen et al., 2004; Mavedzenge, Doyle, & Ross, 2011; Michielsen., 2010; Paul-Ebhohimhen et al., 2008). These programmes focused on HIV prevention and HIV risk-behaviour. The content of interventions was variable, as was the target population and delivery methods.

Overall, there is good evidence to support the effectiveness of HIV-prevention interventions in increasing knowledge about HIV/ AIDS (Albarracin et al., 2005; Gallant et

al., 2004; Huedo-Medina et al., 2010; Jones et al., 2009a; Mavedzenge et al., 2011; Michielsen., 2010; Paul-Ebhohimhen et al., 2008;). Two large meta-analyses provide evidence for the effectiveness of such intervention in promoting attitude, behavioural and skill changes (Albarracin et al., 2005; Johnson et al., 2003). However, reviews of studies that are exclusively school-based, present mixed evidence of effectiveness in relation to attitude, behaviour and skill change; with some studies reporting to positive effects and other studies having no or negative effects on these outcomes (Jones et al., 2009a; Huedo-Medina et al, 2010; Magnussen et al., 2004). Of four reviews set in Sub-Saharan African, one found evidence to suggest that interventions were effective in reducing risky sexual behaviour (Mavendzenge et al., 2011); and three found little or no evidence for the effectiveness of interventions on behaviours (Gallant et al., 2004; Paul-Ebhohimhen et al., 2008; Michielsen et al., 2010). One review found evidence for a reduction in STIs (Johnson et al., 2003); and another review found no evidence to suggest that HIV programmes prevent or reduce pregnancy (Jones et al, 2009a).

One very large systematic review and meta-analysis of 354 studies of HIV-prevention interventions (of which 110 were school-based) provides good quality evidence to support the effectiveness of such interventions in enhancing knowledge, attitudes, behaviours and skills (Albarracin et al., 2005). Meta-analyses revealed that interventions had small but positive impact on a number of knowledge, attitudinal and behavioural outcomes – including condom use. A second meta-analysis of 67 studies of 98 HIV-preventive interventions (of which 49 were school-based), Johnson et al, (2003) found interventions significantly reduced self-reported frequency of sex and number of partners, significantly increased the number of participants reporting abstinence, acquisition and use of condoms, and safe sex communication skills as well as significantly reducing the incidence of STIs (Johnson et al., 2003).

Based on 11 school-based programmes targeting HIV-prevention (including eight RCTs, two NRCT and one Controlled Before and After design), Jones et al., (2009a) reported that these interventions increased knowledge of HIV but had, inconsistent effects on attitudes, with some programmes showing positive, and others negative, effects on the same outcome. Of the few studies that measured behavioural outcomes, programmes had variable impacts on sexual debut and condom use; but appeared to have no impact on sexual activity, numbers of sexual partners, HIV/STI testing, and alcohol or drug use. There was little evidence for the effectiveness of these interventions on pregnancy rates (Jones et al., 2009a). The associated review of interventions targeting children aged 5-11 found positive effects of HIV prevention programmes on knowledge (Jones et al., 2009a).

Four reviews of school-based HIV-prevention interventions implemented in Sub-Saharan Africa (Gallant et al., 2004; Mavedzenge et al., 2011; Michielsen., 2010; Paul-Ebhohimhen et al., 2008) provide similar findings. Whilst all reviews showed positive effects of interventions on knowledge, the impact on attitudes and behaviours were inconsistent and variable. In the review by Gallant et al., (2004) only one of the five studies reporting on sexual behaviour reported a significant reduction; and only one of the four studies assessing condom use found a significant effect of the intervention on condom use. Although 9 of the 11 studies in the review by Mavedzenge et al., (2011) had a positive impact on at least one self-reported behavioural outcome, such effects were inconsistent between studies. A meta-analysis by Michielsen, 2010 found few positive or negative effects on sexual behaviour; although condom use at last sex increased in males only (RR 1.46; 95; CI 1.31–1.64) with little heterogeneity, for the girls, heterogeneity was substantial; thus indicating significant differences in effect sizes between studies.

5. *School-based healthcare centres*

School-based/linked sexual-health services (SBSHS) are "*services or clinics provided in schools; services located near schools that conduct outreach work within those schools; or services located near schools which liaise formally with those schools. The interventions of interest are those delivered to individuals who attend the services on a voluntary basis, and do not include either classroom interventions*" (Owen et al., 2010). Four reviews of studies, largely conducted in the US provide some evidence to suggest that SBSHS do not increase sexual activity (Kirby 2007; Owen et al., 2010). There is, however, also no evidence that SBSHS increase contraceptive use (Blank et al., 2010; Kirby 2007; Mason-Jones et al., 2012; Owen et al., 2010), but some interventions show a statistically significant reduction in sexual activity, numbers of sexual partners, and live births to teenage mothers (Blank et al., 2010; Owen et al., 2010). However, results were variable between studies; with some studies reporting positive effects, and others showing no impact (Mason-Jones et al., 2012). Effectiveness may be increased if these interventions are integrated into whole-school programmes (Kirby, 2007).

A high-quality review was conducted by Owen and colleagues to explore the effectiveness, and factors relating to the effectiveness, of SBSHS (Owen et al., 2010). From 30 studies of 26 interventions, they reported that there was evidence from higher-quality American studies that SBSHs do not increase rates of sexual activity or lower the age of sexual intercourse. They also reported that there was no evidence that SBSHS increased contraceptive use, but that they may be associated with a reduction in the proportion of school students reporting recent sexual activity, high numbers of sexual partners, and live births to teenage mothers. In addition, they found only minimal evidence that these interventions reduced chlamydia prevalence in male students. The authors conclude that further research is needed to explore the effectiveness of SBHCs in UK settings.

Kirby, 2007 reviewed the effectiveness of 59 non-curriculum-based interventions; including school-based clinics and condom availability services. SBHCs and school condom-availability programmes did not increase sexual activity, nor did they appear to increase condom or contraceptive use. Whilst these interventions provided contraceptives to a large number of students, contraceptive use did not appear to change. The authors commented that this may have been because students were simply accessing contraceptives from a new source. The only multi-component programme included in the review included sex education incorporated into all grades in schools in South Carolina. Condoms were made available and young women were taken to family planning clinics. In addition, local media, churches, and other community organisations were involved. This intervention had a significant impact on pregnancy rates for twenty years. Blank et al., (2010) reviewed six studies of school-based health centres, and found similar results. Whilst evidence suggested that the centres increased contraceptive uptake, there was no evidence that this affected sexual behaviour.

So what works to change sexual-risk behaviour patterns in school-based interventions?

Sixteen systematic reviews provide recommendations regarding intervention characteristics that may increase effectiveness. Based on a review of 83 primary evaluations, Kirby's review (cited more than 1,000 times) identified 17 characteristics relevant to intervention development, content, and implementation that may increase the effectiveness of interventions. At the development stage, Kirby (2007) suggest that developers should (i) involve multiple people with varied backgrounds, (ii) assess the needs and resources of the target groups, (iii) use a logic model approach specifying mechanisms of change to plan the intervention content and the implementation of change techniques (iv) ensure continuity between the intervention and the values and resources of the target population (iv) pilot-test all materials.

In relation to intervention content, Kirby (2007) suggested that programmes should (i) specify and focus on clear goals (ii) focus on specific behaviours leading to these health goals and contain clear messages about those behaviours, (iii) address psychosocial risk and protective factors affecting behaviours while in relation to delivery of the intervention the authors recommend (i) creating a safe environment for youth (ii) including multiple strategies and methods to change behaviour (iii) ensure participants are actively involved in intervention activities, (iv) that programmes are appropriate for participants' culture, age, sexual experience, and (v) topics are presented in a logical sequence.

Eight reviews supported and developed Kirby's (2007) recommendations about the content of effective interventions. Robin et al., (2004) concurred with Kirby and colleagues suggesting that effective programmes focus on behaviours and skills that reduce specific sexual-risk behaviour patterns; employ interactive and participatory educational strategies; and are of sufficient duration and intensity. Gallant et al (2004) suggest that the most effective programmes were of the longest duration and include a diverse content and range of activities. Huedo-Medina et al., 2010 suggested that the most effective programmes were those that lasted longer than three hours, and targeted those most at need. DiCenso et al., (1999) suggested that interventions should focus on psychosocial risk and protective factors and actively involve participants. Both DiCenso et al., (1999) and Oringanje et al., (2009) provide data supporting the view that effective programmes employ multiple, as opposed to, single change strategies.

Three reviews recommended content characteristics not featured in Kirby's (2007) recommendations. For example, Harden et al., (2009) and Jackson et al., (2010) both suggested that interventions targeting social disadvantage may be effective in reducing pregnancy. Finally, in relation to mode of delivery, Chin et al. (2012) provided evidence to suggest that group-based comprehensive interventions can promote behavioural and health

improvements; with good quality studies showing a statistically significant reduction in sexual activity (OR=0.84), unprotected sexual activity (OR 0.70) and STI incidence (Odds Ratio = 0.65).

Two reviews (Albarracin et al., 2005; Johnson et al., 2008) offered recommendations on characteristics of interventions that effectively target HIV-prevention which may be relevant to other interventions aiming to reduce STI spread. Using meta-regression analyses, Albarracin et al., (2005) suggested that interventions targeting HIV prevention and condom use should include (i) education (ii) attitudinal arguments (i.e., those that promote positive outcomes of e.g., condom use); (iii) behavioural skills arguments (i.e., arguments that verbally promote recipients behavioural skills for example, by bolstering self-efficacy), and (iv) behavioural skills training (i.e., provide training that enhances participants' competence in, negotiating use of and use of condoms. Albarracin et al., (2005) suggested that the least effective programmes attempted to induce fear of HIV. However, with the exception of condom provision (which was effective for men and women), the authors noted that all other strategies had differential effects on men and women, which suggested tailoring of intervention content to target audiences (Albarracin et al., 2005) was required.

Johnson et al., (2008) conducted a series of meta-regression analyses and suggested that interventions were most successful at reducing the frequency of sexual behaviour when (i) they were implemented with adolescents who were institutionalized, (ii) did not promote abstinence-only (iii) included a greater numbers of intervention sessions, and (iv) participants in comparison conditions did not receive HIV education. They found interventions were most effective in enhancing condom use when they provided a considerable amount of condom skills training or motivational training in each session.

In relation to the implementation of interventions in context, Kirby (2007) recommendations were to (i) ensure there is at least minimal support from authorities, (ii) use

trained educators (iii) recruit youth if necessary, and (iv) ensure fidelity of activities. Robin et al., (2004) also stated that the programme content should be clearly specified so are unlikely to be modified during implementation, and be delivered by trained facilitators. Shepherd et al., 2010 agreed that a lack of effectiveness could be due to imperfect implementation of the interventions, and difficulty in engaging the young people in the intervention (i.e., due to embarrassment or content not meeting their needs). DiCenso et al (1999) and Gallant et al., 2004 reinforce the need to use trained educators to deliver programmes, however, there is a lack of evidence for the effectiveness of peer-led interventions (Kim et al, 2008; Chin et al 2012).

Combining these recommendations, including Kirby's initial 17 characteristics, we can summarise "best bet" suggestions derived from 37 reviews. These are listed in Table 1 which is divided into 5 sections relating to the (i) development, (ii) content, (iii) HIV specific content, (iv) implementation and (v) evaluation of interventions. Recommendations in the development section are all from Kirby. Those presented in the content section are taken from Kirby and supported by eight other reviews (Chin et al., 2012; DiCenso et al., 1999; Gallant et al., 2004; Harden et al., 2009; Huedo-Medina et al., 2010; Jackson et al., 2010; Oringanje et al., 2009; Robin et al., 2004). The HIV-specific content are derived from Albarracin et al, 2005 and Johnson et al., 2003. Recommendations in the implementation section are taken from Kirby and six additional reviews Chin et al., 2012; DiCenso et al., 1999; Kim et al., 2008; Gallant et al., 2004; Robin et al., 2004 and Shepherd et al., 2010. Finally, the evaluation recommendations were supported by nearly all reviews because, across studies, evaluation was generally poor, with lack of long term follow-up and objective measures of health outcomes.

Discussion and conclusions

This is, to date, the most comprehensive review of school-based sexual-health and relationship education interventions. We have systematically identified and synthesised evidence-based recommendations across 37 reviews of 224 relevant trials and quasi-experimental studies, including the highest quality and most highly cited reviews in the field. This compares to 65 and 83 studies included in the largest previous reviews (Jones et al., 2009 & Kir by et al., 2006, respectively). We have also observed how diverse the field is and how little overlap there is between existing reviews. We present evidence for the effectiveness of interventions categorised as (1) abstinence-only (2) comprehensive (3) pregnancy prevention (4) HIV-prevention (5) school-based health services. In Table 1, we summarise this evidence as 32 recommendations in five areas of intervention design, namely, (i) development, (ii) content, (iii) HIV specific content, (iv) implementation and (v) evaluation of interventions. While interventions designers will need to tailor these recommendations to their particular objectives, target populations and objectives, these offer an evidence-based starting point for researchers and designers trying to understand what is known across this diverse literature. We also provide lists of the reviews we found and of the 224 relevant studies they identified.

Many of the reviews highlight how weak and inconsistent the evidence is on behavioural effects but clear conclusions can also be drawn. Sexual-health and relationships interventions that focus exclusively on sexual abstinence- are not effective in changing the sexual behaviour of school students (Jones et al., 2009a; Scher et al., 2006; Underhill, Operario, & Montgomery, 2007). Comprehensive, sexual risk reduction and HIV-prevention interventions were consistently effective in changing knowledge, attitudes and skills (DiCenso et al., 2002; Jones et al., 2009a; Kirby et al., 2005; Kirby, 2007; Lopez et al., 2009; Lopez et al., 2013; Oringanje et al., 2009; Robin et al 2002; Shepherd et al., 2010; Underhill et al., 2008).

Across all categories, interventions provided knowledge on STIs (including HIV and AIDs); contraception and emergency contraception; methods to prevent STIs (including abstinence), and risks and consequences of unprotected sex, pregnancy and STIs, reproductive health. However, definitions of knowledge varied between reviews, and due to lack of detailed descriptions of intervention components in both original articles and subsequent reviews, we are unable to provide specific detail about the most effective forms of information. Interventions targeting life skills, social skills, and skills relating directly to sexual risk prevention were frequently grouped together within reviews. It was often not possible to identify which specific skills had been targeted and for which behaviour patterns there was convincing evidence of change that could be attributed to exposure to an intervention. Some interventions in all five categories led to behaviour change but reviewers reported that positive changes were inconsistent; some interventions leading to change in one outcome, and others failing to replicate this. Two systematic reviews of studies largely conducted in the US, provide some evidence to suggest that school-based Sexual Health Services (SBSHS) do not increase sexual activity or contraceptive use (Kirby, 2007; Owen et al., 2010), but high quality studies have found a statistically significant reduction in sexual activity, numbers of sexual partners, and live births to teenage mothers (Owen et al., 2010).

Implications for research

Systematic reviews frequently stated that the quality of primary trials and quasi-experimental studies was generally poor. Methodological weaknesses identified by many reviewers included; lack of randomisation, insufficient follow-up periods, inappropriate control conditions, small sample sizes, high attrition and lack of replication studies. Whilst length of follow up was generally inadequate, there was evidence that good, long term follow up is possible (Jackson et al., 2010; Jones et al., 2009a). Fidelity of delivery was rarely assessed; only one review commented on the extent to which programme fidelity was

assessed (Lopez et al., 2013). This makes it difficult to attribute ineffectiveness to intervention design and content rather than to failures to implement in accordance with the intervention protocol. It is important to note that this does not mean that trials of such interventions are not of value; indeed, much can be learnt from the extensive literature considered in this review. However, good quality, trials with long term follow up would increase the confidence with which we can claim that such programmes have population level effects.

Some previous reviews have focused on identifying types of change techniques associated with greater intervention effectiveness for specified participants and taxonomies of such techniques have been developed to support categorising and coding of intervention content into types of change techniques (e.g., Abraham & Michie, 2008; Michie et al., 2011). However, in a recent paper, Kok et al. (2016) highlight the need to consider, not just included change techniques, but also the conditions that determine whether or not specified techniques are likely to be effective. These authors helpfully describe a minimal set of characteristics that should be reported on; and suggest that specification of such factors in intervention development will augment the science of behaviour change. Similarly, Abraham (2016) has clarified that inferences drawn from categorisation of change techniques in retrospective deconstruction of intervention content may be misleading if categories include techniques that, while appearing similar, involve different delivery modes and different change mechanisms. Our review highlights how such specificity in identification of intervention content could enhance school-based, sexual-health and relationship education interventions.

We also identified methodological weaknesses in the reviews themselves (Supplement 4). Few reviews made it clear if duplicate study selection and data extraction had been applied, very few reviews considered publication bias, and conflicts of interest were not always stated. We identified 55 systematic reviews of the literature, of which only 10 were

considered to be of high quality. Interestingly, these ten were not the most highly cited with citation rates varying from 1-127 reported by Google Scholar. More highly cited reviews (with at least 200 citations) were generally of lower quality. The reviews considered to be of the highest quality included Cochrane reviews, Campbell collaboration reviews, NICE reports and HTA reports, for which review protocols are readily available and longer reports (than are usually accepted by journals) are presented. Despite this, one such review was cited only once at the time of this review. While citations do not necessarily translate into influence on policy or practice, they do indicate the extent of academic influence of a paper. It is possible, therefore, that reviews of high quality are not as practical or useful to intervention developers as other reviews so limiting their impact on subsequent intervention development and evaluation. Short, readable summaries of recommendations of future research an intervention design could enhance usability.

Lower quality reviews generally received lower AMSTAR scores because reviewers: did not present tables of excluded studies, whether methods of the review had been specified a-priori, did not assess publication bias, or did not declare conflicts of interest. It is possible that many of these apparent quality deficits arose from reporting restrictions, rather than serious bias in the synthesis of results by these reviewers. If scores from the AMSTAR scale are good reflections of the quality of reviews, it is worrying that the many reviews have design flaws. If, on the other hand, low scores reflect reporting inadequacies rather than fundamental methodological flaws, it may be that clearer weighting of the quality assessment features identified in tools such as AMSTAR would help better assess quality.

Strengths and limitations of this review

Our review provides a comprehensive overview of the literature summarising what is known about what does and does not work. However, several challenges limit the utility of reviews in this area. Evaluations of interventions usually relied on self-reported behavioural data with

its inherent limitations, so use of validated measurement techniques is important (Catania et al., 1990; Fenton, Johnson, McManus, & Erens, 2001). Objective measures of health-related outcomes such as pregnancy and STI incidence can be employed but interventions may be effective in generating behaviour change without impacting on such measures, for example, because of the low incidence of sexual intercourse among many school populations. Consequently, such measures may not provide the best assessment of intervention effectiveness.

In line with standard review procedures, we ensured our selection process; data extraction and quality appraisal were checked by a second reviewer. Due to the high rate of agreement, we limited this to 10%. While we may have missed relevant articles we believe this is unlikely due to the high levels of agreement observed.

As we have noted, interventions are categorised inconsistently across reviews and reviews do not substantially overlap in the studies they include because their objectives and research questions differ. We have calculated a measure of the degree of overlap as specified by Pieper et al (2014). This “corrected covered area” is calculated by dividing the frequency of repeated occurrences of a single publication in other reviews by the total number of index publications and reviews, reduced by the number of index publications. This calculation shows that our review has a very low overlap rate of just 3%.

Finally, reviews have identified characteristics related to effectiveness, through content analyses of descriptions of effective programmes which, in general, have not been the subject of meta-regression analyses (only 11 of 37 reviews employed meta-analyses). The relationship between inclusion of these characteristics and effectiveness cannot, therefore, be viewed as causal and estimation of likely effect sizes is based on a small subset of available studies. Nonetheless, our review of reviews suggests characteristics that designers would be wise to include if possible.

Conclusions

We present the most comprehensive literature review of school-based sexual-health interventions to date. We have considered the evidence in relation to five categories of school-based interventions and we suggest that 32 design, content, implementation and evaluation characteristics are likely to enhance the effectiveness of such interventions.

Conflicts of interest

None declared.

Reference List

Bellis, M. A., Hughes, K., Calafat, A., Juan, M., Ramon, A., Rodriguez, J. A. et al. (2008). Sexual uses of alcohol and drugs and the associated health risks: a cross sectional study of young people in nine European cities. *BMC Public Health, 8,* 155. doi.10.1186/1471-2458-8-155

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. Journal of clinical epidemiology, 46(5), 423-429.

Bandura, A. (1999). Health promotion from the perspective of social cognitive theory. *Psychology and Health, 13,* 623-650. doi.10.1080/08870449808407422

Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin, 108,* 339-362. doi.10.1037//0033-2909.108.3.339

Choudhry, V., Agardh, A., Stafstrom, M., & Ostergren, P. O. (2014). Patterns of alcohol consumption and risky sexual behavior: a cross-sectional study among Ugandan university students. *BMC Public Health, 14,* 128. doi.10.1186/1471-2458-14-128

Fenton, K. A., Johnson, A. M., McManus, S., & Erens, B. (2001). Measuring sexual behaviour: methodological challenges in survey research. *Sex Transmitted Infections, 77,* 84-92. doi.10.1136/sti.77.2.84

Gwet, K. (2002). Inter-Rater Reliability: Dependency on Trait Prevelance and Marginal Homgeneity. *Statistical Methods for Inter-Rater Reliability Assessment Series, 2,* 1-9.

Hamilton, B. E., Martin, J. A., & Ventura, S. J. (2014). Births: Preliminary data for 2009. *National vital statistics reports, 59(3).*

Hoffman, S. D., & Maynard, R. A. (Eds.). (2008). Kids having kids: Economic costs and social consequences of teen pregnancy (2nd ed.). Washington, DC: Urban Institute Press.

Hoffman S. Counting It Up: The Public Costs of Teen Childbearing. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy; 2011

Irvine, J. M. (2004). Talk about sex: The battles over sex education in the United States. University of California Press.

Mercer, C. H., Tanton, C., Prah, P., Erens, B., Sonnenberg, P., Clifton, S., ... & Copas, A. J. (2013). Changes in sexual attitudes and lifestyles in Britain through the life course and over time: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). The Lancet, 382(9907), 1781-1794.

Mercer, C. H., Bailey, J. V., Johnson, A. M., Erens, B., Wellings, K., Fenton, K. A. et al. (2007). Women who report having sex with women: British national probability data on prevalence, sexual behaviors, and health outcomes. *American Journal of Public Health, 97,* 1126-1133. doi.10.2105/AJPH.2006.086439

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology, 62,* 1006-1012. doi.10.2105/AJPH.2006.086439

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., Baird, J. (2014) Process evaluation of complex interventions, UK Medical Research Council Guidance, London: MRC Population Health Science Research Network.

Mosher, W. D., Jones, J., Abma, J.C. (2012). Intended and unintended births in the United States: 1982-2010. National Center for Health Statistics. Vital Health Stat (55). Retrieved January 23, 2015, from http://www.cdc.gov/nchs/data/nhsr/nhsr055.pdf.

Pieper, D., Antoine, S. L., Mathes, T., Neugebauer, E. A., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. Journal of clinical epidemiology, 67(4), 368-375

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M. et al. (2006). Guidence on the conduct of narrative synthesis in systematic reviews. Final report; ESRC Methods Programme: Swindon.

Redgrave, K. & Limmer, M. (2005). *It makes you more up for it. School aged young people's perspectives on alcohol and sexual-health.* Rochdale: Rochdale Teenage Pregnancy Strategy.

Shea, B., Grimshaw, J., Wells, G., Boers, M., Andersson, N., Hamel, C. et al. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology, 7,* 10.

Smith, L. & Foxcroft, D. (2009). *Drinking in the UK: An exploration of trends* York: Joseph Rowntree Foundation.

Spratt, J., Philip, K., Shucksmith, J., Mcnaughton, R, Tucker, J. and van Teijlingen, E., 2009. A Review of sex and relationship activities and resources in primary schools in Scotland (Full Report). Project Report. Edinburgh, Scotland: NHS Health Scotland.

Weinstock, H., Berman, S., & Cates, W., Jr. (2004). Sexually transmitted diseases among American youth: incidence and prevalence estimates, 2000. *Perspectives on Sexual and Reproductive Health, 36, 6-10.*