Appendix 1: Supplementary Methods [posted as supplied by author]

Pubmed Search algorithm

#1 endometrial neoplasm [MeSH]
#2 malign* [tiab] OR cancer*[tiab] OR carcinoma*[tiab] OR tumor*[tiab] OR tumour*[tiab]
#3 endometr* [tiab] OR corpus uteri [tiab] OR uterine [tiab]
#4 #2 AND #3
#5 #1 AND # 4

#6 Ovarian Neoplasms [MeSH]
#7 Ovar* AND (cancer* OR carcinoma* OR neoplasm* OR tumor* OR tumour* OR adenocarcinoma* OR Endometrioid carcinoma* OR cystadenoma* OR cystadenocarcinoma* OR adenoma*)
#8 Androblastom* OR arrhenoblastoma* OR sertoli leydig OR Brenner OR granulosa cell tumor* OR granulosa cell tumour* OR luteoma* OR luteinoma*
#9 #6 OR #7 OR #8

#10 Cervical neoplasms [MeSH]
#11 Cervi* AND (cancer* OR carcinoma* OR neoplasm* OR tumor* OR tumour* OR adenocarcinoma* OR squamous cell carcinoma* OR carcinoma OR carcinosarcoma*)
#12 CIN OR cervical intraepithelial neoplasia OR cervical dysplasia OR cervical precancer*
#13 #10 OR #11 OR #12

#14 Pregnancy [mh] OR Pregnant Women [mh] OR pregnan*[tiab] OR Parturition [mh] OR parturi* [tiab] OR gestation* [tiab] OR Gravidity [mh] OR gravid* [tiab] OR maternal* [tiab] OR puerperium [mh] OR puerperi* [tiab] OR postpartum period [mh] OR postpartum* [tiab] OR pregnancy complications [mh] OR pregnancy outcome [mh]

#15 Polycystic ovary syndrome [mh] OR polycystic ovar*[tiab] OR pco [tiab] OR pcos [tiab] OR pcod [tiab]

#16 fertility[mh] OR infertility [mh] OR Abortion, Spontaneous[mh] OR miscarriage [tiab] OR extrauterine pregnancy [tiab] OR ectopic pregnancy [tiab] OR abortion, induced[mh] OR termination of pregnancy[tiab] OR molar pregnancy[tiab] OR IVF[tiab] OR ICSI[tiab] OR insemination[tiab] OR assisted reproduction[tiab] OR pelvic floor disorders[tiab] OR urinary incontinence[tiab] OR fecal incontinence[tiab] OR pelvic organ prolapse[tiab] OR uterine prolapse[tiab] OR vaginal prolapse[tiab] OR pelvic floor defect[tiab] OR pelvic floor*[tiab] OR vaginal wall*[tiab] OR Reproductive Techniques [mh] OR Genital Disease, Female [mh]OR Pelvic Floor Diseases [mh]

#17 menorrhagia[mh] OR metrorragia[mh] OR menstrual cycle[tiab] OR menstrual disorder*[tiab] OR heavy menstrual bleeding[tiab] OR menstrual bleeding[tiab] OR menstrual pain[tiab] OR menstrual cycle*[tiab]

#18 contracept*[tiab] OR Norpregnanes[mh] OR “Contraceptive Agents” [mh] OR “Contraceptive Agents” [Pharmacological Action] OR “Contraceptive Devices” [mh] OR “Contraception” [mh] OR iud [tiab] OR “Intrauterine Devices” [mh] OR emergency contraception[tiab] OR “Contraception, Postcoital” [mh] OR “Contraceptive Agents” [mh] OR nuvaring [tiab] OR “Desogestrel”[mh] OR “Contraceptive Agents, Female” [mh] OR “Contraceptive Devices, Female” [mh] OR hormonal patch[tiab] OR ortho evra [tiab] OR “Norgestrel” [mh] OR “Contraceptive Devices, Female” [mh]) OR “Contraceptive Agents, Female” [mh]

#19: #5 OR #9 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18

#20 obesity[tiab] OR Obese[tiab] OR overweight[tiab] OR over-weight[tiab] OR BMI [tiab] OR Body Mass index[tiab] OR bariatric surgery[tiab] OR hip circumference[tiab] OR waist circumference[tiab] OR body size[MeSH Terms] OR body size [tiab] OR weight gain[tiab] OR waist to hip ratio[tiab] OR weight[tiab] OR body fat*[tiab] OR anthropometry[tiab] OR diabetes [MeSH]

#21 #19 AND #20

#22 animal[mh] NOT human[mh]

#23 #21 NOT #22

#24 Meta-Analysis[ptyp] OR systematic[sb]

#25 #23 AND #24

Below additionally only to identify meta-analyses of RCTs (651 hits 15.7.2016):

#26 Weight Loss OR Diet, Fat-Restricted OR Diet OR Diet, Protein-Restricted OR Diet, Carbohydrate-Restricted OR Diet [mp] OR Diet, Reducing OR Diet Therapy OR Life Style OR Health Education OR Exercise [mp] OR Exercise OR Exercise, Therapy

#27 #25 AND #26

Cochrane database of systematic reviews

#1"obesity" or "body weight" or "body constitution"
#2 [mh obstetrics] or [mh gynecology]

#3 endometrial or cervical or ovari or birth or maternal or gestati or pregnancy or fertility or menstrual or polycystic or hrt or contracepti

#4 #2 or #3

#5 #4 and #1

in: Cochrane reviews and other published systematic reviews only.

Data extraction

From each eligible meta-analysis of observational studies, we extracted the following data: name of first author and the year of publication, the outcome studied, and the type of exposure; Body Mass Index (BMI), weight, weight gain, gestational weight gain (GWG), waist to hip-ratio (WHR), waist circumference (WC), hip circumference (HC) or bariatric surgery. From each individual study in a meta-analysis, we extracted the first author and the publication year; the epidemiological design (case-control, cohort, cross-sectional, case

series); the number of cases and controls in case-control studies or the number of cases and population or person years in cohort studies; the maximally adjusted relative risk (odds ratio for case-control studies and risk ratio for cohort studies), and the 95% confidence intervals. From each included meta-analysis of interventional studies we extracted the name of first author, publication year, intervention type, comparison group, description of study subjects, the outcome studied, N of participants and events, type of metric and meta-analysis model used, N of included studies, GRADE estimate, and the summary effect estimate and respective confidence interval.

Two investigators (IK, MK) independently performed the literature search, assessed the eligibility of the retrieved papers and performed the data extraction. The two authors then compared the results and disagreements were resolved by discussion. If required, a consensus was reached with the involvement of a third investigator (KT).

Data analysis

Meta-analyses of observational studies

Assessment of summary effect

For each exposure and outcome pair, we calculated the summary effect and the 95% confidence interval using both fixed and inverse variance weighted random effects methods.[1]

Assessment of heterogeneity

The heterogeneity between studies was assessed with the Cochran Q test [2] and the I^2 metric of inconsistency [3]. The I^2 metric ranges between 0% and 100% and could reflect either genuine diversity within the studies, chance or bias, and its 95% confidence intervals were calculated according to the method of Ioannidis and colleagues.[4]

Prediction intervals

We further calculated the 95% prediction intervals for the summary random effect estimates. This estimate further accounts for inter-study heterogeneity enabling direct comparison with clinically relevant thresholds, and has been described as the best way to describe this uncertainty.[5,6] This is different from the confidence interval, which describes the uncertainty around the summary effect size, whereas the prediction interval represents the range where the effect estimate of a possible future study would lie.

Assessment of small study effects

We further examined whether smaller studies gave higher effect estimates than larger studies as an indication of publication, reporting bias or a reflection of other causes of heterogeneity amongst small and large studies. We used the P-value from the Egger's regression asymmetry test[7] (P<0.10) where the random effects summary estimate was further away from the null-value compared to the point estimate of the largest study in a meta-analysis.

Evaluation of excess significance

The excess significance test is a test used to evaluate whether an excess of studies with nominally statistically significant findings (P<0.05) exist in the studies included in each meta-analysis. We used this test to assess whether the observed number of studies with nominally significant results (O) was different from the expected number of studies with significant results (E).[8] The expected number of significant studies in each meta-analysis was calculated from the sum of the statistical power estimates for each component study based on a non-central t-distribution.[8,9] Since the true effect size is not known, the effect of the largest study in a meta-analysis i.e., the study with the smallest standard error, was chosen as the plausible effect size.[9,10] Sensitivity analysis was also conducted using as alternative plausible effect sizes the summary estimates from the fixed and the random effects models. Excess significance for individual meta-analyses was determined as P<0.10 (one sided P<0.05 with N of observed > N of expected).[8]

Credibility ceilings

Every single observational study has a probability *c* (credibility ceiling) that the true effect size of the study is in different direction from the one suggested by the point estimate. Therefore the greatest certainty that each observational study reports the true effect size is (100-c)%. In order to exceed that certainty, more studies should demonstrate the same direction of effect. The sensitivity analysis that uses the 'credibility ceilings' aims to account for these methodological limitations of the observational studies that may lead to spurious results by inflating the variance while the point estimate of a single study stays unaltered.[11] The summary effect sizes and the between-study heterogeneities were re-estimated using these inflated variances with wide range of values.[11,12] For each meta-analysis and given a ceiling of c%, the likelihood ratio of the real effect size being in the direction indicated by the summary estimate for the corresponding unit increase (or level) of exposure was computed.[11]

Meta-analyses of interventional studies

When we identified more than one meta-analysis of controlled trials examining the association between exercise or dietary intervention and a given outcome, we examined the

results for concordance in terms of the direction, level of statistical significance (at P≤0.05), and magnitude (overlapping confidence interval) between the included meta-analyses. When meta-analyses for a given outcome existed for meta-analyses of both observational and interventional studies, we compared the results in terms of direction of the effect, the nominal significance of the effect (p<0.05), and the grade of evidence.

References

1. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials* 1986; **7**(3): 177-88.
2. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954; **10**: 101-29.
3. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**(7414): 557-60.
4. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ* 2007; **335**(7626): 914-6.
5. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549.
6. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ open* 2016; **6**(7): e010247.
7. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**(7109): 629-34.
8. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials* 2007; **4**(3): 245-53.
9. Tsilidis KK, Papatheodorou SI, Evangelou E, Ioannidis JP. Evaluation of excess statistical significance in meta-analyses of 98 biomarker associations with cancer risk. *J Natl Cancer Inst* 2012; **104**(24): 1867-78.
10. Tsilidis KK, Panagiotou OA, Sena ES, et al. Evaluation of excess significance bias in animal studies of neurological diseases. *PLoS Biol* 2013; **11**(7): e1001609.
11. Salanti G, Ioannidis JP. Synthesis of observational studies should consider credibility ceilings. *J Clin Epidemiol* 2009; **62**(2): 115-22.
12. Papatheodorou SI, Tsilidis KK, Evangelou E, Ioannidis JP. Application of credibility ceilings probes the robustness of meta-analyses of biomarkers and cancer risk. *J Clin Epidemiol* 2015; **68**(2): 163-74.