Vitamin D and Clinical Cancer Outcomes: A Review of Meta-Analyses

John D Sluyter,1 JoAnn E Manson,2,3 and Robert Scragg1

1School of Population Health, University of Auckland, Auckland, New Zealand
2Department of Medicine, Brigham and Women’s Hospital, and Harvard Medical School, Boston, MA, USA
3Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

ABSTRACT

The relationship between vitamin D status or supplementation and cancer outcomes has been examined in several meta-analyses. To address remaining knowledge gaps, we conducted a systematic overview and critical appraisal of pertinent meta-analyses. For meta-analyses of trials, we assessed their quality using AMSTAR-2 (A Measurement Tool to Assess Systematic Reviews), strength of associations using umbrella review methodology and credibility of evidence using GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) criteria. Meta-analyses of observational studies reported inverse associations of 25OHD with risk of cancer incidence and cancer mortality and, particularly for colorectal cancer, fulfilled some of Bradford-Hill’s causation criteria. In meta-analyses of trials, vitamin D supplementation did not affect cancer incidence. However, we found credible evidence that vitamin D supplementation reduced total cancer mortality risk, with five out of six meta-analyses reporting a relative risk (RR) reduction of up to 16%: RR, 0.84 (95% CI, 0.74–0.95). The strength of the association, however, was classified as weak. This was true among meta-analyses of high, moderate, and lower quality (AMSTAR-2–rated). Trials did not include large numbers of vitamin D-deficient participants; many tested relatively low doses and lacked sufficiently powered data on site-specific cancers. In conclusion, meta-analyses show that, although observational evidence indicates that low vitamin D status is associated with a higher risk of cancer outcomes, randomized trials show that vitamin D supplementation reduces total cancer mortality, but not cancer incidence. However, trials with larger proportions of vitamin D-insufficient participants and longer durations of follow-up, plus adequately powered data on site-specific common cancers, would provide further insight into the evidence base. © 2020 The Authors. JBMR Plus published by Wiley Periodicals LLC on behalf of American Society for Bone and Mineral Research.

KEY WORDS: CANCER INCIDENCE; CANCER MORTALITY; CIRCULATING 25-HYDROXYVITAMIN D; META-ANALYSIS; RANDOMIZED CONTROLLED TRIAL; VITAMIN D

Introduction

Epidemiological studies of cancer incidence and mortality have shown that sun exposure, especially solar UV B (and hence vitamin D production), is associated with reduced risk of many cancer types.1–3 This supports the hypothesis that vitamin D may have a beneficial impact on cancer outcomes, which is further strengthened by the identification of biological mechanisms that may explain these associations4 and prospective observational studies, which show that dietary intake of vitamin D (including vitamin D supplementation) is associated with a reduced risk of cancers, although confounding is important to consider.5–8

Meta-analyses provide accurate, succinct, credible, and comprehensive summaries of the best available evidence, and act as a key tool for health care professionals to achieve evidence-based decisions.9 Several meta-analyses have synthesized data on the association of vitamin D status with cancer outcomes from observational studies10–14 and the effect of vitamin D supplementation on cancer outcomes in randomized clinical trials.15–22 However, the evidence base is inconsistent and fragmented into various meta-analyses that combine different study populations and outcomes, making assessment of the evidence using a similar methodological framework difficult. Further, as the methodologic quality of vitamin D–cancer meta-analyses may vary, uncritically accepting their results carries risk.9

Thus, we decided to conduct a systematic overview and critical appraisal of pertinent meta-analyses to better characterize the evidence on vitamin D status or supplementation in relation to cancer outcomes. In our critical appraisal, particular attention was given to intervention studies as this study design, as the gold standard for effectiveness research, provides the highest...
relevance for evidence-based decision-making. For this, we systematically assessed the quality of meta-analyses, plus strength and credibility of the evidence from these studies across multiple cancer outcomes, and discussed differences between meta-analyses. Finally, we discussed limitations of the evidence presented and provided some future research directions.

Materials and Methods

We searched Medline and PubMed for articles published up until 12 May 2020, using the following search terms: vitamin D, cancer, and meta-analysis. No language restrictions were applied. This was supplemented by a manual search of reference lists from identified articles.

The quality of meta-analyses was assessed using AMSTAR-2 (A Measurement Tool to Assess Systematic Reviews), a 16-point assessment tool of the methodological quality of systematic reviews. Of the 16 domains, items 2, 4, 7, 9, 11, 13, and 15 are considered “critical” domains (can critically affect the validity of a review and its conclusion). Based on weaknesses in the critical and noncritical domains, the overall confidence in the results of the meta-analysis was classified as “High,” “Moderate,” “Low,” or “Critically low.” AMSTAR-2 has good interrater agreement, test–retest reliability, and content validity.

The strength of associations was evaluated based on umbrella review criteria. For this, small-study effects were evaluated using Egger’s test, and the excess significance test was applied (excess significance was claimed at $p \leq 0.10$). Based on these criteria findings, we classified the strength of the association (effect) as “Convincing,” “Highly suggestive,” “Suggestive,” or “Weak” (Appendix Table AA1). Associations were considered nonsignificant if the $P$ value was >0.05. To provide a practical metric of the efficacy of vitamin D supplementation, the number needed to treat (NNT) was calculated.

The credibility of pooled estimates of meta-analyses was qualitatively assessed using the GRADE (the Grading of Recommendations, Assessment, Development, and Evaluation) method. For each outcome, GRADE produces a credibility of estimate and summary of findings in a table that is easily understandable for study participants, policy makers, researchers, guideline developers, and other interested stakeholders (scoring detailed in Appendix Table AA2). AMSTAR-2 and GRADE assessments were performed by one researcher (JDS) and verified by another (RS), and discrepancies were discussed and resolved by consensus.

Results

Meta-analyses of observational studies

We found 35 meta-analyses that investigated relationships between vitamin D status, as measured by circulating 25OHD and cancer outcomes: 29 on cancer incidence, 3 on cancer mortality, and 3 on both (Table 1).

Cancer incidence

With total cancer incidence as the outcome, one of these meta-analyses combined data from eight prospective cohort studies (70,018 participants and 7511 events). The summary relative risk (RR) estimate of the highest 25OHD category compared with the lowest was 0.86 (95% CI, 0.73–1.02), indicating a marginal relationship in the inverse direction, with a significant between-study heterogeneity ($I^2 = 70.8\%$). Dose–response analysis indicated a 7% reduction in risk (RR, 0.93; 95% CI, 0.91–0.96) per 20 nmol/L increment in 25OHD.

There were numerous meta-analyses (43 from 32 articles) on site-specific cancer outcomes: breast (9 out of 12), colon (n = 3), rectum (n = 3), colorectal (n = 9), colorectal adenomas (n = 4), prostate (n = 4), kidney (n = 1), liver (n = 1), lung (n = 1), non-Hodgkin’s lymphoma (n = 1), ovarian (n = 1), pancreatic (n = 1), and thyroid (n = 2). Most of these were prospective: cohort studies or case–control studies nested within these. The vast majority of associations were in the inverse direction. Of these, significant relationships were observed in most meta-analyses of breast cancer (9 out of 12) and all meta-analyses of colorectal cancer. Generally, heterogeneity was high for breast cancer ($I^2$ up to 91%), but low to moderate for colorectal cancer (eg, $I^2$ was 0% to 9% in three meta-analyses). There were strong associations for some cancers: breast (lowest RR, 0.41), colorectal (lowest RR, 0.49), and rectum (lowest significant RR, 0.50). Sizeable relationships were observed too for colorectal adenoma (lowest odds ratio [OR], 0.59). Dose–response meta-analyses reported inverse trends for breast, colon, colorectal, liver, and lung cancers.

Cancer mortality

A recent meta-analysis had total cancer mortality as the outcome. This report combined data from 16 prospective cohort studies on 101,794 participants without cancer at baseline, 8729 of whom had a cancer-related death. The pooled RR—of a highest-versus-lowest 25OHD group comparison—was 0.81 (95% CI, 0.71–0.93), indicating an inverse association, with moderate heterogeneity ($I^2 = 48.8\%$). Dose–response analysis of studies with three categories of 25OHD revealed that the risk of cancer mortality was reduced by 2% (RR, 0.98; 95% CI, 0.97–0.99) with each 20 nmol/L increment of 25OHD.

Four meta-analysis articles of prospective studies focused on patients with cancer. They reported sizeable inverse associations between 25OHD quantile (highest versus lowest) and cancer-specific mortality: pooled hazard ratios (HRs) were 0.50 to 0.65 (in patients with breast cancer, colorectal cancer, and lymphoma) in one study, summary RR was 0.58 for breast cancer mortality in another study; the pooled RR ranged from 0.57 (for breast cancer mortality) to 0.65 (for colorectal cancer mortality) in another; and the summary HR was 0.73 (for colorectal cancer mortality) in the fourth study. One of these studies reported an inverse dose–response association.

Meta-analyses of randomized clinical trials

We identified eight meta-analyses of clinical trials that evaluated the impact of vitamin D supplementation on cancer outcomes (incidence and mortality; Table 2). Appendix Table AA3 summarizes their study selection criteria, which excluded studies based on several factors such as coadministration of calcium that differed across the treatment groups, treatment with hydroxylated vitamin D or vitamin D analogs, duration of follow-up (<1 year or intervention (<3 years), age (<18 or <60 years), number of outcomes (<10), and pregnancy.

Quality of meta-analyses

The AMSTAR-2 ratings of the eight meta-analyses are summarized in Table 2 and detailed in Appendix Table AA4. Five meta-
analyses had shortcomings in critical domains\(^{(17–20,22)}\); All five did not cite a predefined, registered protocol, four did not cite excluded primary studies,\(^{(18–20,22)}\) three did not use a satisfactory technique to assess risk of bias\(^{(17,18,20)}\), and two used a literature search that was not fully comprehensive.\(^{(17,18)}\) All studies had deficiencies in noncritical domains, with the most common being reporting funding sources of primary studies (absent in seven meta-analyses\(^{(15,17–22)}\)) and explaining selection of study designs for inclusion in the review (absent in six meta-analyses\(^{(15–17,19,21,22)}\)). Based on weaknesses in the critical and noncritical domains, the overall confidence in the eight meta-analyses was deemed critically low or moderate for seven of these.\(^{(15,17–22)}\)

**Cancer incidence**

Six of these meta-analyses had total cancer incidence as an outcome.\(^{(16–20,22)}\) These reports were published between 2014 and 2019, with each comprising 4 to 18 studies, 18,440 to 83,353 participants, and 1061 to 6537 cancer incident events. The pooled RR effects in these reports were similar, ranging from 0.98 (10 studies\(^{(20)}\)) to 1.03 (24 studies\(^{(19)}\)) with all 95% CIs encompassing 1. The \(I^2\) for these effects was 0%\(^{(16–20)}\) or 31%\(^{(22)}\) indicating no significant heterogeneity.

Two of these meta-analyses had specific cancer incidence as an outcome.\(^{(15,16)}\) In these, the pooled RR effects were 0.97 (seven studies) to 1.11 (two studies) for breast cancer, 0.86 (95% CI, 0.69–1.07) for lung cancer (five studies), 1.11 (95% CI, 0.92–1.34) for colorectal cancer, and 0.91 (95% CI, 0.57–1.46) for pancreatic cancer (two studies).

**Cancer mortality**

Six of these meta-analyses had cancer mortality as an outcome.\(^{(16,18–22)}\) The pooled RR effects in these reports were similar and all in the inverse direction, ranging from 0.84 (95% CI, 0.74–0.95; 12 studies) to 0.88 (95% CI, 0.70–1.09; 24 studies). The 95% CIs did not encompass one in five out of six of these meta-analyses\(^{(16,18,20–22)}\) indicating consistently beneficial impacts on cancer mortality.

**Strength and credibility of meta-analysis effects**

Five intervention meta-analyses of total cancer mortality had weak strength of association according to umbrella review criteria\(^{(16,18–20,22)}\) with all scoring high with GRADE and with NNT values ranging from 86 to 381. This was consistent among meta-analyses of high, moderate, and lower quality (AMSTAR-2-rated). The remaining meta-analyses reported nonsignificant effects, with GRADE credibility that was moderate\(^{(16,17)}\) to high\(^{(18–20,22)}\) for total cancer incidence and low\(^{(15,16)}\) to moderate\(^{(16)}\) for site-specific cancer incidence. A breakdown of these scores is provided in Appendix Tables AAS and AA6.

**Differences between trial meta-analyses**

As pooled effects across meta-analyses on the same outcome were mostly similar (with overlapping 95% CIs; Table 2) despite variable study selection criteria (Appendix Table AA3), this enhances the external validity of our findings that vitamin D reduces total cancer mortality (in nearly all meta-analyses), but not cancer incidence. Where pooled effects for the same outcome variations did vary, this can be explained by differences in the number of primary studies in meta-analyses, which is influenced by publication year (Table 2) and study selection criteria. The most important variation occurred for cancer mortality, in which the meta-analysis by Goulão and colleagues\(^{(19)}\) did not find a significant overall effect, but the remaining five meta-analyses did (Table 2). Goulão and colleagues\(^{(19)}\) excluded primary studies in which calcium was coadministered with vitamin D, but not with placebo, which allowed it to disentangle calcium and vitamin D effects. This restriction meant that the Women’s Health Initiative\(^{(59)}\) in which calcium was given with vitamin D in the intervention group only, was not included in this meta-analysis (unlike nearly all other ones\(^{(16,18,20,22)}\)). However, this exclusion may not have been important for validity if calcium did not influence cancer mortality, as suggested by prior studies.\(^{(60)}\) Also, the number of participants \((n = 11202)\) and events \((n = 320)\) in the Goulão and colleagues\(^{(19)}\) meta-analysis was substantially lower than those in the other meta-analyses \((n = 44,290\) to 75,239 and \(n = 939\) to 1192 for number of participants and events, respectively; Table 2). This is because two large \((n > 5000)\) trials in addition to the Women’s Health Initiative\(^{(59)}–\)the VIDA (Vitamin D Assessment)\(^{(61)}\) and VITAL (Vitamin D and Omega-3)\(^{(62)}\) studies—were included in other meta-analyses (showing vitamin D benefits\(^{(20–22)}\)) but not in the meta-analysis by Goulão and colleagues\(^{(19)}\) as they were published after it.\(^{(19)}\) Thus, the Goulão and colleagues\(^{(19)}\) meta-analysis did not capture effects from some large trials and had reduced statistical power.

**Limitations of meta-analyses and their primary studies**

Owing to several limitations, the meta-analyses and their primary studies have boundaries of applicability. These limitations influence the strength of associations and are key targets for future research. We discuss these for observational and intervention studies separately.

**Observational studies**

A limitation of the observational studies we reviewed is that, although they adjusted for multiple confounders, they are vulnerable to residual confounding; more so when the confounders (eg, smoking behavior, BMI, physical activity, diet) are measured less well. Adding to this issue is that vitamin D status is related to multiple diseases, besides cancer.\(^{(63)}\) Thus, 25OHD–cancer associations do not fulfill one of Bradford-Hill’s causation criteria, specificity,\(^{(64)}\) raising the possibility that confounding may be important. On the other hand, confounding may not entirely explain these associations as risk factors can cause more than one disease,\(^{(64)}\) and vitamin D has pleiotropic cellular effects.\(^{(63)}\)

A second limitation is that several observational studies, especially those on breast cancer incidence,\(^{(14,29,30,32,33,35)}\) were case–control investigations—and are thus susceptible to reverse causation because 25OHD measurement was performed in those already diagnosed with cancer, and low vitamin D status may be a consequence of the disease rather than a cause. For instance, during cancer therapy or when symptoms are severe, sunlight exposure, physical activity, and dietary habits are likely to change (because of hospitalizations, disability, or lifestyle changes), and cancer-associated inflammation may depress 25OHD.\(^{(65)}\) However, this problem was avoided in prospective studies where blood samples were collected well before cancer diagnosis. The studies in the meta-analyses we reviewed were mainly prospective (nested case–control or cohort) and we observed inverse associations in them, suggesting that 25OHD
| First author, publication year | Studies & design (n) | Participants (n) | Events (n) | Unit of 25OHD comparison | Pooled association (95% CI) | $I^2$ (%) or $P$ value for heterogeneity$^a$ |
|-------------------------------|---------------------|-----------------|------------|-------------------------|---------------------------|----------------------------------|
| **Cancer incidence**          |                     |                 |            |                         |                           |                                  |
| **All**                       |                     |                 |            |                         |                           |                                  |
| Han, 2019$^{(13)}$            | 8 PC                | 70,018          | 7511       | Highest vs. lowest group | RR, 0.86 (0.73–1.02)      | 71                               |
|                              |                     |                 |            |                         | RR, 0.93 (0.91–0.96)      | NR                               |
| **Breast**                    |                     |                 |            |                         |                           |                                  |
| Chen, 2010$^{(29)}$           | 7 (4 CC, 3 NCC)     | 11,330          | 5489       | Highest vs. lowest quartile | OR, 0.55 (0.38–0.80)      | 86                               |
| Yin, 2010$^{(30)}$            | 9 (5 CC, 4 NCC)     | 12,901          | 6147       | Per 20 ng/mL             | OR, 0.73 (0.60–0.88)      | 84                               |
| Chung, 2011$^{(31)}$          | 4 NCC               | 4726            | 2363       | Per 10 nmol/L (dose–response) | OR, 0.99 (0.97–1.01)      | NR                               |
| Gandini, 2011$^{(32)}$        | 10 (1 PC, 4 NCC, 5 CC) | 29,742          | 6324       | Highest vs. lowest quartile | OR, 0.61 (0.47–0.80)      | 41                               |
| Mohr, 2011$^{(33)}$           | 2 (6 NCC, 5 CC)     | 29,742          | 6324       | Highest vs. lowest quartile | OR, 0.87 (0.77–0.99)      | 54                               |
| Bauer, 2013$^{(34)}$          | Premenopause: 6 PC  | 1613            | 8766       | Per 5 ng/mL (dose–response) | OR, 0.61 (0.47–0.80)      | 41                               |
|                              | Postmenopause: 9 PC | 3929            | 8766       | RR, 0.99 (0.79–1.04)      | RR, 0.97 (0.93–1.00)      | NR                               |
| Chen, 2013$^{(35)}$           | 21 (10 NCC, 1 RSP, 10 CC) | 26,317          | 11771      | Highest vs. lowest quartile | OR, 0.52 (0.40–0.68)      | 89                               |
| Wang, 2013$^{(49)}$           | 14 (1 PC, 13 NCC)   | 25,354          | 9110       | Highest vs. lowest group | RR, 0.84 (0.75–0.95)      | 38                               |
| Kim, 2014$^{(56)}$            | 14 (1 PC, 13 NCC)   | 27,534          | 9526       | Highest vs. lowest group | RR, 0.92 (0.83–1.02)      | 27                               |
| Estébanez, 2018$^{(50)}$      | 29 (14 NCC, 15 CC)  | 58,855          | 18358      | High vs. low group       | OR, 0.66 (0.57–0.76)      | 41                               |
|                              | 14 NCC              | 24,271          | 10266      | High vs. low group       | OR, 0.92 (0.83–1.01)      | 16                               |
|                              | 4 PC                | 16,875          | 3350       | Variable group comparisons | OR, 0.85 (0.74–0.98)      | 4                                |
| Hossain, 2019$^{(55)}$        | 14 (12 NCC, 1 CC, 1 MR) | 123,044         | 25515      | Per 10 ng/mL             | OR, 0.99 (0.98–1.00)      | 79                               |
| Song, 2019$^{(54)}$           | 40 (4 PC, 36 CC)    | 162,322         | 31157      | Per 5 nmol/L             | OR, 0.94 (0.93–0.96)      | 91                               |
| Yin, 2009$^{(36)}$            | 6 NCC               | 2081            | 759        | Per 20 ng/mL             | OR, 0.78 (0.54–1.13)      | 45                               |
| Lee, 2011$^{(37)}$            | 8 PC                | 4578            | 1822       | Highest vs. lowest group | OR, 0.77 (0.56–1.07)      | 48                               |
| Touvier, 2011$^{(6)}$         | 6 NCC               | 3550            | 1477       | Per 100 IU/L (dose–response) | OR, 0.95 (0.92–0.995)    |                                  |
| **Colon**                     |                     |                 |            |                         |                           |                                  |
| Yin, 2009$^{(36)}$            | 4 NCC               | 719             | 258        | Per 20 ng/mL             | OR, 0.41 (0.11–1.49)      | 63                               |
| Lee, 2011$^{(37)}$            | 9 NCC               | 4578            | 868        | Highest vs. lowest group | OR, 0.50 (0.28–0.88)      | 67                               |
| Touvier, 2011$^{(6)}$         | 5 NCC               | 1645            | 721        | Per 100 IU/L             | RR, 0.95 (0.89–1.05)      |                                  |
| Colorectal                     |                     |                 |            |                         |                           |                                  |
| Gorham, 2007$^{(38)}$         | 5 NCC               | 1448            | 535        | Highest vs. lowest group | OR, 0.49 (0.35–0.68)      |                                  |

(Continues)
| First author, publication year | Studies & design (n) | Participants (n) | Events (n) | Unit of 25OHD comparison | Pooled association (95% CI) | I² (%) or P value for heterogeneitya |
|-------------------------------|---------------------|-----------------|------------|--------------------------|-----------------------------|----------------------------------|
| Yin, 2009(36)                | 5 NCC               | 3286            | 1199       | Per 20 ng/mL             | OR, 0.57 (0.43–0.76)        | 9                                |
| Chung, 2011(31)              | 9 NCC               | 2249            | 1127       | Per 10 nmol/L (dose–response) | OR, 0.94 (0.91–0.97)        | NR                               |
| Gandini, 2011(32)            | 9 (1 PC, 7 NCC, 1CC) | 22,948         | 2630       | Per 10 ng/mL (dose–response) | RR, 0.85 (0.79–0.91)        | 55                               |
|                             | 8 (1 PC, 7 NCC)     | 22,870          | 2604       | Per 10 ng/mL (dose–response) | RR, 0.85 (0.79–0.92)        | 59                               |
| Lee, 2011(37)                | 8 NCC               | 4578            | 2690       | Highest vs. lowest group  | OR, 0.66 (0.54–0.81)        | p = .04                          |
| Ma, 2011(39)                 | 9 (7 PC, 2 NCC)     | 6715            | 2767       | Highest vs. lowest group  | RR, 0.67 (0.54–0.80)        | 0                                |
|                             |                     |                 |            | Per 10 ng/mL (dose–response) | RR, 0.74 (0.63–0.89)        | NR                               |
| Touvier, 2011(6)             | 6 NCC               | 5833            | 2370       | Per 100 IU/L              | RR, 0.96 (0.94–0.97)        | 0                                |
| Huang, 2019(57)              | 30 (6 PC, 23 NCC, 1CC) | 204,544      | 13051      | Highest vs. lowest group  | RR, 0.68 (0.60–0.78)        | 56                               |
| Zhang, 2019(48)              | 8 (1 NCC, 7 CC)     | 9594            | 2916       | Highest vs. lowest group  | OR, 0.75 (0.58–0.97)        | 54                               |
|                             |                     |                 |            | Per 16 ng/mL (dose–response) | OR, 0.79 (0.64–0.97)        | 54                               |
| Colorectal adenoma           |                      |                 |            |                          |                             |                                  |
| Wei, 2008(58)                | All adenomas: 7 (1 CS, 3 CC, 3 NCC/PC) | 3787          | 2628       | Highest vs. lowest quintile | OR, 0.70 (0.56–0.87)        | 54                               |
|                             | Advanced adenomas: 2 NCC | 1023          | 2347       | High vs. low groups       | OR, 0.64 (0.45–0.90)        | NR                               |
| Fedirko, 2010(40)            | 3 CC                | 1386            | 616        | Highest vs. lowest quartile | OR, 0.59 (0.41–0.84)        | NR                               |
| Yin, 2011(41)                | Incident events: 9 (5 CC, 1 CS, 3 NCC) | 7654          | 3539       | Per 20 ng/mL              | OR, 0.82 (0.69–0.97)        | 66                               |
|                             | Recurrent events: 3 PC | 2169           | 984        | Per 20 ng/mL              | OR, 0.87 (0.56–1.35)        | 57                               |
| Huang, 2019(57)              | 22 (5 PC, 2 NCC, 14 CC, 1CS) | 13652         | 6445       | Highest vs. lowest group  | RR, 0.80 (0.71–0.89)        | 34                               |
| Prostate                     |                      |                 |            |                          |                             |                                  |
| Yin, 2009(46)                | 10 (1 PC, 9 NCC)    | 7806            | 3124       | Per 10 ng/mL              | OR, 1.03 (0.96–1.11)        | 23                               |
| Chung, 2011(31)              | 8 NCC               | 5609            | 2399       | Per 10 nmol/L (dose–response) | OR, 1.01 (0.99–1.04)        | NR                               |
| Gandini, 2011(32)            | 11 PC               | 26,575          | 3956       | Per 10 ng/mL(dose–response) | RR, 0.99 (0.95–1.03)        | 37                               |
| Gilbert, 2011(47)            | 14 (5 PC, 9 NCC)    | 12,051          | 4353       | Per 10 ng/mL              | OR, 1.04 (0.99–1.10)        | 0                                |
| Kidney                       |                      |                 |            |                          |                             |                                  |
| Gallicchio, 2010(42)         | 8 PC                | 1550            | 775        | 50- < 75 vs. ≥100 nmol/L  | OR, 0.92 (0.44–1.92)        | P-value: NS                      |
| Liver                        |                      |                 |            |                          |                             |                                  |
| Guo, 2020(51)                | 6 (1 PC, 5 NCC)     | 60,811          | 992        | High vs. low group        | RR, 0.78 (0.63–0.95)        | 54                               |
|                             |                     |                 |            | Per 10 nmol/L (dose–response) | RR, 0.92 (0.89–0.95)        | NR                               |
| Lung                         |                      |                 |            |                          |                             |                                  |
| Feng, 2017(11)               | 9 (6 PC, 3 CC)      | 111,148         | 1511       | Variable group comparisons | RR, 0.84 (0.74–0.95)        | 50                               |
|                             |                     |                 |            | Per 10 nmol/L (dose–response) | RR, 0.92 (0.87–0.96)        | NR                               |
| Non-Hodgkin lymphoma         |                      |                 |            |                          |                             |                                  |
| Purdue, 2010(43)             | Males: 6 PC Females: 4 PC | 923           | 733        | >100 vs. 50–75 nmol/L     | OR, 0.67 (0.37–1.20)        | NR                               |
|                             |                     |                 |            | >100 vs. 50–75 nmol/L     | OR, 0.81 (0.39–1.69)        | NR                               |

(Continues)
may affect cancer (Table 1) rather than vice versa. Third, several observational studies are prone to 25OHD measurement error because they did not measure vitamin D status with the gold standard laboratory method for 25OHD measurement: liquid chromatography–tandem mass spectrometry. Fourth, whereas many observational studies were of high quality, as assessed by the Newcastle-Ottawa Scale, several were of suboptimal (medium) quality, which may have attenuated 25OHD–cancer associations. In support of this, a recent meta-analysis reported that 25OHD was inversely associated with total cancer incidence and mortality in studies of high quality, but not in those of medium quality. Fifth, though most meta-analyses had low or moderate between-study heterogeneity (I² < 50% or P values >0.05), some did not; thus, their results should be interpreted with caution. These include meta-analyses on colorectal adenoma and cancer occurrences of the breast, rectum, and all types combined, which reported high between-study heterogeneity (I² ≥ 50%). Finally, most participants in the observational studies were of White ethnicity, which limits generalizability of findings to non-White ethnic groups.

### Randomized clinical trials

An important issue for intervention studies is the growing randomized controlled trial (RCT) evidence that health benefits of vitamin D supplementation are greatest in vitamin D-deficient people. This may also apply to cancer as a recent meta-analysis of cohort studies revealed that 25OHD was inversely associated with cancer incidence at approximately <30 nmol/L only, and had an inverse relationship with cancer mortality that was strongest at low 25OHD levels (especially at approximately <50 nmol/L). Thus, any cancer-related benefits of vitamin D supplementation may be greatest in vitamin D-deficient individuals. However, trials have not contained large numbers of such people. For example, the average baseline 25OHD of 10 trials

---

**Table 1. Continued**

| First author, publication year | Studies & design (n) | Participants (n) | Events (n) | Unit of 25OHD comparison | Pooled association (95% CI) | I² (%) or P value for heterogeneitya |
|-------------------------------|---------------------|------------------|------------|-------------------------|-----------------------------|-------------------------------------|
| Ovarian Yin, 2011(44)         | 10 NCC              | 3373             | 884        | Per 20 ng/mL            | 0.83 (0.63–1.08)            | 0                                   |
| Pancreatic Stolzenberg-Solomon, 2010(45) | 6 NCC              | 833              | 345        | ≥100 vs. 50–75 nmol/L   | OR, 2.14 (0.93–4.92)        | p > .30                             |
| Thyroid Hu, 2018[52]          | (7 CC, 1 CS, 1 RSP) | 7099             | 1172       | <20 vs. ≥20 ng/mL       | OR, 1.42 (1.17–1.73)        | 27                                  |
|                              | 7 (5 CC, 2 CS)      | 6498             | 775        | Cases vs. controls      | SMD, −0.20 (−0.36 to −0.03) | 55                                  |
| Zhao, 2019[53]                | 6 CC                | 6241             | 711        | Deficient vs. non-deficient | OR, 1.30 (1.00–1.69)        | 38                                  |
|                              | 12 CC               | 7278             | 1239       | Cases vs. controls      | SMD, −0.37 (−0.45 to −0.28) | 93                                  |
| Cancer mortality              |                     |                  |            |                         |                             |                                     |
| All Li, 2014[54]              | 4 PC                | 4813             | 661        | Highest vs. lowest quartile | HR, 0.65 (0.44–0.98)      | 45                                  |
|                              | 5 CC, 1 NC (RSP)    | 1234             | 511        | Highest vs. lowest quartile | HR, 0.65 (0.47–0.88)      | 6                                   |
| Han, 2019[13]                 | 16 PC               | 101,794          | 8729       | Highest vs. lowest group | Per 20 nmol/L (dose–response) | RR, 0.81 (0.71–0.93) | 49                                  |
|                              |                     |                  |            |                         |                             | RR, 0.98 (0.97–0.99) | NR                                  |
| Breast                        |                     |                  |            |                         |                             |                                     |
| Kim, 2014[56]                 | 4 PC                | 400              | 4556       | Highest vs. lowest group | Per 10 ng/mL (dose–response) | RR, 0.58 (0.40–0.85) | 27                                  |
|                              | 3 PC                | NR               | NR         |                         |                             | RR, 0.88 (0.79–0.98) | 23                                  |
| Maalmi, 2014[10]              | 3 PC                | 2636             | 194        | High vs. low group      | HR, 0.57 (0.38–0.84)       | 17                                  |
| Colorectal                     |                     |                  |            |                         |                             |                                     |
| Maalmi, 2014[10]              | 3 PC                | 1558             | 566        | High vs. low group      | HR, 0.65 (0.49–0.86)       | 0                                   |
| Xu, 2018[55]                  | 5 (3 PC, 2 NCC)     | 4126             | 982        | High vs. low group      | HR, 0.73 (0.55–0.97)       | 69                                  |
| Huang, 2019[57]               | 12 PC               | 53,910           | 2021       | High vs. low group      | HR, 0.64 (0.56–0.73)       | 3                                   |

aDerSimonian-Laird Q statistic.

CC = Case control; CRC = colorectal cancer; CS = cross sectional; HR = hazard ratio; MR = Mendelian randomization; NCC = nested case–control; NR = not reported; NS = not significant; OR = odds ratio; PC = prospective cohort; POR = Peto odds ratio; RR = relative risk; RSP = retrospective; SMD = standardized mean difference.
examined in the recent meta-analysis by Keum and colleagues, we reviewed was approximately 60 nmol/L.

Insufficient vitamin D dosage may be another limitation. For example, in the Goulão and colleagues’ meta-analysis of cancer incidence, the daily dose equivalent (dose divided by days between each dose) in multiple primary studies (n = 7) was relatively low (<1000 IU/day), raising a question of whether a higher dose may have produced a different (stronger) effect. A related issue is the frequency of the dosing regimens. Most trials have utilized daily dosing; evidence is limited for supplementation administered monthly or weekly.

Third, as carcinogenesis is a long-term and gradual process (often spanning decades), the need for a long follow-up period is particularly great. The importance of this is reflected in the long follow-up periods of 25OHD–cancer cohort studies, such as those in the Han and colleagues’ meta-analysis (of total cancer incidence and mortality), which were 12 to 13 years on average and up to 28 years. However, most of the primary studies included in our meta-analyses of trials had follow-up periods of no more than 5 years; this may have been insufficient to detect effects on cancer. In support of this, the meta-analysis by Zhang and colleagues reviewed found that benefit of supplementation on reduced cancer mortality was observed in trials with longer follow-up (>3 years), but not in those with a shorter follow-up.

A fourth limitation is that, as shown in the meta-analyses for observational studies, 25OHD was more consistently associated with colorectal and breast cancers than other cancers (Table 1). Thus, there may be stronger effects of supplementation against certain cancer types or, possibly, residual confounding may be greater for some cancers than others (eg, confounding by BMI, physical activity, and diet may be of particular relevance to studies of vitamin D and colorectal cancer). However, whereas the cancer outcomes that have dominated our RCT findings are those based on all cancers combined (Table 2), results for site-specific cancers are lacking. Almost all trials that evaluated impacts on site-specific cancers did not include these as primary endpoints, and data on rarer cancers (eg, kidney) are missing.

Fifth, the vast majority of participants in the intervention studies were White, which restricts applicability of findings to non-White populations.

Finally, there were quality-related shortcomings of the meta-analyses of intervention studies we reviewed. The most common ones, detected by AMSTAR-2, were the lack of (i) information on funding sources of primary studies (note-worthy as many vitamin D trials are industry-funded and have a high risk of “for-profit” bias), (ii) an explanation of study design selection, and (iii) a review protocol describing prespecified methodology. Most of these meta-analyses were published prior to the availability of AMSTAR-2 reporting standards (in 2017), which may have contributed to their lower AMSTAR-2 ratings. However, as mentioned above, our vitamin D–cancer findings for meta-analyses of high and moderate quality (AMSTAR-2–rated) were similar to those of lower quality.

Limitations of current review

A potential limitation of this review is that, although a comprehensive and systematic literature search was performed, we may have missed some meta-analyses. Second, our study was a meta-review, and although this provides an overarching perspective on a research topic, we did not provide granular analyses at the primary study level. Third, our review focused on meta-analyses; thus, some primary studies may not have been included either because the meta-analysis did not identify them or they were too recent to be included. Finally, we did not critically appraise the quality of all primary studies individually. This should have been done in each meta-analysis; doing this here was beyond the scope of our review.

Future research

Several areas of future research would strengthen our understanding of vitamin D effects on cancer. First, given the emerging evidence for threshold effects related to vitamin D status, future trials should aim to recruit participants with vitamin D insufficiency (25OHD <50 nmol/L). There are major logistical and practical barriers to doing this in populations that are vitamin D replete, and trials could be undertaken more easily and cheaply in populations with a high prevalence of vitamin D insufficiency. These trials should have longer follow-up periods, include more adequately powered data on site-specific cancers (feasible for common cancers), and study more non-White populations. However, ethical issues can arise with the conduct of long-term trials in vitamin D-deficient participants, as 50% will be randomly assigned to placebo and remain deficient for a prolonged period.

Second, cells that express the cell-surface receptor proteins megalin and cubulin (eg, those in the kidney, lung, thyroid, mammary gland, gall bladder, and thyroid) can internalize 25OHD-bound to vitamin D-binding protein, with subsequent unbinding of 25OHD intracellularly and conversion to 1,25-dihydroxyvitamin D, which can exert anticancer effects by activating the vitamin D receptor. In contrast, 25OHD entry in cells not expressing the megalin–cubulin receptor is proposed to occur via diffusion of unbound, free 25OHD across the cell membrane. However, the studies we reviewed measured total 25OHD, comprising not only free (and bioavailable) 25OHD, but mostly (approximately 90%) 25OHD, which is bound tightly to vitamin D-binding protein, and thus may not get into some cancer cells easily. As evidence of importance, a large, prospective cohort study reported that higher bioavailable, rather than total, 25OHD levels were independently associated with improved survival in patients with hepatocellular carcinoma. Further, in a recent RCT of patients with digestive tract cancer, vitamin D supplementation improved 5-year relapse-free survival in those with low bioavailable 25OHD, but not in those with high bioavailable 25OHD or in those with low total 25OHD (<20 ng/mL). These studies suggest that for some cancer types, free and bioavailable 25OHD may better assess true vitamin D status (and deficiency) than total 25OHD in future trials.

Third, articles that reported inverse, longitudinal associations between 25OHD and cancer mortality (110,54–56) (Table 1) were studies of patients with cancer, and thus suggested a potential role of vitamin D in cancer therapy as opposed to cancer prevention, the focus of the other articles we reviewed. RCT data investigating vitamin D as a cancer treatment are scarce, with one relatively small RCT (n = 139) reporting an improvement in median progression-free survival of death (HR, 0.64) over 22.9 months (median) in patients with colorectal cancer and another (n = 417) reporting an age-adjusted benefit on relapse-free survival (HR, 0.66) over 3.5 years (median) in patients with digestive tract cancers. However, further trials are required in this research area, ideally with longer follow-up including biological measurements would promote our understanding of underlying mechanisms: This requires considering not only anti-neoplastic influences, but broad biological effects too, as 25OHD
| Study          | Studies (n) | Sample size (n) | Events (n) | Pooled RR effect (95% CI) | NNT (95% CI) | I² (%) | Quality of meta-analysis (AMSTAR-2 rating) | Strength of association (umbrella review class) | GRADE credibility of evidence |
|---------------|-------------|----------------|------------|---------------------------|--------------|--------|----------------------------------------|-----------------------------------------------|-------------------------------|
| **Cancer incidence** |             |                |            |                           |              |        |                                        |                                 |                               |
| All           |             |                |            |                           |              |        |                                        |                                 |                               |
| Bolland, 2014(17) | 7           | 48,167         | 3979       | 0.99 (0.93–1.05)          | —            | 0      | Critically low                         | NS                             | Moderate                      |
| Bjelakovic, 2014(16) | 14         | 49,891         | 3851       | 1.00 (0.94–1.06)          | —            | 0      | High                                  | NS                             | Moderate                      |
| Keum, 2014(18)   | 4           | 45,151         | 4333       | 1.00 (0.94–1.06)          | —            | 0      | Critically low                         | NS                             | High                          |
| Goulão, 2018(19) | 24          | 18,440         | 1061       | 1.03 (0.91–1.15)          | —            | 0      | Critically low                         | NS                             | High                          |
| Haykal, 2019(22) | 9           | 42,773         | 3022       | 0.96 (0.86–1.07)          | —            | 31     | Critically low                         | NS                             | High                          |
| Keum, 2019(20)   | 10          | 83,353         | 6537       | 0.98 (0.93–1.03)          | —            | 0      | Critically low                         | NS                             | High                          |
| **Breast**      |             |                |            |                           |              |        |                                        |                                 |                               |
| Sperati, 2013(15) | 2           | 5372           | 91         | 1.11 (0.74–1.68)          | —            | 0      | Moderate                               | NS                             | Low                           |
| Bjelakovic, 2014(16) | 7           | 43,669         | 1135       | 0.97 (0.86–1.09)          | —            | 0      | High                                  | NS                             | Moderate                      |
| **Colorectal**  |             |                |            |                           |              |        |                                        |                                 |                               |
| Bjelakovic, 2014(16) | 5           | 45,598         | 436        | 1.11 (0.92–1.34)          | —            | 0      | High                                  | NS                             | Moderate                      |
| **Lung**        |             |                |            |                           |              |        |                                        |                                 |                               |
| Bjelakovic, 2014(16) | 5           | 45,509         | 329        | 0.86 (0.69–1.07)          | —            | 0      | High                                  | NS                             | Moderate                      |
| **Pancreatic**  |             |                |            |                           |              |        |                                        |                                 |                               |
| Bjelakovic, 2014(16) | 2           | 36,405         | 69         | 0.91 (0.57–1.46)          | —            | 0      | High                                  | NS                             | Moderate                      |
| **Cancer mortality** |           |                |            |                           |              |        |                                        |                                 |                               |
| All           |             |                |            |                           |              |        |                                        |                                 |                               |
| Bjelakovic, 2014(16) | 4           | 44,492         | 1192       | 0.88 (0.78–0.98)          | 292 (159–1751) | 0      | High                                  | Weak                          | High                          |
| Keum, 2014(18)   | 3           | 44,290         | 1190       | 0.88 (0.78–0.98)          | 86 (47–515)  | 0      | Critically low                         | Weak                          | High                          |
| Goulão, 2018(19) | 7           | 11,202         | 320        | 0.88 (0.70–1.09)          | —            | 0      | Critically low                         | NS                             | Moderate                      |
| Haykal, 2019(22) | 5           | 70,547         | 1533       | 0.87 (0.79–0.96)          | 381 (236–1238) | 0      | Low                                   | Weak                          | High                          |
| Keum, 2019(20)   | 5           | 75,239         | 1591       | 0.87 (0.79–0.96)          | 294 (182–957) | 0      | Critically low                         | Weak                          | High                          |
| Zhang, 2019(21)  | 12          | 45,578         | 939        | 0.84 (0.74–0.95)          | 279 (171–892) | 0      | Moderate                               | Weak                          | High                          |

*All were randomized controlled trials.*

AMSTAR = A Measurement Tool to Assess Systematic Reviews; GRADE = Grading of Recommendations, Assessment, Development, and Evaluation; NNT = number needed to treat; NS = nonsignificant association; RR = relative risk.
is inversely related to all-cause mortality in patients with or without cancer.\(^{(75)}\)

Fourth, cancer incidence and mortality, and overall survival, though considered the gold standard endpoint in oncology trials, require a large sample size and long follow-up time to achieve adequate statistical power.\(^{(76)}\) In comparison, other clinical endpoints can be assessed earlier, and thus could be measured in parallel. One of these is health-related quality of life, which is considered an outcome in assessing clinical benefit and has emerged as a primary endpoint in oncology clinical trials.\(^{(56)}\) Encouragingly, observational cohort studies found that, in patients with cancer, vitamin D intake and 25OHD predict improved quality of life,\(^{(77–79)}\) but RCT data are needed to evaluate effects on this outcome. Another is tumor-centered endpoints, such as progression-free survival, disease-free survival, tumor response, and circulating tumor cells.\(^{(76)}\) As mentioned, two RCTs reported beneficial effects on median progression-free survival\(^{(74)}\) and relapse-free survival,\(^{(73)}\) but further RCTs are needed.\(^{(75)}\) Thus, adding these other clinical endpoints (alongside gold standard outcomes) to future vitamin D trials would provide an earlier assessment and more comprehensive evaluation of efficacy.\(^{(76)}\)

Finally, a novel area of research is investigating whether vitamin D pathway genes may alter health effects on vitamin D supplementation. A meta-analysis of eight prospective studies found that colorectal cancer risk was lower in participants with the BB genotype of the BsmI vitamin D receptor single-nucleotide polymorphism.\(^{(60)}\) A recent RCT found that vitamin D receptor genotypes modified the effect of vitamin D supplementation on the prevention of advanced colorectal adenomas.\(^{(80)}\)

Specifically, vitamin D supplementation reduced risk by 64% among those with the rs7968585 genotype and increased risk by 41% among those with one or two G alleles.\(^{(80)}\) Such work helps identify who may benefit from supplementation for cancer prevention based on vitamin D-related genotypes. Given the knowledge gap in investigating vitamin D pathway genotypes as modifiers of effects on cancer outcomes, assessment of these in recent trials (possibly as IPD meta-analyses), and further trials in this research area are both warranted. As recognition of the importance of this, the VITAL trial,\(^{(62)}\) for example, is in the process of conducting such analyses of gene variants (JE Manson, personal communication, 2020).

### Conclusion

Observational studies showed that, in many cases, low vitamin D was inversely associated with cancer outcomes. For this, the associations for some outcomes, particularly colorectal cancer, seem to fulfill some (but not all) of Bradford-Hill’s criteria for causation,\(^{(64)}\) including consistency of findings across different meta-analyses and primary studies, temporality (prospective-study associations), biological gradient (dose–response associations), and strength of associations (strong in some cases).

To our knowledge, this review is the first report to systematically compile and appraise clinical evidence—by concurrently using AMSTAR-2, umbrella review, and GRADE assessment tools—of vitamin D supplementation in relation to cancer outcomes from meta-analyses. We found highly credible RCT evidence that vitamin D supplementation reduces risk of total cancer mortality, but the magnitude of effect was classified as weak. Our finding of a highly credible weak effect on total cancer mortality is in line with that of a 2017 systematic review of meta-analyses, which reported that vitamin D supplementation reduces risk of all-cancer mortality.\(^{(81)}\) We extend that review by including more recent meta-analyses in our assessment,\(^{(13,19–21)}\) and by critically evaluating the evidence for this outcome using the above-mentioned appraisal tools.

The available research, however, is not without limitations. To address these limitations and to provide clearer and further insight into the role of vitamin D in cancer incidence and related mortality, future research should include trials with more vitamin D-insufficient participants and of longer follow-up duration, plus adequately powered data on site-specific cancers (where feasible).

### Disclosures

The authors have no conflicts of interest.

### Acknowledgments

The Health Research Council of New Zealand (HRC) supported JDS with a fellowship.

Authors’ roles: JDS conducted the literature search, created the tables, and drafted the manuscript. RS verified the AMSTAR-2 and GRADE assessments. JEM and RS edited and revised the manuscript. HRC had no role in the design, analysis, interpretation, or presentation of the results.

All authors approved the final version of the manuscript.

### Author Contributions

John Sluyter: Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing-original draft; writing-review and editing. JoAnne Manson: Conceptualization; supervision; writing-review and editing. Robert Scragg: Conceptualization; supervision; validation; writing-review and editing.

### Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1002/jbm4.10420.

### References

1. Garland CF, Garland FC. Do sunlight and vitamin D reduce the likelihood of colon cancer? Int J Epidemiol. 1980;9(3):227–31.
2. Garland FC, Garland CF, Gorham ED, Young JF. Geographic variation in breast cancer mortality in the United States: a hypothesis involving exposure to solar radiation. Prev Med. 1990;19(6):614–22.
3. Grant WB. Ecological studies of the UVB-vitamin D-cancer hypothesis. Anticancer Res. 2012;32(1 PART 2):223–36.
4. Fleet JC, Desmet M, Johnson R, Li Y. Vitamin D and breast cancer: a review of molecular mechanisms. Biochem J. 2012;441(1):61–76.
5. Hossain S, Beydoun MA, Beydoun HA, Chen X, Zonderman AB, Wood RJ. Vitamin D and breast cancer: a systematic review and meta-analysis of observational studies. Clin Nutr ESPEN. 2019;30:170–84.
6. Touvier M, Chan DSM, Lau R, et al. Meta-analyses of vitamin D intake, 25-hydroxyvitamin D status, vitamin D receptor polymorphisms, and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev. 2011;20(5):1003–16.
7. Liu Y, Yu Q, Zhu Z, et al. Vitamin and multiple-vitamin supplement intake and incidence of colorectal cancer: a meta-analysis of cohort studies. Med Oncol. 2015;32(1):1–10.
8. Liu Y, Wang X, Sun X, Lu S, Liu S. Vitamin intake and pancreatic cancer risk reduction. Medicine. 2018;97(13):1–8.

9. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j3581–9.

10. Maalmi H, Ordóñez-Mena JM, Schöttker B, Brenner H. Serum 25-hydroxyvitamin D levels and survival in colorectal and breast cancer patients: systematic review and meta-analysis of prospective cohort studies. Eur J Cancer. 2014;50(8):1510–21.

11. Feng Q, Zhang H, Dong Z, Zhou Y, Ma J. Circulating 25-hydroxyvitamin D and lung cancer risk and survival. Medicine. 2017;96(45):1–6.

12. Hu K, Callen DF, Li J, Zheng H. Circulating vitamin D and overall survival in breast cancer patients: a dose-response meta-analysis of cohort studies. Integr Cancer Ther. 2018;17(2):217–25.

13. Han J, Guo X, Yu X, et al. 25-hydroxyvitamin D and total cancer incidence and mortality: a meta-analysis of prospective cohort studies. Nutrients. 2019;11(10):1–17.

14. Song D, Deng Y, Liu K, et al. Vitamin D intake, blood vitamin D levels, and the risk of breast cancer: a dose-response meta-analysis of observational studies. Aging. 2019;11(24):12708–32.

15. Sperati F, Vici P, Maugeri-Saccà M, et al. Vitamin D supplementation and breast cancer prevention: a systematic and meta-analysis of randomized clinical trials. PLoS One. 2013;8(7):1–9.

16. Bjelakovic G, Gluud LL, Nikolova D, et al. Vitamin D supplementation for prevention of cancer in adults. Cochrane Database Syst Rev. 2014;1:CD007469.

17. Bolland MJ, Grey A, Gamble GD, Reid IR. The effect of vitamin D supplementation on cancer risk in randomized clinical trials. Cochrane Database Syst Rev. 2013;5(6):CD000939.

18. Keum N, Giovannucci E. Vitamin D supplements and cancer incidence and mortality: a meta-analysis. Br J Cancer. 2014;111(5):976–80.

19. Goulão B, Stewart F, Ford JA, Macleennan G, Avenell A. Cancer and vitamin D supplementation: a systematic review and meta-analysis. Am J Clin Nutr. 2018;107(4):652–63.

20. Keum N, Lee DH, Greenwood DC, Manson JE, Giovannucci E. Vitamin D supplementation and total cancer incidence and mortality: a meta-analysis of randomized controlled trials. Ann Oncol. 2019;30(5):733–43.

21. Zhang Y, Fang F, Tang J, et al. Association between vitamin D supplements and colorectal cancer: a systematic review and meta-analysis. BMJ. 2019;366:l4673.

22. Haykal T, Samji V, Zayed Y, et al. The role of vitamin D supplementation for primary prevention of cancer: meta-analysis of randomized controlled trials. J Commun Hosp Intern Med Perspect. 2019;9(6):480–8.

23. Papanicolaou S. Umbrella reviews: what are they and why we need them. Eur J Epidemiol. 2019;34:543–6.

24. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–40.

25. Ioannidis JPA, Trikalinos TA. An exploratory test for an excess of significance in meta-analyses of 98 biomarker associations with cancer risk. J Natl Cancer Inst. 2012;104(24):1867–78.

26. Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0: updated March 2011. The Cochrane Collaboration; 2011 [cited April 2020]. Available from: www.cochrane-handbook.org.

27. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Schünemann HJ. GRADE: what is “quality of evidence” and why is it important to clinicians? BMJ. 2008;336(7651):995–8.

28. Chen P, Hu P, Xie D, Qin Y, Wang F, Wang H. Meta-analysis of vitamin D, calcium and the prevention of breast cancer. Breast Cancer Res Treat. 2010;121(2):469–77.

29. Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: serum vitamin D and breast cancer risk. Eur J Cancer. 2010;46(12):2196–205.

30. Chung M, Lee J, Terasawa T, Lau J, Trikalinos TA. Vitamin D with or without calcium supplementation for prevention of cancer and fractures: an updated meta-analysis for the U.S. preventive services task force. Ann Intern Med. 2011;155(12):827–38.

31. Gandini S, Boniol M, Haukka J, et al. Meta-analysis of observational studies of serum 25-hydroxyvitamin D levels and colorectal, breast and prostate cancer and colorectal adenoma. Int J Cancer. 2011;128(6):1414–24.

32. Mohr SB, Gorham ED, Alcaraz JE, et al. Serum 25-hydroxyvitamin D and prevention of breast cancer: pooled analysis. Anticancer Res. 2011;31(9):2939–48.

33. Bauer SR, Hankinson SE, Bertone-Johnson ER, Ding EL. Plasma vitamin D levels, menopause, and risk of breast cancer: dose-response meta-analysis of prospective studies. Medicine. 2013;92(3):123–31.

34. Chen P, Li M, Gu X, et al. Higher blood 25(OH)D level may reduce the breast cancer risk: evidence from a Chinese population based case-control study and meta-analysis of the observational studies. PloS One. 2013;8(1):1–10.

35. Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: Longitudinal studies of serum vitamin D and colorectal cancer risk. Aliment Pharmacol Ther. 2009;30(2):113–25.

36. Lee JE, Li H, Chan AT, et al. Circulating levels of vitamin D and colon and rectal cancer: the Physicians’ health study I and a meta-analysis of prospective studies of cancer prevention. Cancer Prev Res. 2011;4(5):735–43.

37. Gorham ED, Garland CF, Garland FC, et al. Optimal vitamin D status for colorectal cancer prevention. A quantitative meta analysis. Am J Prev Med. 2007;32(3):210–6.

38. Ma Y, Zhang P, Wang F, Yang J, Liu Z, Qin H. Association between vitamin D and risk of colorectal cancer: a systematic review of prospective studies. J Clin Oncol. 2011;29(28):3775–82.

39. Fedirko V, Bostick RM, Goodman M, Flanders WD, Gross MD. Blood 25-hydroxyvitamin D3 concentrations and incident sporadic colorectal adenoma risk: a pooled case-control study. Am J Epidemiol. 2010;172(5):489–500.

40. Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: serum vitamin D and colorectal adenoma risk. Prev Med. 2011;53(1–2):10–6.

41. Gallicchio L, Moore LE, Stevens VL, et al. Circulating 25-hydroxyvitamin D and risk of kidney cancer: cohort consortium vitamin D pooling project of rarer cancers. Am J Epidemiol. 2010;172(1):47–57.

42. Purdue MP, Freedman DM, Gapstur SM, et al. Circulating 25-hydroxyvitamin D and risk of non-Hodgkin lymphoma: cohort consortium vitamin D pooling project of rarer cancers. Am J Epidemiol. 2010;172(1):58–69.

43. Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: circulating vitamin D and ovarian cancer risk. Gynecol Oncol. 2011;121(2):369–75.

44. Stolzenberg-Solomon RZ, Jacobs EJ, Arslan AA, et al. Circulating 25-hydroxyvitamin D and risk of pancreatic cancer: cohort consortium vitamin D pooling project of rarer cancers. Am J Epidemiol. 2010;172(1):81–93.

45. Yin L, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis of longitudinal studies: serum vitamin D and prostate cancer risk. Cancer Epidemiol. 2009;33(6):435–45.

46. Gilbert R, Martin RM, Beynon R, et al. Associations of circulating and dietary vitamin D with prostate cancer risk: a systematic review and dose-response meta-analysis. Cancer Causes Control. 2011;22(3):319–40.

47. Zhang L, Zou H, Zhao Y, et al. Association between blood circulating vitamin D and colorectal cancer risk in Asian countries: a systematic review and dose-response meta-analysis. BMJ Open. 2019;9(12):1–13.

48. Wang D, Vélez De-La-Paz OT, Zhai JX, Liu DW. Serum 25-hydroxyvitamin D and breast cancer risk: a meta-analysis of prospective studies. Tumor Biol. 2013;34(6):3309–17.
50. Estébanez N, Gómez-Acebo I, Palazuelos C, Llorca J, Dierssen-Sotos T. Vitamin D exposure and risk of breast cancer: a meta-analysis. Sci Rep. 2018;8(1):1–13.
51. Guo XF, Zhao T, Han JM, Li S, Li D. Vitamin D and liver cancer risk: a meta-analysis of prospective studies. Asia Pac J Clin Nutr. 2020;29(1):172–82.
52. Hu MJ, Zhang Q, Liang L, et al. Association between vitamin D deficiency and risk of thyroid cancer: a case–control study and a meta-analysis. J Endocrinol Invest. 2018;41(10):1199–210.
53. Zhao J, Wang H, Zhang Z, et al. Vitamin D deficiency as a risk factor for thyroid cancer: a meta-analysis of case-control studies. Nutrition. 2019;57:5–11.
54. Li M, Chen P, Li J, Chu R, Xie D, Wang H. Review: the impacts of circulating 25-hydroxyvitamin D levels on cancer patient outcomes: a systematic review and meta-analysis. J Clin Endocrinol Metab. 2014;99(7):2327–36.
55. Xu J, Yuan X, Tao J, Yu N, Wu R, Zhang Y. Association of circulating 25-hydroxyvitamin D levels with colorectal cancer: an updated meta-analysis. J Nutr Sci Vitaminol. 2018;64(6):432–44.
56. Kim Y, Je Y. Vitamin D intake, blood 25(OH)D levels, and breast cancer risk or mortality: a meta-analysis. Br J Cancer. 2014;110(11):2772–84.
57. Huang D, Lei S, Wu Y, et al. Additively protective effects of vitamin D and calcium against colorectal adenoma incidence, malignant transformation and progression: a systematic review and meta-analysis. J Clin Endocrinol Metab. 2014;99(7):2327–36.
58. Wei MY, Garland CF, Gorham ED, Mohr SB, Giovannucci E. Vitamin D and cancer risk or mortality: a meta-analysis. Br J Cancer. 2014;110(11):2772–84.
59. Wactawski-Wende J, Morley Ketchen J, Anderson GL, et al. Calcium plus vitamin D supplementation and the risk of colorectal cancer. N Engl J Med. 2006;354(7):684–96.
60. Bristow SM, Bolland MJ, MacLennan GS, et al. Calcium supplements and cancer risk: a meta-analysis of randomised controlled trials. Br J Nutr. 2013;110(8):1384–93.
61. Scruggs R, Khaw KT, Toop L, et al. Monthly high-dose vitamin D supplementation and cancer risk: a post hoc analysis of the vitamin D assessment randomized clinical trial. JAMA Oncol. 2018;4(11):e182178.
62. Manson JE, Cook NR, Lee IM, et al. Vitamin D supplements and prevention of cancer and cardiovascular disease. N Engl J Med. 2019;380(1):33–44.
63. Lai Y-H, Fang T-C. The pleiotropic effect of vitamin D. ISRN Nephrol. 2013;2013:898125.
64. Hill AB. The environment and disease: association or causation? Pr R Soc Med. 1965;58:295–300.
65. Barker T, Fulde G, Moulton B, Nadauld LD, Rhodes T. An elevated neutrophil-to-lymphocyte ratio associates with weight loss and cachexia in cancer. Sci Rep. 2020;10(1):7535.
66. Galior K, Ketha H, Grebe S, Singh RJ. 10 years of 25-hydroxyvitamin-D testing by LC-MS/MS-trends in vitamin-D deficiency and sufficiency. Bone Rep. 2018;8:268–73.
67. Scruggs R. Emerging evidence of thresholds for beneficial effects from vitamin D supplementation. Nutrients. 2018;10(5):1–11.
68. Cuzick J. Long-term follow-up in cancer prevention trials (it ain’t over ‘til it’s over). Cancer Prev Res. 2010;3(6):689–91.
69. Bouillon R. Free or total 25OHD as marker for vitamin D status? J Bone Miner Res. 2016;31(6):1124–7.
70. Chun RF, Shieh A, Gottlieb C, et al. Vitamin D binding protein and the biological activity of vitamin D. Front Endocrinol. 2019;10:1–15.
71. Urashima M, Okuyama M, Akutsu T, Ohdaira H, Kaji M, Suzuki Y. Effect of vitamin D supplementation on survival of digestive tract cancer patients with low bioavailable 25-hydroxyvitamin D levels: a post hoc analysis of the AMATERASU randomized clinical trial. Cancers. 2020;12(2):1–11.
72. Fang AP, Long JA, Zhang YJ, et al. Serum bioavailable, rather than total, 25-hydroxyvitamin D levels are associated with hepatocellular carcinoma survival. Hepatology. 2020;72:169–182.
73. Urashima M, Ohdaira H, Akutsu T, et al. Effect of vitamin D supplementation on relapse-free survival among patients with digestive tract cancers: the AMATERASU randomized clinical trial. JAMA. 2019;321(14):1361–9.
74. Ng K, Nimeiri HS, McCleary NJ, et al. Effect of high-dose vs standard-dose vitamin D3 supplementation on progression-free survival among patients with advanced or metastatic colorectal cancer: the SUNSHINE randomized clinical trial. JAMA. 2019;321(14):1370–9.
75. Barry EL, Passarelli MN, Baron JA. Vitamin D as cancer therapy?: insights from 2 new trials. JAMA. 2019;321(14):1354–5.
76. Fiteni F, Westeel V, Pivot X, Borg C, Vernerey D, Bonnetain F. Endpoints in cancer clinical trials. J Visc Surg. 2014;151(1):17–22.
77. Andersen MR, Sweet E, Hager S, Gaul M, Dowd F, Standish LJ. Effects of vitamin D use on health-related quality of life of breast cancer patients in early survivorship. Integr Cancer Ther. 2019;18:1–12.
78. Anand A, Singh S, Sonkar AA, et al. Expression of vitamin D receptor and vitamin D status in patients with oral neoplasms and effect of vitamin D supplementation on quality of life in advanced cancer treatment. Contemp Oncol. 2017;21(2):145–51.
79. Wang L, Wang C, Wang J, Huang X, Cheng Y. Longitudinal, observational study on associations between postoperative nutritional vitamin D3 supplementation and clinical outcomes in esophageal cancer patients undergoing esophagectomy. Sci Rep. 2016;6:1–10.
80. Barry EL, Peacock JL, Rees JR, et al. Vitamin D receptor genotype, vitamin D3 supplementation, and risk of colorectal adenomas: a randomized clinical trial. JAMA Oncol. 2017;3(5):628–35.
81. 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.
## Table A1. Umbrella Review Assessment Grades

| Strength of association | Criteria |
|-------------------------|----------|
| Convincing (class I)    | >1000 cases<sup>a</sup> |
|                         | Significant summary associations ($p < 10^{-6}$) per random-effects calculations |
|                         | No evidence of small-study effects |
|                         | No evidence of excess of significance bias |
|                         | Prediction intervals not including the null value |
|                         | Largest study nominally significant ($p < .05$) |
|                         | Not large heterogeneity ($I^2 < 50\%$) |
| Highly suggestive (class II) | >1000 cases<sup>a</sup> |
|                         | Significant summary associations ($p < 10^{-6}$) per random-effects calculations |
|                         | Largest study nominally significant ($p < .05$) |
| Suggestive (class III)  | >1000 cases<sup>a</sup> |
|                         | Significant summary associations ($p < 10^{-3}$) per random-effects calculations |
| Weak (class IV)         | Significant summary associations ($p < .05$) per random-effects calculations |
| Nonsignificant association | Nonsignificant summary associations ($p \geq .05$) |

<sup>a</sup>Total for the meta-analysis.

## Table A2. GRADE Assessment Scoring<sup>a</sup>

| Study design                  | Quality of evidence | Lower if                              | Higher if                            |
|-------------------------------|---------------------|---------------------------------------|--------------------------------------|
| Randomized trial              | High                | Risk of bias:                         | Large effect                         |
|                               |                     | -1 Serious                            | +1 Large                             |
|                               | Moderate            |                                       |                                      |
| Observational study           | Low                 | -2 Very serious                       | +2 Very large                        |
|                               |                     | Inconsistency                         | Dose response                        |
|                               |                     | -1 Serious                            | +1 Evidence of a gradient            |
|                               |                     | -2 Very serious                       | All plausible confounding            |
|                               |                     | Indirectness                          | +1 Would reduce a                    |
|                               |                     | -1 Serious                            | demonstrated effect or               |
|                               |                     | -2 Very serious                       | +1 Would suggest a spurious effect   |
|                               |                     | Imprecision                           | when results show no effect          |
|                               |                     | -1 Serious                            |                                       |
|                               |                     | -2 Very serious                       |                                       |
|                               |                     | Publication bias                      |                                       |
|                               |                     | -1 Likely                             |                                       |
|                               |                     | -2 Very likely                        |                                       |

<sup>a</sup>From: Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–94.

GRADE = Grading of Recommendations, Assessment, Development, and Evaluation.
Table A3. Study Selection Criteria of Meta-Analyses of Intervention Studies<sup>a</sup> on the Effect of Vitamin D Supplementation on Cancer Outcomes

| Meta-analysis | Inclusion criteria | Exclusion criteria |
|---------------|-------------------|-------------------|
| Sperati, 2013<sup>(15)</sup> | 1. Compared with placebo/no treatment  
2. Vitamin D as single agent  
3. Combined regimens including supplements & lifestyle modifications if used equally in all groups | Pregnant or lactating women |
| Bjelakovic, 2014<sup>(16)</sup> | 1. RCTs, irrespective of blinding, publication, status, or language.  
2. Any dose, duration, and route of administration  
3. Monotherapy or in combination with calcium  
4. Concomitant interventions if used equally in all intervention groups | 1. Secondary induced osteoporosis (eg, glucocorticoid-induced osteoporosis, thyroidectomy, primary hyperparathyroidism, chronic kidney disease, liver cirrhosis, Crohn disease, gastrointestinal bypass surgery)  
2. Pregnant or lactating women  
3. People with cancer |
| Bolland, 2014<sup>(17)</sup> | Cholecalciferol or ergocalciferol | 1. Cluster randomized trials  
2. Trials of hydroxylated vitamin D or vitamin D analogues  
3. Other interventions only in vitamin D group  
4. Trials of fortified dairy products  
5. Chronic comorbidity other than osteoporosis or frailty |
| Keum, 2014<sup>(18)</sup> | With or without calcium supplementation | 1. Non-English articles  
2. Abstracts & unpublished reports |
| Goulão, 2018<sup>(19)</sup> | 1. Mean or median age of ≥60 years  
2. Follow-up ≤1 year  
3. Any vitamin D or vitamin D analog  
4. Co-administration of other medications (eg, calcium) if the comparator group received the same medication  
5. All languages | 1. Renal impairment, steroid-induced osteoporosis, or psoriasis  
2. Nonmelanoma skin cancers not counted as events |
| Haykal, 2019<sup>(22)</sup> | 1. Primary prevention  
2. Vitamin D compared with placebo  
3. Vitamin D for ≥3 years | |
| Keum, 2019<sup>(20)</sup> | Cholecalciferol or ergocalciferol, with or without other nutrients | 1. Number of outcomes ≤10  
2. Follow-up ≤1 year |
| Zhang, 2019<sup>(21)</sup> | 1. Age ≥ 18 years  
2. Any health conditions  
3. Vitamin D (any dose) vs. placebo or no treatment  
4. Concomitant agents had to be same dose in all groups | 1. Case reports, case series, observational studies  
2. All participants received vitamin D  
3. Pregnant or lactating women  
4. Critically patients  
5. Hydroxylated vitamin D or vitamin D analogues |

<sup>a</sup>All were randomized controlled trials (RCTs).
| AMSTAR-2 item | Description | Bolland, 2014\(^{(17)}\) | Sperati, 2013\(^{(19)}\) | Bjelakovic, 2014\(^{(16)}\) | Keum, 2014\(^{(18)}\) | Goulão, 2018\(^{(19)}\) | Haykal, 2019\(^{(22)}\) | Keum, 2019\(^{(20)}\) | Zhang, 2019\(^{(21)}\) |
|---------------|-------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| 1             | Did the research questions and inclusion criteria include the components of PICO? | ✓✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | |
| 2\(^{a}\)    | Did the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ | ×✓✓✓✓✓✓✓ |
| 3             | Did the review authors explain their selection of the study designs for inclusion in the review? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 4\(^{a}\)    | Did the review authors use a comprehensive literature search strategy? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 5             | Did the review authors perform study selection in duplicate? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 6             | Did the review authors perform data extraction in duplicate? | ✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 7\(^{a}\)    | Did the review authors provide a list of excluded studies and justify the exclusions? | ✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 8             | Did the review authors describe the included studies in adequate detail? | ✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 9\(^{a}\)    | Did the review authors use a satisfactory technique to assess the RoB in studies that were included in the review? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 10            | Did the review authors report on the sources of funding for the studies included in the review? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 11\(^{a}\)   | If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? | ✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 12            | If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis? | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |
| 13\(^{a}\)   | Did the review authors account for RoB in individual studies when | ×✓✓✓✓✓✓✓✓✓✓✓ | | | | | | | | | |

(Continues)
| Item | Description | First author, publication year (citation) |
|------|-------------|------------------------------------------|
| 14   | Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? | ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ |
| 15   | If they performed quantitative synthesis did the review authors investigate publication bias and discuss its likely impact on the results? | ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ |
| 16   | Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review? | ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ |

| Rating of overall confidence in the results of the review | CL | Moderate | High | CL | CL | CL | CL | Moderate |

*Critical domains.

AMSTAR = A Measurement Tool to Assess Systematic Reviews; CL = critically low; PICO components: P = patient population/problem, I = intervention (issue of interest, considered for implementation), C = comparison (comparator or current practice), O = outcome (how this is measured); RoB = risk of bias.
Table A5. Umbrella Review Assessment of Meta-Analyses of Intervention Studies on the Effect of Vitamin D Supplementation on Cancer Outcomes

| Reference | RR Effect (95% CI) | >1000 events (cases) | Significant summary associations per random-effects calculations | No evidence of small-study effects | No evidence of excess of significance | Prediction intervals excluding null value | Largest study nominally significant (p < .05) | Not large heterogeneity (I² <50%) | Umbrella review class |
|-----------|-------------------|---------------------|----------------------------------------------------------------|----------------------------------|-------------------------------------|--------------------------------------|------------------------------------------|---------------------------------------|---------------------|
| All cancers |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Bolland, 2014(17) | 0.99 (0.93–1.05) | ✓                   | x x x x ✓ ✓                                                  | x                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Bjelakovic, 2014(16) | 1.00 (0.94–1.06) | ✓                   | x x x x x x                                                  | x                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Keum, 2014(18) | 1.00 (0.94–1.06) | ✓                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Goulão, 2018(19) | 1.03 (0.91–1.15) | ✓                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Haykal, 2019(22) | 0.96 (0.86–1.07) | ✓                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Keum, 2019(20) | 0.98 (0.93–1.03) | ✓                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Breast cancer |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Sperati, 2013(15) | 1.11 (0.74–1.68) | x                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Bjelakovic, 2014(16) | 0.97 (0.86–1.09) | ✓                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Colorectal cancer |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Bjelakovic, 2014(16) | 1.11 (0.92–1.34) | x                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Lung cancer |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Bjelakovic, 2014(16) | 0.86 (0.69–1.07) | x                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Pancreatic cancer |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Bjelakovic, 2014(16) | 0.91 (0.57–1.46) | x                   | x x x x x x                                                  | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Total cancer mortality |                  |                     |                                                                 |                                  |                                     |                                      |                                          |                                        |                     |
| Bjelakovic, 2014(16) | 0.88 (0.78–0.98) | ✓                   | x x x x ✓ x                                                   | x                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | Weak (class IV)      |
| Keum, 2014(18) | 0.88 (0.78–0.98) | ✓                   | x x x x ✓ x                                                   | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | Weak (class IV)      |
| Goulão, 2018(19) | 0.88 (0.70–1.09) | x                   | x x x x ✓ x                                                   | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | NS association       |
| Haykal, 2019(22) | 0.87 (0.79–1.06) | x                   | x x x x ✓ x                                                   | ✓                               | ✓                                    | x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | ✓ x x x x x x x x ✓                        | Weak (class IV)      |
| Reference      | RR Effect (95% CI) | >1000 events (cases) | Significant summary associations per random-effects calculations | No evidence of small-study effects | No evidence of excess of significance | Prediction intervals excluding null value | Largest study nominally significant (p < .05) | Not large heterogeneity (I² < 50%) | Umbrella review class |
|----------------|---------------------|----------------------|---------------------------------------------------------------|----------------------------------|--------------------------------------|-------------------------------------------|---------------------------------------------|------------------------------------------|------------------------|
| Keum, 2019(20) | 0.87 (0.79–0.96)    | ✓ × × ✓ ✓            |                                               | ✓                                | ✓                                    | ✓                                         | ✓                                          | ×                                       | Weak (class IV)        |
| Zhang, 2019(21)| 0.84 (0.74–0.95)    | × × × ✓ ✓ ×           |                                               | ✓                                | ✓                                    | ✓                                         | ✓                                          | ✓                                       | Weak (class IV)        |

NS = nonsignificant; RCT = randomized controlled trial; RR = relative risk.
| Reference          | Studies (N) | Study design | Risk of bias | Imprecision | Inconsistency | Indirectness | Publication bias | RR Effect (95% CI) | Certainty (GRADE) |
|--------------------|-------------|--------------|--------------|-------------|---------------|--------------|-----------------|---------------------|-------------------|
| **All cancers**    |             |              |              |             |               |              |                 |                     |                   |
| Bolland, 2014 \(^{17}\) | 7           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Possible: Egger’s P value = 0.05 | 0.99 (0.93–1.05) | +++ Moderate       |
| Bjelakovic, 2014 \(^{16}\) | 14          | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Likely: Egger’s P value = 0.007 | 1.00 (0.94–1.06) | +++ Moderate       |
| Keum, 2014 \(^{18}\) | 4           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 1.00 (0.94–1.06) | ++++ High          |
| Goulão, 2018 \(^{19}\) | 24          | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 1.03 (0.91–1.15) | ++++ High          |
| Haykal, 2019 \(^{22}\) | 9           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.96 (0.86–1.07) | ++++ High          |
| Keum, 2019 \(^{20}\) | 10          | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.98 (0.93–1.03) | ++++ High          |
| Goulão, 2018 \(^{19}\) | 7           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Too few studies to assess | 0.91 (0.57–1.46) | +++ Moderate       |
| **Breast cancer**  |             |              |              |             |               |              |                 |                     |                   |
| Sperati, 2013 \(^{15}\) | 2           | RCT          | Not serious  | Serious: wide CI from appreciable benefit to appreciable harm | Not serious | Not serious | Unlikely         | 1.11 (0.74–1.68) | ++ Low             |
| Bjelakovic, 2014 \(^{16}\) | 7           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.97 (0.86–1.09) | +++ Moderate       |
| **Lung cancer**    |             |              |              |             |               |              |                 |                     |                   |
| Bjelakovic, 2014 \(^{16}\) | 5           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.86 (0.69–1.07) | +++ Moderate       |
| **Colorectal cancer** |             |              |              |             |               |              |                 |                     |                   |
| Bjelakovic, 2014 \(^{16}\) | 5           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 1.11 (0.92–1.34) | +++ Moderate       |
| **Pancreatic cancer** |             |              |              |             |               |              |                 |                     |                   |
| Bjelakovic, 2014 \(^{16}\) | 2           | RCT          | Not serious  | Serious: wide CI from appreciable benefit to appreciable harm | Not serious | Not serious | Too few studies to assess | 0.91 (0.57–1.46) | +++ Moderate       |
| **Total cancer mortality** |             |              |              |             |               |              |                 |                     |                   |
| Bjelakovic, 2014 \(^{16}\) | 4           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.88 (0.78–0.98) | ++++ High          |
| Keum, 2014 \(^{14}\) | 3           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.88 (0.78–0.98) | ++++ High          |
| Goulão, 2018 \(^{19}\) | 7           | RCT          | Not serious  | Serious: wide CI from appreciable benefit to small harm | Not serious | Not serious | Unlikely         | 0.88 (0.70–1.09) | +++ Moderate       |
| Haykal, 2019 \(^{22}\) | 5           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.87 (0.79–0.96) | ++++ High          |
| Keum, 2019 \(^{20}\) | 5           | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.87 (0.79–0.96) | ++++ High          |
| Zhang, 2019 \(^{21}\) | 12          | RCT          | Not serious  | Not serious | Not serious   | Not serious  | Unlikely         | 0.84 (0.74–0.95) | ++++ High          |

**Note**: RCT = randomized controlled trial; RR = relative risk.