# BMJ Open

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

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

## 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 (SRs).

**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 predefined inclusion criteria: SRs evaluating vitamin D supplementation in pregnant women and/or examining the association between serum vitamin D levels reporting at least one predefined 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, synthesising evidence from 204 unique primary studies. SRs of randomised controlled trials (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), pre-eclampsia (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), caesarean 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), pre-eclampsia (RR 1.57 (1.21, 2.03) for 25-hydroxy vitamin D (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 caesarean 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 predefined outcome, which had low quality evidence.

## 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.

Credibility of the evidence in this field is compromised by study limitations (in particular, the possibility of confounding among observational studies), inconsistency, imprecision and potential for reporting and publication biases.

## 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 (25 (OH) 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 randomised 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 birth weight. 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 synthesise 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 summarise 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 synthesise 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 prespecified perinatal outcomes and graded the quality of available evidence (ie, 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 2 October 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 and the Health Technology Assessment 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 online 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 an 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: preterm birth, pre-eclampsia, gestational diabetes, small for gestational age, stillbirth, low birth weight and caesarean 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 the 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 an SR. 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 categorised the SRs by quality: low (0–3), medium (4–7) and high (8–11).[12] Given the large number of high-quality SRs, we focused data extraction and analysis on them.

### 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 an 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 (ie, results from RCTs) and (2) the association

between serum vitamin D levels and pregnancy outcomes (ie, results from observational studies). For consistency of rating and based on GRADE recommendations,[9] the 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 predefined outcomes, we reported any subgroup analysis 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 used 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 Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)[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 categorised 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 online supplementary table 2 for the completed PRISMA checklist.

### Description of included SRs

The 13 included reviews were published between 2009 and 2018, with a median AMSTAR score of 8 ranging

from 8 to 11 (online supplementary tables 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] UK,[14] USA[13 32] and Thailand.[33] Four SRs included both RCTs and observational studies,[13 14 26 32] five included only RCTs[23 27 28 31 33] and four 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 12861. For interventional studies there was a wide range of dosing regimens, daily doses ranged from 200 to 5000 international units (IU); weekly doses from 714 to 50000 IU; up to 60000 IU monthly and bolus doses ranging from 35000 to 1200000 (600 000×2) IU. Only two reviews reported subgroup analyses based on dose ranges.[27 28] One review had a subgroup for neonatal mortality and small for gestational age for high (>2000 IU/day) and low (≤2000 IU/day),[27] and the other review presented subgroups 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 with 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
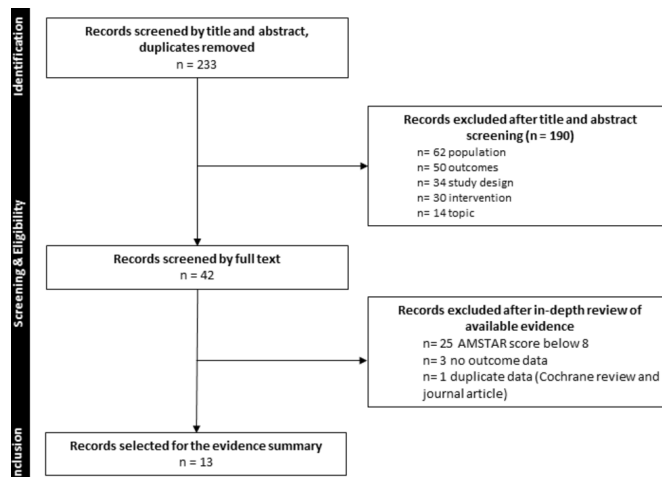


**Figure 1** Flow diagram of screening decisions.

**Table 1** Summary of results from systematic reviews of randomised controlled studies

| Review | Number studies/ individuals | Effect size (CI) risk ratio, random effects | Heterogeneity ($I^2$) | Significance (p value)* | Level of evidence (GRADE) |
|---|---|---|---|---|---|
| **Preterm birth** | | | | | |
| Bi27 | 11/3822 | 0.98 (0.77 to 1.26) | 33% | – (NR) | Moderate |
| De-Regil/Palacios[23 24] | 3/477 | 0.36 (0.14 to 0.93) | 10% | + (0.035) | Very low |
| Perez-Lopez[31] | 3/384 | 1.24 (0.59 to 2.61) | 0% | – (0.56) | Very low |
| Roth [28]† | **14/3757** | **1.00 (0.77 to 1.30)** | **0%** | **– (0.677)** | **High** |
| Zhou26 | 6/1687 | 0.61 (0.34 to 1.07) | 26% | – (0.09) | Low |
| **Pre-eclampsia** | | | | | |
| De-Regil/Palaciosis[23 24] | 2/219 | 0.52 (0.25 to 1.05) | 0% | – (0.069) | Very low |
| Khaing[33] | 3/357 | 0.47 (0.24 to 0.89) | 0% | + (0.02) | Very low |
| Newberry[32] | 1/504 | NR; by group for individual study | NR | + (n=1)‡ | Very low |
| Perez-Lopez[31]† | **3/654** | **0.91 (0.45 to 1.86)** | **24%** | **– (0.80)** | **Low** |
| Roth [28] | 3/706 | 1.09 (0.43 to 2.76) | 66% | – (0.047) | Very low |
| **Gestational diabetes** | | | | | |
| De-Regil/Palacios[23 24] | 2/219 | 0.43 (0.05 to 3.45) | 0% | – (0.43) | Very low |
| Perez-Lopez[31] | 3/384 | 1.05 (0.60 to 1.85) | 0% | – (0.86) | Very low |
| Roth [28] | 5/1030 | 0.65 (0.39 to 1.08) | 45% | – (0.125) | Very low |
| **Small for gestational age** | | | | | |
| Bi27† | **6/1002** | **0.72 (0.52 to 0.99)** | **0%** | **+ (0.04)** | **Low** |
| Harvey[14] | 2/245 | NR; by individual study | NR | – (n=2)‡ | Very low |
| Perez-Lopez[31] | 3/456 | 0.77 (0.46 to 1.30) | 15% | – (0.33) | Very low |
| Roth [28] | 5/741 | 0.60 (0.40 to 0.90) | 0% | + (0.704) | Very low |
| **Low birth weight** | | | | | |
| Bi27 | 4/775 | 0.52 (0.20 to 1.37) | 65% | – (NR) | Very low |
| De-Regil/Palacios[23 24] | 3/493 | 0.4 (0.24 to 0.67) | 4% | + (0.00048) | Very low |
| Perez-Lopez[31] | 4/496 | 0.72 (0.45 to 1.17) | 0% | – (0.19) | Very low |
| Roth [28]† | **7/1156** | **0.74 (0.47 to 1.16)** | **47.3%** | **– (0.077)** | **Low** |
| **Stillbirth** | | | | | |
| De-Regil/Palacios[23 24] | 3/540 | 0.35 (0.06 to 1.99) | 0% | – (0.23) | Low |
| Roth [28]† | **16/4606** | **0.75 (0.50 to 1.12)** | **0%** | **– (0.858)** | **High** |
| **Caesarean section** | | | | | |
| De-Regil/Palaciosis[23 24] | 2/312 | 0.95 (0.69 to 1.31) | 12% | – (0.75) | Low |
| Perez-Lopez[31] | 4/1028 | 0.97 (0.81 to 1.32) | 0% | – (0.75) | Low |
| Roth28† | **16/3240** | **1.02 (0.93 to 1.12)** | **0%** | **– (0.701)** | **High** |

*Significance indicated as positive (+) when p value≤.05 and negative (–) if >0.05.
†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.

online 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; five studies, 1503 participants) and

≥2000 IU/day (RR 1.02, 95% CI 0.76, 1.36; nine studies, 2404 participants).

### Pre-eclampsia

Five SRs of RCTs examined the effectiveness of vitamin D for prevention of pre-eclampsia.[23 28 31–33] The quality of evidence for effectiveness of vitamin D for pre-eclampsia 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 the 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' birth weights 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, the results for a subgroup based on dose were 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 (birth weight <2500 g) (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, 1948 participants) but no difference for ≥2000 IU/day (RR 1.03, 95% CI 0.62, 1.71; 9 studies, 2713 participants).[28]

### Caesarean section

Three SRs examined the effectiveness of vitamin D for caesarean 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, 2303 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 the highest quality of evidence found a significant association with moderate quality of evidence for <50 vs >50 nmol/L 1.13 (95% CI 1.04, 1.23) and non-significant association and low quality of evidence for <75 vs >75 nmol/L 1.03 (95% CI 0.98, 1.08).[26]

### Pre-eclampsia

Five SRs of observational studies examined the association between vitamin D status and pre-eclampsia (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 CIs 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

**Table 2** Summary of results for systematic reviews of observational studies

| Review | Number studies/ individuals | Effect size (CI) risk ratio, random effects | Heterogeneity ($I^2$) | Significance (p value)* | GRADE |
|---|---|---|---|---|---|
| **Preterm birth** | | | | | |
| Harvey[14] | 7/1792 | NR; one individual study showed significance and six others not significant | NR | + (n=1)† − (n=6) | Very low |
| Newberry[32] | 2/371 | NR; by individual study | NR | + (n=1)† − (n=1) | Very low |
| Qin25‡‡ | **10/10 098** | **1.19 (1.08 to 1.31)** | **28%** | **+ (0.004)** | **Moderate** |
| Wei[29] | 4/1111 | 1.27 (1.03 to 1.58) (blood level 25(OH) D<50 nmol/L) | 28% | − (0.03) | Very low |
| | | 1.05 (0.98 to 1.12) (blood level 25(OH) D<75 nmol/L) | 0% | − (0.17) | Very low |
| Zhou26‡ | **16/16 996** | **1.13 (1.04 to 1.23) (<50 vs >50 nmol/L)** | **45%** | **+ (0.003)** | **Moderate** |
| | 15/17 122 | 1.03 (0.98 to 1.08) (<75 vs >75 nmol/L) | 65% | − (0.29) | Low |
| **Pre-eclampsia** | | | | | |
| Chung13 | 1/1189 | 5 (1.7 to 14.1)§ | NR | + (n=1)† | Very low |
| Harvey[14] | 4/642 | 0.75 (0.48 to 1.19)§ | 80.8% | − (0.001) | Very low |
| Newberry[32] | 8/4420 | NR; by individual study | NR | + (n=5) − (n=3) | Very low |
| Tabesh30 | 8/2485 | 2.02 (1.26 to 3.23) | 53% | + (0.04) | Very low |
| Wei[29]‡ | **6/2008** | **1.57 (1.21 to 2.03) (<50 nmol/L)** | **39%** | **+ (0.0006)** | **Low** |
| | 5/1311 | 1.21 (0.99 to 1.46) (<75 nmol/L) | 60% | − (0.06) | Very low |
| **Gestational diabetes** | | | | | |
| Harvey[14] | 8/2668 | NR; by individual study | NR | + (n=3) † − (n=5) | Very low |
| Lu15 | 20/16 515 | 1.45 (1.15 to 1.83)§ | 66.6% | + (0.002) | Low |
| Wei[29]‡ | **10/4126** | **1.12 (1.02 to 1.22) (<50 nmol/L)** | **27%** | **+ (0.02)** | **Moderate** |
| | **8/3840** | **1.09 (1.03 to 1.15) (<75 nmol/L)** | **28%** | **+ (0.002)** | **Moderate** |
| **Small for gestational age** | | | | | |
| Harvey[14] | 7/5660 | NR; by individual study | NR | + (n=2) † − (n=5) | Very low |
| Newberry[32] | 1/412 | NR; by individual study | NR | NR | Very low |
| Wei[29]‡ | **6/6013** | **1.35 (1.18 to 1.54) (<50 nmol/L)** | **15%** | **+ (0.00001)** | **Low** |
| | 5/2283 | 0.99 (0.83 to 1.18) (<75 nmol/L) | 75% | − (0.92) | Very low |
| **Low birth weight** | | | | | |
| Harvey[14] | 3/1676 | NR; by individual study | NR | + (n=1)† − (n=2) | Very low |
| **Caesarean section** | | | | | |
| Harvey[14] | 6/3277 | NR; by individual study | NR | + (n=2) † − (n=4) | Very low |

Continued

| Review | Number studies/ individuals | Effect size (CI) risk ratio, random effects | Heterogeneity (I²) | Significance (p value)* | GRADE |
|--------|------------------------------|----------------------------------------------|---------------------|--------------------------|-------|

**Table 2** Continued

*Significance indicated as positive (+) when p value≤0.05 and negative (−) if >0.05.
†In absence of pooled data, this indicates the number of studies with positive or negative statistical significance.
‡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.
25 (OH)D, 25-hydroxy vitamin D.

significant association for both serum levels examined: <50 nmol/L: 1.12 (95% CI 1.02, 1.22) and <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 birth weights for gestational age (table 2).[14 29 32] The SRs showed mixed findings. One SR included seven studies but did not pool the results as the authors stated that there was substantial variation in methodology and exposure[14]; two studies showed a significant association while five studies showed no significant effect (very low quality of evidence). Another SR included only one 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 the 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.

### Caesarean section

Only one SR of observational studies examined the association between vitamin D status and caesarean section.[14] The SR included six studies but did not pool the 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 predefined outcome—small for gestational age—which had low quality evidence. The differences in findings between these groups of SRs suggest that any apparent association 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 optimise 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 (ie, 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 predefined 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 optimise 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 3 having an AMSTAR

score greater than 7 to be included in the final analysis. Of these three new SRs, only one new primary study was 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 randomised 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] (eg, 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, partly because they can address the problem of confounding as randomisation 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 SR only one-third of the primary studies reported this outcome. Roth *et al* 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 outcome 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 them 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 less number 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 the evidence for one predefined outcome was of low quality. The discrepancy between the observational studies and the RCTs shows that 25(OH)D is lower among women who experience adverse pregnancy outcomes, but supplementation does not appear to alter outcomes. Low 25(OH)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(OH)D may be an acute phase reactant. Future studies need to adequately control for potential confounding (eg, through well-designed randomised 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) synthesising evidence from 204 primary vitamin D studies; further SRs on this topic are wasteful until more well designed and conducted RCTs are completed.

**Author 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, Alberta, Canada
[3]Nutrition Services, Alberta Health Services, Calgary, Alberta, Canada
[4]Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada
[5]Maternal Newborn Child and Youth Strategic Clinical Network, Alberta Health Services, Edmonton, Alberta, Canada
[6]Departments of Pediatrics and Physiology and Pharmacology, Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[7]Faculty of Nursing and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

**ORCID iD**
Liza Bialy http://orcid.org/0000-0002-5861-2769

## 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;5:986–1004.
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;155:245–51.
3. Leffelaar ER, Vrijkotte TGM, 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;104:108–17.
4. Soheilykhah S, Mojibian M, Rashidi M, *et al*. Maternal vitamin D status in gestational diabetes mellitus. *Nutr Clin Pract* 2010;25:524–7.
5. Holmes VA, Barnes MS, Alexander HD, *et al*. Vitamin D deficiency and insufficiency in pregnant women: a longitudinal study. *Br J Nutr* 2009;102:876–81.
6. Bodnar LM, Catov JM, Simhan HN, *et al*. Maternal vitamin D deficiency increases the risk of preeclampsia. *J Clin Endocrinol Metab* 2007;92:3517–22.
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. Chapter V: Overviews of Reviews. In: Higgins JPT, Thomas J, Chandler J, eds. *Cochrane Handbook for systematic reviews of interventions*. London: Cochrane, 2018.
9. Schünemann H, Brozek J, Guyatt G. *GRADE handbook for grading quality of evidence and strength of recommendations*. The GRADE Working Group, 2013. https://gdt.gradepro.org/app/handbook/handbook.html
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;62:1013–20.
12. Canadian Agency for Drugs and Technologies in Health (CADTH). Rapid response summary with critical appraisal: process, 2015. Available: https://www.cadth.ca/about-cadth/what-we-do/products-services/rapid-response-service [Accessed 19 Apr 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;18:1–190.
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;293:959–66.
16. Atkins D, Best D, Briss PA, *et al*. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
17. Guyatt G, Oxman AD, Akl EA, *et al*. Grade guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
18. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. *Grade your evidence and improve your Guideline development in health care*. McMaster University, 2015. https://gradepro.org
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;8:336–41.
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 Gynecol Obstet* 2017;138:242–9.
22. Christensen N, Søndergaard J, Fisker N, *et al*. Infant respiratory tract infections or wheeze and maternal vitamin D in pregnancy. *Pediatr Infect Dis J* 2017;36:384–91.
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;164:148–55.
25. Qin L-L, Lu F-G, Yang S-H, *et al*. Does maternal vitamin D deficiency increase the risk of preterm birth: a meta-analysis of observational studies. *Nutrients* 2016;8:20.
26. Zhou S-S, Tao Y-H, 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;43:247–56.
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;172:635–45.
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 S-Q, Qi H-P, Luo Z-C, *et al*. Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis. *J Matern Fetal Neonatal Med* 2013;26:889–99.
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;98:3165–73.
31. Pérez-López 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;103:1278–88.
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-O, Tantrakul V, *et al*. Calcium and vitamin D supplementation for prevention of preeclampsia: a systematic review and network meta-analysis. *Nutrients* 2017;9:18.
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;12:1654–61.
35. Silva MC, Furlanetto TW. Does serum 25-hydroxyvitamin D decrease during acute-phase response? A systematic review. *Nutr Res* 2015;35:91–6.
36. Patel CJ, Manrai AK. Development of exposome correlation globes to map out environment-wide associations. *Pac Symp Biocomput* 2015:231–42.
37. Guyatt GH, Oxman AD, Montori V, *et al*. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol* 2011;64:1277–82.

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;84:316–7.

39  Devane D, Begley CM, Clarke M, *et al*. Evaluating maternity care: a core set of outcome measures. *Birth* 2007;34:164–72.

40  van 't Hooft J, Duffy JMN, Daly M, *et al*. A core outcome set for evaluation of interventions to prevent preterm birth. *Obstet Gynecol* 2016;127:49–58.

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;12:1495–9.

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;10:28–55.