

Quality Assessment

Risk of Bias assessment results

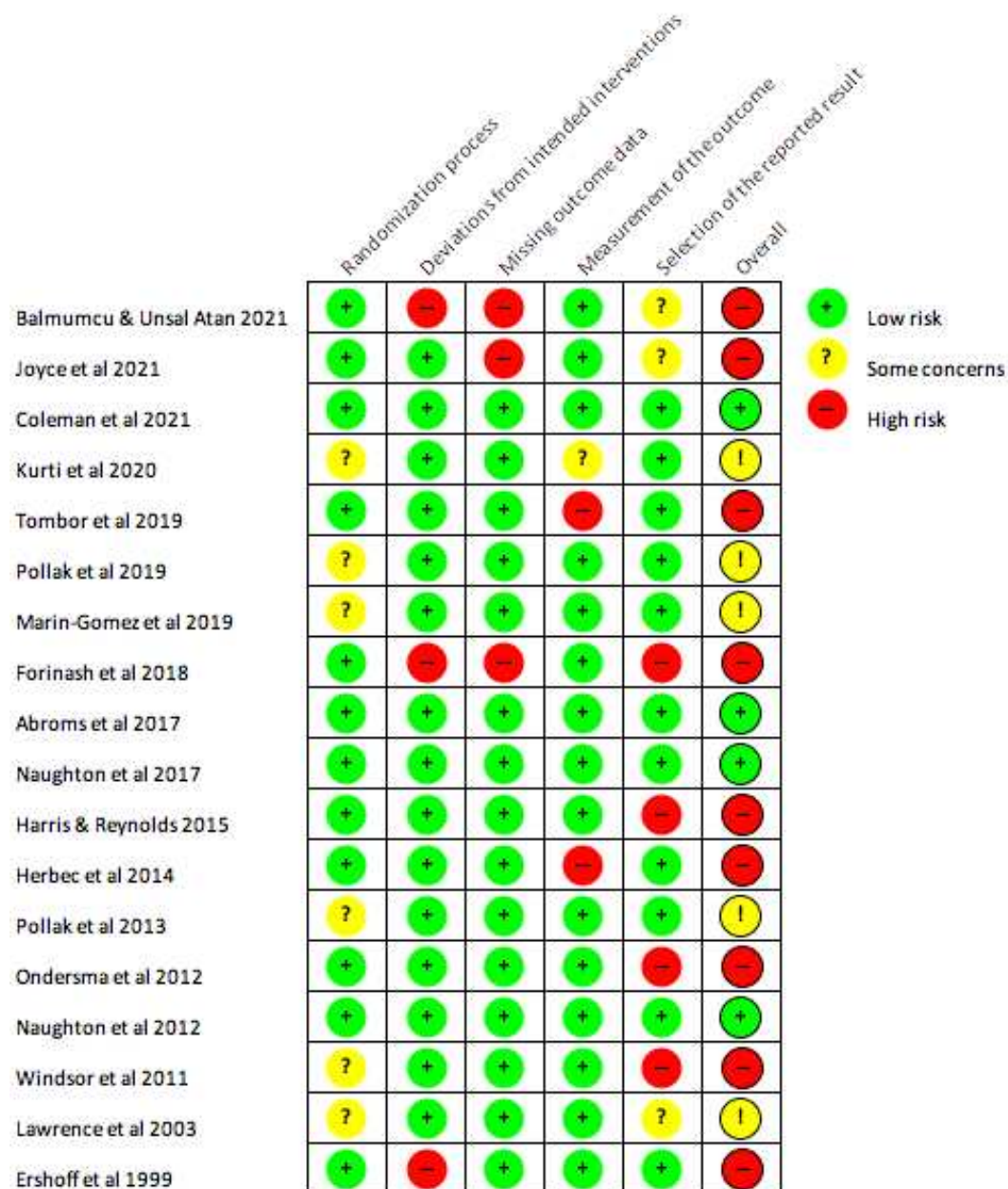


Figure S1: Risk of Bias in studies measuring Digital Health interventions to assist pregnant smokers quit.

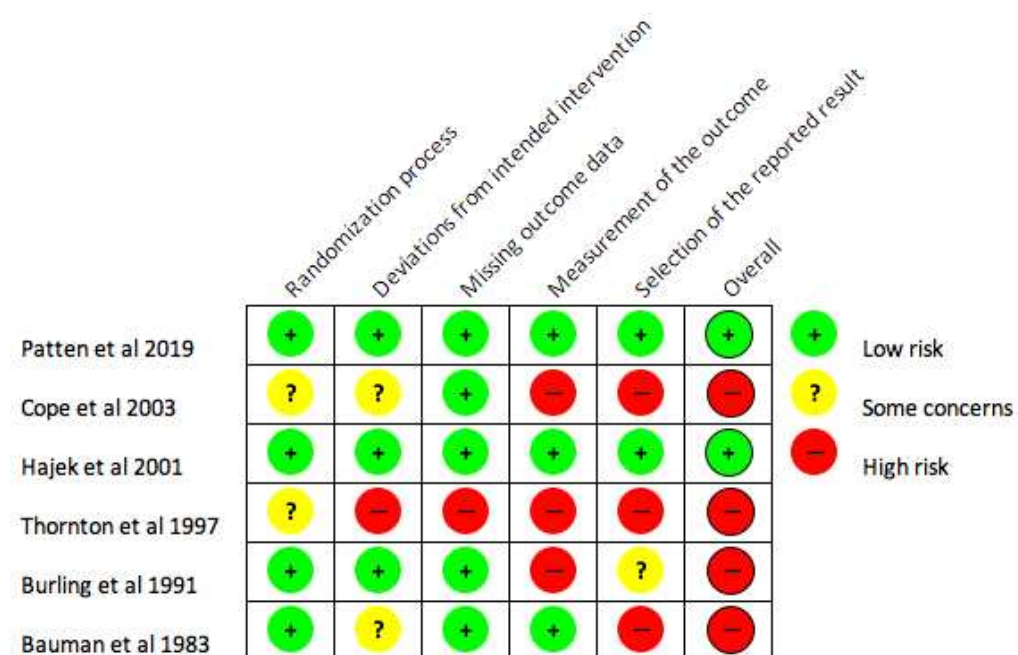


Figure S2: Risk of Bias in studies measuring Biomarker Feedback interventions to assist pregnant smokers quit.

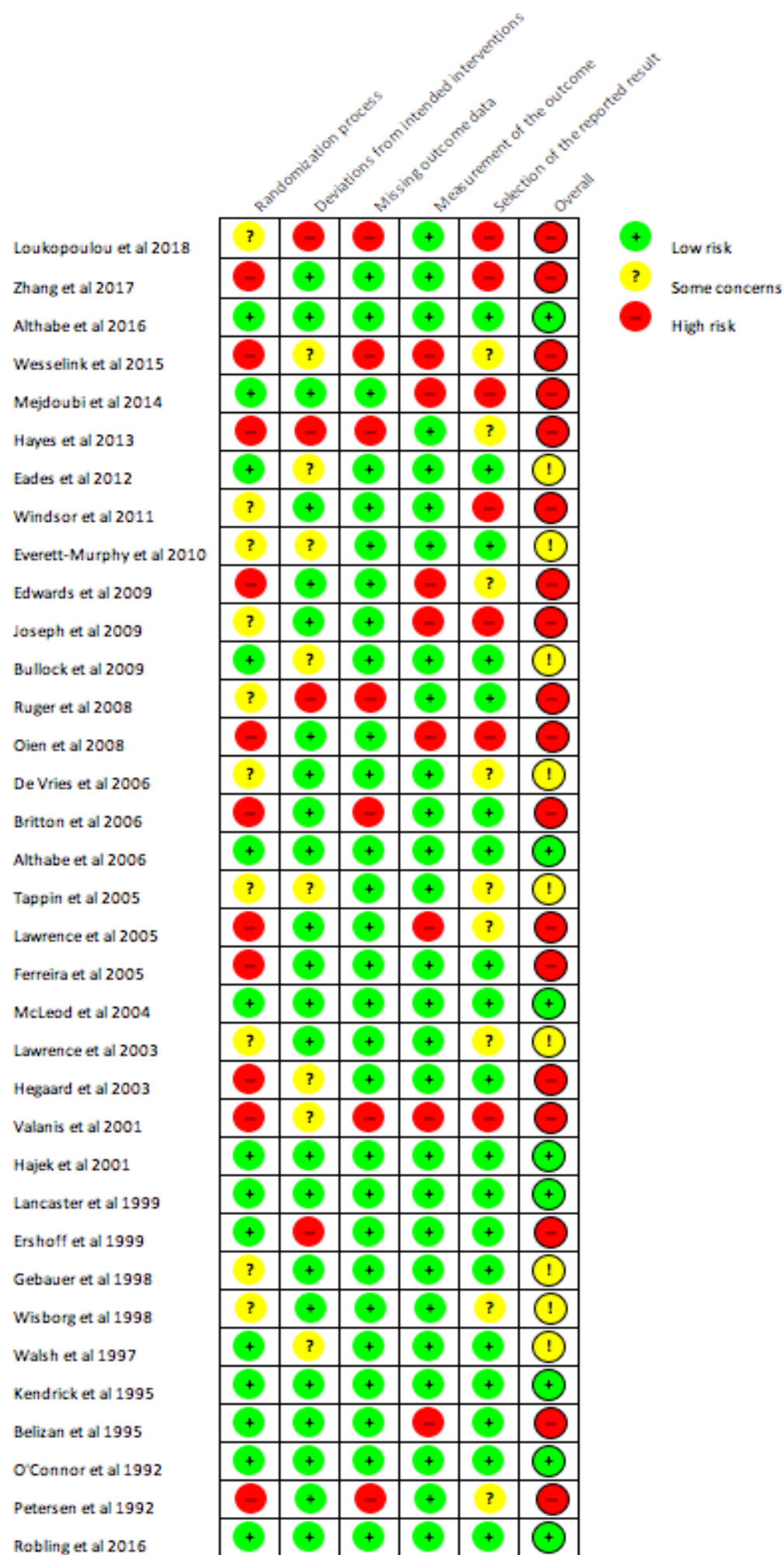


Figure S3: Risk of Bias in studies measuring Nurse or Midwife-led Counselling interventions to assist pregnant smokers quit.

Risk of Bias of included studies

In relation to the DH interventions, 4 studies were at low risk of bias, 5 had some concerns and 9 were at high risk. One study had 3 of the 5 fields rated at high risk due to lack of ITT analysis, potential missing outcomes and selective reporting, which may lead to exaggeration of the effect (89). Another study had 2 fields rated at high risk due to possible deviation from original intended intervention and missing outcome data(88). Further, 2 studies used self-reported abstinence with no biochemical confirmation potentially exaggerating the results and were therefore considered at high risk of bias (82, 96). Three studies did not have information to indicate if their analysis was prespecified(80, 85, 88). Bias concerns were mainly related to randomisation, blinding and allocation concealment including when measuring outcome by assessors (80, 83, 84, 86, 87).

Two of the 6 included studies for BF interventions were considered at low risk of bias, and 4 were considered at high risk. Three did not validate self-reported abstinence outcome biochemically and were at risk of selective reporting of results (95, 100, 102), and one also did not include ITT analysis and had missing outcome data (102).

Similarly, of 35 studies included for NoMC counselling intervention, only 8 were considered at low risk of bias in all fields. Nine studies had some concerns in relation to one or more fields, including blinding, concealment and allocation, application of ITT analysis or selective reporting. Eighteen studies were considered at high risk of bias, 5 due to non-validated self-reported abstinence (103, 119, 122, 125, 127), and 5 had at least 3 fields considered at high risk (103, 120, 125, 130).

Publication Bias Tests and Results

Digital Health interventions

A contour enhanced funnel plot (Figure S4) and Egger's regression test (Table S1) both indicated asymmetry which could be suggestive of publication bias. A trim and fill (Duval and Tweedie) test imputed 8 additional studies and predicted a combined lower overall effect size (Table S2). However, the P-Curve test (Figure S5) and related analysis (Table S3) are arguably more reliable to determine publication bias and estimate the true effect. The P-Curve test indicated the presence of an evidential value (i.e. a true effect) could not be confirmed in this meta-analysis, but equally its absence or a very small effect could not be ruled out either. This may indicate that not enough studies are available to determine if p-hacking or publication bias are pronounced enough and if a true effect is likely.

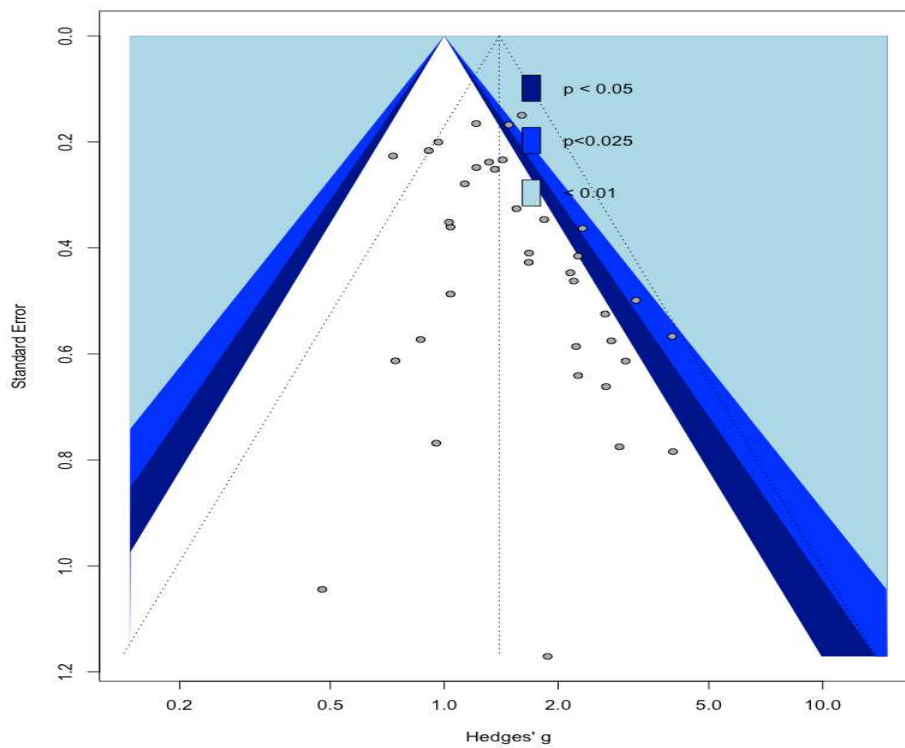


Figure S4: Contour enhanced funnel plot of Digital Health intervention studies to assist pregnant smokers achieve abstinence.

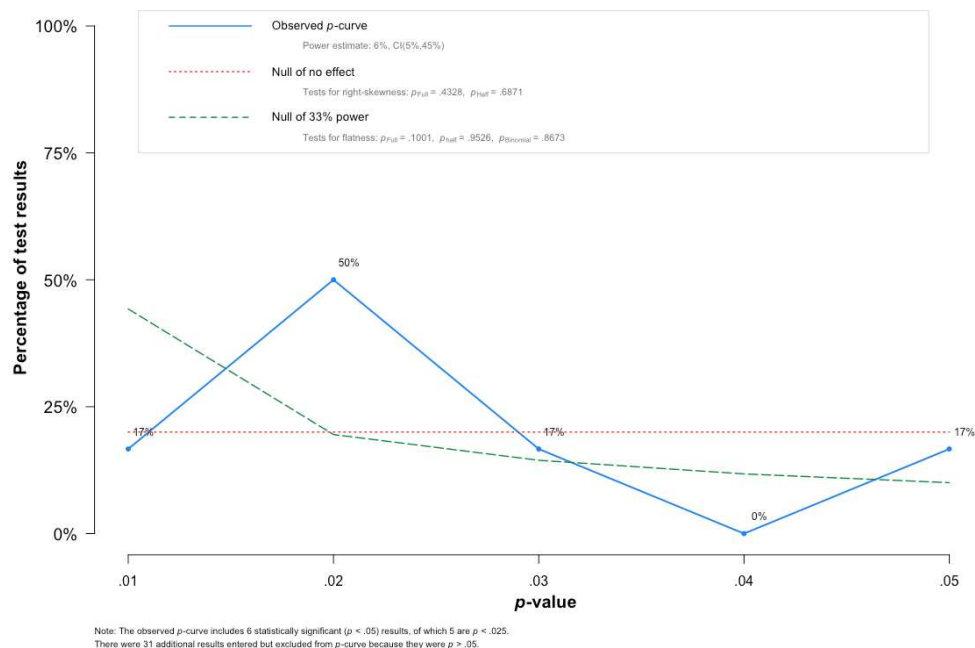


Figure S5: P-Curve test of Digital Health intervention studies to assess risk of publication bias.

Biomarker Feedback interventions

A contour enhanced funnel plot (Figure S6) may have indicated some asymmetry but Egger's regression test (Table S1) indicated no asymmetry and thus publication bias could not be suggested through these tests. A trim and fill (Duval and Tweedie) test imputed only 1 additional study and predicted a combined overall effect size (Table S2). However, the P-Curve test (Figure S7) and related analysis (Table S3) have indicated the presence of an evidential value, and that its absence was less likely. This may indicate that publication bias and p-hacking for this type of interventions may be less pronounced and a true value is more likely.

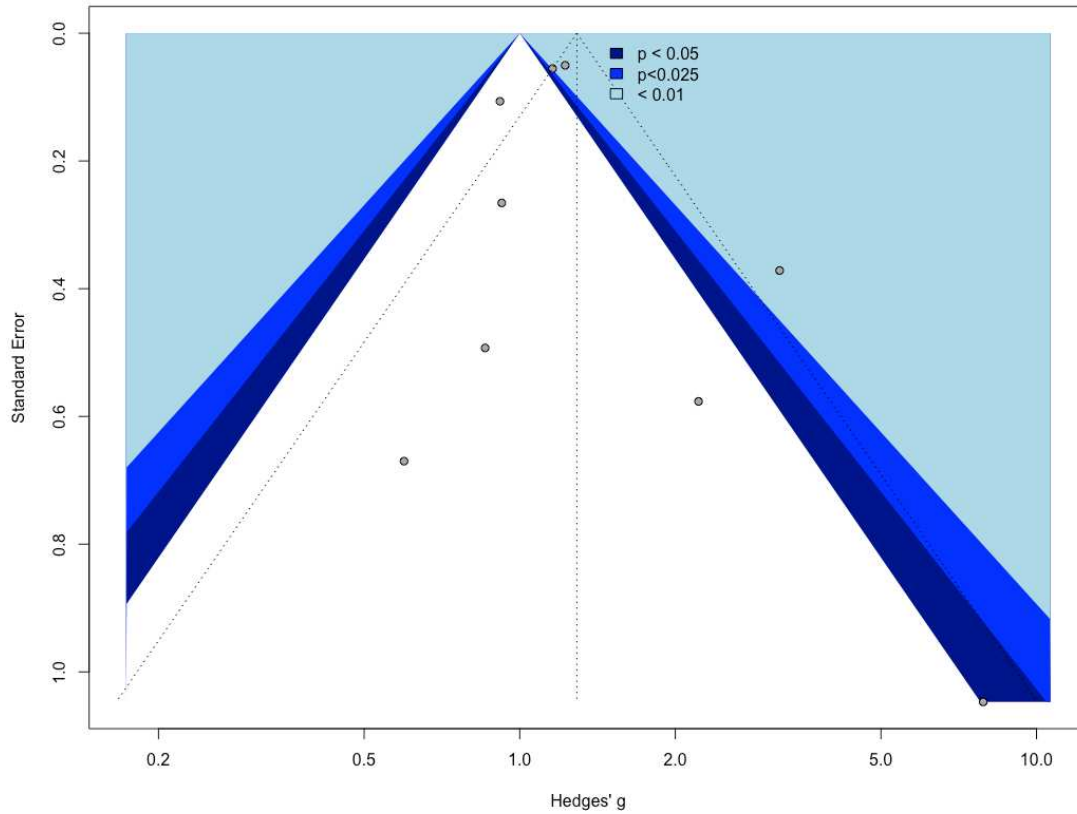


Figure S6: Contour enhanced funnel plot of Biomarker Feedback intervention studies to assist pregnant smokers achieve abstinence.

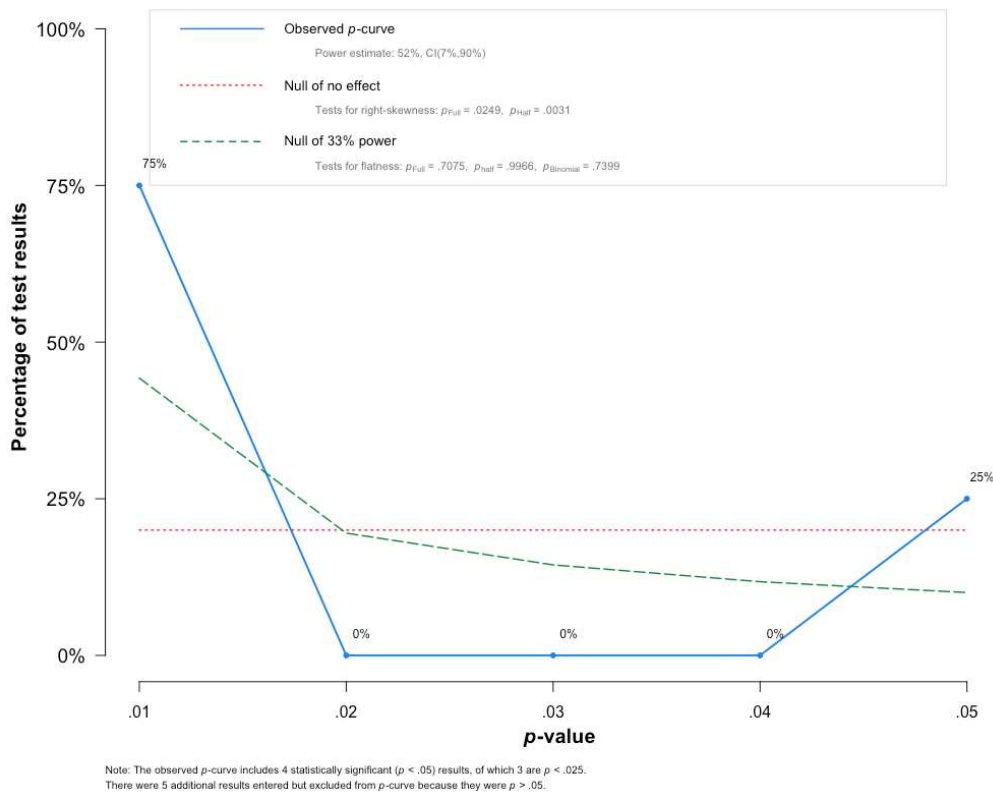


Figure S7: P-Curve test of Biomarker Feedback intervention studies to assess risk of publication bias.

Nurse or Midwife-led Counselling interventions

A contour enhanced funnel plot (Figure S8) and Egger’s regression test (Table S1) both indicated asymmetry which could be suggestive of publication bias. A trim and fill (Duval and Tweedie) imputed 20 studies and predicted a combined lower overall effect size (Table S2). However, the P-Curve test (Figure S9) and related analysis (Table S3) have indicated the presence of an evidential value, and that its absence is less likely. This may indicate that publication bias and p-hacking for this type of intervention may be less pronounced and a true value is more likely.

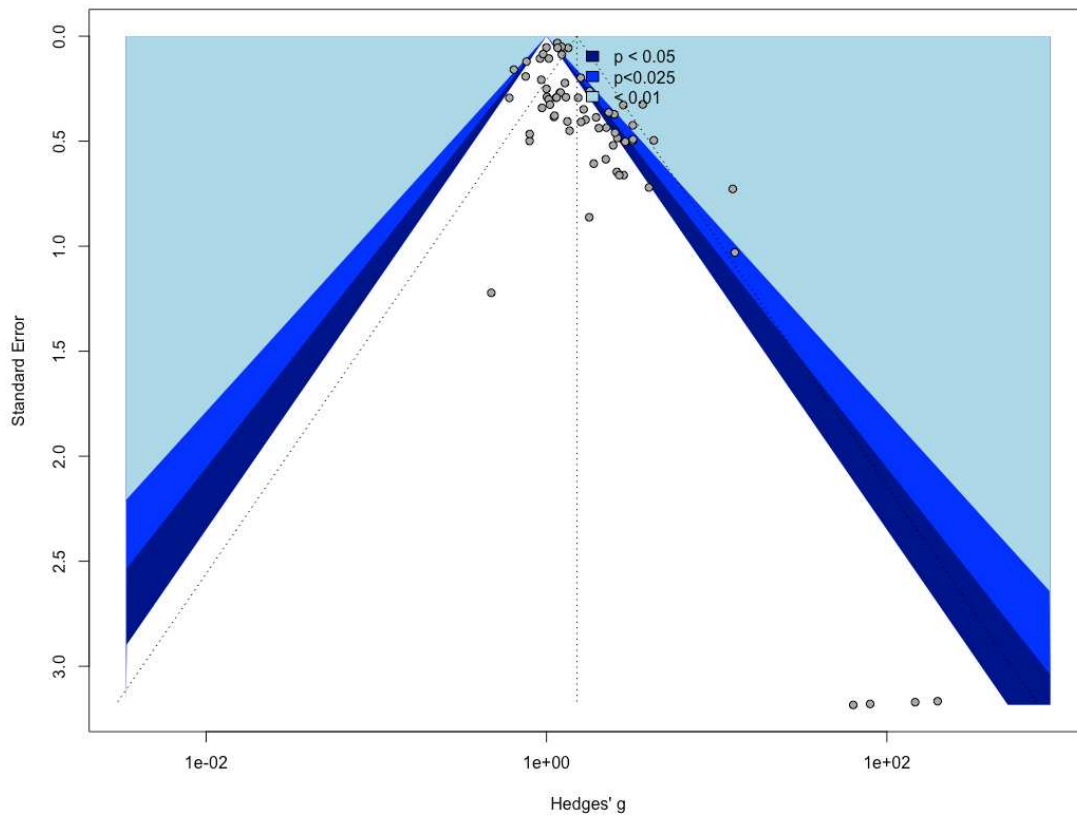


Figure S8: Contour enhanced funnel plot of Nurse or Midwife-led Counselling intervention studies to assist pregnant smokers achieve abstinence.

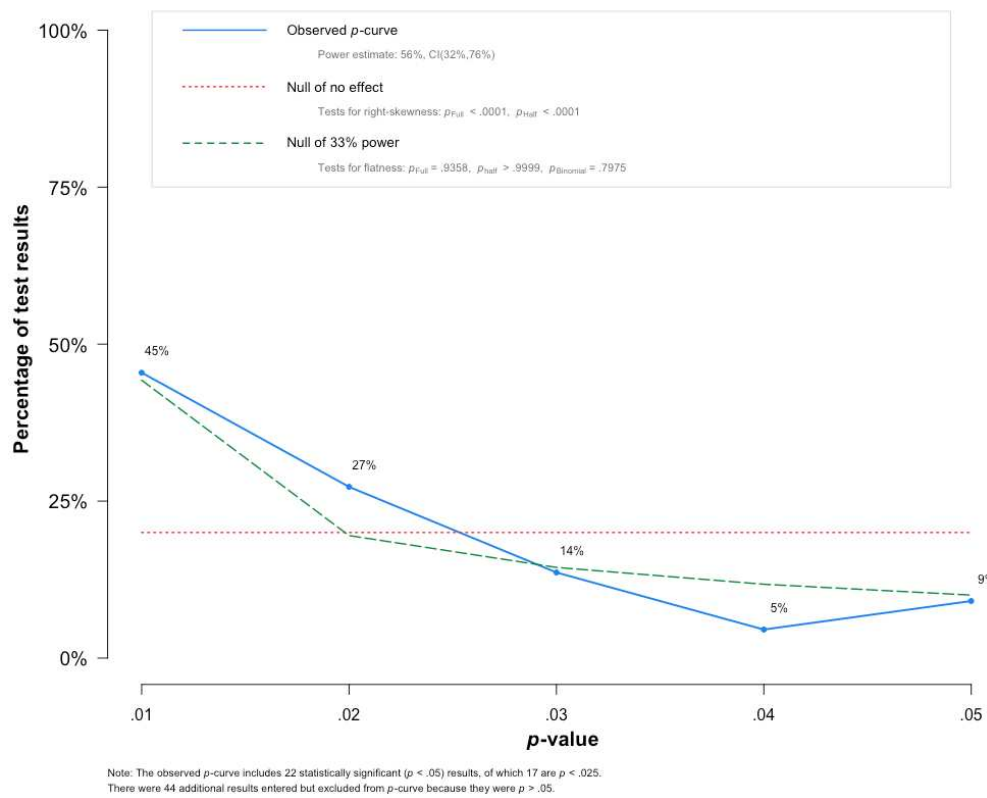


Figure S9: P-Curve test of Nurse or Midwife-led Counselling intervention studies to assess risk of publication bias.

Table S1: Egger's regression test for funnel plot asymmetry by intervention type.

Intervention type	Intercept point	Confidence interval	t value	P value	Evidence of funnel plot asymmetry
Digital health	0.918	0.17 – 1.66	2.417	0.02153944	Yes
Biomarker feedback	0.354	-1.15 – 1.86	0.461	0.658874	No*

Nurse or midwife-led counselling 0.918 0.47 – 1.41 3.923 0.0002158727 Yes

*The meta-analysis contains k = 9 studies. Egger's test may lack the statistical power to detect bias when the number of studies is small (i.e., k<10).

Table S2: Duval and Tweedie's trim and fill test using random effects model for funnel plot asymmetry by intervention type.

	Estimated RR	Confidence interval (95%)	Z value	P value	Imputed studies (k)
Digital health	1.3261	1.0861-16192	2.77	0.0056	8
Biomarker feedback	1.2020	0.6962-2.0753	0.66	0.5090	1
Nurse or midwife-led counselling	1.1819	0.9304-1.5014	1.37	0.1710	20

Table S3: Evidential value testing through p-curve testing of right-skewness and flatness by intervention type.

	pBinomial	Z Full	P Full	Z Half	P Half	Evidential Value present	Evidential value absent/inadequate
<i>Digital health intervention</i>							
Right-skewness test	0.109	-0.169	0.433	0.488	0.687	No	
Flatness test	0.867	-1.281	0.100	1.670	0.953		No
<i>Biomarker feedback intervention</i>							
Right-skewness test	0.312	-1.962	0.025	-2.737	0.003	Yes	
Flatness test	0.740	0.546	0.707	2.706	0.997		No
<i>Nurse or midwife-led counselling intervention</i>							
Right-skewness test	0.008	-4.835	0.000	-3.995	0	Yes	
Flatness test	0.798	1.521	0.936	5.543	1		No

Table S4: Quantitative estimate of overall statistical heterogeneity for all outcomes by intervention type.

	Q-value	DF (Q)	P-value	τ^2	95% CI (τ^2)%	I ² %	95% CI (I ²)%
Digital health intervention	62.26	41	0.0177	0.22	0.003-3.321	34	3.8-54.9
Biomarker feedback Intervention	23.48	9	0.0052	0.56	0.013-2.613	62	23.7-80.7
Nurse or midwife-led counselling intervention	268.29	85	<0.0001	0.99	0.242-0.825	68	60.5-74.6