BMJ Open

Too many systematic reviews of vitamin D and perinatal outcomes: an overview of systematic reviews

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-032626
Article Type:	Research
Date Submitted by the Author:	27-Jun-2019
Complete List of Authors:	Bialy, Liza ; University of Alberta, Pediatrics Fenton, Tanis; University of Calgary, Department of Community Health Sciences, O'Brien Institute of Public Health, Alberta Children's Hospital Research Institute Shulhan-Kilroy, Jocelyn; University of Alberta, Alberta Research Centre for Health Evidence, Department of Pediatrics Johnson, David; Alberta Children's Hospital, McNeil, Deborah A.; University of Calgary, Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine Hartling, Lisa; University of Alberta, Pediatrics
Keywords:	overview of reviews, vitamin D, perinatal

TOO MANY SYSTEMATIC REVIEWS OF VITAMIN D AND PERINATAL OUTCOMES: AN OVERVIEW OF SYSTEMATIC REVIEWS

Liza Bialy,[1] Tanis R. Fenton,[2,3] Jocelyn Shulhan-Kilroy,[4] David W. Johnson,[5,6] Deborah A. McNeil,[6,7] Lisa Hartling[1,4]

AFFILIATIONS:

[1] Alberta SPOR SUPPORT Unit Knowledge Translation Platform, University of Alberta, Edmonton, Alberta, Canada

[2] Community Health Sciences, Institute of Public Health, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3] Nutrition Services, Alberta Health Services, Calgary, AB, Canada

[4] Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada

[5] Departments of Pediatrics and Physiology and Pharmacology, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[6] Maternal Newborn Child and Youth Strategic Clinical Network, Alberta Health Services, Edmonton, AB, Canada

[7] Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

CORRESPONDING AUTHOR:

Liza Bialy, BSc, MPH, 4-486B Edmonton Clinic Health Academy, Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, 11405-87 Avenue NW, Edmonton, AB, T6G 1C9. Email: lbialy@ualberta.ca

WORD COUNT: 3917

ABSTRACT

Objective: To assess effectiveness of vitamin D supplementation during pregnancy and associations of serum vitamin D levels with perinatal outcomes.

Design: Overview of reviews.

Data Sources: Searches conducted in January 2019: Ovid Medline (1946-), Cochrane Library databases.

Study Selection: Two reviewers independently screened titles and abstracts, and full-texts using pre-defined inclusion criteria: systematic reviews (SRs) evaluating vitamin D supplementation in pregnant women and/or examined the association between serum vitamin D levels reporting at least one pre-defined perinatal. Only SRs with high AMSTAR scores were analysed.

Review Methods: Data were extracted independently by one reviewer and checked by a second. Results were assessed for quality independently by two reviewers using GRADE criteria.

Results: Thirteen SRs were included, synthesizing evidence from 204 unique primary studies. SRs of RCTs with the highest level of evidence showed no significant benefit from vitamin D in terms of preterm birth (high quality), preeclampsia (low quality), gestational diabetes (very low quality), stillbirth (high quality), low birth weight (low quality), cesarean section (high quality), with the exception of small-for-gestational age (low quality) that showed a significant difference. SRs of observational studies showed associations between vitamin D levels and preterm birth (moderate quality), preeclampsia (very low quality), and gestational diabetes (moderate quality). SRs showed mixed results for associations between vitamin D and small-for-gestational age (low and very low quality), low birth weight (very low quality), and cesarean section (very low quality).

Conclusion: There is some evidence from SRs of observational studies for associations between vitamin D serum levels and some outcomes, however SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy for any pre-defined outcome. Credibility of the evidence in this field is compromised by the potential for publication and reporting biases and by promotion of low certainty evidence.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- For perinatal outcomes, we provide a comprehensive summary of the existing evidence for the effectiveness of vitamin D and associations with vitamin D.
- A strength of this overview is the rigorous assessment of both the quality and level of evidence using validated measures (AMSTAR and GRADE) and the separation of observational and intervention studies.
- Due to the lack of efficacy of intervention studies on perinatal outcomes, and the differences in findings of observational versus intervention studies, we were unable to make recommendations for the use of vitamin D during pregnancy.

INTRODUCTION

Vitamin D research is an active area of clinical investigation as numerous studies have examined associations between low vitamin D status (low serum 25-hydroxy vitamin D) and many diseases.[1] The evolution of this research began with observational studies examining associations between vitamin D levels and numerous health outcomes. There is now a growing body of randomized controlled trials (RCTs) assessing the effectiveness of vitamin D as an intervention to improve a variety of health outcomes.

Research in pregnancy examining associations between vitamin D with maternal and infant outcomes has also followed this progression. Early studies in this area suggested that low vitamin D levels were associated with undesirable perinatal outcomes, including gestational diabetes, pre-eclampsia, preterm birth and low birthweight RCTs are now available,[2-6] allowing for examination of whether maternal vitamin D supplementation is effective in improving perinatal outcomes.

Given the extensive number of primary studies available on this topic, a number of systematic reviews (SRs) have been conducted to synthesize the evidence in order to guide practice and recommendations regarding perinatal care. However, the SRs vary in their scope, results, and conclusions which poses a challenge for decision-makers in terms of guiding recommendations for the treatment and management of women during pregnancy. The purpose of this study was to conduct an overview of SRs examining 1) the effectiveness of vitamin D supplementation during pregnancy and 2) the association of serum vitamin D levels with adverse pregnancy outcomes, in order to identify, appraise and summarize the SRs to gather the best available evidence in a

single source[7] and clarify variable findings and conclusions across studies and SRs. Overviews are a useful starting point for decision-makers to understand the evidence underlying a specific topic in order "to inform healthcare decision makers' policy options" to improve practice and identify gaps where additional research is needed.[7] Overviews also provide an evidence map to assist decision makers and clinicians with high level conclusions about the topic area.[7]

METHODS

General approach

To synthesize the available evidence in a way that would be most useful to clinicians and decision-makers we conducted a systematic overview of SRs following established methods.[8] In brief, we conducted a comprehensive search for existing SRs, evaluated the SRs in terms of their quality and recency (January 2019), collated the SR results for pre-specified perinatal outcomes, and graded the quality of available evidence (i.e., the certainty of the findings) using the Cochrane Collaboration and GRADE guidance principles.[9] Included SRs were independently assessed for methodological quality using the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist.[10,11]

Literature search strategy

On October 2, 2017, a research librarian with extensive experience conducting SRs carried out searches in Ovid Medline (1946-January 2019) and Wiley Cochrane Library databases (inception-January 2019): Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE), and the Health Technology Assessment (HTA) Database. Searches combined concepts for pregnancy and vitamin D supplementation with the Canadian Agency for

Drugs and Technologies in Health study design filter for SRs (where applicable).[12] No publication date or language filters were applied. The full search was updated in January 2019. The search strategy is available in Supplementary Table 1. Search results were exported to EndNote X7 (Clarivate Analytics) and duplicates removed prior to screening in EndNote.

Eligibility criteria

We included SRs that 1) evaluated vitamin D supplementation in pregnant women of any gestational or chronological age, and/or 2) examined the effect of vitamin D on adverse pregnancy outcomes or the association between serum vitamin D levels and adverse pregnancy outcomes. We defined a SR as a "synthesis of research evidence in which literature searches, inclusion criteria, and critical appraisal methods were explicitly described."[7] We included SRs where vitamin D was administered in any dose or by any route, in comparison with placebo or other doses/forms of vitamin D supplementation. To be included, SRs had to report at least one of the following predefined maternal or neonatal outcomes: pre-term birth, preeclampsia, gestational diabetes, small for gestational age, still birth, low birth weight, and cesarean section. We excluded primary studies.

Selection

Two reviewers (LB, JS-K) independently screened all titles and abstracts and reviewed the full-text of studies that were identified as potentially eligible using standard eligibility criteria. Reviewers compared results and resolved any discrepancies through discussion; where uncertainty remained decisions were made in discussion with the study team.

Assessment of SR quality

Two reviewers (LB, JS-K) independently assessed the methodological quality of all relevant SRs using the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist.[10,11] This reliable and valid tool consists of 11 items regarding the methodological quality of a systematic review. Reviewers compared assessments for each of the 11 items in the AMSTAR checklist and resolved disagreements through discussion or third-party adjudication. Based on the total AMSTAR score (maximum 11 representing highest quality), we categorized the SRs by quality: low (0-3), medium (4-7), high (8-11).[12] Given the large number of high quality SRs, we focused data extraction and analysis on these.

Data collection

One experienced reviewer (LB) extracted data from the SRs using predefined standard forms developed for this overview. For each SR, review level data were extracted on objectives, publication date, country or origin, funding, search date range, inclusion and exclusion criteria, number of included studies, methods of analysis, and quantitative data on included outcomes. For each outcome present in a SR we abstracted study design, intervention, comparator, effect size, and direction of effect. All data were reviewed for accuracy and completeness by a second reviewer (JS-K).

Analysis

We present and discuss the results by SR for each of our predefined outcomes. We present results based on SRs examining: 1) the effectiveness of vitamin D supplementation (i.e., results from randomized controlled trials), and 2) the association between serum vitamin D levels and

pregnancy outcomes (i.e., results from observational studies). For consistency of rating and based on GRADE recommendations[9] results were converted to risk ratios using the random effects model.

Assessing the level of evidence

To assess the certainty of the results, we graded the quality of evidence presented by each SR for each outcome of interest. We followed recommendations of the GRADE Working Group,[13] and assessed the following key domains: risk of bias, inconsistency, indirectness, imprecision, and publication/reporting bias. For SRs of observational studies, we considered the additional domains of magnitude of effect, dose response relationships, and whether all plausible confounding would reduce an effect.[13] For both interventional and observational designs the GRADE assessment started at high quality of evidence, given the designs were appropriate to address questions of effectiveness and association respectively. Two reviewers (LB, LH) independently conducted GRADE assessments and resolved discrepancies through discussion. GRADEpro software was utilized to calculate overall strength of evidence.[9,14] We also used GRADE guidance to classify clinical importance of the observed effects, i.e. risk ratio of 0.5 to 2.0 were interpreted as not large.

Patient involvement

This research was done without patient or public involvement.

RESULTS

Literature search results and study selection

Figure 1 details the flow of information through the stages of this overview using the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)[15] flow diagram. We identified 233 records from the search after removing duplicates. After title and abstract screening 42 records were identified. Three SRs did not report on any of our predefined outcomes and were excluded[16-18], and one SR was represented by both a Cochrane and journal publication reporting the same data.[19,20] Based on the AMSTAR assessment 25 reviews were categorized as low or medium quality and were not included in the data extraction and outcome assessment. In total 13 SRs were included in the final analysis.

Description of included systematic reviews

The 13 included reviews were published between 2009 and 2018, with a median AMSTAR score of 8 ranging from 8 to 11. (Supplementary Table 2 and 3) The literature search dates for these 13 reviews were between September 2014 and May 2018. All 13 SRs were published in English and were from China[21-23], Canada[24-26], Iran[27], Spain[28], Switzerland[19], United Kingdom[29], United States[30,31], and Thailand[32]. Four SRs included both RCTs and observational studies[23,29-31], 5 included only RCTs[19,24,25,28,32], and 4 included only observational studies.[21,22,26,27] All included SRs with the exception of two[30,31] conducted a meta-analysis. Across the 13 SRs there were 204 unique studies (78 RCTs and 126 observational studies).

None of the SRs explicitly searched for low income or high risk populations, the majority described their populations as generally healthy without pre-existing conditions. Individual study sample sizes ranged from 16 to 12,861. For interventional studies there was a wide range of dosing regimens, daily doses ranged from 200 to 5,000 International Units (IU); weekly doses

from 714 to 50,000 IU; up to 60,000 IU monthly and bolus doses ranging from 35,000 to 1,200,000 (600,000 x 2) IU.

Synthesis of results by outcome for SRs examining the effectiveness of vitamin D

Preterm birth. Five SRs of RCTs[19,23-25,28] examined the effectiveness of vitamin D compared to no treatment/placebo or calcium for prevention of preterm birth. Four SRs found no significant difference in preterm birth rates, while one SR found a significant benefit with vitamin D. However, the quality of evidence varied across SRs (see supplementary table 4 for detailed GRADE assessments). One of the SRs had high quality of evidence[25] while the other four were rated as moderate, low and very low quality. The quality of evidence was rated down for the four SRs due to imprecision, risks of bias, and publication bias. The SR with high quality of evidence showed no significant benefit of vitamin D on prevention of preterm birth (RR 1.00, 95% CI 0.77, 1.30).[25] In subgroup analyses, these findings of no effect on preterm birth were robust, not altered when baseline vitamin D status was low (<30 nmol/L), when only studies at low risk of bias were examined, or when the analysis was limited to generally healthy women.(Table 1)

Table 1: Summary of results from SRs of randomized controlled studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	Level of evidence (GRADE)
PRE-TERM BIRTH					
Bi 2018	11/3,822	0.98 (0.77, 1.26)	33%	- (NR)	moderate
De-Regil / Palacios 2016	3 / 477	0.36 (0.14, 0.93)	10%	+ (0.035)	very low
Perez-Lopez 2015	3 / 384	1.24 (0.59, 2.61)	0%	- (0.56)	very low
Roth 2017	**14 / 3,757**	**1.00 (0.77, 1.30)**	**0%**	**- (0.677)**	**high**
Zhou 2017	6 / 1,687	0.61 (0.34, 1.07)	26%	- (0.09)	low
PREECLAMPSIA					
De-Regil / Palaciosis 2016	2 / 219	0.52 (0.25, 1.05)	0%	- (0.069)	very low
Khaing 2017	3 / 357	0.47 (0.24, 0.89)	0%	+ (0.02)	very low
Newberry 2014	1/504	NR; by group for individual study	NR	+ (n=1)	very low
Perez-Lopez 2015	**3 / 654**	**0.91 (0.45, 1.86)**	**24%**	**- (0.80)**	**low**

Roth 2017	3 / 706	1.09 (0.43, 2.76)	66%	- (0.047)	very low
GESTATIONAL DIABETES					
De-Regil / Palacios 2016	2 / 219	0.43 (0.05, 3.45)	0%	- (0.43)	very low
Perez-Lopez 2015	3 / 384	1.05 (0.60, 1.85)	0%	- (0.86)	very low
Roth 2017	5 / 1,030	0.65 (0.39, 1.08)	45%	- (0.125)	very low
SMALL FOR GESTATIONAL AGE					
Bi 2018	**6 / 1002**	**0.72 (0.52, 0.99)**	**0%**	**+ (0.04)**	**low**
Harvey 2014	2 / 245	NR; by individual study	NR	- (n=2)†	very low
Perez-Lopez 2015	3 / 456	0.77 (0.46, 1.30)	15%	- (0.33)	very low
Roth 2017	5 / 741	0.60 (0.40, 0.90)	0%	- (0.704)	very low
LOW BIRTH WEIGHT					
Bi 2018	4/775	0.52 (0.20, 1.37)	65%	- (NR)	very low
De-Regil / Palacios 2016	3 / 493	0.4 (0.24, 0.67)	4%	+ (0.00048)	very low
Perez-Lopez 2015	4 / 496	0.72 (0.45, 1.17)	0%	- (0.19)	very low
Roth 2017	**7 / 1,156**	**0.74 (0.47, 1.16)**	**47.3%**	**- (0.077)**	**low**
STILLBIRTH					
De-Regil / Palacios 2016	3 / 540	0.35 (0.06, 1.99)	0%	- (0.23)	low
Roth 2017	**16 / 4,606**	**0.75 (0.50, 1.12)**	**0%**	**- (0.858)**	**high**
CESAREAN SECTION					
De-Regil / Palaciosis 2016	2 / 312	0.95 (0.69, 1.31)	12%	- (0.75)	low
Perez-Lopez 2015	4 / 1,028	0.97 (0.81, 1.32)	0%	- (0.75)	low
Roth 2017	**16 / 3,240**	**1.02 (0.93, 1.12)**	**0%**	**- (0.701)**	**high**

* for each outcome the review with the highest level of evidence is presented in bold font

† in absence of pooled data this indicates the number of studies with positive or negative significance

± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of RCTs examined the effectiveness of vitamin D for prevention of preeclampsia.[19,25,28,31,32] The quality of evidence for effectiveness of vitamin D for preeclampsia was low and very low; the four SRs that pooled findings from individual studies showed mixed results (Table 1). The SR that provided the highest level of evidence (classified as low quality) found a non-significant risk ratio of 0.91 (95% CI 0.45, 1.86); this SR was downgraded due to imprecision (that is low numbers of studies, participants, and events) and publication bias (only 3 primary studies).[28]

Gestational diabetes. Three SRs of RCTs examined the effectiveness of vitamin D for prevention of gestational diabetes (Table 1).[19,28] None of the SRs found a significant effect with the use of vitamin D in terms of the occurrence of gestational diabetes. The quality of evidence was very low in all SRs due to high risk of bias of the primary studies that contributed data, imprecision due to small numbers of studies and participants and few events (i.e., occurrences of gestational diabetes) overall, and potential for publication and/or reporting bias.

Small for gestational age. Four SRs of RCTs examined the effectiveness of vitamin D in terms of prevention of infants' birthweights being small for gestational age (Table 1).[24,25,28,29] Three of the SR authors conducted meta-analyses to come up with overall effect estimates, while the authors of one SR chose not to pool due to heterogeneity across the two included studies. None of the SRs that pooled data showed a significant effect. The quality of evidence was low or very low due to risk of bias in the primary studies, imprecision, and publication bias.

Low birth weight. Four SRs of RCTs examined the effectiveness of vitamin D to prevent low birth weight (birthweight <2500 grams) (Table 1).[19,24,25,28] One SR found a significant benefit while the other three SRs showed no difference. The SR with the highest quality of evidence (low) showed no significant difference (RR 0.74, 95% CI 0.47, 1.16).[25]

Stillbirth. Two SRs of RCTs examined the effectiveness of vitamin D to prevent stillbirth (Table 1).[19,25] Neither of the SRs found a significant benefit. The SRs had high and low quality of evidence, respectively. The SR with high quality of evidence found a risk ratio of 0.75 (95% CI 0.50, 1.12).[25]

Cesarean section. Three SRs examined the effectiveness of vitamin D for cesarean sections (Table 1).[19,25,28] The quality of evidence ranged from low to high; none of the SRs found a significant effect. The SR providing high quality of evidence found a risk ratio of 1.02 (95% CI 0.93, 1.12).[25]

Synthesis of results by outcome for SRs examining the observational associations of vitamin D with perinatal outcomes

Preterm birth. Five SRs of observational studies examined the association between vitamin D status and preterm birth (Table 2).[22,23,26,29,31] One SR that examined the association between vitamin D and preterm birth found moderate evidence of an association overall 1.27 (1.08, 1.49).[22] Two SRs categorized using two levels of vitamin D: blood level 25(OH)D <50 nmol/L and <75 nmol/L.[23,26] In both SRs the association was slightly greater for the lower serum vitamin D level. The SR with highest quality of evidence found a significant association with moderate quality evidence for <50 vs. >50 nmol/L 1.13 (95% CI 1.04, 1.23) and non-significant association and low quality evidence for <75 vs. >75 nmol/L 1.10 (95% CI 0.89, 1.35).[23] However, the effect sizes were below the cut-off to be considered clinically important.[13,33]

Table 2: Summary of results for SRs of observational studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	GRADE
PRETERM BIRTH					
Harvey 2014	7 / 1,792	NR; 1 individual study showed significance and 6 others not significant	NR	+ (n=1) ‡ - (n=6)	very low
Newberry 2014	2 / 371	NR; by individual study	NR	+ (n=1)‡ - (n=1)	very low
Qin 2016†	**10 / 10,098**	**1.19 (1.08, 1.31)**	**28%**	**+ (0.004)**	**moderate**
Wei 2013	4 / 1,111	1.27 (1.03, 1.58) [blood level 25(OH)D <50nmol/L]	28%	- (0.03)	very low
		1.05 (0.98, 1.12)	0%	- (0.17)	very low

		[blood level 25(OH)D <75nmol/L]			
Zhou 2017	**16 / 16,996**	**1.13 (1.04, 1.23) [<50 vs >50 nmol/L]**	**45%**	**+ (0.003)**	**moderate**
	15 / 17,122	1.03 (0.98, 1.08) [<75 vs >75 nmol/L]	65%	- (0.29)	low
PREECLAMPSIA					
Chung 2009*	1 / 1,189	5 (1.7, 14.1)	NR	+ (n=1) ‡	very low
Harvey 2014*	4 / 642	0.75 (0.48, 1.19)	80.8%	- (0.001)	very low
Newberry 2014	8 / 4420	NR; by individual study	NR	+ (n=5) - (n=3)	very low
Tabesh 2013	8 / 2,485	2.02 (1.26, 3.23)	53%	+ (0.04)	very low
Wei 2013	**6 / 2,008**	**1.57 (1.21, 2.03) [<50 nmol/L]**	**39%**	**+ (0.0006)**	**low**
	5 / 1,311	1.21 (0.99, 1.46) [<75 nmol/L]	60%	- (0.06)	very low
GESTATIONAL DIABETES					
Harvey 2014	8 / 2,668	NR; by individual study	NR	+ (n=3) ‡ - (n=5)	very low
Lu 2016*	**20 / 16,515**	**1.45 (1.15, 1.83)**	**66.6%**	**+ (0.002)**	**low**
Wei 2013	10 / 4,126	1.12 (1.02, 1.22) [<50 nmol/L]	27%	+ (0.02)	moderate
	8 / 3,840	1.09 (1.03, 1.15) [<75 nmol/L]	28%	+ (0.002)	moderate
SMALL FOR GESTATIONAL AGE					
Harvey 2014	7 / 5,660	NR; by individual study	NR	+ (n=2) ‡ - (n=5)	very low
Newberry 2014	1 / 412	NR; by individual study	NR	NR	very low
Wei 2013	**6 / 6,013**	**1.35 (1.18, 1.54) [<50 nmol/L]**	**15%**	**+ (0.00001)**	**low**
	5 / 2,283	0.99 (0.83, 1.18) [<75 nmol/L]	75%	- (0.92)	very low
LOW BIRTH WEIGHT					
Harvey 2014	3 / 1,676	NR; by individual study	NR	+ (n=1) ‡ - (n=2)	very low
CESAREAN SECTION					
Harvey 2014	6 / 3,277	NR; by individual study	NR	+ (n=2) ‡ - (n=4)	very low

*reviews report odds ratios and insufficient data available to convert to risk ratio
† for each outcome the review with the highest level of evidence is presented in bold font
‡ in absence of pooled data this indicates the number of studies with positive or negative significance
± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of observational studies examined the association between vitamin D status and preeclampsia.[26,27,29-31] Three of the five SRs found a significant association, although

in most cases the effect sizes were below the cut-off to be considered clinically important.[26,27,30] One SR assessed different serum levels of vitamin D and found a larger point estimate for <50 nmol/L compared with <75 nmol/L, although the confidence intervals overlapped, so the difference was not statistically significant.[26] (Table 2) The quality of evidence in all cases was very low for the observational studies that examined the association between vitamin D and preeclampsia, due to inconsistency, imprecision, and publication bias in primary studies.

Gestational diabetes. Three SRs of observational studies provided measures of association for vitamin D status and gestational diabetes.[21,26,29] The SR providing the highest quality of evidence showed moderate quality evidence of a significant association for both serum levels examined <50 nmol/L: 1.12 (95% CI 1.02, 1.22), <75 nmol/L: 1.09 (95% CI 1.03, 1.15).[26] (Table 2) The effect sizes were below the cut-off to be considered clinically important.

Small for gestational age. Three SRs of observational studies examined the association between vitamin D status and small birthweights for gestational age.[26,29,31] The SRs showed mixed findings. One SR included 7 studies but did not pool results as the authors stated there was substantial variation in methodology and exposure;[29] 2 studies showed a significant association while 5 studies showed no significant effect (very low quality of evidence). Another SR only included 1 study and could not pool any results.[31] The highest rated (low quality) SR examined the association for different vitamin D serum levels and found a significant association for <50 nmol/L 1.35 (95% CI 1.18, 1.54), but no significant effect for <75 nmol/L 0.99 (95% CI 0.83, 1.18).[26] (Table 2) For both serum vitamin D levels the effect estimates were small and the quality of evidence was low and very low, respectively.

Low birth weight. Only one SR of observational studies examined the association between vitamin D status and low birth weight.[29] The SR included three studies but did not pool results. One study showed a statistically significant result while two studies had non-significant findings. Overall the quality of evidence for this outcome is very low due to due to inconsistency, imprecision, and publication bias.

Stillbirth. There were no SRs of observational studies that examined the association between vitamin D status and stillbirth.

Cesarean section. Only one SR of observational studies examined the association between vitamin D status and cesarean section.[29] The SR included six studies but did not pool results; the authors chose not to combine due to a multitude of factors such as local policies and physician preferences that influence this outcome. Two studies showed a statistically significant association while four studies had non-significant findings. Overall the quality of evidence for this outcome is very low due to inconsistency, imprecision, and publication bias.

DISCUSSION

This overview of SRs found that most of the SRs of randomized trials were of very low quality primarily due to imprecision, risks of bias, and publication bias. All of the highest quality SRs of randomized trials found no significant benefits of vitamin D supplementation for any of the pre-defined pregnancy related outcomes of interest. The findings from the highest quality observational studies observed associations between vitamin D status and the following

outcomes: preterm birth, pre-eclampsia, gestational diabetes and small for gestational age. Of importance, the effect sizes from these studies were of insufficient magnitudes to be above the cut-off to be considered clinically important.

The differences in findings between the observational studies and the randomized controlled trials indicated that there are other, and likely multiple, factors that are associated with both low serum vitamin D levels and poor health outcomes, causing these apparent associations that were not found to be based on cause and effect relationships by the testing in the randomized trials. These findings suggest that low vitamin D levels or deficiencies may be an indicator or marker of poor health status, co-morbidities,[1] or perhaps an acute phase reactant.[34,35] It is likely that pregnant women with these indicators need more attention and care to optimize health outcomes for them and their offspring and not vitamin D supplementation. The current evidence does not support the use of vitamin D supplementation to improve any of these outcomes.

While there were some suggestions of associations between low vitamin D serum levels and preterm birth, preeclampsia, and gestational diabetes in the observational studies, the effect sizes were smaller than required to be considered clinically important. The quality of this observational evidence was almost all low or very low.[13] However, more applicable to clinical practice are the findings from SRs of randomized controlled trials that examined the effectiveness of vitamin D as a treatment to improve pregnancy outcomes. The systematic reviews of randomized trials that provided the highest quality of evidence showed no effect of vitamin D supplementation in pregnancy for any of the pre-defined outcomes of interest. Overall these findings suggest that even if an association exists between vitamin D levels and health

outcomes, vitamin D supplementation in pregnancy would be unlikely to improve these outcomes.

This study provides a methodologically rigorous and comprehensive synthesis of an extensive body of evidence examining vitamin D and perinatal outcomes. The considerable number of primary studies and systematic reviews underscores the importance of this topic as well as the uncertainty about whether and how to manage vitamin D levels to optimize health outcomes. However, the vast number of SRs on this topic is concerning, particularly those of low quality which may propagate inaccurate or biased results and conclusions. Of note, in our update search that captured the most recent publications up to January 2019, we identified 10 new relevant SRs with only three having a score greater than 7 AMSTAR to be included in the final analysis. Of these 3 new studies there was only one new primary study included. We have provided an in-depth analysis by presenting the results of SRs of randomized controlled trials that evaluated the effectiveness of vitamin D as a treatment to improve perinatal outcomes alongside SRs of observational studies that examined the associations between vitamin D levels and health outcomes. Further, we used GRADE's rigorous and transparent method to assess the quality of the body of evidence which provides essential information about the certainty of the effect estimates in order to reconcile findings across individual studies and reviews.

The evidence contributing to the existing SRs varied widely in design and purpose. Observational studies have been used to examine the association between vitamin D levels and health outcomes, and are appropriate for generating hypotheses for testing in randomized trials. One of the limitations of the existing observational studies and synthesis of the same is that

individual studies may or may not sufficiently adjust for confounding[36] (e.g., health status, calcium intake, and social determinants of health). Further, studies that did adjust for confounding differed extensively in the variables they controlled for. Randomized controlled trials represent the highest level of evidence to assess the effectiveness of an intervention, in part because they address the problem of confounding as randomization is intended to equally distribute both known and unknown confounders. It is well documented that early and observational studies often suggest important relationships that do not exist, and that well designed randomized controlled trials are required to fully understand a phenomenon.[37]

An important limitation in this area of investigation is the possibility of reporting and publication bias. While we focused on the highest quality systematic reviews and most indicated that they planned to investigate publication bias, many could not do so because the number of included studies in a given meta-analysis was too small. There remains the possibility that studies, particularly the earlier published studies, showing significant results are more likely to be published while those with non-significant findings remain unpublished.

Also, the potential for selective outcome reporting is important in this body of literature. It is surprising that outcomes that are either routinely collected or relatively easy to ascertain, such as preterm birth, stillbirth, and gestational diabetes were infrequently reported. Selective outcome reporting occurs if researchers focus their reporting on a significant funding and downplay or do not report non-significant results. For example, the most frequently reported outcome was preterm birth; however, among the 48 studies included in one systematic review only one-third of the primary studies reported this outcome. Important efforts have been made to define core

outcomes sets in the area of perinatal research.[38-40] Future studies should focus on critical outcomes for this field. Researchers should also define their outcomes and analyses a priori, register (and ideally publish) study protocols, and ensure clear and transparent reporting.[41,42] Further, researchers should identify all important confounders and address these adequately through appropriate research designs or analytic approaches to ensure valid findings and permit meaningful pooling of data. Currently the credibility of the body of evidence in this important field is compromised due to the potential for confounding, publication bias, reporting bias, and imprecision arising from low numbers of participants.

CONCLUSIONS

While there is some evidence from SRs of observational studies for an association between maternal vitamin D serum levels and some perinatal outcomes, SRs examining effectiveness from randomized controlled trials showed no effect of vitamin D supplementation in pregnancy for any pre-defined outcomes. Credibility of the evidence in this field is compromised by the potential for publication and reporting biases, as well as residual confounding in the observational studies. The discrepancy between the observational and the randomized trials shows that 25-hydroxy vitamin D is lower among people with adverse outcomes, but supplementation does not alter outcomes. Vitamin D is a marker of adverse outcomes rather than a marker of vitamin D status[1] which shows that 25-hydroxy vitamin D, the indicator of vitamin D status, may be an acute phase reactant. Future studies need to adequately control for potential confounding (e.g., through well-designed randomized trials) and include all critical patient-important outcomes. There are currently over 40 published SRs (many of which are low quality) synthesizing evidence from 204 primary vitamin D studies; further systematic reviews on this

topic are wasteful until significantly more well designed and conducted randomized controlled trials are completed and published.

FIGURE LEGENDS

FIGURE 1: Flow diagram of screening decisions

ACKNOWLEDGEMENTS: Research reported in this publication was funded by the Alberta Health Services' Maternal, Newborn, Child & Youth Strategic Clinical Network™ (MNCY SCN™). Additional support was provided by the Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform which is funded by the Canadian Institutes of Health Research and Alberta Innovates. Dr. Hartling is supported by a Canada Research Chair in Knowledge Synthesis and Translation. The content is solely the responsibility of the authors and does not necessarily represent the views of Alberta Health Services. We would like to thank: Tara Landry, MLIS, for peer reviewing the search strategy; Ms. MacKinna Hauff for article retrieval; Robin Featherstone, MLIS, for developing and running the search; and Dr. Seija Kromm for administrative support.

CONTRIBUTORS: LB, LH, TF, DM and DJ designed the study. LB, LH and JKS selected articles, extracted data and performed the assessment of bias. LB supervised study activities and LH wrote the first draft of the manuscript. All authors critically reviewed and revised the manuscript and approved the final version for publication. The corresponding author attests that all listed authors meet authorship criteria and no others meeting the criteria have been omitted.

within and any related or standalone film submitted (the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. I am the sole author of the Contribution and I am one author signing on behalf of all co-owners of this Contribution.

COMPETING INTERESTS: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

DATA SHARING: The dataset is available from the lead author on request.

PATIENT AND PUBLIC INVOLVEMENT: This research was done without patient or public involvement.

REFERENCES

1. Autier P, Mullie P, Macacu A, et al. Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials. *Lancet Diabetes Endocrinol.* 2017;5(12):986-1004.
2. Wagner CL, Baggerly C, McDonnell S, et al. Post-hoc analysis of vitamin D status and reduced risk of preterm birth in two vitamin D pregnancy cohorts compared with South Carolina March of Dimes 2009-2011 rates. *J Steroid Biochem Mol Biol.* 2016;155(Pt B):245-251.
3. Leffelaar ER, Vrijkotte TG, van Eijsden M. Maternal early pregnancy vitamin D status in relation to fetal and neonatal growth: results of the multi-ethnic Amsterdam Born Children and their Development cohort. *Br J Nutr.* 2010;104(1):108-117.
4. Soheilykhah S, Mojibian M, Rashidi M, et al. Maternal vitamin D status in gestational diabetes mellitus. *Nutr Clin Pract.* 2010;25(5):524-527.
5. Holmes VA, Barnes MS, Alexander HD, et al. Vitamin D deficiency and insufficiency in pregnant women: a longitudinal study. *Br J Nutr.* 2009;102(6):876-881.
6. Bodnar LM, Catov JM, Simhan HN, et al. Maternal vitamin D deficiency increases the risk of preeclampsia. *J Clin Endocrinol Metab.* 2007;92(9):3517-3522.
7. Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? *Syst Rev.* 2013;2:26.
8. Pollock M, Fernandes RM, Becker LA, et al. Chapter V: Overviews of Reviews. Draft version (8 October 2018) for inclusion in: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch V (editors). Cochrane Handbook for Systematic Reviews of Interventions. London: Cochrane.
9. Schünemann H, Brozek J, Guyatt G, et al. editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html. Accessed June 3, 2019.
10. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10.
11. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol.* 2009;62(10):1013-1020.
12. Canadian Agency for Drugs and Technologies in Health (CADTH). Rapid response summary with critical appraisal: Process. 2015; https://www.cadth.ca/about-cadth/what-we-do/products-services/rapid-response-service. Accessed April 19, 2018.
13. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ.* 2004;328(7454):1490.

14. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University, 2015 (developed by Evidence Prime, Inc.). Available from: https://gradepro.org. Accessed May 27, 2019.

15. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg.* 2010;8(5):336-341.

16. Yepes-Nunez JJ, Brozek JL, Fiocchi A, et al. Vitamin D supplementation in primary allergy prevention: Systematic review of randomized and non-randomized studies. *Allergy.* 2017;04:04.

17. Zhang H, Huang Z, Xiao L, et al. Meta-analysis of the effect of the maternal vitamin D level on the risk of spontaneous pregnancy loss. *Int J Gynaecol Obstet.* 2017;138(3):242-249.

18. Christensen N, Sondergaard J, Fisker N, et al. Infant Respiratory Tract Infections or Wheeze and Maternal Vitamin D in Pregnancy: A Systematic Review. *Pediatr Infect Dis J.* 2017;36(4):384-391.

19. De-Regil LM, Palacios C, Lombardo LK, et al. Vitamin D supplementation for women during pregnancy. *Cochrane Database Syst Rev.* 2016(1):CD008873.

20. Palacios C, De-Regil LM, Lombardo LK, et al. Vitamin D supplementation during pregnancy: Updated meta-analysis on maternal outcomes. *J Steroid Biochem Mol Biol.* 2016;164:148-155.

21. Lu M, Xu Y, Lv L, et al. Association between vitamin D status and the risk of gestational diabetes mellitus: a meta-analysis. *Arch Gynecol Obstet.* 2016;293(5):959-966.

22. Qin LL, Lu FG, Yang SH, et al. Does Maternal Vitamin D Deficiency Increase the Risk of Preterm Birth: A Meta-Analysis of Observational Studies. *Nutrients.* 2016;8(5):20.

23. Zhou SS, Tao YH, Huang K, et al. Vitamin D and risk of preterm birth: Up-to-date meta-analysis of randomized controlled trials and observational studies. *J Obstet Gynaecol Res.* 2017;43(2):247-256.

24. Bi WG, Nuyt AM, Weiler H, et al. Association Between Vitamin D Supplementation During Pregnancy and Offspring Growth, Morbidity, and Mortality: A Systematic Review and Meta-analysis. *Jama, Pediatr.* 2018;172(7):635-645.

25. Roth DE, Leung M, Mesfin E, et al. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. *BMJ.* 2017;359:j5237.

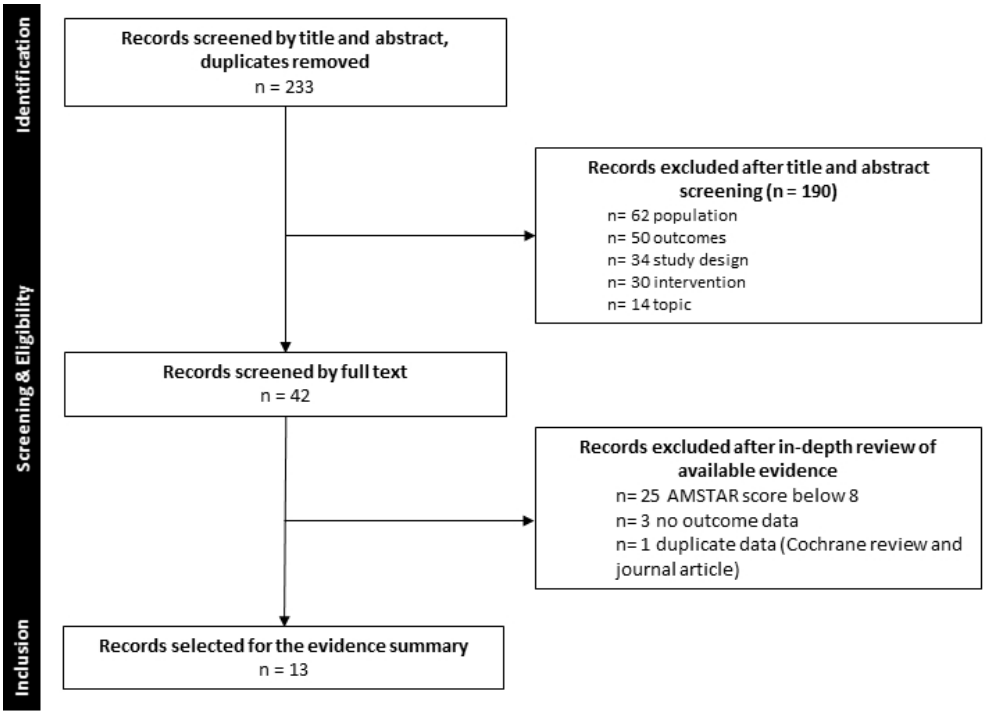
26. Wei SQ, Qi HP, Luo ZC, et al. Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis. *J Matern Fetal Neonatal Med.* 2013;26(9):889-899.

27. Tabesh M, Salehi-Abargouei A, Tabesh M, et al. Maternal vitamin D status and risk of pre-eclampsia: a systematic review and meta-analysis. *J Clin Endocrinol Metab.* 2013;98(8):3165-3173.

28. Perez-Lopez FR, Pasupuleti V, Mezones-Holguin E, et al. Effect of vitamin D supplementation during pregnancy on maternal and neonatal outcomes: a systematic

review and meta-analysis of randomized controlled trials. *Fertil Steril.* 2015;103(5):1278-1288.e1274.

29. Harvey NC, Holroyd C, Ntani G, et al. Vitamin D supplementation in pregnancy: a systematic review. *Health Technol Assess.* 2014;18(45):1-190.
30. Chung M, Balk EM, Brendel M, et al. Vitamin D and calcium: a systematic review of health outcomes. *Evid rep/technol assess.* 2009(183):1-420.
31. Newberry SJ, Chung M, Shekelle PG, et al. Vitamin D and Calcium: A Systematic Review of Health Outcomes (Update). *Evid rep/technol assess.* 2014(217):1-929.
32. Khaing W, Vallibhakara SA, Tantrakul V, et al. Calcium and Vitamin D Supplementation for Prevention of Preeclampsia: A Systematic Review and Network Meta-Analysis. *Nutrients.* 2017;9(10):18.
33. Chen H, Cohen P, Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Commun Stat Simul Comput.* 2010;39:860-864.
34. Madden K, Feldman HA, Chun RF, et al. Critically Ill Children Have Low Vitamin D-Binding Protein, Influencing Bioavailability of Vitamin D. *Ann Am Thorac Soc.* 2015;12(11):1654-1661.
35. Silva MC, Furlanetto TW. Does serum 25-hydroxyvitamin D decrease during acute-phase response? A systematic review. *Nutr Res.* 2015;35(2):91-96.
36. Patel CJ, Manrai AK. Development of exposome correlation globes to map out environment-wide associations. *Pac Symp Biocomput.* 2015:231-242.
37. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol.* 2011;64(12):1277-1282.
38. Molloy EJ, Gale C, Marsh M, et al. Developing core outcome set for women's, newborn, and child health: the CROWN Initiative. *Pediatr Res.* 2018;84(3):316-317.
39. Devane D, Begley CM, Clarke M, et al. Evaluating maternity care: a core set of outcome measures. *Birth.* 2007;34(2):164-172.
40. van 't Hooft J, Duffy JM, Daly M, et al. A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol.* 2016;127(1):49-58.
41. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg.* 2014;12(12):1495-1499.
42. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg.* 2012;10(1):28-55.



190x213mm (96 x 96 DPI)

Supplementary Table 1: Literature search strategy

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Date conducted: 2 October 2017
Strategy:

1 Preconception Care/ (1917)
2 exp Pregnancy/ (855216)
3 exp Pregnancy Complications/ (405775)
4 Pregnant Women/ (6515)
5 Prenatal Care/ (24637)
6 Prenatal Diagnosis/ (35834)
7 (antenatal* or pre-natal* or prenatal*).tw,kf. (120088)
8 (expect* adj2 (female? or mother? or wom#n)).tw,kf. (3728)
9 ((1* or first*) adj2 (tri-mester* or trimester*)).tw,kf. (23473)
10 (pre-conception* or preconception*).tw,kf. (4573)
11 pregnan*.tw,kf. (477757)
12 or/1-11 [Combined MeSH & text words for pregnancy] (1016791)
13 exp Vitamin D/ (54287)
14 Vitamin D Deficiency/ (13412)
15 calcidiol*.tw,kf. (397)
16 calciol*.tw,kf. (20)
17 calcifediol*.tw,kf. (128)
18 cholecalciferol*.tw,kf. (2377)
19 hydroxycholecalciferol*.tw,kf. (1377)
20 hydroxyvitamin D*.tw,kf. (12499)
21 (vitamin D or vitamin D3 or vitamin D$2).tw,kf. (60203)
22 or/13-21 [Combined MeSH & text words for vitamin D] (79566)
23 and/12,22 [Combined concepts for pregnancy & vitamin D] (4365)
24 meta-analysis.pt. (87537)
25 meta-analysis/ or systematic review/ or meta-analysis as topic/ or "meta analysis (topic)"/ or "systematic review (topic)"/ or exp technology assessment, biomedical/ (113240)
26 ((systematic* adj3 (review* or overview*)) or (methodologic* adj3 (review* or overview*))).ti,ab,kf,kw. (127445)
27 ((quantitative adj3 (review* or overview* or synthes*)) or (research adj3 (integrati* or overview*))).ti,ab,kf,kw. (8559)
28 ((integrative adj3 (review* or overview*)) or (collaborative adj3 (review* or overview*)) or (pool* adj3 analy*)).ti,ab,kf,kw. (19993)
29 (data synthes* or data extraction* or data abstraction*).ti,ab,kf,kw. (20992)
30 (handsearch* or hand search*).ti,ab,kf,kw. (7877)
31 (mantel haenszel or peto or der simonian or dersimonian or fixed effect* or latin square*).ti,ab,kf,kw. (21571)
32 (met analy* or metanaly* or technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf,kw. (7548)
33 (meta regression* or metaregression*).ti,ab,kf,kw. (5904)

34 (meta-analy* or metaanaly* or systematic review* or biomedical technology assessment* or bio-medical technology assessment*).mp,hw. (217154)
35 (medline or cochrane or pubmed or medlars or embase or cinahl).ti,ab,hw. (161733)
36 (cochrane or (health adj2 technology assessment) or evidence report).jw. (18083)
37 (meta-analysis or systematic review).mp. [sic – changed .md. to .mp] (200672)
38 (comparative adj3 (efficacy or effectiveness)).ti,ab,kf,kw. (10906)
39 (outcomes research or relative effectiveness).ti,ab,kf,kw. (7938)
40 ((indirect or indirect treatment or mixed-treatment) adj comparison*).ti,ab,kf,kw. (1649)
41 or/24-40 [CADTH SR search filter | Retrieved from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#syst] (358374)
42 and/23,41 [SR filter applied] (187)
43 remove duplicates from 42 (164)

Database: Wiley Cochrane Library
Date conducted: 2 October 2017
Strategy:
#1 [mh ^"Preconception Care"] 103
#2 [mh Pregnancy] 5760
#3 [mh "Pregnancy Complications"] 9364
#4 [mh ^"Pregnant Women"] 156
#5 [mh ^"Prenatal Care"] 1332
#6 [mh ^"Prenatal Diagnosis"] 380
#7 (antenatal* or "pre-natal*" or prenatal):ti,ab,kw 6295
#8 (expect* near/2 (female? or mother? or wom?n)):ti,ab,kw 243
#9 ((1* or first*) near/2 ("tri-mester*" or trimester*)):ti,ab,kw 4141
#10 ("pre-conception*" or preconception*):ti,ab,kw 307
#11 pregnan*:ti,ab,kw 36386
#12 {or #1-#11} 38579
#13 [mh "Vitamin D"] 2941
#14 [mh ^"Vitamin D Deficiency"] 617
#15 calcidiol*:ti,ab,kw 46
#16 calciol*:ti,ab,kw 1
#17 calcifediol*:ti,ab,kw 475
#18 cholecalciferol*:ti,ab,kw 1208
#19 hydroxycholecalciferol*:ti,ab,kw 338
#20 "hydroxyvitamin D*":ti,ab,kw 1931
#21 ("vitamin D" or "vitamin D3" or "vitamin D?"):ti,ab,kw 6774
#22 {or #13-#21} 7581
#23 #11 and #22 354
#24 #11 and #22 in Cochrane Reviews (Reviews and Protocols), Other Reviews and Technology Assessments 14

Supplementary Table 2: Description of included systematic reviews

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Bi Canada May 2018	24 RCTs 5,405 (30 – 965)	Population was healthy, pregnant women without prior vitamin D supplementation of more than 400 IU/d	Vitamin D in the form of cholecalciferol in 22 RCTs and in the form of ergocalciferol in 3 RCTs daily doses: 800 - 5000; weekly doses 35000 or 50000; fortnightly dose 50000; monthly dose 60000; bimonthly dose 60000; and bolus doses 60000 - 200 000	Placebo, no intervention or other dose of vitamin D	Primary: small for gestational age (indicated by birthweight less than the 10th percentile for gestational age, fetal or neonatal mortality Secondary: neonatal (25[OH]D) levels, congenital malformation, admission to a neonatal intensive care unit (NICU), Apgar scores, neonatal calcium levels, birth weight, low birth weight, gestational age, preterm birth, infant growth, asthma, respiratory infection, eczema, and allergy
Khaing Thailand October 2017 Roth Canada September 2017	19 RCTs 28,000 (30 – 9,178) 43 RCTs 8,406 (16 – 1,134)	Pregnant women of any gestational age Participants were pregnant at enrolment or enrolled before pregnancy and then followed-up in pregnancy	Calcium, vitamin D, combined calcium and vitamin D Vitamin D vs. placebo = 3; Calcium + vitamin D vs. calcium = 1 Vitamin D2 or D3, alone or in combination provided the co-intervention is similar in at least one other trial arm Daily doses: 400 – 5000; weekly doses: 714 – 7543; monthly doses: 1645 – 3289; bolus doses: 60000 – 1200000 (600000 x 2)	Placebo, a standard supplementation (e.g., folic acid), or no supplementation Placebo, no vitamin D, or vitamin D up to 600 IU/day (or a less frequent dose that would be about equivalent to 600 IU/day—for example, 4200 IU/week)	Primary: preeclampsia, eclampsia, proteinuria (dipstick urine 2+ or '300 mg/24 h), end-organ dysfunction, or utero-placental dysfunction after 20 weeks of gestation Primary: 25 OHD, preeclampsia, gestational diabetes, gestational hypertension, intra-uterine death, stillbirth, c-section, weight gain, preterm labor, death, adverse events, hospitalizations, birth weight, birth length, head circumference, low birth weight, small for gestational age, gestational age at birth, congenital malformations, neonatal death, respiratory infection, asthma, bone mineral content and density

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Zhou China June 2016	6 RCTs; 9 prospective cohort; 4 nested case-control; 2 cross-sectional; 2 retrospective cohort; 1 case-control 28,391 (50 – 12,861)	Pregnant women without HIV infection	maternal serum 25-OHD or oral supplementation with vitamin D Daily doses of 1,000 to 4,000 IU; weekly doses of 400 daily for 9 weeks; 50,000 for 6 weeks; one time doses starting 60,000 or 2-4 doses of 120,000	no supplementation /placebo, or routine care (ferrous sulfate and calcium, but no vitamin D)	Preterm birth
Qin China August 2015	4 Prospective cohort; 4 Nested case-control; 1 case-control; 1 Retrospective cohort; 1 Cross-sectional 20,608 (134 – 12,861)	Pregnant women without pre-chronic disease or HIV infection, with singleton gestation	NR; measurement of maternal vitamin D levels		Preterm birth
Lu China February 2015	4 Case-control; 7 Cohort; 2 Cross sectional; 7 Nested case control 16,515 (122 – 4,090)	NR	NR; measurement of maternal vitamin D levels		Gestational diabetes
De-Regil / Palacios Switzerland / Puerto Rico February 2015	15 RCTs 2,833 (40 – 990)	Pregnant women of any gestational or chronological age, parity (number of births) and number of fetuses	Vitamin D daily doses: 200 - 2000 Vitamin D single dose: 200,000 – 600,000, and 35,000	No intervention / placebo	Primary: pre-eclampsia, gestational diabetes, vitamin D concentration, adverse effects, preterm birth, low birthweight Secondary: impaired glucose tolerance, c-section, gestational hypertension, maternal death, birth length, head circumference at birth, birthweight, admission to special care, stillbirth, neonatal death, very preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Newberry USA September 2014	2 RCTs; 2 prospective cohorts; 5 nested case-control 4,912 (160 – 1,141)	Primary population of interest is generally healthy people with no known disorders Only including studies for population contributing to pregnancy related outcomes	Vitamin D single doses (for RCT): 2000, 4000 followed by 1 month run-in at 2000	All participants enrolled into one of two vitamin D groups	Preeclampsia, preterm birth, small for gestational age
Perez-Lopez Spain March 2014	13 RCTs 2,299 (40 – 400)	Pregnant women of any gestational or chronologic age and parity, without previous disease history	Vitamin D alone vs. no treatment (placebo); vitamin D + calcium vs. no treatment (placebo); and vitamin D + calcium vs. calcium Daily doses ranged from 400 to 1,000; weekly doses ranged from 35,000 to 50,000; and single doses ranged from 200,000 to 600,000	Active controls, usual treatment without active control, and placebo	Primary: circulating 25-OHD, preeclampsia, gestational diabetes, small for gestational age, low birth weight, preterm birth, birthweight Secondary: birth length, c-section,
Wei Canada October 2012	13 Case-control; 8 cohort; 2 cross-sectional 12,898 (95 – 3,730)	Pregnant women without pre-existing chronic disease or HIV infection	NR; measurement of maternal vitamin D levels		Preeclampsia, gestational diabetes, preterm birth, small for gestational age
Harvey UK June 2012	17 Case-control; 48 cohort/cross-sectional; 9 RCT; 2 intervention studies (non-randomized) NR	Pregnant women or pregnant women and their offspring	vitamin D status [dietary intake, sunlight exposure, circulating 25(OH)D concentration] or supplementation of	For intervention studies: no intervention or placebo	Primary: neonatal hypocalcaemia, rickets in the offspring, offspring bone mass, and maternal osteomalacia Secondary: offspring body composition; offspring preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
			participants with vitamin D or food containing vitamin D (e.g. oily fish)		and later offspring health outcomes; maternal quality of life
Tabesh Iran December 2012	2 Cohort; 4 cross-sectional; 9 case-control 2,936 (32 – 697)	Normal pregnant women	NR; measurement of maternal vitamin D levels		Preeclampsia
Chung USA April 2009	60 RCT; 3 NRCT; 102 cohort or nested case-control; 11 SR NR	Generally healthy people with no known disorders	Vitamin D supplements (no analogues), calcium supplements, and combinations of supplements; food based interventions	NR	Pregnancy-related: preeclampsia, high blood pressure with or without proteinuria, preterm birth or low birth weight, infant mortality

Supplementary Table 3. AMSTAR score by category and individual systematic review

Review	AMSTAR question											
	Q1 A priori design provided	Q2 Duplicate study selection and data extraction	Q3 Comprehensive literature search	Q4 Publication status as inclusion criterion	Q5 List of studies (include and exclude) provided	Q6 Characteristics of the included studies provided	Q7 Quality assess-ment	Q8 Quality used appropriate	Q9 Methods used to combine appropriate	Q10 Publication bias assessed	Q11 Conflict of interest stated	Total
OVERALL HIGH QUALITY												
Bi 2018	n	y	y	n	n	y	y	y	y	y	y	8
Christensen 2017	y	y	y	y	n	y	y	y	y	n	y	9
Chung 2009	y	ca	y	y	y	y	y	y	y	ca	y	9
De-Regil 2016	y	y	y	y	y	y	y	y	y	y	y	11
Harvey 2014	y	y	y	y	n	y	y	y	y	ca	y	9
Khaing 2017	y	y	n	n	n	y	y	y	y	y	y	8
Lu 2016	y	y	y	n	n	y	y	y	y	y	y	9
Newberry 2014	y	ca	y	n	y	y	y	y	y	ca	y	8
Palacios 2016	y	y	y	y	y	y	y	y	y	y	y	11
Perez-Lopez 2015	y	y	y	y	n	y	y	y	ca	ca	y	8
Qin 2016	n	y	y	n	n	y	y	y	y	y	y	8
Roth 2017	y	y	y	y	n	y	y	y	y	ca	y	9
Tabesh 2013	y	y	y	y	n	y	n	n	y	y	y	8
Wei 2013	n	y	y	n	n	y	y	y	y	y	y	8
Yepes-Nunez 2017	n	y	y	y	y	y	y	y	y	y	y	10
Zhang 2017	n	y	y	n	n	y	y	y	y	y	y	8
Zhou 2017	n	y	y	n	n	y	y	y	y	y	y	8
OVERALL MEDIUM AND LOW QUALITY												
Aghajafari 2013	n	y	y	ca	n	y	n	ca	y	y	n	5
Amegah 2017	n	y	y	n	n	y	y	y	y	y	n	6
Amraei 2018	ca	y	y	n	n	y	ca	ca	y	y	y	6
Arain 2015	n	y	ca	n	n	y	n	ca	ca	n	n	2
Chen 2017	n	y	y	n	n	y	y	y	y	y	n	6
Christensen 2012	n	y	n	n	n	y	n	n	n	n	y	3
Fu 2017	n	ca	y	n	n	n	n	n	y	y	y	4
Galthen-Sorensen 2014	n	y	y	n	n	y	y	y	n	n	n	5
Hu 2018	n	y	y	y	n	y	n	ca	y	y	y	7
Hypponen 2014	n	ca	y	n	n	y	ca	n	y	y	y	5
Kamudoni 2016	n	ca	y	y	n	y	n	n	n	n	y	4
Mahomed 2009	y	n	y	y	y	y	ca	ca	n	n	y	6
Martinez-Dominquez 2018	n	ca	y	n	n	y	y	ca	y	y	y	6
Nassar 2011	y	ca	n	n	n	y	n	n	y	n	y	4

Poel 2012	n	ca	y	n	n	y	n	n	ca	y	y	4
Purswani 2017	n	y	n	n	n	y	n	y	y	n	y	5
Santamaria 2018	n	y	n	n	n	y	y	y	y	ca	y	6
Senti 2012	n	y	y	n	n	y	n	n	n	n	y	4
Serrano-Diaz 2018	n	n	y	y	n	y	ca	n	y	y	y	6
Thorne-Lyman 2012	n	n	y	n	n	y	y	y	y	n	n	5
Van der Pligt 2018	n	y	y	n	n	y	y	y	n	n	y	6
Wei 2016	n	y	y	n	n	y	y	ca	y	y	n	6
Yang 2015	n	y	y	n	y	y	y	n	y	y	n	7
Zhang 2018	n	ca	y	ca	n	y	y	y	y	y	y	7
Zhang 2015	n	n	y	n	n	y	y	y	y	y	y	7

* One point was awarded for each item that scored 'yes' (y) and summed for the total score

* 'n' no; 'ca' can't answer

Supplementary Table 4: GRADE tables

Grade Assessments for Preterm Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
11 Bi	RCT	serious	not serious	not serious	not serious	none	moderate
3 De-Regil/Palacios	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
14 Roth	RCT	not serious	not serious	not serious	not serious	none	high
6 Zhou	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low

Grade Assessments for Preeclampsia in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil / Palaciosis	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
3 Khaing	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
1 Newberry	RCT	serious	serious	not serious	serious	publication bias strongly suspected	very low
3 Perez-Lopex	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low
3 Roth	RCT	serious	serious	not serious	serious	publication bias strongly suspected	very low

Grade Assessments for Gestational Diabetes in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low

5 Roth	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low

Grade Assessments for Low Birth Weight in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
4 Bi	RCT	serious	serious	not serious	serious	none	very low
3 De-Regil/ Palaciosis	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
4 Perez-Lopez	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
7 Roth	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low

Grade Assessments for Small for Gestational Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
6 Bi	RCT	serious	not serious	not serious	serious	none	low
2 Harvey	RCT	serious	serious	not serious	serious	none	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low
5 Roth	RCT	serious	not serious	not serious	serious	publication bias strongly suspected	very low

Grade Assessments for Still Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for C-Section Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low
4 Perez-Lopez	RCT	not serious	not serious	not serious	serious	publication bias strongly suspected	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for Preterm Birth in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
2 Newberry	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
10 Qin	OBS	not serious	not serious	serious	not serious	none	moderate
4 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	serious	publication bias strongly suspected	very low
4 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	serious	publication bias strongly suspected	very low
16 Zhou [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
17 Zhou [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	none	low

Grade Assessments for Preeclampsia in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
1 Chung	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
4 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
8 Newberry	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
8 Tabesh	OBS	serious	serious	serious	not serious	none	very low
6 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	publication bias strongly suspected	low
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	publication bias strongly suspected	very low

Grade Assessments for Gestational Diabetes in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
8 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
20 Lu	OBS	not serious	serious	serious	not serious	none	low
10 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
8 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate

Grade Assessments for Low Birth Weight in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low

Grade Assessments for Small for Gestational Age in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
1 Newberry	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low
6 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	publication bias strongly suspected	low
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	publication bias strongly suspected	very low

Grade Assessments for Small for C-Section in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	publication bias strongly suspected	very low

BMJ Open

No high quality evidence supports vitamin d supplementation to improve pregnancy/perinatal outcomes: an overview of 42 systematic reviews

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-032626.R1
Article Type:	Research
Date Submitted by the Author:	18-Nov-2019
Complete List of Authors:	Bialy, Liza ; University of Alberta, Pediatrics Fenton, Tanis; University of Calgary, Department of Community Health Sciences, O'Brien Institute of Public Health, Alberta Children's Hospital Research Institute Shulhan-Kilroy, Jocelyn; University of Alberta, Alberta Research Centre for Health Evidence, Department of Pediatrics Johnson, David; Alberta Children's Hospital, McNeil, Deborah A.; University of Calgary, Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine Hartling, Lisa; University of Alberta, Pediatrics
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Evidence based practice
Keywords:	overview of reviews, vitamin D, perinatal

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

NO HIGH QUALITY EVIDENCE SUPPORTS VITAMIN D SUPPLEMENTATION TO IMPROVE PREGNANCY/PERINATAL OUTCOMES: AN OVERVIEW OF 42 SYSTEMATIC REVIEWS

Liza Bialy,[1] Tanis R. Fenton,[2,3] Jocelyn Shulhan-Kilroy,[4] David W. Johnson,[5,6] Deborah A. McNeil,[6,7] Lisa Hartling[1,4]

AFFILIATIONS:

[1] Alberta SPOR SUPPORT Unit Knowledge Translation Platform, University of Alberta, Edmonton, Alberta, Canada

[2] Community Health Sciences, Institute of Public Health, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3] Nutrition Services, Alberta Health Services, Calgary, AB, Canada

[4] Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada

[5] Departments of Pediatrics and Physiology and Pharmacology, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[6] Maternal Newborn Child and Youth Strategic Clinical Network, Alberta Health Services, Edmonton, AB, Canada

[7] Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

CORRESPONDING AUTHOR:

Liza Bialy, BSc, MPH, 4-486B Edmonton Clinic Health Academy, Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, 11405-87 Avenue NW, Edmonton, AB, T6G 1C9; lbialy@ualberta.ca

WORD COUNT: 4188

ABSTRACT

Objective: To review the evidence to assess effectiveness of vitamin D supplementation during pregnancy and associations of serum vitamin D levels with perinatal outcomes.

Design: Overview of systematic reviews.

Data Sources: Searches conducted in January 2019: Ovid Medline (1946-), Cochrane Library databases.

Eligibility criteria for selecting studies: Two reviewers independently screened titles and abstracts, and full-texts using pre-defined inclusion criteria: systematic reviews (SRs) evaluating vitamin D supplementation in pregnant women and/or examining the association between serum vitamin D levels reporting at least one pre-defined perinatal outcome. Only SRs with high AMSTAR scores were analysed.

Data extraction and synthesis: Data were extracted independently by one reviewer and checked by a second. Results were assessed for quality independently by two reviewers using GRADE criteria.

Results: Thirteen SRs were included, synthesizing evidence from 204 unique primary studies. SRs of RCTs with the highest level of evidence showed no significant benefit from vitamin D in terms of preterm birth [RR 1.00 (95% CI 0.77, 1.30); high quality], preeclampsia [RR 0.91 (0.45, 1.86); low quality], gestational diabetes [RR 0.65 (0.39, 1.08); very low quality], stillbirth [RR 0.75 (0.50, 1.12); high quality], low birth weight [RR 0.74 (0.47, 1.16); low quality], cesarean section [RR 1.02 (0.93, 1.12); high quality]. A significant difference was found for small-for-gestational age [RR 0.72 (0.52, 0.99); low quality]. SRs of observational studies showed associations between vitamin D levels and preterm birth [RR 1.19 (1.08, 1.31); moderate quality], preeclampsia [RR 1.57 (1.21, 2.03) for 25 (OH)D <50 nmol/L subgroup; low quality],

gestational diabetes [RR 1.12 (1.02, 1.22) for 25 (OH)D <50 nmol/L and RR 1.09 (1.03, 1.15) <75 nmol/L; moderate quality], and small-for-gestational age [RR 1.35 (1.18, 1.54) <50 nmol/L; low quality]. SRs showed mixed results for associations between vitamin D and low birth weight (very low quality) and cesarean section (very low quality).

Conclusion: There is some evidence from SRs of observational studies for associations between vitamin D serum levels and some outcomes, however SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy with the exception of one pre-defined outcome, which had low quality evidence. Credibility of the evidence in this field is compromised by study limitations (particularly the possibility of confounding among observational studies), inconsistency, imprecision, and potential for reporting and publication biases.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- We provide a comprehensive summary of the existing evidence for the effectiveness and associations of vitamin D and perinatal outcomes.
- A strength of this overview is the rigorous assessment of the quality of evidence using validated measures (AMSTAR and GRADE).
- The sparsity of high quality evidence for specific outcomes at the primary and systematic review levels currently limits the ability to make strong recommendations for the use of vitamin D during pregnancy.

INTRODUCTION

Vitamin D research is an active area of clinical investigation as numerous studies have examined associations between low vitamin D status (low serum 25-hydroxy vitamin D) and many diseases.[1] The evolution of this research began with observational studies examining associations between vitamin D levels and numerous health outcomes. There is now a growing body of randomized controlled trials (RCTs) assessing the effectiveness of vitamin D as an intervention to improve a variety of health outcomes.

Research in pregnancy examining associations between vitamin D with maternal and infant outcomes has also followed this progression. Early studies in this area suggested that low vitamin D levels were associated with undesirable perinatal outcomes, including gestational diabetes, pre-eclampsia, preterm birth and low birthweight. RCTs are now available,[2-6] allowing for examination of whether maternal vitamin D supplementation is effective in improving perinatal outcomes.

Given the extensive number of primary studies available on this topic, a number of systematic reviews (SRs) have been conducted to synthesize the evidence in order to guide practice and recommendations regarding perinatal care. However, the SRs vary in their scope, results, and conclusions which poses a challenge for decision-makers in terms of guiding recommendations for the treatment and management of women during pregnancy. Overviews are a useful starting point for decision-makers to understand the evidence underlying a specific topic in order "to inform healthcare decision makers' policy options" to improve practice and identify gaps where additional research is needed.[7] Overviews also provide an evidence map to assist decision

makers and clinicians with high level conclusions about the topic area.[7] The purpose of this study was to conduct an overview of SRs examining 1) the effectiveness of vitamin D supplementation during pregnancy and 2) the association of serum vitamin D levels with adverse pregnancy outcomes. We sought to identify, appraise and summarize existing SRs to gather the best available evidence in a single source[7] and clarify variable findings and conclusions across studies and SRs.

METHODS

General approach

To synthesize the available evidence in a way that would be most useful to clinicians and decision-makers we conducted a systematic overview of SRs following established methods.[8] In brief, we conducted a comprehensive search for existing SRs (January 2019), evaluated the SRs in terms of their quality and recency, collated the SR results for pre-specified perinatal outcomes, and graded the quality of available evidence (i.e., the certainty of the findings) using the Cochrane Collaboration and GRADE (Grading of Recommendations Assessment, Development and Evaluation) guidance principles.[9] Included SRs were independently assessed for methodological quality using the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist.[10,11]

Literature search strategy

On October 2, 2017, a research librarian with extensive experience conducting SRs carried out searches in Ovid Medline (1946-January 2019) and Wiley Cochrane Library databases (inception-January 2019): Cochrane Database of Systematic Reviews, Database of Abstracts of

Reviews of Effects (DARE), and the Health Technology Assessment (HTA) Database. Searches combined concepts for pregnancy and vitamin D supplementation with the Canadian Agency for Drugs and Technologies in Health study design filter for SRs (where applicable).[12] No publication date or language filters were applied. The full search was updated in January 2019. The search strategy is available in Supplementary Table 1. Search results were exported to EndNote X7 (Clarivate Analytics) and duplicates removed prior to screening in EndNote.

Eligibility criteria

We included SRs that 1) evaluated vitamin D supplementation in pregnant women of any gestational or chronological age, and/or 2) examined the effect of vitamin D on adverse pregnancy outcomes or the association between serum vitamin D levels and adverse pregnancy outcomes. We defined a SR as a "synthesis of research evidence in which literature searches, inclusion criteria, and critical appraisal methods were explicitly described."[7] We included SRs where vitamin D was administered in any dose or by any route, in comparison with placebo or other doses/forms of vitamin D supplementation. To be included, SRs had to report at least one of the following predefined maternal or neonatal outcomes: pre-term birth, preeclampsia, gestational diabetes, small for gestational age, still birth, low birth weight, and cesarean section. We excluded primary studies.

Selection

Two reviewers (LB, JS-K) independently screened all titles and abstracts and reviewed the full-text of studies that were identified as potentially eligible using standard eligibility criteria.

Reviewers compared results and resolved any discrepancies through discussion; where uncertainty remained decisions were made in discussion with the study team.

Assessment of SR quality

Two reviewers (LB, JS-K) independently assessed the methodological quality of all relevant SRs using the AMSTAR checklist.[10,11] This reliable and valid tool consists of 11 items regarding the methodological quality of a systematic review. Reviewers compared assessments for each of the 11 items in the AMSTAR checklist and resolved disagreements through discussion or third-party adjudication. Based on the total AMSTAR score (maximum 11 representing highest quality), we categorized the SRs by quality: low (0-3), medium (4-7), high (8-11).[12] Given the large number of high quality SRs, we focused data extraction and analysis on these.

Data collection

One experienced reviewer (LB) extracted data from the SRs using predefined standard forms developed for this overview. For each SR, review level data were extracted on objectives, publication date, country of origin, funding, search date range, inclusion and exclusion criteria, number of included studies, methods of analysis, and quantitative data on included outcomes. For each outcome present in a SR we abstracted study design, intervention, comparator, effect size, and direction of effect. All data were reviewed for accuracy and completeness by a second reviewer (JS-K).

Analysis

We present and discuss the results by SR for each of our predefined outcomes. We display results based on SRs examining: 1) the effectiveness of vitamin D supplementation (i.e., results from randomized controlled trials), and 2) the association between serum vitamin D levels and pregnancy outcomes (i.e., results from observational studies). For consistency of rating and based on GRADE recommendations[9] results were converted to risk ratios using the random effects model where possible (in three cases, we had insufficient information to convert the estimates and have reported these as per the original review).[13-15] For each of the pre-defined outcomes we reported any sub-group analyses based on dosage or levels of vitamin D.

Assessing the level of evidence

To assess the certainty of the results, we graded the quality of evidence presented by every SR for each outcome of interest. We followed recommendations of the GRADE Working Group,[16] and assessed the following key domains: risk of bias, inconsistency, indirectness, imprecision, and publication/reporting bias. Rather than rating individual studies GRADE rates individual outcomes across studies; therefore the quality of evidence can differ for different outcomes from the same set of studies or for the same outcomes based on different sets of studies.[17] For SRs of observational studies, we considered the additional domains of magnitude of effect, dose response relationships, and whether all plausible confounding would reduce an effect.[16] For both interventional and observational designs the GRADE assessment started at high quality of evidence, given the designs were appropriate to address questions of effectiveness and association respectively. Two reviewers (LB, LH) independently conducted GRADE assessments and resolved discrepancies through discussion. GRADEpro software was utilized to calculate overall quality of evidence.[9,18]

Patient involvement

This research was done without patient or public involvement.

RESULTS

Literature search results and study selection

Figure 1 details the flow of information through the stages of this overview using the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)[19] flow diagram. We identified 233 records from the search after removing duplicates. After title and abstract screening 42 records were identified. Three SRs did not report on any of our predefined outcomes and were excluded[20-22], and one SR was represented by both a Cochrane and journal publication reporting the same data.[23,24] Based on the AMSTAR assessment 25 reviews were categorized as low or medium quality and were not included in the data extraction and outcome assessment. In total 13 SRs were included in the final analysis. See Supplementary Table 2 for the completed PRISMA checklist.

Description of included systematic reviews

The 13 included reviews were published between 2009 and 2018, with a median AMSTAR score of 8 ranging from 8 to 11 (supplementary table 3 and 4). The literature search dates for these 13 reviews were between September 2014 and May 2018. All 13 SRs were published in English and were from China[15,25,26], Canada[27-29], Iran[30], Spain[31], Switzerland[23], United Kingdom[14], United States[13,32], and Thailand[33]. Four SRs included both RCTs and observational studies[13,14,26,32], 5 included only RCTs[23,27,28,31,33], and 4 included only observational studies.[15,25,29,30] All included

SRs with the exception of two[13,32] conducted a meta-analysis. Across the 13 SRs there were 204 unique studies (78 RCTs and 126 observational studies).

None of the SRs explicitly searched for low income or high risk populations, most studies reported their populations as generally healthy at study entry without pre-existing conditions. Individual study sample sizes ranged from 16 to 12,861. For interventional studies there was a wide range of dosing regimens, daily doses ranged from 200 to 5,000 International Units (IU); weekly doses from 714 to 50,000 IU; up to 60,000 IU monthly and bolus doses ranging from 35,000 to 1,200,000 (600,000 x 2) IU. Only two reviews reported sub-group analyses based on dose ranges.[27,28] One review had a sub-group for neonatal mortality and small for gestational age for high (>2000 IU/day) and low (≤ 2000 IU/day),[27] and the other review presented sub-groups for high (≥ 2000 IU/day) and low (< 2000 IU/day) doses for all outcomes.[28] Two reviews of observational studies presented their analyses based on subgroups of 25 OH(D) levels, <50 nmol/L and <75 nmol/L,[29] and <50 vs >50 nmol/L and <75 vs >75 nmol/L.[26]

Synthesis of results by outcome for SRs examining the effectiveness of vitamin D

Preterm birth. Five SRs of RCTs[23,26-28,31] examined the effectiveness of vitamin D compared to no treatment/placebo or calcium for prevention of preterm birth (Table 1). Four SRs found no significant difference in preterm birth rates, while one SR found a significant benefit with vitamin D. However, the quality of evidence varied across SRs (see supplementary table 5 for detailed GRADE assessments). One of the SRs had high quality of evidence[28] while the other four were rated as moderate, low and very low quality. The SR with high quality of evidence showed no significant benefit of vitamin D on prevention of preterm birth (RR 1.00, 95% CI 0.77, 1.30).[28] In subgroup analyses, these findings of no effect on preterm birth were robust, not

altered when baseline vitamin D status was low (<30 nmol/L), when only studies at low risk of bias were examined, or when the analysis was limited to generally healthy women. There were also no significant differences within subgroups based on the effective daily equivalent dose of vitamin D: <2000 IU/day (RR 0.8, 95% CI 0.40, 1.60; 5 studies, 1,503 participants); ≥2000 IU/day (RR 1.02, 95% CI 0.76, 1.36; 9 studies, 2,404 participants).

Table 1: Summary of results from SRs of randomized controlled studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	Level of evidence (GRADE)
PRE-TERM BIRTH					
Bi 2018	11/3,822	0.98 (0.77, 1.26)	33%	- (NR)	moderate
De-Regil / Palacios 2016	3 / 477	0.36 (0.14, 0.93)	10%	+ (0.035)	very low
Perez-Lopez 2015	3 / 384	1.24 (0.59, 2.61)	0%	- (0.56)	very low
Roth 2017*	**14 / 3,757**	**1.00 (0.77, 1.30)**	**0%**	**- (0.677)**	**high**
Zhou 2017	6 / 1,687	0.61 (0.34, 1.07)	26%	- (0.09)	low
PREECLAMPSIA					
De-Regil / Palaciosis 2016	2 / 219	0.52 (0.25, 1.05)	0%	- (0.069)	very low
Khaing 2017	3 / 357	0.47 (0.24, 0.89)	0%	+ (0.02)	very low
Newberry 2014	1/504	NR; by group for individual study	NR	+ (n=1) †	very low
Perez-Lopez 2015*	**3 / 654**	**0.91 (0.45, 1.86)**	**24%**	**- (0.80)**	**low**
Roth 2017	3 / 706	1.09 (0.43, 2.76)	66%	- (0.047)	very low
GESTATIONAL DIABETES					
De-Regil / Palacios 2016	2 / 219	0.43 (0.05, 3.45)	0%	- (0.43)	very low
Perez-Lopez 2015	3 / 384	1.05 (0.60, 1.85)	0%	- (0.86)	very low
Roth 2017	5 / 1,030	0.65 (0.39, 1.08)	45%	- (0.125)	very low
SMALL FOR GESTATIONAL AGE					
Bi 2018*	**6 / 1002**	**0.72 (0.52, 0.99)**	**0%**	**+ (0.04)**	**low**
Harvey 2014	2 / 245	NR; by individual study	NR	- (n=2)†	very low
Perez-Lopez 2015	3 / 456	0.77 (0.46, 1.30)	15%	- (0.33)	very low
Roth 2017	5 / 741	0.60 (0.40, 0.90)	0%	+ (0.704)	very low
LOW BIRTH WEIGHT					
Bi 2018	4/775	0.52 (0.20, 1.37)	65%	- (NR)	very low
De-Regil / Palacios 2016	3 / 493	0.4 (0.24, 0.67)	4%	+ (0.00048)	very low
Perez-Lopez 2015	4 / 496	0.72 (0.45, 1.17)	0%	- (0.19)	very low
Roth 2017*	**7 / 1,156**	**0.74 (0.47, 1.16)**	**47.3%**	**- (0.077)**	**low**
STILLBIRTH					

De-Regil / Palacios 2016	3 / 540	0.35 (0.06, 1.99)	0%	- (0.23)	low
Roth 2017*	**16 / 4,606**	**0.75 (0.50, 1.12)**	**0%**	**- (0.858)**	**high**
CESAREAN SECTION					
De-Regil / Palaciosis 2016	2 / 312	0.95 (0.69, 1.31)	12%	- (0.75)	low
Perez-Lopez 2015	4 / 1,028	0.97 (0.81, 1.32)	0%	- (0.75)	low
Roth 2017*	**16 / 3,240**	**1.02 (0.93, 1.12)**	**0%**	**- (0.701)**	**high**

* for each outcome the review with the highest level of evidence is presented in bold font
† in absence of pooled data this indicates the number of studies with positive or negative statistical significance
± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of RCTs examined the effectiveness of vitamin D for prevention of preeclampsia.[23,28,31-33] The quality of evidence for effectiveness of vitamin D for preeclampsia was low and very low; the four SRs that pooled findings from individual studies showed mixed results (Table 1). The SR that provided the highest level of evidence (classified as low quality) found a non-significant risk ratio of 0.91 (95% CI 0.45, 1.86).[31] One SR planned subgroup analyses based on dose; all studies reporting the outcome used ≥2000 IU/day, therefore results were the same as the overall pooled estimate, which showed no significant difference (RR 1.09, 95% CI 0.43, 2.76; 3 studies, 706 participants).[28]

Gestational diabetes. Three SRs of RCTs examined the effectiveness of vitamin D for prevention of gestational diabetes (Table 1).[23,31] None of the SRs found a significant effect with the use of vitamin D in terms of the occurrence of gestational diabetes. The quality of evidence was very low in all SRs. One SR conducted subgroup analyses based on dose and found a significant reduction for <2000 IU/day (RR 0.33, 95% 0.13, 0.82) (based on a single study with 87 participants). No significant difference was observed for the subgroup receiving ≥2000 IU/day (RR 0.75, 95% CI 0.44, 1.28; 4 studies, 943 participants).[28]

Small for gestational age. Four SRs of RCTs examined the effectiveness of vitamin D in terms of prevention of infants' birthweights being small for gestational age (Table 1).[14,27,28,31] Three of the SR authors conducted meta-analyses to come up with overall effect estimates, while the authors of one SR chose not to pool due to heterogeneity across the two included studies. The SR with the highest quality of evidence (classified as low) found a significant risk ratio of 0.72 (95% CI 0.52, 0.99). Subgroup analysis in one SR based on dose showed no significant differences for <2000 IU/day (RR 0.63, 95% CI 0.35, 1.11; 3 studies, 352 participants) and ≥2000 IU/day (RR 1.04, 95% CI 0.32, 3.36; 2 studies, 219 participants).[28] In another SR, results for a subgroup based on dose was significant for the lower doses ≤ 2000 IU/day (RR 0.45, 95% CI 0.23, 0.90; 2 studies, 209 participants) with no difference for >2000 IU/day (RR 0.83, 95% CI 0.57, 1.19; 5 studies, 713 participants).[27]

Low birth weight. Four SRs of RCTs examined the effectiveness of vitamin D to prevent low birth weight (birthweight <2500 grams) (Table 1).[23,27,28,31] One SR found a significant benefit while the other three SRs showed no difference. The SR with the highest quality of evidence (low) showed no significant difference (RR 0.74, 95% CI 0.47, 1.16).[28] Subgroup analyses based on dose in this SR showed no significant differences for <2000 IU/day (RR 0.53, 95% CI 0.23, 1.21; 1 study, 126 participants) and ≥2000 IU/day (RR 0.99, 95% CI 0.70, 1.42; 5 studies, 830 participants).[28]

Stillbirth. Two SRs of RCTs examined the effectiveness of vitamin D to prevent stillbirth (Table 1).[23,28] Neither of the SRs found a significant benefit. The SRs had high and low quality of evidence, respectively. The SR with high quality of evidence found a risk ratio of 0.75 (95% CI

0.50, 1.12).[28] Subgroup analyses based on dose from this SR showed a significant difference for <2000 IU/day (RR 0.49, 95% CI 0.27, 0.91; 7 studies, 1,948 participants) but no difference for ≥2000 IU/day (RR 1.03, 95% CI 0.62, 1.71; 9 studies, 2,713 participants).[28]

Cesarean section. Three SRs examined the effectiveness of vitamin D for cesarean sections (Table 1).[23,28,31] The quality of evidence ranged from low to high; none of the SRs found a significant effect. The SR providing high quality of evidence found a risk ratio of 1.02 (95% CI 0.93, 1.12).[28] Subgroup analyses from this SR based on dose showed no significant differences for <2000 IU/day (RR 1.00, 95% CI 0.85, 1.18; 6 studies, 702 participants) or ≥2000 IU/day (RR 1.04, 95% CI 0.91, 1.19; 8 studies, 2,303 participants).[28]

Synthesis of results by outcome for SRs examining associations of vitamin D with perinatal outcomes

Preterm birth. Five SRs of observational studies examined the association between vitamin D status and preterm birth (Table 2).[14,25,26,29,32] One SR that examined the association between vitamin D and preterm birth found moderate evidence of an association overall 1.19 (1.08, 1.31).[25] Two SRs presented their analyses based on subgroups of 25 OH(D) levels: <50 nmol/L and <75 nmol/L,[29] and <50 vs >50 nmol/L and <75 vs >75 nmol/L.[26] In both SRs the association was slightly greater for the lower serum vitamin D level. The SR with highest quality of evidence found a significant association with moderate quality evidence for <50 vs. >50 nmol/L 1.13 (95% CI 1.04, 1.23) and non-significant association and low quality evidence for <75 vs. >75 nmol/L 1.03 (95% CI 0.98, 1.08).[26]

Table 2: Summary of results for SRs of observational studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	GRADE
PRETERM BIRTH					
Harvey 2014	7 / 1,792	NR; 1 individual study showed significance and 6 others not significant	NR	+ (n=1) ‡ - (n=6)	very low
Newberry 2014	2 / 371	NR; by individual study	NR	+ (n=1)‡ - (n=1)	very low
Qin 2016*	**10 / 10,098**	**1.19 (1.08, 1.31)**	**28%**	**+ (0.004)**	**moderate**
Wei 2013	4 / 1,111	1.27 (1.03, 1.58) [blood level 25(OH)D <50nmol/L]	28%	- (0.03)	very low
		1.05 (0.98, 1.12) [blood level 25(OH)D <75nmol/L]	0%	- (0.17)	very low
Zhou 2017*	**16 / 16,996**	**1.13 (1.04, 1.23) [<50 vs >50 nmol/L]**	**45%**	**+ (0.003)**	**moderate**
	15 / 17,122	1.03 (0.98, 1.08) [<75 vs >75 nmol/L]	65%	- (0.29)	low
PREECLAMPSIA					
Chung 2009	1 / 1,189	5 (1.7, 14.1) †	NR	+ (n=1) ‡	very low
Harvey 2014	4 / 642	0.75 (0.48, 1.19) †	80.8%	- (0.001)	very low
Newberry 2014	8 / 4420	NR; by individual study	NR	+ (n=5) - (n=3)	very low
Tabesh 2013	8 / 2,485	2.02 (1.26, 3.23)	53%	+ (0.04)	very low
Wei 2013*	**6 / 2,008**	**1.57 (1.21, 2.03) [<50 nmol/L]**	**39%**	**+ (0.0006)**	**low**
	5 / 1,311	1.21 (0.99, 1.46) [<75 nmol/L]	60%	- (0.06)	very low
GESTATIONAL DIABETES					
Harvey 2014	8 / 2,668	NR; by individual study	NR	+ (n=3) ‡ - (n=5)	very low
Lu 2016	20 / 16,515	1.45 (1.15, 1.83) †	66.6%	+ (0.002)	low
Wei 2013*	**10 / 4,126**	**1.12 (1.02, 1.22) [<50 nmol/L]**	**27%**	**+ (0.02)**	**moderate**
	8 / 3,840	**1.09 (1.03, 1.15) [<75 nmol/L]**	**28%**	**+ (0.002)**	**moderate**
SMALL FOR GESTATIONAL AGE					
Harvey 2014	7 / 5,660	NR; by individual study	NR	+ (n=2) ‡ - (n=5)	very low
Newberry 2014	1 / 412	NR; by individual study	NR	NR	very low
Wei 2013*	**6 / 6,013**	**1.35 (1.18, 1.54) [<50 nmol/L]**	**15%**	**+ (0.00001)**	**low**
	5 / 2,283	0.99 (0.83, 1.18) [<75 nmol/L]	75%	- (0.92)	very low
LOW BIRTH WEIGHT					
Harvey 2014	3 / 1,676	NR; by individual study	NR	+ (n=1) ‡ - (n=2)	very low
CESAREAN SECTION					
Harvey 2014	6 / 3,277	NR; by individual study	NR	+ (n=2) ‡	very low

				- (n=4)	

* for each outcome the review with the highest level of evidence is presented in bold font
† reported as odds ratios as insufficient data available to convert to risk ratio
‡ in absence of pooled data this indicates the number of studies with positive or negative statistical significance
± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of observational studies examined the association between vitamin D status and preeclampsia (Table 2).[13,14,29,30,32] Three of the five SRs found a significant association.[13,29,30] One SR assessed different serum levels of vitamin D and found a larger point estimate for <50 nmol/L compared with <75 nmol/L, although the confidence intervals overlapped.[29] The quality of evidence was low for <50 nmol/L and very low for <75 nmol/L.

Gestational diabetes. Three SRs of observational studies provided measures of association for vitamin D status and gestational diabetes (Table 2).[14,15,29] The SR providing the highest quality of evidence showed moderate quality evidence of a significant association for both serum levels examined: <50 nmol/L: 1.12 (95% CI 1.02, 1.22), <75 nmol/L: 1.09 (95% CI 1.03, 1.15).[29]

Small for gestational age. Three SRs of observational studies examined the association between vitamin D status and small birthweights for gestational age (Table 2).[14,29,32] The SRs showed mixed findings. One SR included 7 studies but did not pool results as the authors stated there was substantial variation in methodology and exposure;[14] 2 studies showed a significant association while 5 studies showed no significant effect (very low quality of evidence). Another SR only included 1 study and could not pool any results.[32] The highest rated (low quality) SR examined the association for different vitamin D serum levels and found a significant association for <50 nmol/L 1.35 (95% CI 1.18, 1.54), but no significant effect for <75 nmol/L 0.99 (95% CI 0.83, 1.18).[29] The quality of evidence was low for <50 nmol/L and very low for <75 nmol/L.

Low birth weight. Only one SR of observational studies examined the association between vitamin D status and low birth weight.[14] The SR included three studies but did not pool results. One study showed a statistically significant result while two studies had non-significant findings. Overall the quality of evidence for this outcome is very low.

Stillbirth. There were no SRs of observational studies that examined the association between vitamin D status and stillbirth.

Cesarean section. Only one SR of observational studies examined the association between vitamin D status and cesarean section.[14] The SR included six studies but did not pool results; the authors chose not to combine due to a multitude of factors such as local policies and physician preferences that influence this outcome. Two studies showed a statistically significant association while four studies had non-significant findings. Overall the quality of evidence for this outcome is very low.

DISCUSSION

This overview provides a comprehensive analysis of SRs examining vitamin D and pregnancy outcomes. We grouped and reported results separately for SRs of RCTs and SRs of observational studies. SRs of observational studies showed evidence of associations between vitamin D serum levels and some outcomes, however SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy with the exception of one pre-defined outcome—small for gestational age—which had low quality evidence. The differences in findings between

these groups of SRs suggest that any apparent associations may not be based on causal relationships. They suggest that low vitamin D levels or deficiencies may be an indicator or marker of poor health status, co-morbidities,[1] or perhaps an acute phase reactant.[34,35] It is likely that pregnant women with these indicators need more attention and care to optimize health outcomes for them and their offspring and not vitamin D supplementation. The current evidence does not support the use of vitamin D supplementation to improve any of these outcomes.

While there were some suggestions of associations between low vitamin D serum levels and some outcomes in the observational studies (i.e., preterm birth, preeclampsia, gestational diabetes, and small for gestational age), the effect sizes may be considered not clinically important.[9] The quality of this observational evidence was almost all low or very low.[16] However, more applicable to clinical practice are the findings from SRs of RCTs that examined the effectiveness of vitamin D as a treatment to improve pregnancy outcomes. The SRs of RCTs that provided the highest quality of evidence showed no effect of vitamin D supplementation in pregnancy for all but one of the pre-defined outcomes of interest. Overall these findings suggest that even if an association exists between vitamin D levels and health outcomes, vitamin D supplementation in pregnancy may be unlikely to improve these outcomes.

This study provides a methodologically rigorous and comprehensive synthesis of an extensive body of evidence examining vitamin D and perinatal outcomes. The considerable number of primary studies and SRs underscores the importance of this topic as well as the uncertainty about whether and how to manage vitamin D levels to optimize health outcomes. However, the vast number of SRs on this topic is concerning, particularly those of low quality which may propagate

inaccurate or biased results and conclusions. Of note, in our update search that captured the most recent publications up to January 2019, we identified 10 new relevant SRs with only three having an AMSTAR score greater than 7 to be included in the final analysis. Of these 3 new SRs there was only one new primary study included. We have provided an in-depth analysis by presenting the results of SRs of RCTs that evaluated the effectiveness of vitamin D as a treatment to improve perinatal outcomes alongside SRs of observational studies that examined the associations between vitamin D levels and health outcomes. Further, we used GRADE's rigorous and transparent method to assess the quality of the body of evidence which provides essential information about the certainty of the effect estimates in order to reconcile findings across individual studies and reviews.

The evidence contributing to the existing SRs varied widely in design and purpose (to examine associations vs. effectiveness). Observational studies have been used to examine the association between vitamin D levels and health outcomes, and are appropriate for generating hypotheses for testing in randomized trials. One of the limitations of the existing observational studies and synthesis of the same is that individual studies may or may not sufficiently adjust for confounding[36] (e.g., health status, calcium intake, and social determinants of health). Further, studies that did adjust for confounding differed in the variables they included and controlled for. RCTs, when well-designed, represent a higher level of evidence to assess the effectiveness of an intervention, in part because they can address the problem of confounding as randomization is intended to equally distribute both known and unknown confounders. It is well documented that early and observational studies often suggest important relationships that do not exist, and that well designed RCTs are often needed to fully understand a phenomenon.[37]

An important limitation in this area of investigation is the possibility of reporting and publication bias. While we focused on the highest quality SRs and most indicated that they planned to investigate publication bias, many could not do so because the number of included studies in a given meta-analysis was too small. There remains the possibility that studies, particularly the earlier published studies, showing significant results are more likely to be published while those with non-significant findings remain unpublished.

Also, the potential for selective outcome reporting is important in this body of literature. It is surprising that outcomes that are either routinely collected or relatively easy to ascertain, such as preterm birth, stillbirth, and gestational diabetes were infrequently reported. Selective outcome reporting occurs if researchers focus their reporting on a significant finding and downplay or do not report non-significant results. For example, the most frequently reported outcome was preterm birth; however, among the 48 studies included in one systematic review only one-third of the primary studies reported this outcome. Roth et al. also found that "missing data on clinical outcomes was the norm rather than exception" in this body of literature which could lead to "potentially biased meta-analyses based on small non-representative subsets of trials and participants".[28] Important efforts have been made to define core outcomes sets in the area of perinatal research.[38-40] Future studies should focus on critical outcomes for this field. Researchers should also define their outcomes and analyses a priori, register (and ideally publish) study protocols, and ensure clear and transparent reporting.[41,42] Further, researchers should identify all important confounders and address these adequately through appropriate research designs or analytic approaches to ensure valid findings and permit meaningful pooling of data. Currently

the credibility of the body of evidence in this important field is compromised due to the potential for confounding, publication bias, reporting bias, and imprecision arising from low numbers of participants.

CONCLUSIONS

While there is some evidence from SRs of observational studies for an association between maternal vitamin D serum levels and some perinatal outcomes, SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy for all but one pre-defined outcome, the evidence for the one outcome was low quality. The discrepancy between the observational studies and the RCTs shows that 25-hydroxy vitamin D is lower among women who experience adverse pregnancy outcomes, but supplementation does not appear to alter outcomes. Low 25-hydroxy vitamin D, the indicator of vitamin D status, is a marker of adverse outcomes rather than a marker of vitamin D status[1] which shows that 25-hydroxy vitamin D may be an acute phase reactant. Future studies need to adequately control for potential confounding (e.g., through well-designed randomized trials)[42] and include all outcomes that are considered critical to this field. There are currently over 40 published SRs (many of which are low quality) synthesizing evidence from 204 primary vitamin D studies; further SRs on this topic are wasteful until more well designed and conducted RCTs are completed.

FIGURE LEGENDS

FIGURE 1: Flow diagram of screening decisions

ACKNOWLEDGEMENTS: Research reported in this publication was funded by the Alberta Health Services' Maternal, Newborn, Child & Youth Strategic Clinical Network™ (MNCY SCN™). Additional support was provided by the Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform which is funded by the Canadian Institutes of Health Research and Alberta Innovates. Dr. Hartling is supported by a Canada Research Chair in Knowledge Synthesis and Translation. The content is solely the responsibility of the authors and does not necessarily represent the views of Alberta Health Services. We would like to thank: Tara Landry, MLIS, for peer reviewing the search strategy; Ms. MacKinna Hauff for article retrieval; Robin Featherstone, MLIS, for developing and running the search; and Dr. Seija Kromm for administrative support.

CONTRIBUTORS: LB, LH, TF, DM and DJ designed the study. LB, LH and JKS selected articles, extracted data and performed the assessment of bias. LB supervised study activities and LH wrote the first draft of the manuscript. All authors critically reviewed and revised the manuscript and approved the final version for publication. The corresponding author attests that all listed authors meet authorship criteria and no others meeting the criteria have been omitted.

within and any related or standalone film submitted (the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. I am the sole author of the Contribution and I am one author signing on behalf of all co-owners of this Contribution.

COMPETING INTERESTS: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

DATA SHARING: The dataset is available from the lead author on request.

PATIENT AND PUBLIC INVOLVEMENT: This research was done without patient or public involvement.

REFERENCES

1. Autier P, Mullie P, Macacu A, et al. Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials. *Lancet Diabetes Endocrinol*. 2017;5(12):986-1004.
2. Wagner CL, Baggerly C, McDonnell S, et al. Post-hoc analysis of vitamin D status and reduced risk of preterm birth in two vitamin D pregnancy cohorts compared with South Carolina March of Dimes 2009-2011 rates. *J Steroid Biochem Mol Biol*. 2016;155(Pt B):245-251.
3. Leffelaar ER, Vrijkotte TG, van Eijsden M. Maternal early pregnancy vitamin D status in relation to fetal and neonatal growth: results of the multi-ethnic Amsterdam Born Children and their Development cohort. *Br J Nutr*. 2010;104(1):108-117.
4. Soheilykhah S, Mojibian M, Rashidi M, et al. Maternal vitamin D status in gestational diabetes mellitus. *Nutr Clin Pract*. 2010;25(5):524-527.
5. Holmes VA, Barnes MS, Alexander HD, et al. Vitamin D deficiency and insufficiency in pregnant women: a longitudinal study. *Br J Nutr*. 2009;102(6):876-881.
6. Bodnar LM, Catov JM, Simhan HN, et al. Maternal vitamin D deficiency increases the risk of preeclampsia. *J Clin Endocrinol Metab*. 2007;92(9):3517-3522.
7. Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? *Syst Rev*. 2013;2:26.
8. Pollock M Fernandes RM, Becker LA, et al. Chapter V: Overviews of Reviews. Draft version (8 October 2018) for inclusion in: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch V (editors). Cochrane Handbook for Systematic Reviews of Interventions. London: Cochrane.
9. Schünemann H, Brozek J, Guyatt G, et al, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html. Accessed June 3, 2019.
10. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10.
11. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62(10):1013-1020.
12. Canadian Agency for Drugs and Technologies in Health (CADTH). Rapid response summary with critical appraisal: Process. 2015; https://www.cadth.ca/about-cadth/what-we-do/products-services/rapid-response-service. Accessed April 19, 2018.
13. Chung M, Balk EM, Brendel M, et al. Vitamin D and calcium: a systematic review of health outcomes. *Evid rep/technol assess*. 2009(183):1-420.

14. Harvey NC, Holroyd C, Ntani G, et al. Vitamin D supplementation in pregnancy: a systematic review. *Health Technol Assess*. 2014;18(45):1-190.
15. Lu M, Xu Y, Lv L, et al. Association between vitamin D status and the risk of gestational diabetes mellitus: a meta-analysis. *Arch Gynecol Obstet*. 2016;293(5):959-966.
16. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490.
17. Guyatt G, Oxman AD, Elie A, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394.
18. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University, 2015 (developed by Evidence Prime, Inc.). Available from: https://gradepro.org. Accessed May 27, 2019.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;8(5):336-341.
20. Yepes-Nunez JJ, Brozek JL, Fiocchi A, et al. Vitamin D supplementation in primary allergy prevention: Systematic review of randomized and non-randomized studies. *Allergy*. 2017;04:04.
21. Zhang H, Huang Z, Xiao L, et al. Meta-analysis of the effect of the maternal vitamin D level on the risk of spontaneous pregnancy loss. *Int J Gynaecol Obstet*. 2017;138(3):242-249.
22. Christensen N, Sondergaard J, Fisker N, et al. Infant Respiratory Tract Infections or Wheeze and Maternal Vitamin D in Pregnancy: A Systematic Review. *Pediatr Infect Dis J*. 2017;36(4):384-391.
23. De-Regil LM, Palacios C, Lombardo LK, et al. Vitamin D supplementation for women during pregnancy. *Cochrane Database Syst Rev*. 2016(1):CD008873.
24. Palacios C, De-Regil LM, Lombardo LK, et al. Vitamin D supplementation during pregnancy: Updated meta-analysis on maternal outcomes. *J Steroid Biochem Mol Biol*. 2016;164:148-155.
25. Qin LL, Lu FG, Yang SH, et al. Does Maternal Vitamin D Deficiency Increase the Risk of Preterm Birth: A Meta-Analysis of Observational Studies. *Nutrients*. 2016;8(5):20.
26. Zhou SS, Tao YH, Huang K, et al. Vitamin D and risk of preterm birth: Up-to-date meta-analysis of randomized controlled trials and observational studies. *J Obstet Gynaecol Res*. 2017;43(2):247-256.
27. Bi WG, Nuyt AM, Weiler H, et al. Association Between Vitamin D Supplementation During Pregnancy and Offspring Growth, Morbidity, and Mortality: A Systematic Review and Meta-analysis. *JAMA Pediatr*. 2018;172(7):635-645.
28. Roth DE, Leung M, Mesfin E, et al. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. *BMJ*. 2017;359:j5237.
29. Wei SQ, Qi HP, Luo ZC, et al. Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis. *J Matern Fetal Neonatal Med*. 2013;26(9):889-899.

30. Tabesh M, Salehi-Abargouei A, Tabesh M, et al. Maternal vitamin D status and risk of pre-eclampsia: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2013;98(8):3165-3173.

31. Perez-Lopez FR, Pasupuleti V, Mezones-Holguin E, et al. Effect of vitamin D supplementation during pregnancy on maternal and neonatal outcomes: a systematic review and meta-analysis of randomized controlled trials. *Fertil Steril*. 2015;103(5):1278-1288.e1274.

32. Newberry SJ, Chung M, Shekelle PG, et al. Vitamin D and Calcium: A Systematic Review of Health Outcomes (Update). *Evid rep/technol assess*. 2014(217):1-929.

33. Khaing W, Vallibhakara SA, Tantrakul V, et al. Calcium and Vitamin D Supplementation for Prevention of Preeclampsia: A Systematic Review and Network Meta-Analysis. *Nutrients*. 2017;9(10):18.

34. Madden K, Feldman HA, Chun RF, et al. Critically Ill Children Have Low Vitamin D-Binding Protein, Influencing Bioavailability of Vitamin D. *Ann Am Thorac Soc*. 2015;12(11):1654-1661.

35. Silva MC, Furlanetto TW. Does serum 25-hydroxyvitamin D decrease during acute-phase response? A systematic review. *Nutr Res*. 2015;35(2):91-96.

36. Patel CJ, Manrai AK. Development of exposome correlation globes to map out environment-wide associations. *Pac Symp Biocomput*. 2015:231-242.

37. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol*. 2011;64(12):1277-1282.

38. Molloy EJ, Gale C, Marsh M, et al. Developing core outcome set for women's, newborn, and child health: the CROWN Initiative. *Pediatr Res*. 2018;84(3):316-317.

39. Devane D, Begley CM, Clarke M, et al. Evaluating maternity care: a core set of outcome measures. *Birth*. 2007;34(2):164-172.

40. van 't Hooft J, Duffy JM, Daly M, et al. A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol*. 2016;127(1):49-58.

41. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg*. 2014;12(12):1495-1499.

42. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg*. 2012;10(1):28-55.

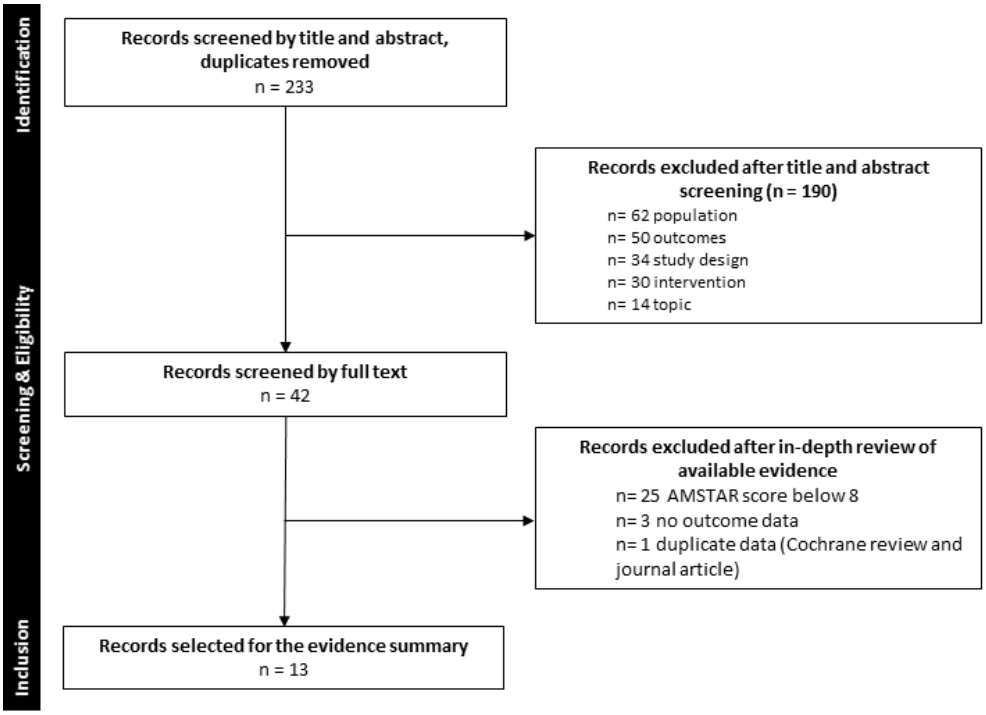


Figure 1: Study flow diagram

Supplementary Table 1: Literature search strategy

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Date conducted: 2 October 2017
Strategy:

1 Preconception Care/ (1917)
2 exp Pregnancy/ (855216)
3 exp Pregnancy Complications/ (405775)
4 Pregnant Women/ (6515)
5 Prenatal Care/ (24637)
6 Prenatal Diagnosis/ (35834)
7 (antenatal* or pre-natal* or prenatal*).tw,kf. (120088)
8 (expect* adj2 (female? or mother? or wom#n)).tw,kf. (3728)
9 ((1* or first*) adj2 (tri-mester* or trimester*)).tw,kf. (23473)
10 (pre-conception* or preconception*).tw,kf. (4573)
11 pregnan*.tw,kf. (477757)
12 or/1-11 [Combined MeSH & text words for pregnancy] (1016791)
13 exp Vitamin D/ (54287)
14 Vitamin D Deficiency/ (13412)
15 calcidiol*.tw,kf. (397)
16 calciol*.tw,kf. (20)
17 calcifediol*.tw,kf. (128)
18 cholecalciferol*.tw,kf. (2377)
19 hydroxycholecalciferol*.tw,kf. (1377)
20 hydroxyvitamin D*.tw,kf. (12499)
21 (vitamin D or vitamin D3 or vitamin D$2).tw,kf. (60203)
22 or/13-21 [Combined MeSH & text words for vitamin D] (79566)
23 and/12,22 [Combined concepts for pregnancy & vitamin D] (4365)
24 meta-analysis.pt. (87537)
25 meta-analysis/ or systematic review/ or meta-analysis as topic/ or "meta analysis (topic)"/ or "systematic review (topic)"/ or exp technology assessment, biomedical/ (113240)
26 ((systematic* adj3 (review* or overview*)) or (methodologic* adj3 (review* or overview*))).ti,ab,kf,kw. (127445)
27 ((quantitative adj3 (review* or overview* or synthes*)) or (research adj3 (integrati* or overview*))).ti,ab,kf,kw. (8559)
28 ((integrative adj3 (review* or overview*)) or (collaborative adj3 (review* or overview*)) or (pool* adj3 analy*)).ti,ab,kf,kw. (19993)
29 (data synthes* or data extraction* or data abstraction*).ti,ab,kf,kw. (20992)
30 (handsearch* or hand search*).ti,ab,kf,kw. (7877)
31 (mantel haenszel or peto or der simonian or dersimonian or fixed effect* or latin square*).ti,ab,kf,kw. (21571)
32 (met analy* or metanaly* or technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf,kw. (7548)
33 (meta regression* or metaregression*).ti,ab,kf,kw. (5904)

34 (meta-analy* or metaanaly* or systematic review* or biomedical technology assessment* or bio-medical technology assessment*).mp,hw. (217154)
35 (medline or cochrane or pubmed or medlars or embase or cinahl).ti,ab,hw. (161733)
36 (cochrane or (health adj2 technology assessment) or evidence report).jw. (18083)
37 (meta-analysis or systematic review).mp. [sic – changed .md. to .mp] (200672)
38 (comparative adj3 (efficacy or effectiveness)).ti,ab,kf,kw. (10906)
39 (outcomes research or relative effectiveness).ti,ab,kf,kw. (7938)
40 ((indirect or indirect treatment or mixed-treatment) adj comparison*).ti,ab,kf,kw. (1649)
41 or/24-40 [CADTH SR search filter | Retrieved from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#syst] (358374)
42 and/23,41 [SR filter applied] (187)
43 remove duplicates from 42 (164)

Database: Wiley Cochrane Library
Date conducted: 2 October 2017
Strategy:
#1 [mh ^"Preconception Care"] 103
#2 [mh Pregnancy] 5760
#3 [mh "Pregnancy Complications"] 9364
#4 [mh ^"Pregnant Women"] 156
#5 [mh ^"Prenatal Care"] 1332
#6 [mh ^"Prenatal Diagnosis"] 380
#7 (antenatal* or "pre-natal*" or prenatal):ti,ab,kw 6295
#8 (expect* near/2 (female? or mother? or wom?n)):ti,ab,kw 243
#9 ((1* or first*) near/2 ("tri-mester*" or trimester*)):ti,ab,kw 4141
#10 ("pre-conception*" or preconception*):ti,ab,kw 307
#11 pregnan*:ti,ab,kw 36386
#12 {or #1-#11} 38579
#13 [mh "Vitamin D"] 2941
#14 [mh ^"Vitamin D Deficiency"] 617
#15 calcidiol*:ti,ab,kw 46
#16 calciol*:ti,ab,kw 1
#17 calcifediol*:ti,ab,kw 475
#18 cholecalciferol*:ti,ab,kw 1208
#19 hydroxycholecalciferol*:ti,ab,kw 338
#20 "hydroxyvitamin D*":ti,ab,kw 1931
#21 ("vitamin D" or "vitamin D3" or "vitamin D?"):ti,ab,kw 6774
#22 {or #13-#21} 7581
#23 #11 and #22 354
#24 #11 and #22 in Cochrane Reviews (Reviews and Protocols), Other Reviews and Technology Assessments 14

PRISMA 2009 Checklist

Supplementary Table 2: PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3-4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6-7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary Table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	NA
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	9

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10-11, Supplementary Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	10
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-13, 16-17
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	12-13, 16-17
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Supplementary Table 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	23

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Supplementary Table 3: Description of included systematic reviews

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Bi Canada May 2018	24 RCTs 5,405 (30 – 965)	Population was healthy, pregnant women without prior vitamin D supplementation of more than 400 IU/d	Vitamin D in the form of cholecalciferol in 22 RCTs and in the form of ergocalciferol in 3 RCTs daily doses: 800 - 5000; weekly doses 35000 or 50000; fortnightly dose 50000; monthly dose 60000; bimonthly dose 60000; and bolus doses 60000 - 200 000	Placebo, no intervention or other dose of vitamin D	Primary: small for gestational age (indicated by birthweight less than the 10th percentile for gestational age, fetal or neonatal mortality Secondary: neonatal (25[OH]D) levels, congenital malformation, admission to a neonatal intensive care unit (NICU), Apgar scores, neonatal calcium levels, birth weight, low birth weight, gestational age, preterm birth, infant growth, asthma, respiratory infection, eczema, and allergy
Khaing Thailand October 2017 Roth Canada September 2017	19 RCTs 28,000 (30 – 9,178) 43 RCTs 8,406 (16 – 1,134)	Pregnant women of any gestational age Participants were pregnant at enrolment or enrolled before pregnancy and then followed-up in pregnancy	Calcium, vitamin D, combined calcium and vitamin D Vitamin D vs. placebo = 3; Calcium + vitamin D vs. calcium = 1 Vitamin D2 or D3, alone or in combination provided the co-intervention is similar in at least one other trial arm Daily doses: 400 – 5000; weekly doses: 714 – 7543; monthly doses: 1645 – 3289; bolus doses: 60000 – 1200000 (600000 x 2)	Placebo, a standard supplementation (e.g., folic acid), or no supplementation Placebo, no vitamin D, or vitamin D up to 600 IU/day (or a less frequent dose that would be about equivalent to 600 IU/day—for example, 4200 IU/week)	Primary: preeclampsia, eclampsia, proteinuria (dipstick urine 2+ or '300 mg/24 h), end-organ dysfunction, or utero-placental dysfunction after 20 weeks of gestation Primary: 25 OHD, preeclampsia, gestational diabetes, gestational hypertension, intra-uterine death, stillbirth, c-section, weight gain, preterm labor, death, adverse events, hospitalizations, birth weight, birth length, head circumference, low birth weight, small for gestational age, gestational age at birth, congenital malformations, neonatal death, respiratory infection, asthma, bone mineral content and density

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Zhou China June 2016	6 RCTs; 9 prospective cohort; 4 nested case-control; 2 cross-sectional; 2 retrospective cohort; 1 case-control 28,391 (50 – 12,861)	Pregnant women without HIV infection	maternal serum 25-OHD or oral supplementation with vitamin D Daily doses of 1,000 to 4,000 IU; weekly doses of 400 daily for 9 weeks; 50,000 for 6 weeks; one time doses starting 60,000 or 2-4 doses of 120,000	no supplementation /placebo, or routine care (ferrous sulfate and calcium, but no vitamin D)	Preterm birth
Qin China August 2015	4 Prospective cohort; 4 Nested case-control; 1 case-control; 1 Retrospective cohort; 1 Cross-sectional 20,608 (134 – 12,861)	Pregnant women without pre-chronic disease or HIV infection, with singleton gestation	NR; measurement of maternal vitamin D levels		Preterm birth
Lu China February 2015	4 Case-control; 7 Cohort; 2 Cross sectional; 7 Nested case control 16,515 (122 – 4,090)	NR	NR; measurement of maternal vitamin D levels		Gestational diabetes
De-Regil / Palacios Switzerland / Puerto Rico February 2015	15 RCTs 2,833 (40 – 990)	Pregnant women of any gestational or chronological age, parity (number of births) and number of fetuses	Vitamin D daily doses: 200 - 2000 Vitamin D single dose: 200,000 – 600,000, and 35,000	No intervention / placebo	Primary: pre-eclampsia, gestational diabetes, vitamin D concentration, adverse effects, preterm birth, low birthweight Secondary: impaired glucose tolerance, c-section, gestational hypertension, maternal death, birth length, head circumference at birth, birthweight, admission to special care, stillbirth, neonatal death, very preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Newberry USA September 2014	2 RCTs; 2 prospective cohorts; 5 nested case-control 4,912 (160 – 1,141)	Primary population of interest is generally healthy people with no known disorders Only including studies for population contributing to pregnancy related outcomes	Vitamin D single doses (for RCT): 2000, 4000 followed by 1 month run-in at 2000	All participants enrolled into one of two vitamin D groups	Preeclampsia, preterm birth, small for gestational age
Perez-Lopez Spain March 2014	13 RCTs 2,299 (40 – 400)	Pregnant women of any gestational or chronologic age and parity, without previous disease history	Vitamin D alone vs. no treatment (placebo); vitamin D + calcium vs. no treatment (placebo); and vitamin D + calcium vs. calcium Daily doses ranged from 400 to 1,000; weekly doses ranged from 35,000 to 50,000; and single doses ranged from 200,000 to 600,000	Active controls, usual treatment without active control, and placebo	Primary: circulating 25-OHD, preeclampsia, gestational diabetes, small for gestational age, low birth weight, preterm birth, birthweight Secondary: birth length, c-section,
Wei Canada October 2012	13 Case-control; 8 cohort; 2 cross-sectional 12,898 (95 – 3,730)	Pregnant women without pre-existing chronic disease or HIV infection	NR; measurement of maternal vitamin D levels		Preeclampsia, gestational diabetes, preterm birth, small for gestational age
Harvey UK June 2012	17 Case-control; 48 cohort/cross-sectional; 9 RCT; 2 intervention studies (non-randomized) NR	Pregnant women or pregnant women and their offspring	vitamin D status [dietary intake, sunlight exposure, circulating 25(OH)D concentration] or supplementation of	For intervention studies: no intervention or placebo	Primary: neonatal hypocalcaemia, rickets in the offspring, offspring bone mass, and maternal osteomalacia Secondary: offspring body composition; offspring preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
			participants with vitamin D or food containing vitamin D (e.g. oily fish)		and later offspring health outcomes; maternal quality of life
Tabesh Iran December 2012	2 Cohort; 4 cross-sectional; 9 case-control 2,936 (32 – 697)	Normal pregnant women	NR; measurement of maternal vitamin D levels		Preeclampsia
Chung USA April 2009	60 RCT; 3 NRCT; 102 cohort or nested case-control; 11 SR NR	Generally healthy people with no known disorders	Vitamin D supplements (no analogues), calcium supplements, and combinations of supplements; food based interventions	NR	Pregnancy-related: preeclampsia, high blood pressure with or without proteinuria, preterm birth or low birth weight, infant mortality

Supplementary Table 4. AMSTAR score by category and individual systematic review

Review	AMSTAR question											
	Q1 A priori design provided	Q2 Duplicate study selection and data extraction	Q3 Comprehensive literature search	Q4 Publication status as inclusion criterion	Q5 List of studies (include and exclude) provided	Q6 Characteristics of the included studies provided	Q7 Quality assess-ment	Q8 Quality used appropriate	Q9 Methods used to combine appropriate	Q10 Publication bias assessed	Q11 Conflict of interest stated	Total
OVERALL HIGH QUALITY												
Bi 2018	n	y	y	n	n	y	y	y	y	y	y	8
Christensen 2017	y	y	y	y	n	y	y	y	y	n	y	9
Chung 2009	y	ca	y	y	y	y	y	y	y	ca	y	9
De-Regil 2016	y	y	y	y	y	y	y	y	y	y	y	11
Harvey 2014	y	y	y	y	n	y	y	y	y	ca	y	9
Khaing 2017	y	y	n	n	n	y	y	y	y	y	y	8
Lu 2016	y	y	y	n	n	y	y	y	y	y	y	9
Newberry 2014	y	ca	y	n	y	y	y	y	y	ca	y	8
Palacios 2016	y	y	y	y	y	y	y	y	y	y	y	11
Perez-Lopez 2015	y	y	y	y	n	y	y	y	ca	ca	y	8
Qin 2016	n	y	y	n	n	y	y	y	y	y	y	8
Roth 2017	y	y	y	y	n	y	y	y	y	ca	y	9
Tabesh 2013	y	y	y	y	n	y	n	n	y	y	y	8
Wei 2013	n	y	y	n	n	y	y	y	y	y	y	8
Yepes-Nunez 2017	n	y	y	y	y	y	y	y	y	y	y	10
Zhang 2017	n	y	y	n	n	y	y	y	y	y	y	8
Zhou 2017	n	y	y	n	n	y	y	y	y	y	y	8
OVERALL MEDIUM AND LOW QUALITY												
Aghajafari 2013	n	y	y	ca	n	y	n	ca	y	y	n	5
Amegah 2017	n	y	y	n	n	y	y	y	y	y	n	6
Amraei 2018	ca	y	y	n	n	y	ca	ca	y	y	y	6
Arain 2015	n	y	ca	n	n	y	n	ca	ca	n	n	2
Chen 2017	n	y	y	n	n	y	y	y	y	y	n	6
Christensen 2012	n	y	n	n	n	y	n	n	n	n	y	3
Fu 2017	n	ca	y	n	n	n	n	n	y	y	y	4
Galthen-Sorensen 2014	n	y	y	n	n	y	y	y	n	n	n	5
Hu 2018	n	y	y	y	n	y	n	ca	y	y	y	7
Hypponen 2014	n	ca	y	n	n	y	ca	n	y	y	y	5
Kamudoni 2016	n	ca	y	y	n	y	n	n	n	n	y	4
Mahomed 2009	y	n	y	y	y	y	ca	ca	n	n	y	6
Martinez-Dominquez 2018	n	ca	y	n	n	y	y	ca	y	y	y	6
Nassar 2011	y	ca	n	n	n	y	n	n	y	n	y	4

Poel 2012	n	ca	y	n	n	y	n	n	ca	y	y	4
Purswani 2017	n	y	n	n	n	y	n	y	y	n	y	5
Santamaria 2018	n	y	n	n	n	y	y	y	y	ca	y	6
Senti 2012	n	y	y	n	n	y	n	n	n	n	y	4
Serrano-Diaz 2018	n	n	y	y	n	y	ca	n	y	y	y	6
Thorne-Lyman 2012	n	n	y	n	n	y	y	y	y	n	n	5
Van der Pligt 2018	n	y	y	n	n	y	y	y	n	n	y	6
Wei 2016	n	y	y	n	n	y	y	ca	y	y	n	6
Yang 2015	n	y	y	n	y	y	y	n	y	y	n	7
Zhang 2018	n	ca	y	ca	n	y	y	y	y	y	y	7
Zhang 2015	n	n	y	n	n	y	y	y	y	y	y	7

* One point was awarded for each item that scored 'yes' (y) and summed for the total score
* 'n' no; 'ca' can't answer

Supplementary Table 5: GRADE tables

Grade Assessments for Preterm Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
11 Bi	RCT	serious	not serious	not serious	not serious	none	moderate
3 De-Regil/Palacios	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
14 Roth	RCT	not serious	not serious	not serious	not serious	none	high
6 Zhou	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low

Grade Assessments for Preeclampsia in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil / Palaciosis	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Khaing	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
1 Newberry	RCT	serious	serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopex	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
3 Roth	RCT	serious	serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Gestational Diabetes in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
5 Roth	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Low Birth Weight in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
4 Bi	RCT	serious	serious	not serious	serious	none	very low
3 De-Regil/ Palaciosis	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
4 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
7 Roth	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low

Grade Assessments for Small for Gestational Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
6 Bi	RCT	serious	not serious	not serious	serious	none	low
2 Harvey	RCT	serious	serious	not serious	serious	none	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
5 Roth	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Still Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for C-Section Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
4 Perez-Lopez	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for Preterm Birth in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
2 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
10 Qin	OBS	not serious	not serious	serious	not serious	none	moderate
4 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	serious	potential for publication / reporting bias	very low

4 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	serious	potential for publication / reporting bias	very low
16 Zhou [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
17 Zhou [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	none	low

Grade Assessments for Preeclampsia in Observational Studies

# studies Author	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Overall Certainty**
1 Chung	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
4 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
8 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
8 Tabesh	OBS	serious	serious	serious	not serious	none	very low
6 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	potential for publication / reporting bias	low
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	potential for publication / reporting bias	very low

Grade Assessments for Gestational Diabetes in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
8 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
20 Lu	OBS	not serious	serious	serious	not serious	none	low
10 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
8 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate

Grade Assessments for Low Birth Weight in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Small for Gestational Age in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
1 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
6 Wei	OBS	not serious	not serious	serious	not serious	potential for publication / reporting bias	low

[blood level 25(OH)D <50nmol/L]							
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	potential for publication / reporting bias	very low

Grade Assessments for Small for C-Section in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low

BMJ Open

Vitamin D supplementation to improve pregnancy and perinatal outcomes: an overview of 42 systematic reviews

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-032626.R2
Article Type:	Research
Date Submitted by the Author:	02-Jan-2020
Complete List of Authors:	Bialy, Liza ; University of Alberta, Pediatrics Fenton, Tanis; University of Calgary, Department of Community Health Sciences, O'Brien Institute of Public Health, Alberta Children's Hospital Research Institute Shulhan-Kilroy, Jocelyn; University of Alberta, Alberta Research Centre for Health Evidence, Department of Pediatrics Johnson, David; Alberta Children's Hospital, McNeil, Deborah A.; University of Calgary, Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine Hartling, Lisa; University of Alberta, Pediatrics
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Evidence based practice
Keywords:	overview of reviews, vitamin D, perinatal

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

VITAMIN D SUPPLEMENTATION TO IMPROVE PREGNANCY AND PERINATAL OUTCOMES: AN OVERVIEW OF 42 SYSTEMATIC REVIEWS

Liza Bialy,[1] Tanis R. Fenton,[2,3] Jocelyn Shulhan-Kilroy,[4] David W. Johnson,[5,6] Deborah A. McNeil,[6,7] Lisa Hartling[1,4]

AFFILIATIONS:

[1] Alberta SPOR SUPPORT Unit Knowledge Translation Platform, University of Alberta, Edmonton, Alberta, Canada

[2] Community Health Sciences, Institute of Public Health, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3] Nutrition Services, Alberta Health Services, Calgary, AB, Canada

[4] Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada

[5] Departments of Pediatrics and Physiology and Pharmacology, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[6] Maternal Newborn Child and Youth Strategic Clinical Network, Alberta Health Services, Edmonton, AB, Canada

[7] Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

CORRESPONDING AUTHOR:

Liza Bialy, BSc, MPH, 4-486B Edmonton Clinic Health Academy, Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, 11405-87 Avenue NW, Edmonton, AB, T6G 1C9; lbialy@ualberta.ca

WORD COUNT: 4188

ABSTRACT

Objective: To review the evidence to assess effectiveness of vitamin D supplementation during pregnancy and associations of serum vitamin D levels with perinatal outcomes.

Design: Overview of systematic reviews.

Data Sources: Searches conducted in January 2019: Ovid Medline (1946-), Cochrane Library databases.

Eligibility criteria for selecting studies: Two reviewers independently screened titles and abstracts, and full-texts using pre-defined inclusion criteria: systematic reviews (SRs) evaluating vitamin D supplementation in pregnant women and/or examining the association between serum vitamin D levels reporting at least one pre-defined perinatal outcome. Only SRs with high AMSTAR scores were analysed.

Data extraction and synthesis: Data were extracted independently by one reviewer and checked by a second. Results were assessed for quality independently by two reviewers using GRADE criteria.

Results: Thirteen SRs were included, synthesizing evidence from 204 unique primary studies. SRs of RCTs with the highest level of evidence showed no significant benefit from vitamin D in terms of preterm birth [RR 1.00 (95% CI 0.77, 1.30); high quality], preeclampsia [RR 0.91 (0.45, 1.86); low quality], gestational diabetes [RR 0.65 (0.39, 1.08); very low quality], stillbirth [RR 0.75 (0.50, 1.12); high quality], low birth weight [RR 0.74 (0.47, 1.16); low quality], cesarean section [RR 1.02 (0.93, 1.12); high quality]. A significant difference was found for small-for-gestational age [RR 0.72 (0.52, 0.99); low quality]. SRs of observational studies showed associations between vitamin D levels and preterm birth [RR 1.19 (1.08, 1.31); moderate quality], preeclampsia [RR 1.57 (1.21, 2.03) for 25 (OH)D <50 nmol/L subgroup; low quality],

gestational diabetes [RR 1.12 (1.02, 1.22) for 25 (OH)D <50 nmol/L and RR 1.09 (1.03, 1.15) <75 nmol/L; moderate quality], and small-for-gestational age [RR 1.35 (1.18, 1.54) <50 nmol/L; low quality]. SRs showed mixed results for associations between vitamin D and low birth weight (very low quality) and cesarean section (very low quality).

Conclusion: There is some evidence from SRs of observational studies for associations between vitamin D serum levels and some outcomes, however SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy with the exception of one pre-defined outcome, which had low quality evidence. Credibility of the evidence in this field is compromised by study limitations (particularly the possibility of confounding among observational studies), inconsistency, imprecision, and potential for reporting and publication biases.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- We provide a comprehensive summary of the existing evidence for the effectiveness and associations of vitamin D and perinatal outcomes.
- A strength of this overview is the rigorous assessment of the quality of evidence using validated measures (AMSTAR and GRADE).
- The sparsity of high quality evidence for specific outcomes at the primary and systematic review levels currently limits the ability to make strong recommendations for the use of vitamin D during pregnancy.

INTRODUCTION

Vitamin D research is an active area of clinical investigation as numerous studies have examined associations between low vitamin D status (low serum 25-hydroxy vitamin D) and many diseases.[1] The evolution of this research began with observational studies examining associations between vitamin D levels and numerous health outcomes. There is now a growing body of randomized controlled trials (RCTs) assessing the effectiveness of vitamin D as an intervention to improve a variety of health outcomes.

Research in pregnancy examining associations between vitamin D with maternal and infant outcomes has also followed this progression. Early studies in this area suggested that low vitamin D levels were associated with undesirable perinatal outcomes, including gestational diabetes, pre-eclampsia, preterm birth and low birthweight. RCTs are now available,[2-6] allowing for examination of whether maternal vitamin D supplementation is effective in improving perinatal outcomes.

Given the extensive number of primary studies available on this topic, a number of systematic reviews (SRs) have been conducted to synthesize the evidence in order to guide practice and recommendations regarding perinatal care. However, the SRs vary in their scope, results, and conclusions which poses a challenge for decision-makers in terms of guiding recommendations for the treatment and management of women during pregnancy. Overviews are a useful starting point for decision-makers to understand the evidence underlying a specific topic in order "to inform healthcare decision makers' policy options" to improve practice and identify gaps where additional research is needed.[7] Overviews also provide an evidence map to assist decision

makers and clinicians with high level conclusions about the topic area.[7] The purpose of this study was to conduct an overview of SRs examining 1) the effectiveness of vitamin D supplementation during pregnancy and 2) the association of serum vitamin D levels with adverse pregnancy outcomes. We sought to identify, appraise and summarize existing SRs to gather the best available evidence in a single source[7] and clarify variable findings and conclusions across studies and SRs.

METHODS

General approach

To synthesize the available evidence in a way that would be most useful to clinicians and decision-makers we conducted a systematic overview of SRs following established methods.[8] In brief, we conducted a comprehensive search for existing SRs (January 2019), evaluated the SRs in terms of their quality and recency, collated the SR results for pre-specified perinatal outcomes, and graded the quality of available evidence (i.e., the certainty of the findings) using the Cochrane Collaboration and GRADE (Grading of Recommendations Assessment, Development and Evaluation) guidance principles.[9] Included SRs were independently assessed for methodological quality using the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist.[10,11]

Literature search strategy

On October 2, 2017, a research librarian with extensive experience conducting SRs carried out searches in Ovid Medline (1946-January 2019) and Wiley Cochrane Library databases (inception-January 2019): Cochrane Database of Systematic Reviews, Database of Abstracts of

Reviews of Effects (DARE), and the Health Technology Assessment (HTA) Database. Searches combined concepts for pregnancy and vitamin D supplementation with the Canadian Agency for Drugs and Technologies in Health study design filter for SRs (where applicable).[12] No publication date or language filters were applied. The full search was updated in January 2019. The search strategy is available in Supplementary Table 1. Search results were exported to EndNote X7 (Clarivate Analytics) and duplicates removed prior to screening in EndNote.

Eligibility criteria

We included SRs that 1) evaluated vitamin D supplementation in pregnant women of any gestational or chronological age, and/or 2) examined the effect of vitamin D on adverse pregnancy outcomes or the association between serum vitamin D levels and adverse pregnancy outcomes. We defined a SR as a "synthesis of research evidence in which literature searches, inclusion criteria, and critical appraisal methods were explicitly described."[7] We included SRs where vitamin D was administered in any dose or by any route, in comparison with placebo or other doses/forms of vitamin D supplementation. To be included, SRs had to report at least one of the following predefined maternal or neonatal outcomes: pre-term birth, preeclampsia, gestational diabetes, small for gestational age, still birth, low birth weight, and cesarean section. We excluded primary studies.

Selection

Two reviewers (LB, JS-K) independently screened all titles and abstracts and reviewed the full-text of studies that were identified as potentially eligible using standard eligibility criteria.

Reviewers compared results and resolved any discrepancies through discussion; where uncertainty remained decisions were made in discussion with the study team.

Assessment of SR quality

Two reviewers (LB, JS-K) independently assessed the methodological quality of all relevant SRs using the AMSTAR checklist.[10,11] This reliable and valid tool consists of 11 items regarding the methodological quality of a systematic review. Reviewers compared assessments for each of the 11 items in the AMSTAR checklist and resolved disagreements through discussion or third-party adjudication. Based on the total AMSTAR score (maximum 11 representing highest quality), we categorized the SRs by quality: low (0-3), medium (4-7), high (8-11).[12] Given the large number of high quality SRs, we focused data extraction and analysis on these.

Data collection

One experienced reviewer (LB) extracted data from the SRs using predefined standard forms developed for this overview. For each SR, review level data were extracted on objectives, publication date, country of origin, funding, search date range, inclusion and exclusion criteria, number of included studies, methods of analysis, and quantitative data on included outcomes. For each outcome present in a SR we abstracted study design, intervention, comparator, effect size, and direction of effect. All data were reviewed for accuracy and completeness by a second reviewer (JS-K).

Analysis

We present and discuss the results by SR for each of our predefined outcomes. We display results based on SRs examining: 1) the effectiveness of vitamin D supplementation (i.e., results from randomized controlled trials), and 2) the association between serum vitamin D levels and pregnancy outcomes (i.e., results from observational studies). For consistency of rating and based on GRADE recommendations[9] results were converted to risk ratios using the random effects model where possible (in three cases, we had insufficient information to convert the estimates and have reported these as per the original review).[13-15] For each of the pre-defined outcomes we reported any sub-group analyses based on dosage or levels of vitamin D.

Assessing the level of evidence

To assess the certainty of the results, we graded the quality of evidence presented by every SR for each outcome of interest. We followed recommendations of the GRADE Working Group,[16] and assessed the following key domains: risk of bias, inconsistency, indirectness, imprecision, and publication/reporting bias. Rather than rating individual studies GRADE rates individual outcomes across studies; therefore the quality of evidence can differ for different outcomes from the same set of studies or for the same outcomes based on different sets of studies.[17] For SRs of observational studies, we considered the additional domains of magnitude of effect, dose response relationships, and whether all plausible confounding would reduce an effect.[16] For both interventional and observational designs the GRADE assessment started at high quality of evidence, given the designs were appropriate to address questions of effectiveness and association respectively. Two reviewers (LB, LH) independently conducted GRADE assessments and resolved discrepancies through discussion. GRADEpro software was utilized to calculate overall quality of evidence.[9,18]

Patient involvement

This research was done without patient or public involvement.

RESULTS

Literature search results and study selection

Figure 1 details the flow of information through the stages of this overview using the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)[19] flow diagram. We identified 233 records from the search after removing duplicates. After title and abstract screening 42 records were identified. Three SRs did not report on any of our predefined outcomes and were excluded[20-22], and one SR was represented by both a Cochrane and journal publication reporting the same data.[23,24] Based on the AMSTAR assessment 25 reviews were categorized as low or medium quality and were not included in the data extraction and outcome assessment. In total 13 SRs were included in the final analysis. See Supplementary Table 2 for the completed PRISMA checklist.

Description of included systematic reviews

The 13 included reviews were published between 2009 and 2018, with a median AMSTAR score of 8 ranging from 8 to 11 (supplementary table 3 and 4). The literature search dates for these 13 reviews were between September 2014 and May 2018. All 13 SRs were published in English and were from China[15,25,26], Canada[27-29], Iran[30], Spain[31], Switzerland[23], United Kingdom[14], United States[13,32], and Thailand[33]. Four SRs included both RCTs and observational studies[13,14,26,32], 5 included only RCTs[23,27,28,31,33], and 4 included only observational studies.[15,25,29,30] All included

SRs with the exception of two[13,32] conducted a meta-analysis. Across the 13 SRs there were 204 unique studies (78 RCTs and 126 observational studies).

None of the SRs explicitly searched for low income or high risk populations, most studies reported their populations as generally healthy at study entry without pre-existing conditions. Individual study sample sizes ranged from 16 to 12,861. For interventional studies there was a wide range of dosing regimens, daily doses ranged from 200 to 5,000 International Units (IU); weekly doses from 714 to 50,000 IU; up to 60,000 IU monthly and bolus doses ranging from 35,000 to 1,200,000 (600,000 x 2) IU. Only two reviews reported sub-group analyses based on dose ranges.[27,28] One review had a sub-group for neonatal mortality and small for gestational age for high (>2000 IU/day) and low (≤ 2000 IU/day),[27] and the other review presented sub-groups for high (≥ 2000 IU/day) and low (< 2000 IU/day) doses for all outcomes.[28] Two reviews of observational studies presented their analyses based on subgroups of 25 OH(D) levels, <50 nmol/L and <75 nmol/L,[29] and <50 vs >50 nmol/L and <75 vs >75 nmol/L.[26]

Synthesis of results by outcome for SRs examining the effectiveness of vitamin D

Preterm birth. Five SRs of RCTs[23,26-28,31] examined the effectiveness of vitamin D compared to no treatment/placebo or calcium for prevention of preterm birth (Table 1). Four SRs found no significant difference in preterm birth rates, while one SR found a significant benefit with vitamin D. However, the quality of evidence varied across SRs (see supplementary table 5 for detailed GRADE assessments). One of the SRs had high quality of evidence[28] while the other four were rated as moderate, low and very low quality. The SR with high quality of evidence showed no significant benefit of vitamin D on prevention of preterm birth (RR 1.00, 95% CI 0.77, 1.30).[28] In subgroup analyses, these findings of no effect on preterm birth were robust, not

altered when baseline vitamin D status was low (<30 nmol/L), when only studies at low risk of bias were examined, or when the analysis was limited to generally healthy women. There were also no significant differences within subgroups based on the effective daily equivalent dose of vitamin D: <2000 IU/day (RR 0.8, 95% CI 0.40, 1.60; 5 studies, 1,503 participants); ≥2000 IU/day (RR 1.02, 95% CI 0.76, 1.36; 9 studies, 2,404 participants).

Table 1: Summary of results from SRs of randomized controlled studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	Level of evidence (GRADE)
PRE-TERM BIRTH					
Bi 2018	11/3,822	0.98 (0.77, 1.26)	33%	- (NR)	moderate
De-Regil / Palacios 2016	3 / 477	0.36 (0.14, 0.93)	10%	+ (0.035)	very low
Perez-Lopez 2015	3 / 384	1.24 (0.59, 2.61)	0%	- (0.56)	very low
Roth 2017*	**14 / 3,757**	**1.00 (0.77, 1.30)**	**0%**	**- (0.677)**	**high**
Zhou 2017	6 / 1,687	0.61 (0.34, 1.07)	26%	- (0.09)	low
PREECLAMPSIA					
De-Regil / Palaciosis 2016	2 / 219	0.52 (0.25, 1.05)	0%	- (0.069)	very low
Khaing 2017	3 / 357	0.47 (0.24, 0.89)	0%	+ (0.02)	very low
Newberry 2014	1/504	NR; by group for individual study	NR	+ (n=1) †	very low
Perez-Lopez 2015*	**3 / 654**	**0.91 (0.45, 1.86)**	**24%**	**- (0.80)**	**low**
Roth 2017	3 / 706	1.09 (0.43, 2.76)	66%	- (0.047)	very low
GESTATIONAL DIABETES					
De-Regil / Palacios 2016	2 / 219	0.43 (0.05, 3.45)	0%	- (0.43)	very low
Perez-Lopez 2015	3 / 384	1.05 (0.60, 1.85)	0%	- (0.86)	very low
Roth 2017	5 / 1,030	0.65 (0.39, 1.08)	45%	- (0.125)	very low
SMALL FOR GESTATIONAL AGE					
Bi 2018*	**6 / 1002**	**0.72 (0.52, 0.99)**	**0%**	**+ (0.04)**	**low**
Harvey 2014	2 / 245	NR; by individual study	NR	- (n=2)†	very low
Perez-Lopez 2015	3 / 456	0.77 (0.46, 1.30)	15%	- (0.33)	very low
Roth 2017	5 / 741	0.60 (0.40, 0.90)	0%	+ (0.704)	very low
LOW BIRTH WEIGHT					
Bi 2018	4/775	0.52 (0.20, 1.37)	65%	- (NR)	very low
De-Regil / Palacios 2016	3 / 493	0.4 (0.24, 0.67)	4%	+ (0.00048)	very low
Perez-Lopez 2015	4 / 496	0.72 (0.45, 1.17)	0%	- (0.19)	very low
Roth 2017*	**7 / 1,156**	**0.74 (0.47, 1.16)**	**47.3%**	**- (0.077)**	**low**
STILLBIRTH					

De-Regil / Palacios 2016	3 / 540	0.35 (0.06, 1.99)	0%	- (0.23)	low
Roth 2017*	**16 / 4,606**	**0.75 (0.50, 1.12)**	**0%**	**- (0.858)**	**high**
CESAREAN SECTION					
De-Regil / Palaciosis 2016	2 / 312	0.95 (0.69, 1.31)	12%	- (0.75)	low
Perez-Lopez 2015	4 / 1,028	0.97 (0.81, 1.32)	0%	- (0.75)	low
Roth 2017*	**16 / 3,240**	**1.02 (0.93, 1.12)**	**0%**	**- (0.701)**	**high**

* for each outcome the review with the highest level of evidence is presented in bold font
† in absence of pooled data this indicates the number of studies with positive or negative statistical significance
± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of RCTs examined the effectiveness of vitamin D for prevention of preeclampsia.[23,28,31-33] The quality of evidence for effectiveness of vitamin D for preeclampsia was low and very low; the four SRs that pooled findings from individual studies showed mixed results (Table 1). The SR that provided the highest level of evidence (classified as low quality) found a non-significant risk ratio of 0.91 (95% CI 0.45, 1.86).[31] One SR planned subgroup analyses based on dose; all studies reporting the outcome used ≥2000 IU/day, therefore results were the same as the overall pooled estimate, which showed no significant difference (RR 1.09, 95% CI 0.43, 2.76; 3 studies, 706 participants).[28]

Gestational diabetes. Three SRs of RCTs examined the effectiveness of vitamin D for prevention of gestational diabetes (Table 1).[23,31] None of the SRs found a significant effect with the use of vitamin D in terms of the occurrence of gestational diabetes. The quality of evidence was very low in all SRs. One SR conducted subgroup analyses based on dose and found a significant reduction for <2000 IU/day (RR 0.33, 95% 0.13, 0.82) (based on a single study with 87 participants). No significant difference was observed for the subgroup receiving ≥2000 IU/day (RR 0.75, 95% CI 0.44, 1.28; 4 studies, 943 participants).[28]

Small for gestational age. Four SRs of RCTs examined the effectiveness of vitamin D in terms of prevention of infants' birthweights being small for gestational age (Table 1).[14,27,28,31] Three of the SR authors conducted meta-analyses to come up with overall effect estimates, while the authors of one SR chose not to pool due to heterogeneity across the two included studies. The SR with the highest quality of evidence (classified as low) found a significant risk ratio of 0.72 (95% CI 0.52, 0.99). Subgroup analysis in one SR based on dose showed no significant differences for <2000 IU/day (RR 0.63, 95% CI 0.35, 1.11; 3 studies, 352 participants) and ≥2000 IU/day (RR 1.04, 95% CI 0.32, 3.36; 2 studies, 219 participants).[28] In another SR, results for a subgroup based on dose was significant for the lower doses ≤ 2000 IU/day (RR 0.45, 95% CI 0.23, 0.90; 2 studies, 209 participants) with no difference for >2000 IU/day (RR 0.83, 95% CI 0.57, 1.19; 5 studies, 713 participants).[27]

Low birth weight. Four SRs of RCTs examined the effectiveness of vitamin D to prevent low birth weight (birthweight <2500 grams) (Table 1).[23,27,28,31] One SR found a significant benefit while the other three SRs showed no difference. The SR with the highest quality of evidence (low) showed no significant difference (RR 0.74, 95% CI 0.47, 1.16).[28] Subgroup analyses based on dose in this SR showed no significant differences for <2000 IU/day (RR 0.53, 95% CI 0.23, 1.21; 1 study, 126 participants) and ≥2000 IU/day (RR 0.99, 95% CI 0.70, 1.42; 5 studies, 830 participants).[28]

Stillbirth. Two SRs of RCTs examined the effectiveness of vitamin D to prevent stillbirth (Table 1).[23,28] Neither of the SRs found a significant benefit. The SRs had high and low quality of evidence, respectively. The SR with high quality of evidence found a risk ratio of 0.75 (95% CI

0.50, 1.12).[28] Subgroup analyses based on dose from this SR showed a significant difference for <2000 IU/day (RR 0.49, 95% CI 0.27, 0.91; 7 studies, 1,948 participants) but no difference for ≥2000 IU/day (RR 1.03, 95% CI 0.62, 1.71; 9 studies, 2,713 participants).[28]

Cesarean section. Three SRs examined the effectiveness of vitamin D for cesarean sections (Table 1).[23,28,31] The quality of evidence ranged from low to high; none of the SRs found a significant effect. The SR providing high quality of evidence found a risk ratio of 1.02 (95% CI 0.93, 1.12).[28] Subgroup analyses from this SR based on dose showed no significant differences for <2000 IU/day (RR 1.00, 95% CI 0.85, 1.18; 6 studies, 702 participants) or ≥2000 IU/day (RR 1.04, 95% CI 0.91, 1.19; 8 studies, 2,303 participants).[28]

Synthesis of results by outcome for SRs examining associations of vitamin D with perinatal outcomes

Preterm birth. Five SRs of observational studies examined the association between vitamin D status and preterm birth (Table 2).[14,25,26,29,32] One SR that examined the association between vitamin D and preterm birth found moderate evidence of an association overall 1.19 (1.08, 1.31).[25] Two SRs presented their analyses based on subgroups of 25 OH(D) levels: <50 nmol/L and <75 nmol/L,[29] and <50 vs >50 nmol/L and <75 vs >75 nmol/L.[26] In both SRs the association was slightly greater for the lower serum vitamin D level. The SR with highest quality of evidence found a significant association with moderate quality evidence for <50 vs. >50 nmol/L 1.13 (95% CI 1.04, 1.23) and non-significant association and low quality evidence for <75 vs. >75 nmol/L 1.03 (95% CI 0.98, 1.08).[26]

Table 2: Summary of results for SRs of observational studies

Review	Number studies / individuals	Effect size (CI) risk ratio, random effects	Heterogeneity (I^2)	Significance (p-value) ±	GRADE
PRETERM BIRTH					
Harvey 2014	7 / 1,792	NR; 1 individual study showed significance and 6 others not significant	NR	+ (n=1) ‡ - (n=6)	very low
Newberry 2014	2 / 371	NR; by individual study	NR	+ (n=1)‡ - (n=1)	very low
Qin 2016*	**10 / 10,098**	**1.19 (1.08, 1.31)**	**28%**	**+ (0.004)**	**moderate**
Wei 2013	4 / 1,111	1.27 (1.03, 1.58) [blood level 25(OH)D <50nmol/L]	28%	- (0.03)	very low
		1.05 (0.98, 1.12) [blood level 25(OH)D <75nmol/L]	0%	- (0.17)	very low
Zhou 2017*	**16 / 16,996**	**1.13 (1.04, 1.23) [<50 vs >50 nmol/L]**	**45%**	**+ (0.003)**	**moderate**
	15 / 17,122	1.03 (0.98, 1.08) [<75 vs >75 nmol/L]	65%	- (0.29)	low
PREECLAMPSIA					
Chung 2009	1 / 1,189	5 (1.7, 14.1) †	NR	+ (n=1) ‡	very low
Harvey 2014	4 / 642	0.75 (0.48, 1.19) †	80.8%	- (0.001)	very low
Newberry 2014	8 / 4420	NR; by individual study	NR	+ (n=5) - (n=3)	very low
Tabesh 2013	8 / 2,485	2.02 (1.26, 3.23)	53%	+ (0.04)	very low
Wei 2013*	**6 / 2,008**	**1.57 (1.21, 2.03) [<50 nmol/L]**	**39%**	**+ (0.0006)**	**low**
	5 / 1,311	1.21 (0.99, 1.46) [<75 nmol/L]	60%	- (0.06)	very low
GESTATIONAL DIABETES					
Harvey 2014	8 / 2,668	NR; by individual study	NR	+ (n=3) ‡ - (n=5)	very low
Lu 2016	20 / 16,515	1.45 (1.15, 1.83) †	66.6%	+ (0.002)	low
Wei 2013*	**10 / 4,126**	**1.12 (1.02, 1.22) [<50 nmol/L]**	**27%**	**+ (0.02)**	**moderate**
	8 / 3,840	**1.09 (1.03, 1.15) [<75 nmol/L]**	**28%**	**+ (0.002)**	**moderate**
SMALL FOR GESTATIONAL AGE					
Harvey 2014	7 / 5,660	NR; by individual study	NR	+ (n=2) ‡ - (n=5)	very low
Newberry 2014	1 / 412	NR; by individual study	NR	NR	very low
Wei 2013*	**6 / 6,013**	**1.35 (1.18, 1.54) [<50 nmol/L]**	**15%**	**+ (0.00001)**	**low**
	5 / 2,283	0.99 (0.83, 1.18) [<75 nmol/L]	75%	- (0.92)	very low
LOW BIRTH WEIGHT					
Harvey 2014	3 / 1,676	NR; by individual study	NR	+ (n=1) ‡ - (n=2)	very low
CESAREAN SECTION					
Harvey 2014	6 / 3,277	NR; by individual study	NR	+ (n=2) ‡	very low

				- (n=4)	

* for each outcome the review with the highest level of evidence is presented in bold font
† reported as odds ratios as insufficient data available to convert to risk ratio
‡ in absence of pooled data this indicates the number of studies with positive or negative statistical significance
± significance indicated as positive (+) when p-value <= 0.05 and negative (-) if > 0.05

Preeclampsia. Five SRs of observational studies examined the association between vitamin D status and preeclampsia (Table 2).[13,14,29,30,32] Three of the five SRs found a significant association.[13,29,30] One SR assessed different serum levels of vitamin D and found a larger point estimate for <50 nmol/L compared with <75 nmol/L, although the confidence intervals overlapped.[29] The quality of evidence was low for <50 nmol/L and very low for <75 nmol/L.

Gestational diabetes. Three SRs of observational studies provided measures of association for vitamin D status and gestational diabetes (Table 2).[14,15,29] The SR providing the highest quality of evidence showed moderate quality evidence of a significant association for both serum levels examined: <50 nmol/L: 1.12 (95% CI 1.02, 1.22), <75 nmol/L: 1.09 (95% CI 1.03, 1.15).[29]

Small for gestational age. Three SRs of observational studies examined the association between vitamin D status and small birthweights for gestational age (Table 2).[14,29,32] The SRs showed mixed findings. One SR included 7 studies but did not pool results as the authors stated there was substantial variation in methodology and exposure;[14] 2 studies showed a significant association while 5 studies showed no significant effect (very low quality of evidence). Another SR only included 1 study and could not pool any results.[32] The highest rated (low quality) SR examined the association for different vitamin D serum levels and found a significant association for <50 nmol/L 1.35 (95% CI 1.18, 1.54), but no significant effect for <75 nmol/L 0.99 (95% CI 0.83, 1.18).[29] The quality of evidence was low for <50 nmol/L and very low for <75 nmol/L.

Low birth weight. Only one SR of observational studies examined the association between vitamin D status and low birth weight.[14] The SR included three studies but did not pool results. One study showed a statistically significant result while two studies had non-significant findings. Overall the quality of evidence for this outcome is very low.

Stillbirth. There were no SRs of observational studies that examined the association between vitamin D status and stillbirth.

Cesarean section. Only one SR of observational studies examined the association between vitamin D status and cesarean section.[14] The SR included six studies but did not pool results; the authors chose not to combine due to a multitude of factors such as local policies and physician preferences that influence this outcome. Two studies showed a statistically significant association while four studies had non-significant findings. Overall the quality of evidence for this outcome is very low.

DISCUSSION

This overview provides a comprehensive analysis of SRs examining vitamin D and pregnancy outcomes. We grouped and reported results separately for SRs of RCTs and SRs of observational studies. SRs of observational studies showed evidence of associations between vitamin D serum levels and some outcomes, however SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy with the exception of one pre-defined outcome—small for gestational age—which had low quality evidence. The differences in findings between

these groups of SRs suggest that any apparent associations may not be based on causal relationships. They suggest that low vitamin D levels or deficiencies may be an indicator or marker of poor health status, co-morbidities,[1] or perhaps an acute phase reactant.[34,35] It is likely that pregnant women with these indicators need more attention and care to optimize health outcomes for them and their offspring and not vitamin D supplementation. The current evidence does not support the use of vitamin D supplementation to improve any of these outcomes.

While there were some suggestions of associations between low vitamin D serum levels and some outcomes in the observational studies (i.e., preterm birth, preeclampsia, gestational diabetes, and small for gestational age), the effect sizes may be considered not clinically important.[9] The quality of this observational evidence was almost all low or very low.[16] However, more applicable to clinical practice are the findings from SRs of RCTs that examined the effectiveness of vitamin D as a treatment to improve pregnancy outcomes. The SRs of RCTs that provided the highest quality of evidence showed no effect of vitamin D supplementation in pregnancy for all but one of the pre-defined outcomes of interest. Overall these findings suggest that even if an association exists between vitamin D levels and health outcomes, vitamin D supplementation in pregnancy may be unlikely to improve these outcomes.

This study provides a methodologically rigorous and comprehensive synthesis of an extensive body of evidence examining vitamin D and perinatal outcomes. The considerable number of primary studies and SRs underscores the importance of this topic as well as the uncertainty about whether and how to manage vitamin D levels to optimize health outcomes. However, the vast number of SRs on this topic is concerning, particularly those of low quality which may propagate

inaccurate or biased results and conclusions. Of note, in our update search that captured the most recent publications up to January 2019, we identified 10 new relevant SRs with only three having an AMSTAR score greater than 7 to be included in the final analysis. Of these 3 new SRs there was only one new primary study included. We have provided an in-depth analysis by presenting the results of SRs of RCTs that evaluated the effectiveness of vitamin D as a treatment to improve perinatal outcomes alongside SRs of observational studies that examined the associations between vitamin D levels and health outcomes. Further, we used GRADE's rigorous and transparent method to assess the quality of the body of evidence which provides essential information about the certainty of the effect estimates in order to reconcile findings across individual studies and reviews.

The evidence contributing to the existing SRs varied widely in design and purpose (to examine associations vs. effectiveness). Observational studies have been used to examine the association between vitamin D levels and health outcomes, and are appropriate for generating hypotheses for testing in randomized trials. One of the limitations of the existing observational studies and synthesis of the same is that individual studies may or may not sufficiently adjust for confounding[36] (e.g., health status, calcium intake, and social determinants of health). Further, studies that did adjust for confounding differed in the variables they included and controlled for. RCTs, when well-designed, represent a higher level of evidence to assess the effectiveness of an intervention, in part because they can address the problem of confounding as randomization is intended to equally distribute both known and unknown confounders. It is well documented that early and observational studies often suggest important relationships that do not exist, and that well designed RCTs are often needed to fully understand a phenomenon.[37]

An important limitation in this area of investigation is the possibility of reporting and publication bias. While we focused on the highest quality SRs and most indicated that they planned to investigate publication bias, many could not do so because the number of included studies in a given meta-analysis was too small. There remains the possibility that studies, particularly the earlier published studies, showing significant results are more likely to be published while those with non-significant findings remain unpublished.

Also, the potential for selective outcome reporting is important in this body of literature. It is surprising that outcomes that are either routinely collected or relatively easy to ascertain, such as preterm birth, stillbirth, and gestational diabetes were infrequently reported. Selective outcome reporting occurs if researchers focus their reporting on a significant finding and downplay or do not report non-significant results. For example, the most frequently reported outcome was preterm birth; however, among the 48 studies included in one systematic review only one-third of the primary studies reported this outcome. Roth et al. also found that "missing data on clinical outcomes was the norm rather than exception" in this body of literature which could lead to "potentially biased meta-analyses based on small non-representative subsets of trials and participants".[28] Important efforts have been made to define core outcomes sets in the area of perinatal research.[38-40] Future studies should focus on critical outcomes for this field. Researchers should also define their outcomes and analyses a priori, register (and ideally publish) study protocols, and ensure clear and transparent reporting.[41,42] Further, researchers should identify all important confounders and address these adequately through appropriate research designs or analytic approaches to ensure valid findings and permit meaningful pooling of data. Currently

the credibility of the body of evidence in this important field is compromised due to the potential for confounding, publication bias, reporting bias, and imprecision arising from low numbers of participants.

CONCLUSIONS

While there is some evidence from SRs of observational studies for an association between maternal vitamin D serum levels and some perinatal outcomes, SRs examining effectiveness from RCTs showed no effect of vitamin D supplementation in pregnancy for all but one pre-defined outcome, the evidence for the one outcome was low quality. The discrepancy between the observational studies and the RCTs shows that 25-hydroxy vitamin D is lower among women who experience adverse pregnancy outcomes, but supplementation does not appear to alter outcomes. Low 25-hydroxy vitamin D, the indicator of vitamin D status, is a marker of adverse outcomes rather than a marker of vitamin D status[1] which shows that 25-hydroxy vitamin D may be an acute phase reactant. Future studies need to adequately control for potential confounding (e.g., through well-designed randomized trials)[42] and include all outcomes that are considered critical to this field. There are currently over 40 published SRs (many of which are low quality) synthesizing evidence from 204 primary vitamin D studies; further SRs on this topic are wasteful until more well designed and conducted RCTs are completed.

FIGURE LEGENDS

FIGURE 1: Flow diagram of screening decisions

ACKNOWLEDGEMENTS: Research reported in this publication was funded by the Alberta Health Services' Maternal, Newborn, Child & Youth Strategic Clinical Network™ (MNCY SCN™). Additional support was provided by the Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform which is funded by the Canadian Institutes of Health Research and Alberta Innovates. Dr. Hartling is supported by a Canada Research Chair in Knowledge Synthesis and Translation. The content is solely the responsibility of the authors and does not necessarily represent the views of Alberta Health Services. We would like to thank: Tara Landry, MLIS, for peer reviewing the search strategy; Ms. MacKinna Hauff for article retrieval; Robin Featherstone, MLIS, for developing and running the search; and Dr. Seija Kromm for administrative support.

CONTRIBUTORS: LB, LH, TF, DM and DJ designed the study. LB, LH and JKS selected articles, extracted data and performed the assessment of bias. LB supervised study activities and LH wrote the first draft of the manuscript. All authors critically reviewed and revised the manuscript and approved the final version for publication. The corresponding author attests that all listed authors meet authorship criteria and no others meeting the criteria have been omitted.

within and any related or standalone film submitted (the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. I am the sole author of the Contribution and I am one author signing on behalf of all co-owners of this Contribution.

COMPETING INTERESTS: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

DATA SHARING: The dataset is available from the lead author on request.

PATIENT AND PUBLIC INVOLVEMENT: This research was done without patient or public involvement.

REFERENCES

1. Autier P, Mullie P, Macacu A, et al. Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials. *Lancet Diabetes Endocrinol*. 2017;5(12):986-1004.
2. Wagner CL, Baggerly C, McDonnell S, et al. Post-hoc analysis of vitamin D status and reduced risk of preterm birth in two vitamin D pregnancy cohorts compared with South Carolina March of Dimes 2009-2011 rates. *J Steroid Biochem Mol Biol*. 2016;155(Pt B):245-251.
3. Leffelaar ER, Vrijkotte TG, van Eijsden M. Maternal early pregnancy vitamin D status in relation to fetal and neonatal growth: results of the multi-ethnic Amsterdam Born Children and their Development cohort. *Br J Nutr*. 2010;104(1):108-117.
4. Soheilykhah S, Mojibian M, Rashidi M, et al. Maternal vitamin D status in gestational diabetes mellitus. *Nutr Clin Pract*. 2010;25(5):524-527.
5. Holmes VA, Barnes MS, Alexander HD, et al. Vitamin D deficiency and insufficiency in pregnant women: a longitudinal study. *Br J Nutr*. 2009;102(6):876-881.
6. Bodnar LM, Catov JM, Simhan HN, et al. Maternal vitamin D deficiency increases the risk of preeclampsia. *J Clin Endocrinol Metab*. 2007;92(9):3517-3522.
7. Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? *Syst Rev*. 2013;2:26.
8. Pollock M Fernandes RM, Becker LA, et al. Chapter V: Overviews of Reviews. Draft version (8 October 2018) for inclusion in: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch V (editors). Cochrane Handbook for Systematic Reviews of Interventions. London: Cochrane.
9. Schünemann H, Brozek J, Guyatt G, et al, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html. Accessed June 3, 2019.
10. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10.
11. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62(10):1013-1020.
12. Canadian Agency for Drugs and Technologies in Health (CADTH). Rapid response summary with critical appraisal: Process. 2015; https://www.cadth.ca/about-cadth/what-we-do/products-services/rapid-response-service. Accessed April 19, 2018.
13. Chung M, Balk EM, Brendel M, et al. Vitamin D and calcium: a systematic review of health outcomes. *Evid rep/technol assess*. 2009(183):1-420.

14. Harvey NC, Holroyd C, Ntani G, et al. Vitamin D supplementation in pregnancy: a systematic review. *Health Technol Assess*. 2014;18(45):1-190.
15. Lu M, Xu Y, Lv L, et al. Association between vitamin D status and the risk of gestational diabetes mellitus: a meta-analysis. *Arch Gynecol Obstet*. 2016;293(5):959-966.
16. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490.
17. Guyatt G, Oxman AD, Elie A, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394.
18. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University, 2015 (developed by Evidence Prime, Inc.). Available from: https://gradepro.org. Accessed May 27, 2019.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;8(5):336-341.
20. Yepes-Nunez JJ, Brozek JL, Fiocchi A, et al. Vitamin D supplementation in primary allergy prevention: Systematic review of randomized and non-randomized studies. *Allergy*. 2017;04:04.
21. Zhang H, Huang Z, Xiao L, et al. Meta-analysis of the effect of the maternal vitamin D level on the risk of spontaneous pregnancy loss. *Int J Gynaecol Obstet*. 2017;138(3):242-249.
22. Christensen N, Sondergaard J, Fisker N, et al. Infant Respiratory Tract Infections or Wheeze and Maternal Vitamin D in Pregnancy: A Systematic Review. *Pediatr Infect Dis J*. 2017;36(4):384-391.
23. De-Regil LM, Palacios C, Lombardo LK, et al. Vitamin D supplementation for women during pregnancy. *Cochrane Database Syst Rev*. 2016(1):CD008873.
24. Palacios C, De-Regil LM, Lombardo LK, et al. Vitamin D supplementation during pregnancy: Updated meta-analysis on maternal outcomes. *J Steroid Biochem Mol Biol*. 2016;164:148-155.
25. Qin LL, Lu FG, Yang SH, et al. Does Maternal Vitamin D Deficiency Increase the Risk of Preterm Birth: A Meta-Analysis of Observational Studies. *Nutrients*. 2016;8(5):20.
26. Zhou SS, Tao YH, Huang K, et al. Vitamin D and risk of preterm birth: Up-to-date meta-analysis of randomized controlled trials and observational studies. *J Obstet Gynaecol Res*. 2017;43(2):247-256.
27. Bi WG, Nuyt AM, Weiler H, et al. Association Between Vitamin D Supplementation During Pregnancy and Offspring Growth, Morbidity, and Mortality: A Systematic Review and Meta-analysis. *JAMA Pediatr*. 2018;172(7):635-645.
28. Roth DE, Leung M, Mesfin E, et al. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. *BMJ*. 2017;359:j5237.
29. Wei SQ, Qi HP, Luo ZC, et al. Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis. *J Matern Fetal Neonatal Med*. 2013;26(9):889-899.

30. Tabesh M, Salehi-Abargouei A, Tabesh M, et al. Maternal vitamin D status and risk of pre-eclampsia: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2013;98(8):3165-3173.

31. Perez-Lopez FR, Pasupuleti V, Mezones-Holguin E, et al. Effect of vitamin D supplementation during pregnancy on maternal and neonatal outcomes: a systematic review and meta-analysis of randomized controlled trials. *Fertil Steril*. 2015;103(5):1278-1288.e1274.

32. Newberry SJ, Chung M, Shekelle PG, et al. Vitamin D and Calcium: A Systematic Review of Health Outcomes (Update). *Evid rep/technol assess*. 2014(217):1-929.

33. Khaing W, Vallibhakara SA, Tantrakul V, et al. Calcium and Vitamin D Supplementation for Prevention of Preeclampsia: A Systematic Review and Network Meta-Analysis. *Nutrients*. 2017;9(10):18.

34. Madden K, Feldman HA, Chun RF, et al. Critically Ill Children Have Low Vitamin D-Binding Protein, Influencing Bioavailability of Vitamin D. *Ann Am Thorac Soc*. 2015;12(11):1654-1661.

35. Silva MC, Furlanetto TW. Does serum 25-hydroxyvitamin D decrease during acute-phase response? A systematic review. *Nutr Res*. 2015;35(2):91-96.

36. Patel CJ, Manrai AK. Development of exposome correlation globes to map out environment-wide associations. *Pac Symp Biocomput*. 2015:231-242.

37. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol*. 2011;64(12):1277-1282.

38. Molloy EJ, Gale C, Marsh M, et al. Developing core outcome set for women's, newborn, and child health: the CROWN Initiative. *Pediatr Res*. 2018;84(3):316-317.

39. Devane D, Begley CM, Clarke M, et al. Evaluating maternity care: a core set of outcome measures. *Birth*. 2007;34(2):164-172.

40. van 't Hooft J, Duffy JM, Daly M, et al. A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol*. 2016;127(1):49-58.

41. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg*. 2014;12(12):1495-1499.

42. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg*. 2012;10(1):28-55.

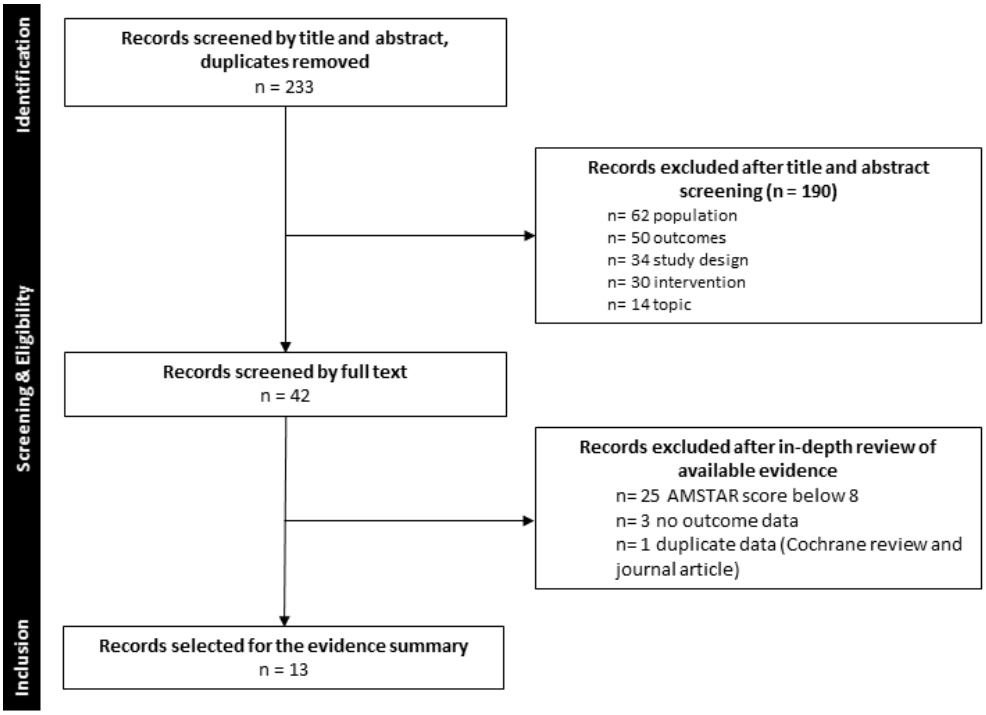


Figure 1: Study flow diagram

Supplementary Table 1: Literature search strategy

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Date conducted: 2 October 2017
Strategy:

1 Preconception Care/ (1917)
2 exp Pregnancy/ (855216)
3 exp Pregnancy Complications/ (405775)
4 Pregnant Women/ (6515)
5 Prenatal Care/ (24637)
6 Prenatal Diagnosis/ (35834)
7 (antenatal* or pre-natal* or prenatal*).tw,kf. (120088)
8 (expect* adj2 (female? or mother? or wom#n)).tw,kf. (3728)
9 ((1* or first*) adj2 (tri-mester* or trimester*)).tw,kf. (23473)
10 (pre-conception* or preconception*).tw,kf. (4573)
11 pregnan*.tw,kf. (477757)
12 or/1-11 [Combined MeSH & text words for pregnancy] (1016791)
13 exp Vitamin D/ (54287)
14 Vitamin D Deficiency/ (13412)
15 calcidiol*.tw,kf. (397)
16 calciol*.tw,kf. (20)
17 calcifediol*.tw,kf. (128)
18 cholecalciferol*.tw,kf. (2377)
19 hydroxycholecalciferol*.tw,kf. (1377)
20 hydroxyvitamin D*.tw,kf. (12499)
21 (vitamin D or vitamin D3 or vitamin D$2).tw,kf. (60203)
22 or/13-21 [Combined MeSH & text words for vitamin D] (79566)
23 and/12,22 [Combined concepts for pregnancy & vitamin D] (4365)
24 meta-analysis.pt. (87537)
25 meta-analysis/ or systematic review/ or meta-analysis as topic/ or "meta analysis (topic)"/ or "systematic review (topic)"/ or exp technology assessment, biomedical/ (113240)
26 ((systematic* adj3 (review* or overview*)) or (methodologic* adj3 (review* or overview*))).ti,ab,kf,kw. (127445)
27 ((quantitative adj3 (review* or overview* or synthes*)) or (research adj3 (integrati* or overview*))).ti,ab,kf,kw. (8559)
28 ((integrative adj3 (review* or overview*)) or (collaborative adj3 (review* or overview*)) or (pool* adj3 analy*)).ti,ab,kf,kw. (19993)
29 (data synthes* or data extraction* or data abstraction*).ti,ab,kf,kw. (20992)
30 (handsearch* or hand search*).ti,ab,kf,kw. (7877)
31 (mantel haenszel or peto or der simonian or dersimonian or fixed effect* or latin square*).ti,ab,kf,kw. (21571)
32 (met analy* or metanaly* or technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf,kw. (7548)
33 (meta regression* or metaregression*).ti,ab,kf,kw. (5904)

34 (meta-analy* or metaanaly* or systematic review* or biomedical technology assessment* or bio-medical technology assessment*).mp,hw. (217154)
35 (medline or cochrane or pubmed or medlars or embase or cinahl).ti,ab,hw. (161733)
36 (cochrane or (health adj2 technology assessment) or evidence report).jw. (18083)
37 (meta-analysis or systematic review).mp. [sic – changed .md. to .mp] (200672)
38 (comparative adj3 (efficacy or effectiveness)).ti,ab,kf,kw. (10906)
39 (outcomes research or relative effectiveness).ti,ab,kf,kw. (7938)
40 ((indirect or indirect treatment or mixed-treatment) adj comparison*).ti,ab,kf,kw. (1649)
41 or/24-40 [CADTH SR search filter | Retrieved from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#syst] (358374)
42 and/23,41 [SR filter applied] (187)
43 remove duplicates from 42 (164)

Database: Wiley Cochrane Library
Date conducted: 2 October 2017
Strategy:
#1 [mh ^"Preconception Care"] 103
#2 [mh Pregnancy] 5760
#3 [mh "Pregnancy Complications"] 9364
#4 [mh ^"Pregnant Women"] 156
#5 [mh ^"Prenatal Care"] 1332
#6 [mh ^"Prenatal Diagnosis"] 380
#7 (antenatal* or "pre-natal*" or prenatal):ti,ab,kw 6295
#8 (expect* near/2 (female? or mother? or wom?n)):ti,ab,kw 243
#9 ((1* or first*) near/2 ("tri-mester*" or trimester*)):ti,ab,kw 4141
#10 ("pre-conception*" or preconception*):ti,ab,kw 307
#11 pregnan*:ti,ab,kw 36386
#12 {or #1-#11} 38579
#13 [mh "Vitamin D"] 2941
#14 [mh ^"Vitamin D Deficiency"] 617
#15 calcidiol*:ti,ab,kw 46
#16 calciol*:ti,ab,kw 1
#17 calcifediol*:ti,ab,kw 475
#18 cholecalciferol*:ti,ab,kw 1208
#19 hydroxycholecalciferol*:ti,ab,kw 338
#20 "hydroxyvitamin D*":ti,ab,kw 1931
#21 ("vitamin D" or "vitamin D3" or "vitamin D?"):ti,ab,kw 6774
#22 {or #13-#21} 7581
#23 #11 and #22 354
#24 #11 and #22 in Cochrane Reviews (Reviews and Protocols), Other Reviews and Technology Assessments 14

PRISMA 2009 Checklist

Supplementary Table 2: PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3-4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6-7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary Table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	NA
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	9

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10-11, Supplementary Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	10
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-13, 16-17
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	12-13, 16-17
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Supplementary Table 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	23

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Supplementary Table 3: Description of included systematic reviews

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Bi Canada May 2018	24 RCTs 5,405 (30 – 965)	Population was healthy, pregnant women without prior vitamin D supplementation of more than 400 IU/d	Vitamin D in the form of cholecalciferol in 22 RCTs and in the form of ergocalciferol in 3 RCTs daily doses: 800 - 5000; weekly doses 35000 or 50000; fortnightly dose 50000; monthly dose 60000; bimonthly dose 60000; and bolus doses 60000 - 200 000	Placebo, no intervention or other dose of vitamin D	Primary: small for gestational age (indicated by birthweight less than the 10th percentile for gestational age, fetal or neonatal mortality Secondary: neonatal (25[OH]D) levels, congenital malformation, admission to a neonatal intensive care unit (NICU), Apgar scores, neonatal calcium levels, birth weight, low birth weight, gestational age, preterm birth, infant growth, asthma, respiratory infection, eczema, and allergy
Khaing Thailand October 2017 Roth Canada September 2017	19 RCTs 28,000 (30 – 9,178) 43 RCTs 8,406 (16 – 1,134)	Pregnant women of any gestational age Participants were pregnant at enrolment or enrolled before pregnancy and then followed-up in pregnancy	Calcium, vitamin D, combined calcium and vitamin D Vitamin D vs. placebo = 3; Calcium + vitamin D vs. calcium = 1 Vitamin D2 or D3, alone or in combination provided the co-intervention is similar in at least one other trial arm Daily doses: 400 – 5000; weekly doses: 714 – 7543; monthly doses: 1645 – 3289; bolus doses: 60000 – 1200000 (600000 x 2)	Placebo, a standard supplementation (e.g., folic acid), or no supplementation Placebo, no vitamin D, or vitamin D up to 600 IU/day (or a less frequent dose that would be about equivalent to 600 IU/day—for example, 4200 IU/week)	Primary: preeclampsia, eclampsia, proteinuria (dipstick urine 2+ or '300 mg/24 h), end-organ dysfunction, or utero-placental dysfunction after 20 weeks of gestation Primary: 25 OHD, preeclampsia, gestational diabetes, gestational hypertension, intra-uterine death, stillbirth, c-section, weight gain, preterm labor, death, adverse events, hospitalizations, birth weight, birth length, head circumference, low birth weight, small for gestational age, gestational age at birth, congenital malformations, neonatal death, respiratory infection, asthma, bone mineral content and density

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Zhou China June 2016	6 RCTs; 9 prospective cohort; 4 nested case-control; 2 cross-sectional; 2 retrospective cohort; 1 case-control 28,391 (50 – 12,861)	Pregnant women without HIV infection	maternal serum 25-OHD or oral supplementation with vitamin D Daily doses of 1,000 to 4,000 IU; weekly doses of 400 daily for 9 weeks; 50,000 for 6 weeks; one time doses starting 60,000 or 2-4 doses of 120,000	no supplementation /placebo, or routine care (ferrous sulfate and calcium, but no vitamin D)	Preterm birth
Qin China August 2015	4 Prospective cohort; 4 Nested case-control; 1 case-control; 1 Retrospective cohort; 1 Cross-sectional 20,608 (134 – 12,861)	Pregnant women without pre-chronic disease or HIV infection, with singleton gestation	NR; measurement of maternal vitamin D levels		Preterm birth
Lu China February 2015	4 Case-control; 7 Cohort; 2 Cross sectional; 7 Nested case control 16,515 (122 – 4,090)	NR	NR; measurement of maternal vitamin D levels		Gestational diabetes
De-Regil / Palacios Switzerland / Puerto Rico February 2015	15 RCTs 2,833 (40 – 990)	Pregnant women of any gestational or chronological age, parity (number of births) and number of fetuses	Vitamin D daily doses: 200 - 2000 Vitamin D single dose: 200,000 – 600,000, and 35,000	No intervention / placebo	Primary: pre-eclampsia, gestational diabetes, vitamin D concentration, adverse effects, preterm birth, low birthweight Secondary: impaired glucose tolerance, c-section, gestational hypertension, maternal death, birth length, head circumference at birth, birthweight, admission to special care, stillbirth, neonatal death, very preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
Newberry USA September 2014	2 RCTs; 2 prospective cohorts; 5 nested case-control 4,912 (160 – 1,141)	Primary population of interest is generally healthy people with no known disorders Only including studies for population contributing to pregnancy related outcomes	Vitamin D single doses (for RCT): 2000, 4000 followed by 1 month run-in at 2000	All participants enrolled into one of two vitamin D groups	Preeclampsia, preterm birth, small for gestational age
Perez-Lopez Spain March 2014	13 RCTs 2,299 (40 – 400)	Pregnant women of any gestational or chronologic age and parity, without previous disease history	Vitamin D alone vs. no treatment (placebo); vitamin D + calcium vs. no treatment (placebo); and vitamin D + calcium vs. calcium Daily doses ranged from 400 to 1,000; weekly doses ranged from 35,000 to 50,000; and single doses ranged from 200,000 to 600,000	Active controls, usual treatment without active control, and placebo	Primary: circulating 25-OHD, preeclampsia, gestational diabetes, small for gestational age, low birth weight, preterm birth, birthweight Secondary: birth length, c-section,
Wei Canada October 2012	13 Case-control; 8 cohort; 2 cross-sectional 12,898 (95 – 3,730)	Pregnant women without pre-existing chronic disease or HIV infection	NR; measurement of maternal vitamin D levels		Preeclampsia, gestational diabetes, preterm birth, small for gestational age
Harvey UK June 2012	17 Case-control; 48 cohort/cross-sectional; 9 RCT; 2 intervention studies (non-randomized) NR	Pregnant women or pregnant women and their offspring	vitamin D status [dietary intake, sunlight exposure, circulating 25(OH)D concentration] or supplementation of	For intervention studies: no intervention or placebo	Primary: neonatal hypocalcaemia, rickets in the offspring, offspring bone mass, and maternal osteomalacia Secondary: offspring body composition; offspring preterm birth

First Author Country Last assessed up-to-date	Number of studies Sample size (range)	Population	Intervention Doses in IU	Comparison	Outcomes for which data are reported
			participants with vitamin D or food containing vitamin D (e.g. oily fish)		and later offspring health outcomes; maternal quality of life
Tabesh Iran December 2012	2 Cohort; 4 cross-sectional; 9 case-control 2,936 (32 – 697)	Normal pregnant women	NR; measurement of maternal vitamin D levels		Preeclampsia
Chung USA April 2009	60 RCT; 3 NRCT; 102 cohort or nested case-control; 11 SR NR	Generally healthy people with no known disorders	Vitamin D supplements (no analogues), calcium supplements, and combinations of supplements; food based interventions	NR	Pregnancy-related: preeclampsia, high blood pressure with or without proteinuria, preterm birth or low birth weight, infant mortality

Supplementary Table 4. AMSTAR score by category and individual systematic review

Review	AMSTAR question											
	Q1 A priori design provided	Q2 Duplicate study selection and data extraction	Q3 Comprehensive literature search	Q4 Publication status as inclusion criterion	Q5 List of studies (include and exclude) provided	Q6 Characteristics of the included studies provided	Q7 Quality assess-ment	Q8 Quality used appropriate	Q9 Methods used to combine appropriate	Q10 Publication bias assessed	Q11 Conflict of interest stated	Total
OVERALL HIGH QUALITY												
Bi 2018	n	y	y	n	n	y	y	y	y	y	y	8
Christensen 2017	y	y	y	y	n	y	y	y	y	n	y	9
Chung 2009	y	ca	y	y	y	y	y	y	y	ca	y	9
De-Regil 2016	y	y	y	y	y	y	y	y	y	y	y	11
Harvey 2014	y	y	y	y	n	y	y	y	y	ca	y	9
Khaing 2017	y	y	n	n	n	y	y	y	y	y	y	8
Lu 2016	y	y	y	n	n	y	y	y	y	y	y	9
Newberry 2014	y	ca	y	n	y	y	y	y	y	ca	y	8
Palacios 2016	y	y	y	y	y	y	y	y	y	y	y	11
Perez-Lopez 2015	y	y	y	y	n	y	y	y	ca	ca	y	8
Qin 2016	n	y	y	n	n	y	y	y	y	y	y	8
Roth 2017	y	y	y	y	n	y	y	y	y	ca	y	9
Tabesh 2013	y	y	y	y	n	y	n	n	y	y	y	8
Wei 2013	n	y	y	n	n	y	y	y	y	y	y	8
Yepes-Nunez 2017	n	y	y	y	y	y	y	y	y	y	y	10
Zhang 2017	n	y	y	n	n	y	y	y	y	y	y	8
Zhou 2017	n	y	y	n	n	y	y	y	y	y	y	8
OVERALL MEDIUM AND LOW QUALITY												
Aghajafari 2013	n	y	y	ca	n	y	n	ca	y	y	n	5
Amegah 2017	n	y	y	n	n	y	y	y	y	y	n	6
Amraei 2018	ca	y	y	n	n	y	ca	ca	y	y	y	6
Arain 2015	n	y	ca	n	n	y	n	ca	ca	n	n	2
Chen 2017	n	y	y	n	n	y	y	y	y	y	n	6
Christensen 2012	n	y	n	n	n	y	n	n	n	n	y	3
Fu 2017	n	ca	y	n	n	n	n	n	y	y	y	4
Galthen-Sorensen 2014	n	y	y	n	n	y	y	y	n	n	n	5
Hu 2018	n	y	y	y	n	y	n	ca	y	y	y	7
Hypponen 2014	n	ca	y	n	n	y	ca	n	y	y	y	5
Kamudoni 2016	n	ca	y	y	n	y	n	n	n	n	y	4
Mahomed 2009	y	n	y	y	y	y	ca	ca	n	n	y	6
Martinez-Dominquez 2018	n	ca	y	n	n	y	y	ca	y	y	y	6
Nassar 2011	y	ca	n	n	n	y	n	n	y	n	y	4

Poel 2012	n	ca	y	n	n	y	n	n	ca	y	y	4
Purswani 2017	n	y	n	n	n	y	n	y	y	n	y	5
Santamaria 2018	n	y	n	n	n	y	y	y	y	ca	y	6
Senti 2012	n	y	y	n	n	y	n	n	n	n	y	4
Serrano-Diaz 2018	n	n	y	y	n	y	ca	n	y	y	y	6
Thorne-Lyman 2012	n	n	y	n	n	y	y	y	y	n	n	5
Van der Pligt 2018	n	y	y	n	n	y	y	y	n	n	y	6
Wei 2016	n	y	y	n	n	y	y	ca	y	y	n	6
Yang 2015	n	y	y	n	y	y	y	n	y	y	n	7
Zhang 2018	n	ca	y	ca	n	y	y	y	y	y	y	7
Zhang 2015	n	n	y	n	n	y	y	y	y	y	y	7

* One point was awarded for each item that scored 'yes' (y) and summed for the total score
* 'n' no; 'ca' can't answer

Supplementary Table 5: GRADE tables

Grade Assessments for Preterm Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
11 Bi	RCT	serious	not serious	not serious	not serious	none	moderate
3 De-Regil/Palacios	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
14 Roth	RCT	not serious	not serious	not serious	not serious	none	high
6 Zhou	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low

Grade Assessments for Preeclampsia in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil / Palaciosis	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Khaing	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
1 Newberry	RCT	serious	serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopex	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
3 Roth	RCT	serious	serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Gestational Diabetes in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
5 Roth	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Low Birth Weight in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
4 Bi	RCT	serious	serious	not serious	serious	none	very low
3 De-Regil/ Palaciosis	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
4 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
7 Roth	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low

Grade Assessments for Small for Gestational Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
6 Bi	RCT	serious	not serious	not serious	serious	none	low
2 Harvey	RCT	serious	serious	not serious	serious	none	very low
3 Perez-Lopez	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low
5 Roth	RCT	serious	not serious	not serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Still Birth in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for C-Section Age in RCT's

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
2 De-Regil/ Palaciosis	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
4 Perez-Lopez	RCT	not serious	not serious	not serious	serious	potential for publication / reporting bias	low
16 Roth	RCT	not serious	not serious	not serious	not serious	none	high

Grade Assessments for Preterm Birth in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
2 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
10 Qin	OBS	not serious	not serious	serious	not serious	none	moderate
4 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	serious	potential for publication / reporting bias	very low

4 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	serious	potential for publication / reporting bias	very low
16 Zhou [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
17 Zhou [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	none	low

Grade Assessments for Preeclampsia in Observational Studies

# studies Author	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Overall Certainty**
1 Chung	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
4 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
8 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
8 Tabesh	OBS	serious	serious	serious	not serious	none	very low
6 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	potential for publication / reporting bias	low
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	potential for publication / reporting bias	very low

Grade Assessments for Gestational Diabetes in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
8 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
20 Lu	OBS	not serious	serious	serious	not serious	none	low
10 Wei [blood level 25(OH)D <50nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate
8 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	not serious	serious	not serious	none	moderate

Grade Assessments for Low Birth Weight in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low

Grade Assessments for Small for Gestational Age in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
7 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
1 Newberry	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low
6 Wei	OBS	not serious	not serious	serious	not serious	potential for publication / reporting bias	low

[blood level 25(OH)D <50nmol/L]							
5 Wei [blood level 25(OH)D <75nmol/L]	OBS	not serious	serious	serious	not serious	potential for publication / reporting bias	very low

Grade Assessments for Small for C-Section in Observational Studies

# studies Author	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall Certainty
3 Harvey	OBS	not serious	serious	serious	serious	potential for publication / reporting bias	very low