Supplementary Online Content

Viswanathan M, Treiman KA, Kish-Doto J, Middleton JC, Coker-Schwimmer EJL, Nicholson WK. Folic acid supplementation for the prevention of neural tube defects: an evidence report and systematic review for the US Preventive Services Task Force. *JAMA*. doi:10.1001/jama.2016.19193

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Systematic Review Literature Search Strategies

PubMed Benefits Search

Search	Query
#1	Search (Pregnancy[MeSH] OR pregnancy[tw] OR pregnant[tw] OR "pregnant women"[MeSH])
#2	Search (("folic acid"[MeSH] OR "vitamin b9"[tw] OR "vitamin m"[tw] or "Pteroylglutamic Acid"[tw] OR "folvite"[tw] OR "folacin"[tw] OR "folate"[tw] OR "folic acid"[tw] OR "5-Me-THF" OR "5-Me-H4F" OR "5-methyltetrahydrofolate" OR "5-methyltetrahydropteroylpentaglutamate"[Supplementary Concept] OR "5-methyltetrahydrofolate triglutamate"[Supplementary Concept]))
#3	Search ((multivitamin[all fields] OR "prenatal vitamin"[all fields] OR multivitamins[all fields] OR "prenatal vitamins"[all fields] OR "vitamin supplement"[all fields] OR "vitamin supplements"[all fields]))
#4	Search (("neural tube defects"[MeSH Terms] OR "spina bifida"[All Fields] OR "neural tube damage"[All Fields] OR "neural tube defect"[All Fields] OR "neural tube defects"[All Fields] OR "neural tube disorders"[All Fields] OR "Neural tube defect, folate-sensitive"[Supplementary Concept] OR Craniorachischisis[tw] OR Craniorachischises[tw] OR Diastematomyelia[tw] OR Diastematomyelias[tw] OR "Tethered Cord Syndrome"[tw] OR "Tethered Cord Syndromes"[tw] OR "Occult Spinal Dysraphism Sequence"[tw] OR "Tethered Spinal Cord Syndrome"[tw] OR "Occult Spinal Dysraphism"[tw] OR "Occult Spinal Dysraphisms"[tw] OR Iniencephaly[tw] OR Iniencephalies[tw] OR "Neurenteric Cyst"[tw] OR "Neurenteric Cysts"[tw] OR "Neuroenteric Cyst"[tw] OR "Neuroenteric Cysts"[tw] OR "Spinal Cord Myelodysplasia"[tw] OR "Spinal Cord Myelodysplasias"[tw] OR Acrania[tw] OR Acranias[tw] OR Exencephaly[tw] OR Exencephalies[tw]))
#5	Search ((#1 AND (#2 OR #3) AND #4))
#6	Search ((#1 AND (#2 OR #3) AND #4)) Filters: Humans
#7	Search ((#1 AND (#2 OR #3) AND #4)) Filters: Other Animals
#8	Search ((#7 NOT #6))
#9	Search ((#5 NOT #8))
#10	Search (("retraction"[All Fields] OR "Retracted Publication"[pt] OR Duplicate Publication [PT] OR Erratum[All Fields]))
#11	Search ((#9 and #10))

PubMed Harms Search

Search	Query
#1	Search (("folic acid"[mesh] OR "folic acid"[tiab] OR "folvite"[tiab] OR "folacin"[tiab] OR 5-Me-THF OR 5-Me-H4F OR 5-methyltetrahydrofolate OR "5-methyltetrahydropteroylpentaglutamate" [Supplementary Concept] OR "5-methyltetrahydrofolate triglutamate" [Supplementary Concept]))
#2	Search ((Pregnancy[MeSH] OR pregnancy[tw] OR pregnant[tw] OR "pregnant women"[MeSH]))
#3	Search ((#1 and #2))
#4	Search (((("Colorectal Neoplasms"[Mesh] OR ("Vitamin B 12 Deficiency"[Mesh] OR

	(“vitamin b 12”[MeSH Terms] AND deficien*[Text Word])) OR (“Vitamin B 6 Deficiency”[Mesh] OR (“vitamin b 6”[MeSH Terms] AND deficien*[Text Word])) OR (“Drug-Related Side Effects and Adverse Reactions”[Majr] OR “Patient Harm”[Majr] OR harm[tiab] OR harms[tiab] OR “adverse effects” [Subheading] OR “adverse effect”[tiab] OR “adverse effects”[tiab] OR “adverse event”[tiab] OR “adverse events”[tiab] OR complication[tiab] OR complications[tiab]) OR “Twins”[mesh] OR “Pregnancy, Twin”[mesh] OR twinning OR twins OR asthma[Mesh] OR asthma)))
#5	Search ((#3 and #4))
#6	Search ((#3 and #4)) Filters: Humans
#7	Search ((#3 and #4)) Filters: Other Animals
#8	Search ((#7 NOT #6))
#9	Search ((#5 not #8))
#11	Search ((“retraction”[All Fields] OR “Retracted Publication”[pt] OR Duplicate Publication [PT] OR Erratum[All Fields]))
#12	Search ((#5 and #11))

Cochrane Benefits Search

ID	Search
#1	[mh Pregnancy] or pregnancy or pregnant or [mh “pregnant women”]
#2	[mh “folic acid”] or “vitamin b9” or “vitamin m” or “Pteroylglutamic Acid” or “folvite” or “folacin” or “folate” or “folic acid” or 5-Me-THF or 5-Me-H4F or 5-methyltetrahydrofolate 4051
#3	Multivitamin or “prenatal vitamin” or multivitamins or “prenatal vitamins” or “vitamin supplement” or “vitamin supplements”
#4	[mh “neural tube defects”] or “spina bifida” or “neural tube damage” or “neural tube defect” or “neural tube defects” or “neural tube disorders” or “Neural tube defect, folate-sensitive” or Craniorachischisis or Craniorachischises or Diastematomyelia or Diastematomyelias or “Tethered Cord Syndrome” or “Tethered Cord Syndromes” or “Occult Spinal Dysraphism Sequence” or “Tethered Spinal Cord Syndrome” or “Occult Spinal Dysraphism” or “Occult Spinal Dysraphisms” or Iniencephaly or Iniencephalies or “Neurenteric Cyst” or “Neurenteric Cysts” or “Neuroenteric Cyst” or “Neuroenteric Cysts” or “Spinal Cord Myelodysplasia” or “Spinal Cord Myelodysplasias” or Acrania or Acranias or Exencephaly or Exencephalies

Cochrane Harms Search

#1	[mh “folic acid”] or “folic acid” or folvite or folacin 5-Me-THF or 5-Me-H4F or 5-methyltetrahydrofolate or “5-methyltetrahydropteroylpentaglutamate” or “5-methyltetrahydrofolate triglutamate”
#2	[mh Pregnancy] or pregnancy or pregnant or [mh “pregnant women”]
#3	#1 and #2
#4	[mh Twins] or [mh “Pregnancy, Twin”] or twinning or twins or [mh “Colorectal Neoplasms”] or ([mh “Vitamin B 12 Deficiency”] or ([mh “vitamin b 12”] and deficien*)) or ([mh “Vitamin B 6 Deficiency”] or ([mh “vitamin b 6”] and deficien*)) or ([mh “Drug-Related Side Effects and Adverse Reactions” [mj]] or [mh “Patient Harm” [mj]] or harm or harms or [mh /AE] or “adverse effect” or “adverse effects” or

	“adverse event” or “adverse events” or complication or complications or [mh Asthma] or asthma)

EMBASE Benefits Search

#9
#8
#4 NOT #7
#7
#6 NOT #5
#6
#4 AND [animals]/lim
#5
#4 AND [humans]/lim
#4
#1 AND #2 AND #3
#3
‘spinal dysraphism’/exp OR ‘spinal dysraphism’ OR ‘neural tube damage’ OR ‘neural tube defect’/exp OR ‘neural tube defect’ OR ‘neural tube defects’/exp OR ‘neural tube defects’ OR ‘neural tube disorders’ OR ‘neural tube defect, folate-sensitive’ OR craniorachischisis OR craniorachischises OR ‘diastematomyelia’ OR ‘diastematomyelia’/exp OR diastematomyelia OR diastematomyelias OR ‘tethered cord syndrome’/exp OR ‘tethered cord syndrome’ OR ‘tethered cord syndromes’ OR ‘occult spinal dysraphism sequence’ OR ‘tethered spinal cord syndrome’ OR ‘occult spinal dysraphism’/exp OR ‘occult spinal dysraphism’ OR ‘occult spinal dysraphisms’ OR ‘iniencephaly’ OR ‘iniencephaly’/exp OR iniencephaly OR iniencephalies OR ‘neurenteric cyst’ OR ‘neurenteric cysts’ OR ‘neuroenteric cyst’ OR ‘neuroenteric cysts’ OR ‘spinal cord myelodysplasia’ OR ‘spinal cord myelodysplasias’ OR acrania OR acranias OR ‘exencephaly’ OR ‘exencephaly’/exp OR exencephaly OR exencephalies
#2
‘vitamin b9’ OR ‘vitamin m’/exp OR ‘vitamin m’ OR ‘pteroylglutamic acid’/exp OR ‘pteroylglutamic acid’ OR ‘folvite’/exp OR ‘folvite’ OR ‘folacin’/exp OR ‘folacin’ OR ‘folate’/exp OR ‘folate’ OR ‘folic acid’/exp OR ‘folic acid’ OR ‘multivitamin’ OR ‘multivitamin’/exp OR multivitamin OR ‘prenatal vitamin’ OR ‘multivitamins’ OR ‘multivitamins’/exp OR multivitamins OR ‘prenatal vitamins’ OR ‘vitamin supplement’ OR ‘vitamin supplements’ OR ‘5 me thf’ OR ‘5 me h4f’ OR ‘5 methyltetrahydrofolate’/exp OR ‘5 methyltetrahydrofolate’ OR ‘5 methyltetrahydropteroylpentaglutamate’ OR ‘5-methyltetrahydrofolate triglutamate’
#1
‘pregnant women’/exp OR ‘pregnant women’ OR ‘pregnancy’/exp OR ‘pregnancy’ OR pregnant

EMBASE Harms Search

#12
#11
#7 NOT #10
#10
#9 NOT #8
#9
#7 AND [animals]/lim

#8
#7 AND [humans]/lim
#7
#5 NOT #6
#6
#5 AND [medline]/lim
#5
#3 AND #4
#4
‘colorectal tumor’/exp OR ‘cyanocobalamin deficiency’/exp OR (b12 AND deficien*) OR ‘pyridoxine deficiency’/exp OR (b6 AND deficien*) OR ‘adverse drug reaction’/exp/mj OR ‘patient harm’/exp/mj OR harm OR harms OR ‘adverse drug reaction’/exp OR ‘adverse drug reaction’ OR ‘adverse effect’ OR ‘adverse effects’ OR ‘adverse event’ OR ‘adverse events’ OR complication OR complications OR ‘twin pregnancy’/exp OR ‘twins’/exp OR twin OR twins OR twinning OR ‘asthma’/exp OR asthma
#3
#1 AND #2
#2
‘pregnancy’/exp OR ‘pregnant women’/exp OR ‘pregnancy’/de OR pregnant
#1
‘folic acid’/exp OR ‘folic acid’ OR ‘folic acid’:ti OR ‘folic acid’:ab OR folvite:ti OR folvite:ab OR folacin:ti OR folacin:ab OR ‘5 me thf’ OR ‘5 me h4f’ OR ‘5 methyltetrahydrofolate’/exp OR ‘5 methyltetrahydrofolate’ OR ‘5 methyltetrahydropteroylpentaglutamate’ OR ‘5-methyltetrahydrofolate triglutamate’

Gray Literature Benefits Search (ClinicalTrials.gov, HSRProj, the World Health Organization’s International Clinical Trials Registry Platform, and NIH Reporter)

Benefits search strategy:
(Pregnancy OR pregnant OR “pregnant women”) AND (“folic acid” OR “vitamin b9” OR “vitamin m” OR “Pteroylglutamic Acid” OR folvite OR folacin OR folate OR “folic acid” OR multivitamin OR “prenatal vitamin” OR multivitamins OR “prenatal vitamins” OR “vitamin supplement” OR “vitamin supplements” OR 5-Me-THF OR 5-Me-H4F OR 5-methyltetrahydrofolate) AND (“neural tube defects” OR “spina bifida” OR “neural tube damage” OR “neural tube defect” OR “neural tube disorders” OR Craniorachischisis OR Craniorachischises OR Diastematomyelia OR Diastematomyelias OR “Tethered Cord Syndrome” OR “Tethered Cord Syndromes” OR “Tethered Spinal Cord Syndrome” OR “Occult Spinal Dysraphism” OR “Occult Spinal Dysraphisms” OR Iniencephaly OR Iniencephalies OR “Neurenteric Cyst” OR “Neurenteric Cysts” OR “Neuroenteric Cyst” OR “Neuroenteric Cysts” OR “Spinal Cord Myelodysplasia” OR “Spinal Cord Myelodysplasias” OR Acrania OR Acranias OR Exencephaly OR Exencephalies)

Gray Literature HARMS Search (ClinicalTrials.gov, HSRProj, the World Health Organization's International Clinical Trials Registry Platform, and NIH Reporter)

I. SPECIFIC HARMS

("folic acid" OR folvite OR folacin) AND ((twin OR twin OR twinning) OR "colorectal neoplasms" OR "colorectal cancer" OR "colorectal tumor" OR "vitamin b12 deficiency" OR "vitamin b6 deficiency") AND ("drug-related side effects" OR "adverse reaction" OR harm OR harms OR "adverse effect" OR "adverse effects" OR "adverse event" OR "adverse events" OR Complication OR Complications OR asthma)

II. GENERAL HARMS

(folic acid OR folvite OR folacin) AND ("drug-related side effects" OR "adverse reaction" OR harm OR harms OR "adverse effect" OR "adverse effects" OR "adverse event" OR "adverse events" OR Complication OR Complications)

eTable1. Inclusion/Exclusion Criteria

	Include	Exclude
Populations	KQ 1: Women of childbearing age (postmenarchal and premenopausal; women with the potential for or planning childbearing) KQ 2: Women of childbearing age (postmenarchal and premenopausal, women with the potential for or planning childbearing), fetus, neonate, or child from index pregnancy	KQ 1: Prepubertal girls, men, women without the potential for childbearing (e.g., women who are postmenopausal or have genetic uterine or ovarian abnormalities)
Interventions	Folic acid supplementation, with or without food fortification or naturally occurring folate, for the prevention of neural tube defects and other birth defects Supplementation with micronutrients (e.g., multivitamins, iron) in combination with folic acid for the prevention of neural tube defects only	Food fortification only Naturally occurring folate only Counseling to improve dietary supplementation Supplementation with micronutrients (e.g., multivitamins, iron) in combination with folic acid for the prevention of harms only
Comparisons	KQs 1a, 1b, 2a: Placebo or no treatment, dietary supplementation only, supplementation with prenatal vitamins without folic acid, iron supplements without folic acid KQs 1b, 1c, 2b: All of the above plus folic acid supplementation of varying dosages	KQs 1a, 1b, 2a: Lower or higher doses of folic acid supplementation, folic acid vs. other active comparators KQs 1c, 2b: Folic acid vs. other active comparators (e.g., multivitamins)
Outcomes	Neonatal outcomes: Neural tube defects Harms from treatment: Twins Colorectal cancer or other reported types of cancer Vitamin B12 deficiency Vitamin B6 deficiency Other reported child, neonatal, fetal, or maternal harms	Benefits not specified in inclusion criteria
Timing	KQs 1a, 1b: Supplementation initiated before index pregnancy or in the first trimester KQs 1c, 2a, 2b: All timing	KQs 1a, 1b: Supplementation initiated after the first trimester of pregnancy
Settings	Developed countries categorized as "Very High" on the Human Development Index (as defined by the United Nations Development Programme)	Countries not categorized as "Very High" on the Human Development Index

	Include	**Exclude**
Study designs	Efficacy (KQ 1): Randomized clinical trials; controlled clinical trials; cohort or case-control studies Harms (KQ 2): Randomized clinical trials; controlled clinical trials; or observational studies (case-control, cohort, registry data)	Commentaries, editorials, case reports
Sample size	More than 50 participants	50 participants or fewer
Quality	Good and fair quality	Poor quality
Language	English	Non-English studies

eTable 2. Quality Assessment Criteria

Quality Assessment Criteria by Study Design	
Systematic Reviews	Was a comprehensive literature search performed? Was the status of publication (i.e., grey literature) used as an inclusion criterion? Was an "a priori" design provided? Was a list of studies (included and excluded) provided? Were there explicit inclusion/exclusion criteria for the selection of studies? Were the characteristics of the included studies provided? Was the likelihood of publication bias assessed? Was there duplicate study selection and data extraction? Was the scientific quality of the included studies assessed and documented? Was the conflict of interest included? Was the scientific quality of the included studies used appropriately in formulating conclusions? Were the methods used to combine the findings of studies appropriate? Were the authors' conclusions supported by the evidence they presented?
RCTs	Were eligibility criteria described clearly? Were the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants? Was the intervention or exposure clearly defined across all study participants? Was randomization adequate? Was allocation concealment adequate? Were groups similar at baseline? Did the study control for baseline differences between groups? Were the participants and the administrators of the intervention blinded to the intervention or exposure status of participants? Were the outcome assessors blinded to the outcome status of participants? Was intervention fidelity adequate? Was there a risk of recall bias? Did the study focus on a time period of interest? Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? Did variation from the study protocol compromise the conclusions of the study? What was the overall attrition? What was the differential attrition? Did the study have high attrition or low response rate raising concern for bias? Did the study use an intention-to-treat analysis? Did the study have crossovers or contamination raising concern for bias? Were outcomes prespecified/defined and adequately described? Were outcome measures valid and reliable? Were all important outcomes considered? Was the duration of followup adequate to assess the outcome?

Quality Assessment Criteria by Study Design	
Case-Control Studies	Were eligibility criteria described clearly? Were the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants? Was the symptom status of subjects determined using valid and reliable methods, implemented across all study participants? Was the intervention or exposure clearly defined across all study participants? Was the strategy for recruiting participants into the study the same across study groups? Were groups similar at baseline? Did the study control for baseline differences between groups? Were outcome assessors blinded to the exposure? Was intervention fidelity adequate? Was there a risk of recall bias? Did the study focus on a time period of interest? Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? What was the overall response rate? Did the study have high attrition or low response rate raising concern for bias? Did the analysis adjust for potential confounders? Were outcomes prespecified/defined and adequately described? Was an appropriate method used to handle missing data?
Quality Assessment Criteria by Study Design	
Cohort Studies	Were eligibility criteria described clearly? Were the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants? Was the intervention or exposure clearly defined across all study participants? Was the strategy for recruiting participants into the study the same across study groups? Did start of followup and start of intervention coincide? Were groups similar at baseline? Did the study control for baseline differences between groups? Were the participants and the administrators of the intervention blinded to the intervention or exposure status of participants? Were outcome assessors blinded to the outcome status of participants? Was adherence to the intervention adequate? Was there a risk of recall bias? Did the study focus on a time period of interest? Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? What was the overall attrition? What was the differential attrition? Did the study have high attrition or low response rate raising concern for bias? Did the analysis adjust for potential confounders? Did the study have crossovers or contamination raising concern for bias?

	Were outcomes prespecified/defined and adequately described? Were outcome measures valid and reliable? Were all important outcomes considered? Was the duration of followup adequate to assess the outcome? Was an appropriate method used to handle missing data?

eTable 3. Quality Rating Criteria

Definition of Quality Ratings for Systematic Reviews	
Good	Recent, relevant review with comprehensive sources and search strategies; explicit and relevant selection criteria; standard appraisal of included studies; and valid conclusions
Fair	Recent, relevant review that is not clearly biased but lacks comprehensive sources and search strategies
Poor	Outdated, irrelevant, or biased review without systematic search for studies, explicit selection criteria, or standard appraisal of studies
Definition of Quality Ratings for RCTs	
Good	Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (followup ≥80%); reliable and valid measurement instruments are used and applied equally to all groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention is given to confounders in analysis. In addition, intention-to-treat analysis is used.
Fair	Studies are graded "fair" if any or all of the following problems occur without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially, but some question remains whether some (although not major) differences occurred with followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is used.
Poor	Studies are graded "poor" if any of the following fatal flaws exist: Groups assembled initially are not close to being comparable or maintained throughout the study, unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment), and key confounders are given little or no attention. Intention-to-treat analysis is lacking.
Definition of Quality Ratings for Case-Control Studies	
Good	Appropriate ascertainment of cases and nonbiased selection of case and control participants, exclusion criteria applied equally to cases and controls, response rate equal to or greater than 80%, accurate diagnostic procedures and measurements applied equally to cases and controls, and appropriate attention to confounding variables
Fair	Recent, relevant, and without major apparent selection or diagnostic workup bias, but response rate less than 80% or attention to some but not all important confounding variables
Poor	Major selection or diagnostic workup bias, response rate less than 50%, or inattention to confounding variables
Definition of Quality Ratings for Cohort Studies	
Good	Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (followup ≥80%), reliable and valid measurement instruments are used and applied equally to all groups, interventions are spelled out clearly, all important outcomes are considered, and appropriate attention to

	confounders in analysis
Fair	Studies are graded "fair" if any or all of the following problems occur without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially, but some question remains whether some (although not major) differences occurred with followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for.
Poor	Studies are graded "poor" if any of the following fatal flaws exist: Groups assembled initially are not close to being comparable or maintained throughout the study, unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment), and key confounders are given little or no attention.

[a] US Preventive Services Task Force, Procedure Manual, Appendix VI http://www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes. December 2015

eTable 4. Quality Ratings of Included Systematic Reviews

First Author, Year	Comprehensive literature search performed?	Was the status of publication used as an inclusion criterion?	Was an 'a priori' design provided?	Was a list of studies provided?	Was there explicit inclusion/exclusion criteria for the selection of studies?	Were the characteristics of the included studies provided?	Was the likelihood of publication bias assessed?	Was there duplicate study selection and data extraction?
Brown et al., 2014[1]	No	Unclear	No	No	No	Yes	No	Unclear
Crider et al, 2013[2]	Yes	No	Unclear	No	Yes	Yes	Yes	Yes
Goh et al., 2006[3]	Yes	No	No	No	No	No	Yes	Yes
Wang et al., 2015 [4]	Yes	No	Yes	Yes	Unclear	Yes	Yes	Yes
Wolff et al., 2009[5] Wolff et al., 2009[6]	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes
Yang et al., 2014[7]	Yes	No	Unclear	No	Yes	Yes	Yes	Unclear

First Author, Year	Was the scientific quality of the included studies assessed and documented?	Was the conflict of interest included?	Was the scientific quality of the included studies used appropriately in formulating conclusions?	Were the methods used to combine the findings of studies appropriate?	Was the authors' conclusions supported by the evidence they presented?	Quality Rating	Comments
Brown et al., 2014[1]	No	No	NA	Unclear	No	Poor	Although the review is titled a systematic review the cruical aspects of a systematic review including details on searches, review, risk of bias appraisal, and synthesis are NR.
Crider et al, 2013[2]	Yes	Unclear	Yes	Yes	Yes	Good	
Goh et al., 2006[3]	No	No	No	Yes	Yes	Poor	The quality of studies is not assessed. The characteristics of studies included in the meta-analysis were not presented; Includes an appropriate synthesis and statistical testing, but does not include a discussion of publication bias.
Wang et al., 2015 [4]	Yes	Yes	Yes	Yes	Yes	Fair	A list of excluded studies and the inclusion criteria was not provided.
Wolff et al., 2009[5] Wolff et al., 2009[6]	Yes	no	Yes	NA	Yes	Good	
Yang et al., 2014[7]	Yes	Unclear	Unclear	Yes	Yes	Fair	Unclear how authors used risk of bias assessments in the analysis. The study noted that they included high quality studies.

eTable 5.1. Quality Ratings of Included Studies

First Author, Year	Were eligibility criteria described clearly?	Are the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants?	Was the symptom status of subjects determined using valid and reliable methods, implemented across all study participant?	Was the intervention or exposure clearly defined, across all study participant?	Was method of randomization adequate?	Was allocation concealment adequate?
Abe et al., 2014[8]	No	Unclear	NA	Unclear	NA	NA
Abe et al., 2013[9]	No	Unclear	NA	No	NA	NA
Abe et al., 2015[10]	No	Unclear	NA	Unclear	NA	NA
Agopian et al., 2013[11]	Yes	Yes	Yes	No	NA	NA
Ahrens et al., 2011[12]	Yes	Yes	Unclear	Yes	NA	NA
Berry et al., 2004[13]	Yes	Yes	Yes	No	NA	NA
Botto et al., 2002[14]	Yes	Yes	Yes	Yes	NA	NA
Bower et al., 1989[15] Bower, 1992[16]	Yes	Yes	Yes	No	NA	NA
Brescianini et al., 2012[17]	No	Unclear	Yes	Unclear	NA	NA
Carmichael et al., 2010[18]	Yes	Yes	Yes	No	NA	NA
Chandler et al., 2012[19]	Yes	Yes	Yes	No	NA	NA
Charles et al., 2004[20] Charles et al., 2005[21] Taylor et al., 2015[22]	Yes	Yes	Yes	Yes	No	Yes
Correa et al., 2012[23]	Yes	Yes	Yes	No	NA	NA

First Author, Year	Were eligibility criteria described clearly?	Are the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants?	Was the symptom status of subjects determined using valid and reliable methods, implemented across all study participant?	Was the intervention or exposure clearly defined, across all study participant?	Was method of randomization adequate?	Was allocation concealment adequate?
Czeizel et al., 2004[24]	Yes	Yes	Yes	Yes	NA	NA
Czeizel et al., 2004[25]	Yes	Yes	NA	Yes	NA	NA
Czeizel et al., 1992[26]; Czeizel et al., 1993[27] Czeizel et al., 1993[28] Czeizel et al., 1994[29] Czeizel et al., 1994[30] Czeizel et al., 1996[31] Czeizel et al., 1998[32]	Yes	Yes	NA	Yes	Yes	Unclear
Czeizel et al., 1996[33]	Yes	Yes	Yes	Yes	NA	NA
De Marco et al., 2011[34]	Yes	Yes	Yes	Yes	NA	NA
DeSoto et al., 2012[35]	Yes	Yes	Yes	No	NA	NA
Ericson et al., 2001[36]	No	No	NA	No	NA	NA
Gildestad et al., 2013[37]	Unclear	Unclear	Unclear	Unclear	NA	NA
Haberg et al., 1994[38]	Unclear	Unclear	NA	Unclear	NA	NA
Hernandez et al., 2001[39]	Yes	Yes	Yes	Unclear	NA	NA

First Author, Year	Were eligibility criteria described clearly?	Are the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants?	Was the symptom status of subjects determined using valid and reliable methods, implemented across all study participant?	Was the intervention or exposure clearly defined, across all study participant?	Was method of randomization adequate?	Was allocation concealment adequate?
Kallen et al., 2004[40]	Yes	Yes	Yes	No	NA	NA
Kallen et al., 2007[41]	Yes	Yes	NA	No	NA	NA
Kondo et al., 2015[42]	No	Yes	Unclear	No	NA	NA
Medvezky et al., 2003[43]	Yes	Yes	Yes	No	NA	NA
Mills et al., 1989[44]	No	Yes	Yes	Yes	NA	NA
Moore et al., 2003[45] Milunsky et al., 1989[46]	Yes	Yes	NA	Yes	NA	NA
Mosley 2009 et al.,[47]	Yes	Yes	Yes	Yes	NA	NA
Mulinare et al., 1988[48]	Yes	Yes	Yes	Yes	NA	NA
Ohya et al., 2011[49]	No	Unclear	NA	Unclear	NA	NA
Shaw et al., 2002[50]	Yes	Yes	Yes	No	NA	NA
Shaw et al., 1995[51]	Yes	Yes	Yes	Yes	NA	NA

First Author, Year	Were eligibility criteria described clearly?	Are the inclusion/exclusion criteria measured using valid and reliable measures, implemented across all study participants?	Was the symptom status of subjects determined using valid and reliable methods, implemented across all study participant?	Was the intervention or exposure clearly defined, across all study participant?	Was method of randomization adequate?	Was allocation concealment adequate?
Suarez et al., 2000[52]	Yes	Yes	Unclear	Yes	NA	NA
Veeranki et al., 2014[53]	Yes	Yes	NA	No	NA	NA
Veeranki et al., 2014[54]	No	Unclear	NA	No	NA	NA
Veeranki et al., 2015[55]	Yes	Yes	NA	No	NA	NA
Vollset et al., 2005[56]	Yes	Yes	Yes	No	NA	NA
Werler et al., 1993[57]	Yes	Unclear	Yes	Yes	NA	NA

eTable 5.2. Quality Ratings of Included Studies

First Author, Year	Was the strategy for recruiting participants into the study the same across study groups?	Do start of follow-up and start of intervention coincide?	Are baseline characteristics similar between groups?	Did the study control for baseline differences between groups?	Were the participants and the administrators of the intervention blinded to the intervention or exposure status of participants?	Were the outcome assessors blinded to the outcome status of participants?
Abe et al., 2014[8]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Abe et al., 2013[9]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Abe et al., 2015[10]	Unclear	Unclear	Unclear	Unclear	No	Unclear
Agopian et al., 2013[11]	No	NA	No	Yes	NA	NA
Ahrens et al., 2011[12]	Unclear	NA	No	Yes	NA	NA
Berry et al., 2004[13]	NA (registry)	No	Unclear	Unclear	NA	Unclear
Botto et al., 2002[14]	No	NA	Yes	Yes	NA	Yes
Bower et al., 1989[15] Bower et al., 1992[16]	No	NA	Unclear	Unclear	NA	NA
Brescianini et al., 2012[17]	Unclear	No	Unclear	Unclear	NA	NA
Carmichael et al., 2010[18]	Yes	NA	No	Yes	NA	NA
Chandler et al., 2012[19]	No	NA	No	Yes	NA	NA
Charles et al., 2004[20] Charles et al., 2005[21] Taylor et al., 2015[22]	NA	NA	No	Yes	No	Unclear
Correa et al., 2012[23]	No	NA	No	Yes	NA	NA
Czeizel et al., 2004[24]	Yes	No	No	No	NA	NA

First Author, Year	Was the strategy for recruiting participants into the study the same across study groups?	Do start of follow-up and start of intervention coincide?	Are baseline characteristics similar between groups?	Did the study control for baseline differences between groups?	Were the participants and the administrators of the intervention blinded to the intervention or exposure status of participants?	Were the outcome assessors blinded to the outcome status of participants?
Czeizel et al., 2004[25]	No	no	No	Yes	No	Unclear
Czeizel et al., 1992[26]; Czeizel et al., 1993[27] Czeizel et al., 1993[28] Czeizel et al., 1994[29] Czeizel et al., 1994[30] Czeizel et al., 1996[31] Czeizel et al., 1998[32]	NA	Yes	Yes	NA	Unclear	Unclear
Czeizel et al., 1996[33]	Yes	No	Unclear	Unclear	NA	NA
De Marco et al., 2011[34]	Yes	NA	No	Yes	NA	NA
DeSoto et al., 2012[35]	Yes	NA	Yes	Yes	NA	NA
Ericson et al., 2001[36]	Unclear	No	Unclear	Unclear	Unclear	Unclear
Gildestad et al., 2013[37]	Unclear	No	Unclear	Unclear	Unclear	Unclear
Haberg et al., 1994[38]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Hernandez et al., 2001[39]	Yes	NA	No	Yes	NA	Unclear
Kallen et al., 2004[40]	NA (registry)	No	Unclear	Unclear	NA	Unclear
Kallen et al., 2007[41]	Yes	No	No	Yes	Unclear	Unclear

First Author, Year	Was the strategy for recruiting participants into the study the same across study groups?	Do start of follow-up and start of intervention coincide?	Are baseline characteristics similar between groups?	Did the study control for baseline differences between groups?	Were the participants and the administrators of the intervention blinded to the intervention or exposure status of participants?	Were the outcome assessors blinded to the outcome status of participants?
Kondo et al., 2015[42]	No	NA	No	Yes	NA	NA
Medvezky et al., 2003[43]	Yes	NA	Unclear	Unclear	NA	NA
Mills et al., 1989[44]	No	NA	No	Yes	NA	Unclear
Moore et al., 2003[45] Milunsky et al., 1989[46]	Yes	No	No	Yes	No	No
Mosley et al., 2009[47]	Yes	NA	Yes	Yes	NA	Unclear
Mulinare et al., 1988[48]	No	NA	Unclear	Unclear	NA	Unclear
Ohya et al., 2011[49]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Shaw et al., 2002[50]	No	NA	Unclear	Unclear	NA	NA
Shaw et al., 1995[51]	No	NA	No	Yes	NA	NA
Suarez et al., 2000[52]	No	NA	Unclear	Yes	NA	NA
Veeranki et al., 2014[53]	Yes	No	No	Yes	Unclear	Unclear
Veeranki et al., 2014[54]	Yes	No	Unclear	Unclear	No	Unclear
Veeranki et al., 2015[55]	Yes	No	No	Yes	No	Yes
Vollset et al., 2005[56]	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Werler et al., 1993[57]	Unclear	Yes	Unclear	No	NA	NA

eTable 5.3. Quality Ratings of Included Studies

First Author, Year	Were outcome assessors blinded to the exposure?	Was intervention fidelity adequate?	Was there a risk of recall bias?	Did the study focus on the time period that we are interested in?	Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?	Did variation from the study protocol compromise the conclusions of the study?
Abe et al., 2014[8]	NA	Unclear	Unclear	Unclear	Unclear	Unclear
Abe et al., 2013[9]	NA	Unclear	Yes	Unclear	No	NA
Abe et al., 2015[10]	NA	Unclear	Yes	Yes	Unclear	NA
Agopian et al., 2013[11]	Unclear	Unclear	Unclear	Yes	Yes	NA
Ahrens et al., 2011[12]	Yes	Unclear	Unclear	Yes	Yes	NA
Berry et al., 2004[13]	NA	Unclear	Unclear	No	No	NA
Botto et al., 2002[14]	Unclear	Yes	Unclear	Yes	Yes	NA
Bower et al., 1989[15] Bower et al., 1992[16]	Unclear	Unclear	Unclear	Unclear	Yes	NA
Brescianini et al., 2012[17]	Unclear	Unclear	Unclear	Unclear	No	NA
Carmichael et al., 2010[18]	Unclear	Unclear	Unclear	No	Yes	NA
Chandler et al., 2012[19]	Unclear	Unclear	Yes	Yes	Yes	NA
Charles et al., 2004[20] Charles et al., 2005[21] Taylor et al., 2015[22]	NA	Yes	No	Yes	No	No
Correa et al., 2012[23]	Unclear	Unclear	Yes	No	No	NA
Czeizel et al., 2004[24]	Unclear	Unclear	Unclear	Unclear	No	NA
Czeizel et al., 2004[25]	Unclear	Unclear	No	Yes	No	No

First Author, Year	Were outcome assessors blinded to the exposure?	Was intervention fidelity adequate?	Was there a risk of recall bias?	Did the study focus on the time period that we are interested in?	Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?	Did variation from the study protocol compromise the conclusions of the study?
Czeizel et al., 1992[26]; Czeizel et al., 1993[27] Czeizel et al., 1993[28] Czeizel et al., 1994[29] Czeizel et al., 1994[30] Czeizel et al., 1996[31] Czeizel et al., 1998[32]	NA	Yes	No	Yes	No	No
Czeizel et al., 1996[33]	Unclear	Unclear	Unclear	Yes	No	NA
De Marco et al., 2011[34]	Unclear	Yes	Unclear	No	No	NA
DeSoto et al., 2012[35]	Unclear	Unclear	Yes	No	No	NA
Ericson et al., 2001[36]	NA	Yes	Unclear	Unclear	No	NA
Gildestad et al., 2013[37]	NA	Unclear	Unclear	Unclear	No	NA
Haberg, 1994[38]	NA	Unclear	Unclear	Yes	Unclear	NA
Hernandez et al., 2001[39]	Unclear	Yes	Unclear	No	No	NA
Kallen et al., 2004[40]	NA	Unclear	Unclear	No	No	NA
Kallen et al., 2007[41]	NA	Yes	Unclear	Unclear	No	NA
Kondo et al., 2015[42]	No	Unclear	Yes	Unclear	No	NA
Medvezky et al., 2003[43]	Unclear	Unclear	Unclear	No	Unclear	NA
Mills et al., 1989[44]	Yes	Yes	Unclear	Yes	Yes	NA

First Author, Year	Were outcome assessors blinded to the exposure?	Was intervention fidelity adequate?	Was there a risk of recall bias?	Did the study focus on the time period that we are interested in?	Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?	Did variation from the study protocol compromise the conclusions of the study?
Moore et al., 2003[45] Milunsky et al., 1989[46]	NA	Unclear	Unclear	No	Yes	NA
Mosley et al., 2009[47]	Unclear	Yes	Unclear	Yes	Yes	NA
Mulinare et al., 1988[48]	Yes	Unclear	Yes	No	No	NA
Ohya et al., 2011[49]	NA	Unclear	Unclear	Unclear	Unclear	Unclear
Shaw et al., 2002[50]	Unclear	Unclear	Yes	No	Yes	NA
Shaw et al., 1995[51]	Yes	Unclear	Unclear	Yes	Yes	NA
Suarez et al., 2000[52]	Unclear	Yes	Unclear	Yes	Unclear	NA
Veeranki et al., 2014[53]	NA	Unclear	No	Unclear	No	NA
Veeranki et al., 2014[54]	NA	Unclear	No	Yes	No	NA
Veeranki et al., 2015[55]	NA	Unclear	No	Unclear	No	NA
Vollset et al., 2005[56]	NA	Unclear	No	Unclear	No	NA
Werler et al., 1993[57]	Unclear	Yes	Yes	Yes	Yes	NA

eTable 5.4. Quality Ratings of Included Studies

First Author, Year	What was the overall attrition/overall response rate?	What was the overall differential attrition?	Did the study have high attrition or low response rate raising concern for bias?	Is the analysis conducted on an intention-to-treat (ITT) basis?	Did the analysis adjust for potential confounders?	Did the study have cross-overs or contamination raising concern for bias?
Abe et al., 2014[8]	Unclear	Unclear	Unclear	Na	Unclear	Unclear
Abe et al., 2013[9]	Unclear	Unclear	Unclear	NA	Unclear	Unclear
Abe et al., 2015[10]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Agopian et al., 2013[11]	Response Rate (overall sample in original study; see Yoon et al., 2001 companion article) G1 + G2: About 74 (NR out of 7,470) G2: About 63 (NR out of 3,821)	NA	Unclear	NA	Unclear	NA
Ahrens et al., 2011[12]	G1: 66% G2: 53%	NA	no	NA	Yes	NA
Berry et al., 2004[13]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Botto et al,, 2002[14]	Overall sample G1: 69% G2: 71% NTD analysis G1: NR G2: NR	NA	Unclear	NA	Unclear	NA
Bower et al., 1989[15] Bower et al., 1992[16]	G1: 93% G2: 88% G3: 84%	Response Rate G1: 93 (77/83) G2: 88 (77/87) G3: 84 (154/183)	No	NA	No	NA

First Author, Year	What was the overall attrition/overall response rate?	What was the overall differential attrition?	Did the study have high attrition or low response rate raising concern for bias?	Is the analysis conducted on an intention-to-treat (ITT) basis?	Did the analysis adjust for potential confounders?	Did the study have cross-overs or contamination raising concern for bias?
Brescianini et al., 2012[17]	G1: 93% G2: 88% G3: 84%	NA	No	NA	No	NA
Carmichael et al., 2010[18]	G1: 73 (146/200) G2: 79 (191/241) G3: 80% (626/786)	NA	No	NA	Yes	NA
Chandler et al., 2012[19]	62% for anencephaly; 76% spina bifida and 71% controls	NA	No	NA	Unclear	NA
Charles et al., 2004[20] Charles et al., 2005[21] Taylor et al., 2015 [22]	Unclear	Unclear	Unclear	No	NA	No
Correa et al., 2012[23]	Overall, 70% among mothers of case infants and 67% among mothers of control infants. Response rate not reported for NTD mothers	NA	Unclear	NA	Yes	NA
Czeizel et al., 2004[24]	Unclear	NA	Unclear	NA	Yes	NA
Czeizel et al., 2004[25]	Overall attrition Unclear	G1: 3069/3981 (77.1%) G2: Unclear	Yes	No	Yes	Unclear

First Author, Year	What was the overall attrition/overall response rate?	What was the overall differential attrition?	Did the study have high attrition or low response rate raising concern for bias?	Is the analysis conducted on an intention-to-treat (ITT) basis?	Did the analysis adjust for potential confounders?	Did the study have cross-overs or contamination raising concern for bias?
Czeizel et al., 1992[26]; Czeizel et al., 1993[27] Czeizel et al., 1993[28] Czeizel et al., 1994[29] Czeizel et al., 1994[30] Czeizel et al., 1996[31] Czeizel et al., 1998[32]	1%	0.10%	No	NA	NA	No
Czeizel et al., 1996[33]	63% for negative controls, rate for positive controls NR	NA	Unclear	Na	No	NA
De Marco et al., 2011[34]	Response Rate G1: 92 (133/145) G2: 82 (273/332)	NA	No	NA	Unclear	NA
DeSoto et al., 2012[35]	Response Rate G1: 48.1% (321/668) G2: 31.7% (774/2444)	NA	Yes	NA	Unclear	NA
Ericson et al., 2001[36]	Unclear	Unclear	Unclear	NA	Yes	No
Gildestad et al., 2013[37]	Unclear	NA	Unclear	NA	Unclear	Unclear
Haberg et al., 1994[38]	Unclear	Unclear	NA	Unclear	Unclear	Unclear

First Author, Year	What was the overall attrition/overall response rate?	What was the overall differential attrition?	Did the study have high attrition or low response rate raising concern for bias?	Is the analysis conducted on an intention-to-treat (ITT) basis?	Did the analysis adjust for potential confounders?	Did the study have cross-overs or contamination raising concern for bias?
Hernandez et al., 2001[39]	Response Rate G1: 84% (1,242/NR) G2: 83% (6,600/NR) G3: 80% (1,626/NR) G4: NR (2,138/NR)	NA	No	NA	No	NA
Kallen et al., 2004[40]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Kallen et al., 2007[41]	Unknown	Unclear	Unclear	NA	Yes	No
Kondo et al., 2015[42]	Response Rate G1: 79% G2: 56%	NA	No	NA	Yes	NA
Medvezky et al., 2003[43]	96.9% cases; 96% other non-NTD cases; 83.1% controls	NA	Yes (was Unclear)	NA	Unclear	NA
Mills et al., 1989[44]	Response Rate G1: 64.8%-82% (571/NR) G2: NR (546/NR) G3: NR (573/NR)	NA	No	NA	Unclear	NA

First Author, Year	What was the overall attrition/overall response rate?	What was the overall differential attrition?	Did the study have high attrition or low response rate raising concern for bias?	Is the analysis conducted on an intention-to-treat (ITT) basis?	Did the analysis adjust for potential confounders?	Did the study have cross-overs or contamination raising concern for bias?
Moore et al., 2003[45] Milunsky et al., 1989[46]	3% (715/23,491)	Unclear	No	NA	Unclear	No
Mosley et al., 2009[47]	62% anencephaly; 76% SB; 71% controls	NA	No	NA	Yes	NA
Mulinare, 1988[48]	G1: 347/519 (66.9%) G2: 2829/4043 (69.9%)	NA	No	NA	Yes	NA
Ohya et al., 2011[49]	Unclear	Unclear	Unclear	NA	Unclear	Unclear
Shaw et al., 2002[50]	Response Rate G1 (NTD only): 84% G2: 76% from both control cohorts	NA	No	NA	No	NA
Shaw et al., 1995[51]	88% both groups	NA	No	NA	Yes	NA
Suarez et al., 2000[52]	72% cases; 53% controls	NA	No	NA	Yes	NA
Veeranki et al., 2014[53]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Veeranki et al., 2014[54]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Veeranki et al., 2015[55]	Unclear	Unclear	Unclear	NA	Yes	Unclear

Vollset et al., 2005[56]	Unclear	Unclear	Unclear	NA	Yes	Unclear
Werler et al., 1993[57]	G1: 567-436/567=76.9% G2: 3672-2615/3672=71.2%	NA	No	NA	Yes	NA

eTable 5.5. Quality Ratings of Included Studies

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Abe et al., 2014[8]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.
Abe et al., 2013[9]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.
Abe et al., 2015[10]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.
Agopian et al., 2013[11]	Yes	NA	NA	NA	Unclear	Fair	Most of the major confounders were adjusted for in the analyses. However, there was still a possibility of residual confounding because some NTD-specific variables not controlled for, specifically, previous NTD pregnancy or having or having partner(s) with NTDs. Cases and controls selected from different populations. Specifically, cases could be stillborn infants or therapeutic abortions, while all controls were liveborn infants. Definition of exposure not defined clearly. Positive response to folic acid supplement use question could have indicated any frequency of usage.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Ahrens et al., 2011[12]	Yes	NA	NA	NA	Unclear	Fair	Ascertainment of cases for non-live births is not routine. Unclear how missing data was handled (although "all women were included" in the analysis).
Berry et al., 2004[13]	Yes	Yes	Yes	Yes	Unclear	Poor	The study itself does not describe the data source well, but it cites other studies that do. Based on these other studies, we infer that the exposure period starts before pregnancy and extends to first attendance (in 90% of cases before the end of the first trimester and usually around week 10) with lack of clarity on degree of exposure, difficult to clearly distinguish exposure from non-exposure (defined as any vs. no supplement), degree of adherence is unclear; the study does not appear to account for fetal deaths, the resulting selection bias would serve to mute rather than exaggerate the effect of FA on twinning; study uses probabilistic simulations to assess bias caused by misclassification of the use of IVF.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Botto, 2002[14]	Yes	NA	NA	NA	Unclear	Poor	Cases and controls selected from different populations. Cases could be liveborn or stillborn infants, while all controls were liveborn infants. No information about 1) timing of data collection given to assess recall bias; 2) whether potential NTD-specific confounders were measured or adjusted for in analyses; 3) exposure to dietary folate; 4) response rates for patients analyzed for analysis of NTD outcomes; or 5) whether MVs contained similar and clinically effective doses of FA. Missing data not accounted for in an ITT analysis; covers a 6 month period of exposure (3 months before to 3 months after pregnancy); no controls for concurrent interventions such as exposure to dietary folate, however, this sample was drawn before dietary supplementation; significant risk of recall bias because women asked to remember for a period ranging from 1968-1980. Only still births, no information on terminations. Cases included live and still-borns controls only

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Bower et al., 1989[15] Bower et al., 1992[16]	Yes	NA	NA	NA	Unclear	Poor	Cases and controls selected from different populations. Cases could be stillborn infants or elective terminations following antenatal NTD diagnosis, while Control group 1 included liveborn infants and terminated pregnancies with non-NTD malformations and Control group 2 included live born infants only. Possibility of residual confounding because no NTD-specific variables controlled for in analyses. Definition of exposure not defined clearly. Positive response to folic acid supplement use question could have indicated any frequency of usage. Risk of recall bias because mothers interviewed up to 99 weeks after last menstrual period. Also, minor risk of interviewer bias during 5 interviews because interviewers unintentionally learned case-control status.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Brescianini et al., 2012[17]	Unclear	NA	NA	NA	Unclear	Unclear	Meeting abstract with very little information to base judgment on most domains. Although study does not appear to account for fetal deaths, the resulting selection bias would serve to mute rather than exaggerate the effect of folic acid on twinning.
Carmichael et al., 2010[18]	Yes	NA	NA	NA	Unclear	Poor	Unclear how missing data was handled; time period extends to 2 months before and 2 months after and with lack of clarity on degree of exposure, difficult to clearly distinguish exposure from non-exposure; authors did not specify definition of exposure clearly (defined as any vs. no supplement), so the degree of adherence is Unclear. The control group only looked at live births and not fetal deaths.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Chandler et al., 2012[19]	Yes	NA	NA	NA	Unclear	Fair	See Yoon et al. (2001) for recruitment information. Cases and controls selected from different populations. Cases could be stillborn infants or therapeutic abortions, while all controls were liveborn infants. Also, definition of exposure not defined clearly. Positive response to folic acid supplement use question could have indicated any frequency of usage. Possibility of residual confounding because some NTD-specific variables not controlled for, specifically, previous NTD pregnancy or having or having partner(s) with NTDs. In addition, analyses not adjusted for comparisons across different centers.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Charles et al., 2004[20] Charles et al., 2005[21] Taylor et al., 2015[22]	Yes	Yes	Yes	Yes	NA	Poor	Randomization was inadequate: "tablets were kept in numbered drawers and distributed in sequence." Unclear whether administrators or outcome assessors were blinded because "The patients' notes were marked with a sticker the same colour as the tablets they were receiving." Some proportion of deaths not linked to patient files (occuring before 1980) but followup N not reported, so rate of overall attrition and differential attrition is Unclear. Because N at followup is not reported, it does not appear that the analysis used an intention-to-treat analysis for the missing data.
Correa et al., 2012[23]	Yes	NA	NA	NA	No	Poor	Unclear how missing data was handled but appears to have been excluded in some tables. Definition of exposure does not account for adherence to meds ("any use during the month before conception or during the first 3 months of pregnancy) so exposed vs. non-exposed not clearly demarcated; additionally, the time period of recall is 1 month before conception to 3 months after, so folic acid supplementation after becoming aware of pregnancy would be misclassified as pre-pregnancy exposure; mothers interviewed up to 24 months after birth so risk of recall bias; Authors did not specify definition of exposure clearly, so the degree of adherence is Unclear. Controls were only live borns.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Czeizel et al., 2004[24]	Yes	NA	NA	NA	Unclear	Poor	Does not mention how missing data was handled. Unclear how authors determined who were clearly users of folic acid. Unable to determine response rate due to limited information. Women continued folic acid usage until at least 3rd trimester. There was only one product of folic acid at the time of the study (3mg) and required a prescription. Unclear on baseline differences and if they were controlled for.
Czeizel et al., 2004[25]	Yes	Yes	Yes	Yes	Unclear	Fair	The trial recruited patients to each arm differently. Supplemented women were recruited before pregnancy and asked to take vitamins and were followed up for several months. This recruitment before exposure and continuous measurement would have meant that all pregnancies and terminations would have been counted. Unsupplemented women were identified at 8–12 weeks of pregnancy, by which time, early pregnancy losses would have occurred (possibly due to lack of folic acid). Because one arm differentially identified women, this could have potentially led to a high and differential risk of selection bias, but the study restricted the analysis for supplemented cases with a pregnancy at 14 weeks. Thus the risk of differential selection bias was reduced but the risk of attrition bias was increased. A second potential source of bias arises from the residual confounding effects of having a higher proportion in the supplemented group of previous fetal deaths and in fact mortality because of congenital abnormalities.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Czeizel et al., 1992[26]; Czeizel et al., 1993 Czeizel et al., 1993[28] Czeizel et al., 1994[29] Czeizel et al., 1994[30] Czeizel et al., 1996[31] Czeizel et al., 1998[32]	Yes	Yes	Yes	Yes	NA	Fair	Does not include fetal death in the analysis (but provides data for that calculation; Unclear allocation concealment and blinding processes (participants blinded, Unclear whether administrators or outcome assessors were blinded); does not conduct ITT but dropout extremely low so risk of bias low from dropout. Study did not consider diet and did not mention how missing data was handled.
Czeizel et al., 1996[33]	Yes	NA	NA	NA	No	Poor	Does not include fetal deaths; study does not control for dietary intake but this study predates food fortification; relevant analysis (supplementation vs. no supplementation in NTD in critical period vs. healthy births) does not control for confounding; does not measure adherence. Unclear on how missing data was handled.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
De Marco et al., 2011[34]	Yes	NA	NA	NA	Unclear	Poor	Well-defined outcomes; exposure is slightly outside our period of interest (i.e., 3 months prior to pregnancy, rather 2 months prior to pregnancy). Possibility of residual confounding because some NTD-specific variables not controlled for, specifically, use of antiepileptic drugs or having or having partner(s) with NTDs. Investigators measured percentages of mothers with previous NTD-affected pregnancies, but relationship to outcomes of interest unlikely (see pg. 1080). In addition, dietary folate intake either not assessed or not taken into account in analyses. Risk of recall bias because mothers completed interviews 18-24 months after childbirth. Unclear how missing data affected findings.
DeSoto et al., 2012[35]	Yes	Yes	NA	NA	Unclear	Poor	Definition of folic acid exposure not defined clearly. Positive response to folic acid supplement use question could have indicated any frequency of usage. High risk of recall bias because mothers asked about FA use 6-13 years after childbirth. In addition, dietary folate not taken into account. Unclear if important confounders related to ASD included in statistical analyses.

First Author, Year		Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Ericson et al., 2001[36]		Yes	Yes	Unclear	Yes	Unclear	Poor	Participation rates not reported; unclear whether the two groups are actually comparable-one group included multivitamin use only but whether folic acid was in vitamins and the amount is unclear; also unclear whether those taking folic acid tablets were or were not also taking multivitamins with folic acid.
Gildestad et al., 2013[37]		Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.
Haberg et al., 1994[38]		Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.
Hernandez et al., 2001[39]		Yes	NA	NA	NA	Unclear	Fair	Possibility of residual confounding because numerous NTD-specific variables not controlled for, specifically, diabetes, family history of NTDs, prior NTD-affected pregnancy, or having or having partner(s) with NTDs. Dietary folate intake not accounted for in analysis of interest. Authors make point in Discussion that unaccounted effects of folate intake would reduce the magnitude of their findings, which were statistically significant. Outcome assessors blind to study hypothesis, but that does not mean they were blind to women's' case-control status.

First Author, Year		Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Kallen et al., 2004[40]		Yes	Yes	Yes	Yes	Unclear	Poor	Exposure period starts before pregnancy and extends to first attendance (in 90% of cases before the end of the first trimester and usually around week 10) with lack of clarity on degree of exposure, difficult to clearly distinguish exposure from non-exposure; authors did not specify definition of exposure clearly (defined as any vs. no supplement), so the degree of adherence is Unclear; although study does not appear to account for fetal deaths, the resulting selection bias would serve to mute rather than exaggerate the effect of FA on twinning; study does not control for dietary intake but this study predates food fortification; study does not control for dietary intake but this study predates food fortification; study controls for use of ovarian stimulation drugs, but this control variable is insufficient because fertility treatment includes several other options

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Kallen et al., 2007[41]	Yes	Yes	Yes	Yes	Unclear	Poor	Response rate Unclear - percentage of women approached who agreed to participate is Unclear. Also, total number of eligible women is Unclear. Women were recruited at the first antenatal visit, but the total number of women presenting for care during the time period is Unclear. Also, the extent of folic acid not ascertained from participants.
Kondo et al., 2015[42]	Yes	NA	NA	NA	No	Poor	Selection bias from being limited to live births. Definition of folic acid exposure not defined clearly. Positive response to folic acid supplement use question could have indicated any frequency of usage. High risk of recall bias because half of the control and case mothers asked about folic acid use 6–12 years after childbirth. In addition, dietary folate not taken into account. Controls and cases not matched on year or place of birth, statistically significant differences in knowledge of FA benefits.
Medvezky et al., 2003[43]	Yes	NA	NA	NA	Unclear	Poor	Unclear how missing data handled. Exposure not well defined or quantified.
Mills et al., 1989[44]	Yes	NA	NA	NA	Unclear	Fair	Eligibility criteria for cases Unclear because they do not clarify whether infants or fetuses were eligible if stillborns or had been aborted. Possibility of residual confounding because numerous NTD-specific variables not controlled for, specifically, diabetes, family history of NTDs, prior NTD-affected pregnancy, having or having partner(s) with NTDs, or treatment with folic acid antagonists. Unclear how high response rates for control groups were. Unclear how missing data was handled.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Moore et al., 2003[45] Milunsky et al., 1989[46]	Yes	Unclear	Yes		Unclear	Fair	Physicians provided 76.5% of the outcome data, but if physicians did not respond, mothers completed the outcome questionnaires; information provided by mothers may not have been entirely accurate in terms of prenatal test results, presence of birth defects or chromosomal abnormalities, complications of pregnancy or delivery, complications of the newborn, or perinatal maternal illnesses. Treatment fidelity not entirely clear, specifically weekly frequency of folic acid supplementation. Possibility of residual confounding because use of folic acid antagonists not taken into account in analyses. Not enough information provided to calculate differential attrition, but overall attrition rate was very low for full sample and therefore unlikely to bias findings
Mosley et al., 2009[47]	Yes	NA	NA	NA	Yes	Fair	Exposure not well defined or quantified. Response rates didn't approached 80% but relevant, without major apparent selection or diagnostic work-up bias.
Mulinare et al., 1988[48]	Yes	NA	NA	NA	No	Poor	Missing data not accounted for in an ITT analysis; covers a 6 month period of exposure (3 months before to 3 months after pregnancy); No controls for concurrent interventions such as exposure to dietary folate, however, this sample was drawn before dietary supplementation; significant risk of recall bias because women asked to remember for a period ranging from 1968-1980. Only still births, no information on terminations. Cases included live an still-borns controls only.
Ohya et al., 2011[49]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Not enough information in the publication to assess quality.

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Shaw et al., 2002[50]	Yes	NA	NA	NA	Unclear	Poor	Unclear how missing data was handle, did not adjust for potential NTD confounders; wide range between those who actually used vitamin (did they use vitamin or mineral supplements during 4 month period but does not say how they divided group out), study does consider cereal usage but not other forms of dietary folate. Potential recall bias - 3.7 to 3.9 years later. Does not mention miscarriages and stillbirths.
Shaw et al., 1995[51]	Yes	NA	NA	NA	Unclear	Fair	3 months before and 3 months after; did not control for all of the confounders we are interested in like folate antagonist medications; does not mention miscarriages and stillbirths. Differences based on ethnicity, age, and education but study controlled for. No mention of how missing data was handled.
Suarez et al., 2000[52]	Yes	NA	NA	NA	Unclear	Fair	1) Response rate less than 80%; 2) some selection bias as controls do not include recruitment at all of the same centers; 3) study collects data on dietary folate intake, but no data shown of association of folic acid with NTDs, adjusted for dietary intake. Strength is adjustment for other confounders. Several issues: (1) very low prevalence of folic acid supplements, limited power (2) food frequency questionnaire doesn't distinguish 3 months prior to conception (3) differential recall period produced by not matching case and control infants/fetuses for gestational age (control women recalling exposures further in past than case women);

							(4) different response rate between case (72%) and control (53%)

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Veeranki et al., 2014[53]	Yes	Yes	Yes	Yes	Unclear	Poor	Exposure to multivitamin supplement defined as filling rather than consuming prenatal vitamins; assumes no misclassification from consumption of over-the-counter supplements; also needed only 1 day of fill in period of exposure to count as exposed; differences between exposed (G1) and non-exposed (G3) in prenatal care, maternal asthma, number of siblings: does not separate out the potential effect of folic acid specifically from other micronutrients that may have an independent effect on respiratory outcomes; does not control for dietary folate exposure; does not control for environmental exposure; excludes stillbirths and miscarriages so risk of selection bias; because of definition of exposure, Unclear whether time period of fill for first trimester covers the first month of pregnancy); loss through poor response rate or missing data NR so cannot judge attrition bias
Veeranki et al., 2014[54]	Unclear	No	Unclear	Yes	Unclear	Poor	Exposure to multivitamin supplement defined as filling rather than consuming prenatal vitamins; assumes no misclassification from consumption of over-the-counter supplements; does not separate out the potential effect of folic acid specifically from other micronutrients that may have an independent effect on allergic rhinitis; does not control for dietary folate exposure; does not control for environmental exposure; excludes stillbirths and miscarriages so risk of selection bias; because of definition of exposure, Unclear whether time period of fill for first trimester covers the first month of

							pregnancy); may not include all cases of allergic rhinitis

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Veeranki et al., 2015[55]	Yes	No	Yes	Yes	Unclear	Poor	Exposure to multivitamin supplement defined as filling rather than consuming prenatal vitamins; assumes no misclassification from consumption of over-the-counter supplements; also needed only 1 day of fill in period of exposure to count as exposed; differences between exposed (G1) and nonexposed (G3) in prenatal care, maternal asthma, number of siblings: does not separate out the potential effect of folic acid specifically from other micronutrients that may have an independent effect on respiratory outcomes; does not control for dietary folate exposure; does not control for environmental exposure; excludes stillbirths and miscarriages and preterm, so risk of selection bias; because of definition of exposure, Unclear whether time period of fill for first trimester covers the first month of pregnancy; loss through poor response rate or missing data NR so cannot judge attrition bias; may not include all cases of asthma
Vollset et al., 2005[56]	Yes	Yes	Yes	Yes	Unclear	Fair	Risk of recall bias in original data assumed to be high - based on estimates of underreporting of folate use, they estimated that 45% of women who took folate before conception were registered as nonusers and adjusted potential misclassification as a result; however, periconceptional use not defined, as a result, cannot tell if exposed vs. non-exposed is clearly defined; looks at pregnancies, but Unclear how stillbirths and terminations were handled; no adjustment for dietary folate; because of lack of definition of exposure, Unclear whether time period of fill for first trimester covers the first month of pregnancy); loss through poor response rate or missing data NR so cannot judge attrition bias

First Author, Year	Were outcomes pre-specified/ defined and adequately described?	Were outcome measures valid and reliable?	Were all important outcomes considered?	Was the duration of follow-up adequate to assess the outcome?	Was an appropriate method used to handle missing data?	Quality Rating	Comments
Werler et al., 1993[57]	Yes	NA	NA	NA	No	Fair	Cases and controls identified by "systematic contact" at tertiary and birth hospitals, details NR, as a result, unable to determine whether cases and controls came from similar populations and had a similar chance of selection; looked at the effect of dietary intake of folate, but only for those with no use of supplements, so does not fully control for concurrent interventions

G1 = group 1; G3 = group 3; G4 = group 4; ITT = intent to treat; NA = not applicable; NR = not reported; NTD = neutral tube defect

eTable 6. Variations in the Effect of Folic Acid Supplementation on Neural Tube Defects by Race and Ethnicity (Key Question 1b)

First Author, Year Study Name	Design	Subgroup	N	Results
Ahrens et al., 2011[12] Slone Birth Defects Study	Case-control	White, non-Hispanic Black, non-Hispanic Hispanic	White, non-Hispanic Cases: 128 Controls: 4,535 Black, non-Hispanic Cases: 22 Controls: 459 Hispanic Cases: 39 Controls: 892	White, non-Hispanic Crude OR (95% CI) Consistent users: 0.78 (0.49–1.25) Early pregnancy initiators: 0.63 (0.38–1.06) Adjusted OR (95% CI) Consistent users: 0.93 (0.56–1.54) Early pregnancy initiators: 0.68 (0.40–1.16) Black, non-Hispanic Crude OR (95% CI) Consistent users: 1.11 (0.34–3.61) Early pregnancy initiators: 0.77 (0.29–2.02) Adjusted OR (95% CI) Consistent users: Not calculated, too few cases Early pregnancy initiators: 0.86 (0.32–2.30) Hispanic Crude OR (95% CI) Consistent users: 1.81 (0.85–3.84) Early pregnancy initiators: 0.61 (0.27–1.38) Adjusted OR (95% CI) Consistent users: 2.20 (0.98–4.92) Early pregnancy initiators: 0.74 (0.32–1.70)

eTable 6. Variations in the Effect of Folic Acid Supplementation on Neural Tube Defects by Race and Ethnicity (Key Question 1b) (continued)

First Author, Year Study Name	Design	Subgroup	N	Results
Mosley et al., 2009[47] National Birth Defects Prevention Study	Case-control	White, black, Hispanic	Anencephaly cases White, non-Hispanic: 83 Black, non-Hispanic: 18 Hispanic: 67 Spina bifida cases White, non-Hispanic: 191 Black, non-Hispanic: 42 Hispanic: 134 Controls White, non-Hispanic: 2,173 Black, non-Hispanic: 431 Hispanic: 865	Anencephaly White, non-Hispanic Crude OR (95% CI) 3 months before pregnancy: 1.2 (0.7–2.1) First month of pregnancy: 1.5 (0.9–2.6) Black, non-Hispanic Crude OR (95% CI) 3 months before pregnancy: 2.8 (0.8–10.4) 1st month of pregnancy: 3.9 (1.3–11.5) Hispanic Crude OR (95% CI) 3 months before pregnancy: 0.7 (0.2–2.2) 1st month of pregnancy: 1.4 (0.8–2.5) Spina bifida White, non-Hispanic Crude OR (95% CI) 3 months before pregnancy: 1.3 (0.9–1.9) 1st month of pregnancy: 1.1 (0.8–1.7) Black, non-Hispanic Crude OR (95% CI) 3 months before pregnancy: 1.2 (0.5–2.8) 1st month of pregnancy: 0.6 (0.3–1.6) Hispanic Crude OR (95% CI) 3 months before pregnancy: 0.4 (0.2–1.2) 1st month of pregnancy: 1.3 (0.9–2.0)

eTable 6. Variations in the Effect of Folic Acid Supplementation on Neural Tube Defects by Race and Ethnicity (Key Question 1b) (continued)

First Author, Year Study Name	Design	Subgroup	N	Results
Shaw et al., 1995[51] California Birth Defects Monitoring Program	Case-control	Hispanic, non-Hispanic white, black, other	Hispanic Cases: 265 Controls: 196 Non-Hispanic white Cases: 217 Controls: 272 Black Cases: 27 Controls: 31 Other Cases: 28 Controls: 39	OR for NTD from maternal use of a folic acid–containing vitamin in 3 months before conception (95% CI) Hispanic: 0.96 (0.44–2.10) Non-Hispanic: 0.62 (0.3 to -1.10) African American: 0.54 (0.09–3.20) Other: 4.3 (0.23–145) OR for NTD from maternal use of a folic acid–containing vitamin in first 3 months postconception OR (95% CI) Hispanic: 0.73 (0.49–1.10) Non-Hispanic: 0.58 (0.36–0.94) African American: 0.29 (0.08–1.10) Other: 1.9 (0.57–6.30)

CI = confidence interval; G = group; N = number; NTD = neural tube defect; OR = odds ratio.

eTable 7. Variations in Effect of Folic Acid Supplementation on Neural Tube Defects by Dosage (Key Question 1c)

First Author, Year	Design	Subgroup	N	Results
Milunsky et al., 1989[46] Moore et al., 2003[45]	Cohort	Dietary folate equivalent groups: 0, 1–399, 400–799, ≥800	Dietary folate equivalents (cases/N total in group) 0: 37/13,431 1–399: 2/2,489 400–799: 2/1,812 ≥800: 8/5,494	Relative risk of NTD (95% CI) dietary folate from supplements dietary folate equivalent/day (weeks 1–5) 1–399: 0.29 (0.07–1.2) 400–799: 0.41 (0.10–1.7) ≥800: 0.56 (0.24–1.3)
Shaw et al., 1995[51] California Birth Defects Monitoring Program	Case-control	Groups of average amount used/day of a folic acid–containing vitamin supplement: Any, <0.4, 0.4–0.9, ≥1.0	Use in 3 months before conception (cases/controls) <0.4: 53/56 0.4–0.9: 29/32 ≥1.0: 5/6 Use in 3 months after conception (cases/controls) <0.4: 37/27 0.4–0.9: 243/322 ≥1.0: 42/33 None: 207/149 Unknown: 4/2	OR for NTDs with maternal use of a folic acid–containing vitamin supplement in the 3 months before conception OR (95%) <0.4: 0.68 (0.43–1.10) 0.4–0.9: 0.65 (0.37–1.20) ≥1.0: 0.60 (0.16–2.30) OR for NTDs with maternal use of a folic acid–containing vitamin supplement 3 months after conception OR (95%) <0.4: 0.99 (0.56–1.80) 0.4–0.9: 0.54 (0.41–0.72) ≥1.0: 0.92 (0.54–1.60)
Werler et al., 1993[57] Slone Birth Defects Study	Case-control	Doses of folic acid supplement: ≥1 mg, 0.5–0.9 mg, 0.4 mg, <0.4 mg	Daily dose: 34/339 Less than daily dose: 41/234	Calculated OR for daily vs. less than daily dose 0.57; 95% CI, 0.35–0.93

eTable 7. Variations in Effect of Folic Acid Supplementation on Neural Tube Defects by Dosage (Key Question 1c) (continued)

First Author, Year	Design	Subgroup	N	Results
Mills et al., 1989[44] National Institute of Child Health and Human Development Neural Tube Defects Study (data from California and Illinois)	Case-control	RDA of folic acid supplements or more vs. none Any amount vs. none	RDA or more NTD cases: 86 Cases with other major malformations: 70 Controls: 84 Less than RDA NTD cases: 15 Cases with other major malformations: 17 Controls: 27 None NTD cases: 464 Cases with other major malformations: 451 Controls: 456	Calculated OR of NTDS with RDA or more vs. less than RDA: 1.84; 95% CI, 0.92–3.71

CI = confidence interval; mg = milligrams; N = number; NTD = neural tube defect; RDA = recommended daily allowance; OR = odds ratio; vs. = versus.

eTable 8. Variations in Effect of Folic Acid Supplementation on Neural Tube Defects by Timing (Key Question 1c)

First Author, Year	Design	Subgroup	N	Results
Ahrens et al., 2011[12] Slone Birth Defects Study	Case-control	Consistent users (4 or more days per week) 2 of 3 periconceptional months vs. initiating in the first month (4 or more days per week)	Spina bifida (cases/controls) Consistent users: 83/2,573 Initiating in the first month: 60/2,293	Calculated OR, 1.23, 95% CI, 0.88–1.73
Mosley et al., 2008[47] National Birth Defects Prevention Study	Case-control	Consistent users 3 months before pregnancy through first month of pregnancy vs. initiating in the first month	Anencephaly (cases/controls) Consistent users: 38/61 Initiating in the first month: 965/948	Anencephaly, calculated OR 0.61; 95% CI, 0.40–0.93
			Spina bifida (cases/controls) Consistent users: 97/100 Initiating in the first month: 965/948	Spina bifida, calculated OR 0.95; 95% CI, 0.71–1.28
Milunsky et al., 1989[46] Moore et al., 2003[45]	Cohort	Women who did not use multivitamins after conception with women who used multivitamins in the first 6 weeks of pregnancy and women who started multivitamin use only after week 6	Use in weeks 1–6 (cases/total N): 10/10,731 Use in weeks 7 and later (cases/total N): 25/7,795	Calculated OR, 0.29, 95% CI, 0.14–0.60
Suarez et al., 2000[52] Texas Department of Health's Neural Tube Defect Project	Case-control	Preconceptional use vs. postconceptional use	Preconceptional use (cases/controls): 8/5 Postconceptional use: 74/85	Calculated OR, 1.84; 95% CI, 0.58–5.86

eTable 8. Variations in Effect of Folic Acid Supplementation on Neural Tube Defects by Timing (Key Question 1c) (continued)

First Author, Year	Design	Subgroup	N	Results
Shaw et al., 1995[51] California Birth Defects Monitoring Program	Case-control	Use in 3 months before conception vs. use in 3 months after conception	Use in 3 months before or after conception (cases/controls) Any use in 3 months before conception: 88/98 Any use in 3 months after conception: 322/384	Calculated OR, 1.07; 95% CI, 0.77–1.48

CI = confidence interval; N = number; OR = odds ratio; vs. = versus.

eTable 9. Variation in Harms of Folic Acid Supplementation by Dosage (Key Question 2b)

First Author, Design	Design	Intervention Groups (n)	Timing of Measurement Exposure	Outcomes	Odds Ratio (95% CI)
Crider, 2013[2] (citing Dunstan, 2012[58])	Systematic review	G1: >0.5 mg/day G2: 0.2–0.499 mg/day G3: <0.2 mg/day	Third trimester	Any allergic disease, sensitization, recurrent wheeze, eczema, food reactions, IgE-mediated food allergy, and sensitization to food allergens	12 associations overlap line of no difference. 2 span line of no difference for G3 vs. G1 (OR: 1.5; 95% CI, 1.0–2.5) and G2 vs. G1 (OR: 1.7; 95% CI, 1.0–2.8) for eczema.

CI = confidence interval; G = group; IgE = immunoglobulin E; mg = milligram; n = number; OR = odds ratio; vs. = versus.

eTable 10. Variation in Harms of Folic Acid Supplementation by Timing (Key Question 2b)

First Author, Design	Design	Intervention Groups (n)	Timing of Measurement Exposure	Outcomes	Relative Risk (95% CI)
Crider, 2013[2]	Systematic review	G1. Second or third trimester use of folic acid (NR) G2: No use (NR)	Second or third trimester	Asthma or wheezing	Wheeze in infants/toddler RR: 1.20 (95% CI: 1.04–1.39) in 1 study[59] 14 other associations for asthma or wheezing not statistically significantly different[59,60]
				Other allergic outcomes	38 associations for other allergic outcomes not statistically significant

CI = confidence interval; G = group; n = number; NR = not reported; RR = relative risk.

eReferences

1. Brown SB, Reeves KW, Bertone-Johnson ER. Maternal folate exposure in pregnancy and childhood asthma and allergy: a systematic review. Nutr Rev. 2014 Jan;72(1):55-64. PMID: 24551950.
2. Crider KS, Cordero AM, Qi YP, et al. Prenatal folic acid and risk of asthma in children: a systematic review and meta-analysis. Am J Clin Nutr. 2013 Nov;98(5):1272-81. PMID: 24004895.
3. Goh YI, Bollano E, Einarson TR, et al. Prenatal multivitamin supplementation and rates of congenital anomalies: a meta-analysis. J Obstet Gynaecol Can. 2006 Aug;28(8):680-9. PMID: 17022907.
4. Wang T, Zhang HP, Zhang X, et al. Is folate status a risk factor for asthma or other allergic diseases? Allergy Asthma Immunol Res. 2015 Nov;7(6):538-46. PMID: 26333700.
5. Wolff T, Witkop CT, Miller T, et al. Folic Acid Supplementation for the Prevention of Neural Tube Defects: An Update of the Evidence for the U.S. Preventive Services Task Force. U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews Rockville, MD: Agency for Healthcare Research and Quality; 2009.
6. Wolff T, Witkop CT, Miller T, et al. Folic acid supplementation for the prevention of neural tube defects: an update of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2009 May 5;150(9):632-9. PMID: 19414843.
7. Yang L, Jiang L, Bi M, et al. High dose of maternal folic acid supplementation is associated to infant asthma. Food Chem Toxicol. 2015 Jan;75:88-93. PMID: 25449200.
8. Abe M. Folic acid supplementation during pregnancy affects food allergy risk in offspring depending on parental allergies. Allergy. 2014;69:152.
9. Abe M. The effect of maternal folic acid intake during pregnancy on incidence of child food allergy. Allergy. 2013;68:559-60.
10. Abe M. The effect of folic acid supplementation before conception on the onset of food allergies in infants. Allergy. 2015;70:266.
11. Agopian AJ, Tinker SC, Lupo PJ, et al. Proportion of neural tube defects attributable to known risk factors. Birth Defects Res A Clin Mol Teratol. 2013 Jan;97(1):42-6. PMID: 23427344.
12. Ahrens K, Yazdy MM, Mitchell AA, et al. Folic acid intake and spina bifida in the era of dietary folic acid fortification. Epidemiology. 2011 Sep;22(5):731-7. PMID: 21659881.
13. Berry RJ, Kihlberg R, Devine O. Impact of misclassification of in vitro fertilisation in studies of folic acid and twinning: modelling using population based Swedish vital records. BMJ. 2005 Apr 9;330(7495):815. PMID: 15722370.
14. Botto LD, Erickson JD, Mulinare J, et al. Maternal fever, multivitamin use, and selected birth defects: evidence of interaction? Epidemiology. 2002 Jul;13(4):485-8. PMID: 12094106.
15. Bower C, Stanley FJ. Dietary folate as a risk factor for neural-tube defects: evidence from a case-control study in Western Australia. Med J Aust. 1989 Jun 5;150(11):613-9. PMID: 2725375.

16. Bower C, Stanley FJ. Periconceptional vitamin supplementation and neural tube defects; evidence from a case-control study in Western Australia and a review of recent publications. J Epidemiol Community Health. 1992 Apr;46(2):157-61. PMID: 1583432.
17. Brescianini S, Cotichini R, Nistico L, et al. Folic acid supplementation before conception and chance of a twin pregnancy: Preliminary results from an italian case-control study. Twin Res Hum Genet. 2012;15(2):177.
18. Carmichael SL, Yang W, Shaw GM. Periconceptional nutrient intakes and risks of neural tube defects in California. Birth Defects Res A Clin Mol Teratol. 2010 Aug;88(8):670-8. PMID: 20740594.
19. Chandler AL, Hobbs CA, Mosley BS, et al. Neural tube defects and maternal intake of micronutrients related to one-carbon metabolism or antioxidant activity. Birth Defects Res A Clin Mol Teratol. 2012 Nov;94(11):864-74. PMID: 22933447.
20. Charles D, Ness AR, Campbell D, et al. Taking folate in pregnancy and risk of maternal breast cancer. BMJ. 2004 Dec 11;329(7479):1375-6. PMID: 15591563.
21. Charles DH, Ness AR, Campbell D, et al. Folic acid supplements in pregnancy and birth outcome: re-analysis of a large randomised controlled trial and update of Cochrane review. Paediatr Perinat Epidemiol. 2005 Mar;19(2):112-24. PMID: 15787886.
22. Taylor CM, Atkinson C, Penfold C, et al. Folic acid in pregnancy and mortality from cancer and cardiovascular disease: further follow-up of the Aberdeen folic acid supplementation trial. J Epidemiol Community Health. 2015 Aug;69(8):789-94. PMID: 25855124.
23. Correa A, Gilboa SM, Botto LD, et al. Lack of periconceptional vitamins or supplements that contain folic acid and diabetes mellitus-associated birth defects. Am J Obstet Gynecol. 2012 Mar;206(3):218.e1-13. PMID: 22284962.
24. Czeizel AE, Vargha P. Periconceptional folic acid/multivitamin supplementation and twin pregnancy. Am J Obstet Gynecol. 2004 Sep;191(3):790-4. PMID: 15467542.
25. Czeizel AE, Dobo M, Vargha P. Hungarian cohort-controlled trial of periconceptional multivitamin supplementation shows a reduction in certain congenital abnormalities. Birth Defects Res A Clin Mol Teratol. 2004 Nov;70(11):853-61. PMID: 15523663.
26. Czeizel AE, Dudas I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. N Engl J Med. 1992 Dec 24;327(26):1832-5. PMID: 1307234.
27. Czeizel AE. Prevention of congenital abnormalities by periconceptional multivitamin supplementation. BMJ. 1993 Jun 19;306(6893):1645-8. PMID: 8324432.
28. Czeizel AE. Controlled studies of multivitamin supplementation on pregnancy outcomes. Ann N Y Acad Sci. 1993 Mar 15;678:266-75. PMID: 8257482.
29. Czeizel AE, Dudas I, Metneki J. Pregnancy outcomes in a randomised controlled trial of periconceptional multivitamin supplementation. Final report. Arch Gynecol Obstet. 1994;255(3):131-9. PMID: 7979565.
30. Czeizel AE, Metneki J, Dudas I. The higher rate of multiple births after periconceptional multivitamin supplementation: an analysis of causes. Acta Genet Med Gemellol (Roma). 1994;43(3-4):175-84. PMID: 8588492.
31. Czeizel AE. Reduction of urinary tract and cardiovascular defects by periconceptional multivitamin supplementation. Am J Med Genet. 1996 Mar 15;62(2):179-83. PMID: 8882400.

32. Czeizel AE. Periconceptional folic acid containing multivitamin supplementation. Eur J Obstet Gynecol Reprod Biol. 1998 Jun;78(2):151-61. PMID: 9622312.
33. Czeizel AE, Toth M, Rockenbauer M. Population-based case control study of folic acid supplementation during pregnancy. Teratology. 1996 Jun;53(6):345-51. PMID: 8910980.
34. De Marco P, Merello E, Calevo MG, et al. Maternal periconceptional factors affect the risk of spina bifida-affected pregnancies: an Italian case-control study. Childs Nerv Syst. 2011 Jul;27(7):1073-81. PMID: 21207040.
35. Desoto MC, Hitlan RT. Synthetic folic acid supplementation during pregnancy may increase the risk of developing autism. J Pediatr Biochem. 2012;2(4):251-61.
36. Ericson A, Kallen B, Aberg A. Use of multivitamins and folic acid in early pregnancy and multiple births in Sweden. Twin Res. 2001 Apr;4(2):63-6. PMID: 11665336.
37. Gildestad T, Nordtveit TI, Nilsen RM, et al. Maternal folic acid and multivitamin supplementation and risk of neural tube defects: a population-based registry study. Eur J Epidemiol. 2013;28(1):S242-S3.
38. Haberg SE. Folic acid supplements in pregnancy and respiratory health in early childhood. J Allergy Clin Immunol. 2009;123(2):S18.
39. Hernandez-Diaz S, Werler MM, Walker AM, et al. Neural tube defects in relation to use of folic acid antagonists during pregnancy. Am J Epidemiol. 2001 May 15;153(10):961-8. PMID: 11384952.
40. Kallen B. Use of folic acid supplementation and risk for dizygotic twinning. Early Hum Dev. 2004 Nov;80(2):143-51. PMID: 15500994.
41. Kallen B. Congenital malformations in infants whose mothers reported the use of folic acid in early pregnancy in Sweden. A prospective population study. Congenit Anom (Kyoto). 2007 Dec;47(4):119-24. PMID: 17988253.
42. Kondo A, Morota N, Date H, et al. Awareness of folic acid use increases its consumption, and reduces the risk of spina bifida. Br J Nutr. 2015 Jul 14;114(1):84-90. PMID: 25999131.
43. Medveczky E, Puho E. Parental employment status and neural-tube defects and folic acid/multivitamin supplementation in Hungary. Eur J Obstet Gynecol Reprod Biol. 2004 Aug 10;115(2):178-84. PMID: 15262352.
44. Mills JL, Rhoads GG, Simpson JL, et al. The absence of a relation between the periconceptional use of vitamins and neural-tube defects. National Institute of Child Health and Human Development Neural Tube Defects Study Group. N Engl J Med. 1989 Aug 17;321(7):430-5. PMID: 2761577.
45. Moore LL, Bradlee ML, Singer MR, et al. Folate intake and the risk of neural tube defects: an estimation of dose-response. Epidemiology. 2003 Mar;14(2):200-5. PMID: 12606886.
46. Milunsky A, Jick H, Jick SS, et al. Multivitamin/folic acid supplementation in early pregnancy reduces the prevalence of neural tube defects. JAMA. 1989 Nov 24;262(20):2847-52. PMID: 2478730.
47. Mosley BS, Cleves MA, Siega-Riz AM, et al. Neural tube defects and maternal folate intake among pregnancies conceived after folic acid fortification in the United States. Am J Epidemiol. 2009 Jan 1;169(1):9-17. PMID: 18953063.
48. Mulinare J, Cordero JF, Erickson JD, et al. Periconceptional use of multivitamins and the occurrence of neural tube defects. JAMA. 1988 Dec 2;260(21):3141-5. PMID: 3184392.

49. Ohya Y, Yonemoto J, Ogata T, et al. Influence of environmental chemicals and drugs taken before and during pregnancy on onset of childhood asthma and eczema. Allergy. 2011;66:554.
50. Shaw GM, Nelson V, Carmichael SL, et al. Maternal periconceptional vitamins: interactions with selected factors and congenital anomalies? Epidemiology. 2002 Nov;13(6):625-30. PMID: 12410002.
51. Shaw GM, Schaffer D, Velie EM, et al. Periconceptional vitamin use, dietary folate, and the occurrence of neural tube defects. Epidemiology. 1995 May;6(3):219-26. PMID: 7619926.
52. Suarez L, Hendricks KA, Cooper SP, et al. Neural tube defects among Mexican Americans living on the US-Mexico border: effects of folic acid and dietary folate. Am J Epidemiol. 2000 Dec 1;152(11):1017-23. PMID: 11117610.
53. Veeranki SP, Gebretsadik T, Dorris SL, et al. Association of folic acid supplementation during pregnancy and infant bronchiolitis. Am J Epidemiol. 2014 Apr 15;179(8):938-46. PMID: 24671071.
54. Veeranki SP, Gebretsadik T, Mitchel EF, et al. Timing of prenatal folic acid supplementation and risk of allergic rhinitis in early childhood. Am J Respir Crit Care Med. 2014;189.
55. Di Renzo GC, Spano F, Giardina I, et al. Iron deficiency anemia in pregnancy. Womens Health (Lond Engl). 2015 Oct 16. PMID: 26472066.
56. Vollset SE, Gjessing HK, Tandberg A, et al. Folate supplementation and twin pregnancies. Epidemiology. 2005 Mar;16(2):201-5. PMID: 15703534.
57. Werler MM, Shapiro S, Mitchell AA. Periconceptional folic acid exposure and risk of occurrent neural tube defects. JAMA. 1993 Mar 10;269(10):1257-61. PMID: 8437302.
58. Dunstan JA, West C, McCarthy S, et al. The relationship between maternal folate status in pregnancy, cord blood folate levels, and allergic outcomes in early childhood. Allergy. 2012 Jan;67(1):50-7. PMID: 21923665.
59. Bekkers MB, Elstgeest LE, Scholtens S, et al. Maternal use of folic acid supplements during pregnancy, and childhood respiratory health and atopy. Eur Respir J. 2012 Jun;39(6):1468-74. PMID: 22034647.
60. Haberg SE, London SJ, Stigum H, et al. Folic acid supplements in pregnancy and early childhood respiratory health. Arch Dis Child. 2009 Mar;94(3):180-4. PMID: 19052032.