Interventions to improve participation amongst underserved population groups in young person and adult national screening programmes in the UK: a systematic review

APPENDIX 4

Review of reviews, narrative and tables

Review of Reviews

The search strategy also included systematic reviews and economic analyses. Previous systematic reviews cannot be directly incorporated into a new systematic review (unless the new review is explicitly an update of an existing review) but it is good practice to consider what has been found by others considering similar questions.

We produced a shortlist of systematic reviews and economic analyses which addressed a relevant question and included at least one UK trial which would be eligible for this review (or a UK perspective for economic evaluations). Many of these studies did not report results for any underserved groups in detail and few had a strong relevance to the UK context. The funder opted to consider only those reviews or economic analyses which had a strong UK perspective and which considered at least one underserved group in some detail.

The five included studies (three systematic reviews and two economic evaluations) are summarised below with more detailed summary tables, including risk of bias assessments, provided at the end of this document.

Study	Question(s)	RoB
Systematic reviews		
Jepson 2000	The determinants of screening uptake and interventions for increasing uptake: a systematic review	Low
Sokal 2010	A critical review of the literature on the uptake of cervical and breast screening in British South Asian women	Unclear
Myers 2019	Ways to use interventions to increase participation in mail-out bowel cancer screening: a systematic review and meta-analysis	High
Economic evaluations		
Brown 2006	A Bayesian approach to analysing the cost-effectiveness of two primary care interventions aimed at improving attendance for breast screening	Low
Asaria 2015	Distributional cost-effectiveness analysis of health care programmes – a methodological case study of the UK bowel cancer screening programme	Low

Jepson 2000

The determinants of screening uptake and interventions for increasing uptake: a systematic review

Risk of bias: **LOW**

This well-conducted systematic review included two components: a review of the determinants of screening and a review of intervention studies (including non-randomised controlled trials) reporting uptake. We have reviewed only the second component (interventions to improve uptake) here but the first component was used to help establish which groups might be considered underserved for each screening programme by sex and age.

The review had no restrictions on the screening programmes to be included, regardless of their applicability to the UK context and used an extensive search strategy, using 23 databases (including grey literature) and no language restrictions. The review of interventions to improve uptake included 190 trials, including 130 RCTs. The methods were strong and the results interpreted appropriately.

This review is now twenty years old, conducted very early on in the establishment of population-based screening programmes and, especially, the availability of RCT evidence from such programmes; many of the comparators were opportunistic screening (as were some of the interventions). The review of evidence for interventions to improve uptake did not report results for underserved groups separately but we compare the findings to ours here as we found few strong interactions within subgroups reported.

Excluding those results which compare (or test) opportunistic screening, they report promising evidence for fixed appointments compared to open invitations, as also reported by Stead 1998 (BSP), Allgood 2017 (BSP) and Kitchener 2018 (CSP) in this review, and the removal of financial barriers (we did not find any UK RCTs of this intervention, although much of this evidence in Jepson 2000 derived from private healthcare contexts). They also report evidence that the use of rewards and incentives are ineffective, as also reported by Judah 2018 (DES) in this review.

They found some evidence in favour of pre-appointment reminders, in common with Kerrison 2015 (BSP) and Allgood 2016 (BSP), for text and postal pre-appointment reminders in this review. They also found evidence that the use of reminders for missed appointments appeared effective (with evidence only from mammography), a result which is overwhelmingly confirmed by trials included in this review although we have not reported those control arms which did not offer any reminder for screening programmes where reminders are now standard.

They report that educational home visits may be effective, in common with McAvoy 1991 (CSP), but this trial was conducted before systematic screening. Sharp 1996 (BSP) found that home visits did not appear more effective than a GP letter and Shankleman 2014 (BCSP) did not suggest any benefit of home visits compared to a telephone call, in common with Hoare 1994 (BSP).

They found that reminders for physicians were effective but with some evidence that office systems or the use of audit and feedback might also be useful. These results are

in line with Bankhead 2001 (BSP) and Richards 2001 (BSP), parallel trials which tested the use of flags in (pre-computerised) notes.

The authors recommend that studies of uptake include measures of informed choice; we concur but would also note the difficulties in doing trials of uptake whilst also asking participants to complete questionnaires, introducing both a (self-) selection bias and possibly reducing uptake by asking too much of those randomised..

As we have also noted, they report finding little evidence addressing barriers for minority ethnic groups. The three UK trials that they found addressing this question are also included in this review: Hoare 1994 (BSP), McAvoy 1991 (CSP) and Atri 1997 (BSP). They also included Lancaster 1992 (CSP) but do not note the retrospective Asian subgroup in that study. Also included in both reviews are Turner 1994 (BSP), Meldrum 1994 (BSP), Sharp 1996 (BSP), Stead 1998 (BSP), and O'Connor 1998 (BSP).

Sokal 2010

Risk of bias: **UNCLEAR**

A critical review of the literature on the uptake of cervical and breast screening in British South Asian women

This paper describes itself as a critical (and also scoping) review rather than systematic review but the methods (including a reproducible search strategy) are strong enough to include and it addresses a difficult sub-topic.

The only randomised evidence found for uptake was Atri 1997 (BSP) which is also included in this review. McAvoy 1991 (CSP) and Hoare 1994 (BSP) were excluded because the date range of the searches was for 1996 onwards to focus on more recent evidence. The other two studies of uptake identified were observational studies and reported general practice based interventions providing support via translation, transport and link workers, and a community development programme, respectively.

Mixed evidence of the relationship between ethnicity and uptake was found, with the author noting the difficulties for drawing conclusions from observational studies where ethnicity and socioeconomic status are confounded and difficult to untangle.

Myers 2020

Risk of bias: **HIGH**

Ways to use interventions to increase participation in mail-out bowel cancer screening: a systematic review and meta-analysis

This review used standard systematic review methods for identification and quality assessment of the trials, which included a small number of non-randomised controlled trials with the majority of studies being RCTs.

Unfortunately the methods used to combine the results are non-standard and statistically naive. Cluster trials (which included the four very large ASCEND trials) were not adjusted for clustering; raw data was used for all analyses. This will introduce a substantial bias as cluster trials have a much smaller effective sample size than the raw numbers suggest and are also difficult to balance well with respect to demographics.

There is also a serious problem with the approach to examining subgroup effects (for sex, age and socioeconomic status). They produced binary groups for men/women, younger/older and deprived/less deprived, extracted data for each subgroup, pooled the subgroups across trials, and then conducted a test for interaction across the pooled subgroups. This approach might be somewhat justifiable if the trials were clinically and statistically homogeneous but this was not the case for most of the analyses reported. Most of the subgroups were very heterogeneous and the intervention “simplified testing” pooled trials of no dietary restrictions with those asking for fewer samples in the completed test kit. These interventions are very different and the trials pooled in this group cannot be considered clinically homogeneous. This problem is magnified by the failure to adjust for clustering and by differing international contexts.

The review also attempted to examine the value of combined interventions but found that there were too many combinations to deal with analytically. A simple analysis of extra printed materials vs no extra printed materials (without accounting for subgroup effects) suggested that additional printed materials reduced uptake, which is not out of line with what some of our trial results suggest.

Several trials included in our review were also in this one: the four ASCEND trials, Libby 2011, Hirst 2016 and O’Carroll 2015. Lo 2014 is missing despite reporting all three subgroups of interest (age, sex, socioeconomic status) and Shankleman 2014 is also missing (possibly because no data were reported for the only subgroup of interest to this review and the authors were unable to provide us with that data).

Cost-effectiveness studies

Brown 2006

A Bayesian approach to analysing the cost-effectiveness of two primary care interventions aimed at improving attendance for breast screening

This cost-effectiveness analysis was done as part of two parallel trials included in this review: Bankhead 2001 (BSP) and Richards 2001 (BSP), comparing two primary care interventions (GP letters and flags in notes) against no further action (standard practice

in this context) using a factorial design. The paper scores highly on the CHEC list, a tool used for critical appraisal of economic evaluations.[1]

Bankhead 2001 randomised individual women from GP lists who had recently missed a screening appointment; Richards 2001 cluster-randomised general practices to the same interventions for all women due to be invited for screening. More details of these trials are in the trial summary tables in Appendix 2 of this review. Both trials reported that both interventions were effective at increasing uptake, but there was no clear evidence that combining the two interventions was more effective than each one alone.

Neither trial reported detailed results for subgroups by screening history and the authors were unable to provide more information, but both reported that effect sizes were similar across groups defined by screening history (with first-time invitees and previous non-attenders being of particular interest to this review). Bankhead 2001 was a trial of recent non-attenders and so the results for the whole trial are directly relevant for this review.

Data from these two trials, which included a study on resource use by the practices, were used to populate a probabilistic decision analytic model, using a Markov Chain Monte Carlo simulation, to calculate the incremental cost-effectiveness ratio (ICER) for the cost per additional screen attended. The probabilistic element of the model allows the uncertainty in the parameters to be directly reflected in the results.

Costs were estimated from the perspective of the NHS and included: resources (including stationary, printing, staff time and on-costs) used to produce the letters, activate and retrieve the flags, and additional or longer GP consultations resulting from the interventions. A random sample (20% in Bankhead 2001, 10% in Richards 2001) was used to estimate number of consultations with duration estimated from encounter forms. The results are reported using prices in 1998-9. Detailed tables of costs associated with both interventions, and the sources of the estimates, are reported in Tables 1 to 3 of the published paper, with Tables 4 and 5 (reproduced below) summarising total estimated costs and additional attendances, respectively.

Figure 1 Total costs for an average size GP practice (reproduced from Brown 2006)

Table 4. The total additional cost of each intervention compared with none

	Trial 1 setting: all invited women ($n = 256$)		Trial 2 setting: non-attenders ($n = 89$)	
Intervention	Total NHS cost (£)	Total (%) GP cost (£)	Total NHS cost (£)	Total (%) GP cost (£)
Letter	690	174 (25.2)	143	61 (42.5)
Flag	1122	685 (61.0)	244	162 (66.5)
Both	1487	859 (57.7)	354	223 (63.0)

[1] CHEC list - Consensus Health Economic Criteria

Figure 2 Total additional attendances for an average size GP practice (reproduced from Brown 2006)

Table 5. The total additional number of attendances for screening compared with none

Intervention	Trial 1 setting: all invited women ($n = 256$)	Trial 2 setting: non-attenders ($n = 89$)
Letter	25.2	3.5
Flag	22.3	4.1
Both	29.9	5.8

The way these results are summarised in the text is not straightforward but the ICERs are easily calculated from the tables, with a little rounding error. The cost per screen attended is the (additional) cost of the intervention divided by the number of (additional) screens.

No underserved subgroups were reported for Richards 2011 (“trial 1”). For recent non-attenders (Bankhead 2011, “trial 2”), the ICER for the letter alone is £143/3.5 = £40.86 (reported as £40.92) and for the flag alone is £59.51 (not reported). For the combined intervention it is £61.03 compared to no intervention, £64.70 compared to flag alone and £91.74 compared to the letter alone (reported as £90.06).

Note that these are 1998-9 prices and would need to be replaced by modern costings. These trials explicitly excluded computerised practices and so these results may not be directly relevant today but at least some of the costs are easily updateable and the basis for calculating costs of additional or longer appointments may still be relevant.

Asaria 2015

Distributional cost-effectiveness analysis of health care programmes – a methodological case study of the UK bowel cancer screening programme

Using the BCSP as a case study, this paper proposes a method for "distributional cost-effectiveness analysis", for the allocation of (limited) resources which takes account of "intervention-generated inequality", a phenomenon associated with screening where overall population health may be improved while worsening health inequalities, due to different rates of uptake associated with socioeconomic status. The aim of the model is to assist decision-makers in balancing the twin aims of maximising population health while minimising health inequalities.

The CHEC list is not strictly relevant to this study as it does not model real-world data or produce real-world ICERs and so has not been completed for this paper. The model is described in more detail below.

Four different screening strategies are considered:

- no screening
- 'standard' screening as implemented at the establishment of the BCSP in 2006
- targeted reminders consisting of a GP-signed letter and tailored information sent only to the most deprived 40% as measured by IMD or with large South Asian populations, approximately half the total population invited for screening, with an assumed cost per person of £7 and increase in uptake of 12%
- universal reminders consisting of a GP-endorsed reminder sent to all non-responders, with an assumed cost of £3.50 per person and an estimated increase in uptake of 6%

Note that the costs and estimated increase in uptake for the two reminder strategies are chosen to assign equal total costs and produce equal impact on total screening uptake. These are constructed for the purposes of highlighting the trade-offs but are loosely based on previous empirical work.

The probabilistic model simulates one million 30 year olds through their lifetimes, with invitations for BCSP sent every two years from age 60-74. Baseline health and life expectancy is defined by level of socioeconomic deprivation, age and sex using Office of National Statistics (ONS) data from 2007 and health-related quality of life data by age and sex based on UK norms for EQ-5D[2] adjusted for socioeconomic status. No variation by ethnicity was included as relevant data were not available.

Opportunity costs due to use of funding that cannot be used elsewhere are assumed to be distributed equally across all population groups in the base case, valued at one quality-adjusted life year (QALY) per £20,000 (a figure commonly cited as the approximate ICER threshold for NICE). Sensitivity analysis at the extremes assumed that all of the opportunity cost falls on the healthiest and least healthy subgroups: women in the least deprived and ethnically diverse areas, and men living in the most

[2] EQ-5D instruments (EuroQoL)

deprived and ethnically diverse areas. The burden of disease was assumed to be equal across population subgroups, due to data limitations (and limited evidence that this variation is small).

Several different measures of inequality from the modelled health distributions were considered, ranging from simple relative or absolute changes across the population distribution to more sophisticated measures which summarise across the distribution and allow different assumptions about relative inequality aversion to be incorporated.

Social welfare indices are incorporated, trading off increases in mean population health with greater equality in the distribution of health, allowing outcomes to be ranked for different levels of "inequality aversion", with further refinements allowing for different social value judgements to be applied to different sources of variation in health to control for perceived 'fair' variation in health while accounting for those deemed 'unfair'.

We reproduce the base case results here for illustration but the empirical results are not directly applicable as the cost and effectiveness parameters of the model were constructed for illustrative purposes and are not based on real-world estimates. This paper has been reviewed here as a potentially useful framework for decision-analytic models where health inequalities are an important consideration alongside maximising the overall health of the population.

Figure 3 Base case cost-effectiveness results (reproduced from Asaria 2015)

Table I. Standard cost-effectiveness results

	Bowel cancer-related cost (£)	Life years	QALYs	Incremental net health benefit (QALYs)[a]
No screening	278 793 874	50 577 384	41 762 818	—
Standard screening	350 872 069	50 634 273	41 806 794	40 372
Screening + targeted reminder	400 936 962	50 639 192	41 810 506	41 581
Screening + universal reminder	385 268 692	50 639 452	41 810 784	42 642

Results based on a lifetime model for a cohort comprising of 1 million 30 year olds.
[a]Incremental to no screening.

Figure 4 Incremental population QALYs compared to no screening (a) and standard screening (b) (reproduced from Asaria 2015)

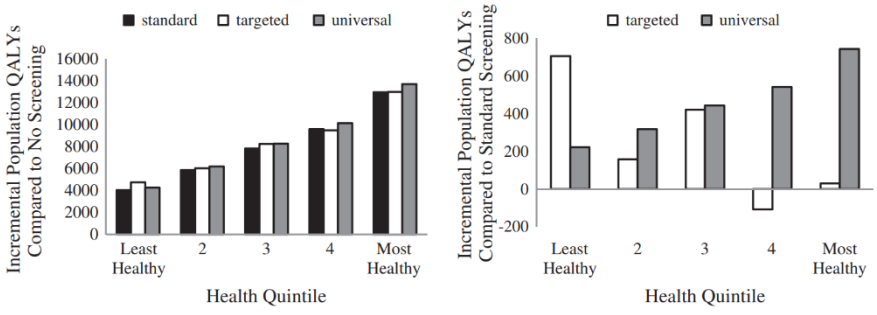

Figure 3. (a) Health compared with no screening (per million of population invited for screened). (b) Health compared with standard screening (per million of population invited for screening)

Jepson (2000)

Primary reference	Jepson et al, The determinants of screening uptake and interventions for increasing uptake: a systematic review		
Additional resources	inc Appendices (published as a separate document)		
Review questions	*(1) What factors (i.e. determinants) were associated with uptake of screening for different diseases?* [not reviewed in detail here] (2) What interventions were shown to increase uptake of screening programmes (or informed uptake) within populations?	**SPs included**	All screening regardless of application to the UK context (in 2000) in primary, secondary and tertiary prevention
Population (s)		**Intervention(s)**	All interventions in trials of screening uptake in any screening context
Study designs included	RCTs, quasi-RCTs, controlled trials reporting uptake	**Comparator(s)**	All identified, including opportunistic screening
Methods	Narrative synthesis (including forest plots to visualise results)	**Outcome(s)**	Primary: uptake Secondary: informed choice
Results	190 eligible trials, including 130 RCTs, reporting comparisons of uptake were identified. 65% were conducted in the USA or Canada, with different age ranges and screening intervals. Underserved groups were the focus of the review but were not addressed in detail within the uptake review. (Please note that the italicised text below is taken from the report summary with minor editing of punctuation): ***Promising interventions*** *included: invitation appointments, letters (less effective for mammography) and telephone calls; telephone counselling, and removal of financial barriers (eg transport and postage costs).* ***Interventions that may be effective*** *included: educational home visits; opportunistic screening; multicomponent community interventions; simpler procedures; combination of different components aimed at individuals; reminders for non-attenders (for mammography only); and invitation follow-up prompts.* ***Interventions that were found to have limited effectiveness*** *included: printed and audiovisual educational materials; educational sessions; risk-factor questionnaires; and face-to-face counselling.* ***Interventions that were shown to be ineffective*** *included: the use of rewards or incentives.* *There was either no good-quality evidence or insufficient evidence to evaluate the effectiveness of other interventions.* *Reminder interventions were found to be effective for physicians. Further interventions that may be effective included office systems or the use of audit and feedback to increase uptake. For physician education interventions there was insufficient good-quality evidence to assess their effectiveness. Of those interventions aimed at both physicians and individuals, a combination of physician reminders and patient invitations was found to be effective. When comparing interventions aimed at individuals with those for physicians, there was a small but beneficial effect for the interventions targeting individuals.*		

	*When assessing **informed uptake**, only four of the 190 intervention studies (all for antenatal screening) reported giving information on the risks and benefits of screening, and included knowledge as an outcome. Only one study evaluated the effect of this information and knowledge on the decision-making process. Whether informed uptake affects actual levels of uptake, therefore, has yet to be fully evaluated.* **Recommendations** include: a strong recommendation for the inclusion of informed uptake in trials of uptake and for a review of informed decision-making; further research into reducing barriers for minority ethnic groups; research into additional factors influencing uptake, and a plea for researchers to report all potential determinants of uptake, not just those which produce 'statistically significant' results.
Comments	This review has two parts, one review of the determinants of uptake and one review of interventions to improve uptake. Only the second review, of interventions, is reviewed in detail here (but the first review was used to assist us in identifying underserved groups for age and sex).

Quality assessment (RoBIS)

Domain 1: study eligibility criteria		Domain 2: identification & selection of studies		Domain 3: data collection & study appraisal		Domain 4: synthesis & findings	
RoBIS 1.1	Y	RoBIS 2.1	23 databases including grey literature, bibliographies and contact with specialists, with no language restrictions (46k abstracts screened) Y	RoBIS 3.1	Single data extraction checked by a second reviewer with disagreements resolved through discussion with a third Y	RoBIS 4.1	Where data reported Y
RoBIS 1.2	Y	RoBIS 2.2	As above Y	RoBIS 3.2	Around 20% of studies reported too little information for at least one quality domain assessed (Table 12); quality of most of the studies was difficult to ascertain and few reported the method of randomisation or explicitly that ITT had been used PN	RoBIS 4.2	Y
RoBIS 1.3	Y	RoBIS 2.3	Comprehensive search strategies, reproduced in the appendices Y	RoBIS 3.3	Only where available from published reports PN	RoBIS 4.3	Y
RoBIS 1.4	Eligible studies had to report uptake Y	RoBIS 2.4	No restrictions beyond reporting uptake Y	RoBIS 3.4	7 point checklist, reported in detail (not simplified to a score) Y	RoBIS 4.4	Tests for heterogeneity reported for any pooled results, random effects model used, and high prevalence of heterogeneity emphasised Y
RoBIS 1.5	No language restriction, grey literature included Y	RoBIS 2.5	One reviewer screened abstracts with a 5% random sample of included/excluded checked by a second	RoBIS 3.5	As above Y	RoBIS 4.5	No formal methods used but these would be of limited use with such a clinically heterogeneous group of studies; interpretation was sound

			reviewer; full papers screened by two reviewers Y				PY
						RoBIS 4.6	Carefully reported with unadjusted cluster trials de-emphasised in the plots Y
Concerns	Low	Concerns	Low	Concerns	Low	Concerns	Low
Rationale	Very sound methods.	**Rationale**	Some compromises on single-screening of title and abstracts (with double-screening of papers) but quality checks used and a reasonable compromise given the size of the review.	**Rationale**	They encountered the usual difficulties with poorly reported trials but handled this well, with cautious interpretation. They did not attempt to correct unadjusted cluster trials but excluded the confidence intervals from the plots to avoid presenting misleading results.	**Rationale**	This difficult review was very well done and appropriately interpreted.
RoB A	Y	RoB B	Y	RoB C	Y	**Overall**	Low
Comments	This review was an extraordinary undertaking covering a very large volume of evidence. Some compromises were made on single-screening of titles and abstracts and for single-data extraction (both with appropriate quality checks) but these are not critical concerns. It is very well reported with strong methods and interpretation.						

Sokal (2010)

Primary reference	Sokal, Rachel. 'A Critical Review of the Literature on the Uptake of Cervical and Breast Screening in British South Asian Women'. Quality in Primary Care 18, no. 4 (2010): 251–261.		
Additional resources			
Aims of review	To review the literature on uptake, and barriers to uptake, of cervical and breast cancer screening for British South Asian women.		
Population (s)	British South Asian women	**Intervention(s)**	Any
Study designs included	Any (no exclusion criteria defined)	**Comparator(s)**	Any
Methods	Narrative	**Outcome(s)**	Uptake (for review of interventions)
Results	Only 3 studies (1 RCT) found for uptake, Atri 1997 which is also included in our review. The cut-off date of 1996 excluded two of the studies we found, McAvoy 1991 and Hoare 1994. One single cohort study reported a set of interventions around translated materials, transport and link workers, another reported a community development programme around education and empowerment.		
Comments	This review reports on barriers to uptake as well as interventions to improve uptake; only the latter is reviewed here (and includes only Atri 1997).		

Quality assessment (RoBIS)

Domain 1: study eligibility criteria		**Domain 2: identification & selection of studies**		**Domain 3: data collection & study appraisal**		**Domain 4: synthesis & findings**	
RoBIS 1.1	No eligibility criteria stated beyond those implicit in the search strategy PY	RoBIS 2.1	Medline, EMBASE, BNI and CAB, date-limited from 1996 onwards and UK-only for relevance; no attempt to find unpublished or grey literature PN	RoBIS 3.1	Single author, no strategies reported PN	RoBIS 4.1	PY
RoBIS 1.2	PY	RoBIS 2.2	Reference lists and citation searches Y	RoBIS 3.2	Narrative is brief but includes useful considerations on study design and interpretation PY	RoBIS 4.2	None defined, aims were met PY
RoBIS 1.3	No explicit criteria PN	RoBIS 2.3	Not a particularly comprehensive set of search terms but likely to achieve reasonable coverage PN	RoBIS 3.3	No formal synthesis, no evidence that relevant studies were excluded PY	RoBIS 4.3	Narrative is fine Y

RoBIS 1.4	PY	RoBIS 2.4	PY	RoBIS 3.4	N	RoBIS 4.4	Not applicable NI
RoBIS 1.5	PY	RoBIS 2.5	Only one reviewer PN	RoBIS 3.5	Not applicable NI	RoBIS 4.5	Not applicable NI
						RoBIS 4.6	Addressed in narrative Y
Concerns	Unclear	**Concerns**	Unclear	**Concerns**	Unclear	**Concerns**	Low
Rationale	Eligibility criteria not stated separately from search strategy but approach appears reasonable given the aims	**Rationale**	The methods are limited but not inappropriate to the purpose of the review	**Rationale**	Useful summaries tabulated, no data extracted for analysis purposes and this is fine given the aims of the review; statistically naive in places but not terrible	**Rationale**	The narrative was brief but included useful comments on biases and interpretation
RoB A	PY	**RoB B**	Y	**RoB C**	PN	**Overall**	Unclear
Comments	This review does not claim to be a systematic review and is included because it has enough formal components of a systematic review (in particular, a reproducible search strategy across a reasonable number of sources) to be included here, given the focus on an under-researched underserved group. It is not high quality when judged against requirements for a systematic review but it is a useful piece of work which achieved the stated aims.						

Myers (2020)

Primary reference	Myers et al, 2019. Ways to use interventions to increase participation in mail-out bowel cancer screening: a systematic review and meta-analysis. TBM 2019 (now Myers L, Goodwin B, March S, et al. Ways to use interventions to increase participation in mail-out bowel cancer screening: a systematic review and meta-analysis. Transl Behav Med 2020; 10: 384–393)		
Additional resources			
Aims of review	To investigate effects of targeting interventions at particular subgroups and the potential benefit of combining interventions for particular subgroups.		
Population (s)	BCSP home test kits. Particularly focused on men, socioeconomically deprived and younger USGs, and on single vs combined interventions	**Intervention(s)**	Any (with a particular interest in combined interventions)
Study designs included	Not stated or defined in eligibility criteria (RCTs and CCTs in results)	**Comparator(s)**	Any
Methods	Meta-analysis (of pooled binarised subgroups, by age, sex and SES)	**Outcome(s)**	Uptake, by subgroups of age, sex and SES
Results	Results reported by subgroups of sex, age and SES for: • Pre-notification letters (2 RCTs) with substantial heterogeneity within subgroups for men, younger age and higher SES • GP endorsement (5 RCTs, one CCT) with substantial heterogeneity within all subgroups with Egger's test showing an influence of small studies. • Simplified testing (no dietary restrictions or fewer samples) (4 RCTs, 2 CCTs but 7 studies mentioned). These studies are not clinically homogeneous and the pooled results are especially uninterpretable. Substantial heterogeneity within all subgroups with Egger's test showing an influence of small studies. • Additional print material (7 RCTs) with substantial heterogeneity within all subgroups with Egger's test showing an influence of small studies. • Collection paper (1 RCT), only age and sex available, no interactions found. • Financial incentive (1 RCT), only sex and SES available, no interactions found. • Community drop-off location (1 CCT), only sex and age available, a weak interaction found for both. Only one strong subgroup effect was found, for pre-notification letters having a larger effect in men. Weak interactions by sex and age in the single (non-randomised) controlled study of community drop-off. Evidence for combining interventions (8 RCTs, 1CCT). Limited ability to pool given heterogeneity of combinations with a great deal of heterogeneity. Some evidence that additional printed material reduces uptake. These results are generally uninterpretable due to the pooling across trials despite substantial evidence of heterogeneity (which is accounted for but in no way resolved by the random effects model).		
Comments	Some strange methodological decisions: pooling subgroups across trials to produce a pooled subgroup effect, using the random effects model for clinically heterogeneous studies but the fixed effects model for clinically homogeneous (if there is no statistical heterogeneity both models will		

	produce the same result so this approach is not necessary or justifiable in the presence of statistical heterogeneity despite apparent clinical homogeneity). Pooling some very different interventions (eg no dietary restrictions with fewer samples required).

Quality assessment (RoBIS)

Domain 1: study eligibility criteria		Domain 2: identification & selection of studies		Domain 3: data collection & study appraisal		Domain 4: synthesis & findings	
RoBIS 1.1	Y	RoBIS 2.1	Pubmed, Scopus, PsycInfo, CINAHL, and ProQuest Dissertations and Theses Y	RoBIS 3.1	PY	RoBIS 4.1	PY
RoBIS 1.2	Y	RoBIS 2.2	Google Scholar and references Y	RoBIS 3.2	Y	RoBIS 4.2	Y
RoBIS 1.3	Y	RoBIS 2.3	Y	RoBIS 3.3	All reported results included and additional information from authors where available PY	RoBIS 4.3	Odd approach to analysis, pooling heterogeneous subgroups and testing interaction on the pooled data; no adjustment for cluster trials N
RoBIS 1.4	Y	RoBIS 2.4	Excluded studies which failed to report quantitative details of uptake despite aiming to improve it; unlikely to matter but this can introduce publication bias by excluding studies which could have, but did not, report key outcomes. English language only (not unreasonable given UK focus and BCSP context) PY	RoBIS 3.4	RoB and RoBINS-I Y	RoBIS 4.4	Yes, but results reported anyway Y
RoBIS 1.5	Y	RoBIS 2.5	NI	RoBIS 3.5	PY	RoBIS 4.5	Funnel plot and Eggers regression test used

							PY
						RoBIS 4.6	In the narrative
Concerns	Low	Concerns	Low	Concerns	Low	Concerns	High
Rationale		**Rationale**	No statement of who did what (or how many) but overall the methods appear reasonably sound	**Rationale**	Basic SR methods appear fine, reporting a little cursory in places	**Rationale**	The statistical approach is very strange and the results of the evidence synthesis uninterpretable.
RoB A	Y	RoB B	The included studies are fine but the analysis does not illuminate much Y	RoB C	N	**Overall**	High
Comments	The basic review methods are fine and reasonably well reported (including the appendices) but the approach to evidence synthesis is statistically naive and the results uninterpretable. The authors have included the data for cluster ransomised trials as unadjusted raw numbers. This is inappropriate for cluster randomised trials. There are too many errors in the methods to place any reliance on the results and conclusions.						

Brown 2006

	CHEC-list	YES	NO
1.	Is the study population clearly described?	Y	
2.	Are competing alternatives clearly described?	Y	
3.	Is a well-defined research question posed in answerable form?	Y	
4.	Is the economic study design appropriate to the stated objective?	Y	
5.	Is the chosen time horizon appropriate in order to include relevant costs and consequences?	Y	
6.	Is the actual perspective chosen appropriate?	Y	
7.	Are all important and relevant costs for each alternative identified?	Y	
8.	Are all costs measured appropriately in physical units?	Y	
9.	Are costs valued appropriately?	Y	
10.	Are all important and relevant outcomes for each alternative identified?	Y	
11.	Are all outcomes measured appropriately?	Y	
12.	Are outcomes valued appropriately?	Y	
13.	Is an incremental analysis of costs and outcomes of alternatives performed?	Y	
14.	Are all future costs and outcomes discounted appropriately?	Y	
15.	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	Y	
16.	Do the conclusions follow from the data reported?	Y	
17.	Does the study discuss the generalizability of the results to other settings and patient/client groups?		N
18.	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	Y	
19.	Are ethical and distributional issues discussed appropriately?		N