BMJ Open

Serum 25-hydroxyvitamin D and risk of cancer in a large community population under investigation for cardiovascular disease

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056543
Article Type:	Original research
Date Submitted by the Author:	17-Aug-2021
Complete List of Authors:	Yang, Jaeun; University of Alberta Faculty of Medicine & Dentistry, Department of Anesthesiology & Pain Medicine Dong, Yuan; University of Calgary Cumming School of Medicine, Community Health Sciences Naugler, Christopher; University of Calgary Cumming School of Medicine, Pathology and Laboratory Medicine de Koning, Lawrence; University of Calgary Cumming School of Medicine, Pathology and Laboratory Medicine
Keywords:	Epidemiology < ONCOLOGY, Clinical chemistry < PATHOLOGY, STATISTICS & RESEARCH METHODS

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

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

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

Serum 25-hydroxyvitamin D and risk of cancer in a large community population under investigation for cardiovascular disease

Jaeun Yang[1], Yuan Dong[2], Christopher T Naugler[3,4] and Lawrence de Koning[3,4]*

[1] Department of Anesthesiology & Pain Medicine, University of Alberta
2-150 Clinical Sciences Building (CSB) 8440 112 St NW, Edmonton, Alberta, Canada T6G 2G3.

[2] Department of Community Health Sciences, Cumming School of Medicine, Teaching, Research and Wellness Building, 5E25-3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6.

[3] Department of Pathology and Laboratory Medicine, Cumming School of Medicine, Health Sciences Centre, Foothills Campus, University of Calgary, 3030 Hospital Drive NW, Calgary, AB, Canada, T2N 4N1

[4] Alberta Precision Laboratories, Diagnostic and Scientific Research Centre, #9 3535 Research Way NW, Calgary, AB, Canada, T2L 2K8

*Address for correspondence:

Lawrence de Koning, PhD DABCC FAACC FCACB
Alberta Children's Hospital, Room B3-724, 28 Oki Drive NW
Calgary, Alberta, Canada
T3B6A8

Tel: 403-955-2277

Fax: 403-955-2321

Email: abldekon@ucalgary.ca

Twitter: **@ldekoning**

Word count: 2601

Tables/Figures: 5

Keywords: Vitamin D, cancer, epidemiology, risk factors

Conflict of interest/competing interests statement: The authors declare no potential conflicts of interest.

Role of the funding source: This work was partially supported by funding from the MSI foundation of Alberta (grant #871). The funding source had no role in this work.

Data sharing statement: Please send any requests to the corresponding author (abldekon@ucalgary.ca).

Contributors: JY, YD and LdK had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: LdK, CN. Acquisition, analysis or interpretation of data: JY, YD, LdK. Drafting the manuscript: YD, LdK. Critical revision of the manuscript: All authors. Analysis: JY, YD, LdK. Supervision: LdK.

Abstract:

Objectives: It remains unclear whether vitamin D status is related to cancer risk. We examined this relationship in a retrospective cohort study using laboratory, administrative and survey data.

Design: Retrospective cohort study

Setting: All care settings within Calgary, Alberta, Canada and surrounding rural communities.

Participants: Patients tested for serum 25-hydroxyvitamin D from 2009-2013 in without a past cancer diagnosis but with an electrocardiogram and body mass index (BMI) +/- 3 months from testing were included. Age, sex, mean hours of daylight during month of testing were linked to census dissemination area-level indicators of socioeconomic status (SES) measured in 2011.

Primary and secondary outcome measures: Hospital discharge diagnoses for any cancer, major cancer [colorectal, breast, lung, prostate, skin], and other cancers >3 months from testing from 2009-2016. Cox proportional hazard models were used to examine associations with incident cancer after adjusting for potential confounders. Interactions were tested using multiplicative terms.

Results: Among 72 171 patients, there were 3439 cancer diagnoses over a median of 5.9 years. After adjustment, increasing quartile of serum 25-OH vitamin D was significantly associated with an increased risk of any cancer and major cancer however this was completely driven by an increased risk of skin cancer. (Q4 vs Q1: HR = 2.56, 95% CI: 1.69-3.83, p for linear trend < 0.01). This association was strengthened among individuals residing in communities with higher proportions of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking an official Canadian language (English or French) at home.

Conclusions: Higher vitamin D status was associated with a greater risk of skin cancer in a large community population under investigation for cardiovascular disease. This association was likely due to sun exposure and may be modified by community variation in vitamin D supplementation.

Article summary:

Strengths and limitations of this study:

- We assembled a large retrospective cohort study of community patients by linking patient laboratory, national survey, and hospital administrative data during a time of high public and medical interest in vitamin D deficiency, and a commensurately high testing rate at our laboratory for serum 25-OH vitamin D.
- Linkage allowed us to control for body mass index, a confounder of 25-OH vitamin D-chronic disease relationships, and to explore variation in associations according to community-level socioeconomic factors correlated with vitamin D supplement use.
- As our study was restricted to patients who had received an electrocardiogram, it may not be generalizable to all patients.

1. Introduction:

Vitamin D deficiency, defined as a serum 25-OH (hydroxy) vitamin D concentration < 50 nmol/L, [1] is relatively common, especially in the Northern latitudes where people spend more time indoors due to cold and experience prolonged periods of darkness during winter [2]. In Canada, 33% of residents may be vitamin D deficient.[3] As such, there remains significant interest in whether vitamin D deficiency is related to the risk of disease – particularly cancer.

Meta-analysis of prospective cohort studies suggest that vitamin D deficiency is associated with an increased risk of multiple types of cancer, including all cancers [4], colorectal cancer [5], bladder cancer [6] head and neck cancer [7], liver cancer [8], and also death due to cancer [9]. These associations have been explained by *in vitro* and *in vivo* effects of the active form vitamin D, 1,25 OH_2 (dihydroxy) vitamin D, which promotes cellular differentiation, decreases cancer cell growth, stimulates cell death (apoptosis), and reduces angiogenesis.[10] Despite these plausible mechanisms, however, associations may also be explained by the presence of confounding factors associated with vitamin D deficiency and a higher risk of cancer.

For example, adiposity is a sink for and diluent of serum 25-hydroxyvitamin D [11] as well as a risk factor for several types of cancer [12]. As such, adjustment for some measure of adiposity (e.g. body mass index [BMI]) is generally recognized as essential to control for bias in epidemiologic studies. [1] Interestingly, low socioeconomic status (SES), while also a strong a risk factor for vitamin D deficiency [13] and cancer [14], is infrequently controlled for – probably because it is uncommonly measured in epidemiologic studies.

Historic uncertainty in the validity of epidemiologic findings have thus lead the Institute of Medicine (IOM) in 2011 to indicate that evidence of a relationship between vitamin D status and non-skeletal chronic diseases does not meet criteria for establishing cause-and-effect. [1] However, vitamin D status could still be a useful and convenient cancer risk marker – especially if its association is independent of other commonly measured factors and is observed in a large population of free-living individuals.

Our objectives were therefore to (i) examine the relationship between serum 25-hydroxyvitamin D (the major circulating form of vitamin D) and risk of cancer in a large community-based population, (ii) adjust for important confounders such as adiposity and socioeconomic status (SES), and (iii) test whether associations are modified by these and other covariates.

2. Materials and Methods:

2.1 Ethics statement

This study was approved by the University of Calgary Conjoint Health Review Ethics Board (Ethics ID 25065). Research in this article was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2 *Patient and public involvement statement*

Due to the design of the study and because we did not collect the primary data, we did not involve patients or the public in the design, conduct or reporting of our research.

2.3 Population, primary exposure variable and covariates

We used the Cerner (Kansas City, MO, USA) Millennium laboratory information system (LIS) to identify those who had a serum 25-hydroxyvitamin D result with a test date from December 8 2009 to April 1 2013. This database contained all laboratory results on patients tested in Calgary, Alberta, Canada (population 1.4 million) as well as surrounding rural communities. During the time of this study, 25-hydroxyvitamin D testing was available to any ordering physician for any reason – and high test volumes (~16k / month) reflected a strong public and medical interest in vitamin D deficiency. For these patients, we retained data only for those who had an ECG because these patients had a self-reported height and weight entered into the LIS to calculate body mass index (BMI; weight in kg / (height in meters)^2). We then extracted age, sex, and the provincial health care number (PHN) to link to hospital administrative outcome data and postal codes. As vitamin D status is related to sun exposure, we also calculated mean hours of daylight during month of 25-hydroxyvitamin D testing for each person based on publicly available data [15]. This was done so that we could adjust for short-term variation in 25-hydroxyvitamin D related to seasonal changes in sun exposure at the time of testing. All serum 25-hydroxyvitamin D tests were performed on the DiaSorin (Saluggia, Italy) Liaison total 25-hydroxyvitamin D automated immunoassay platform, which predominantly detects 25-hydroxyvitamin D3.[16]

Census-dissemination area (CDA)-level SES covariates were obtained from the 2011 Canadian National Household Survey (NHS) after postal code to CDA conversion. We extracted proportion of CDA residents who were (i) Canadian citizens, (ii) recent

immigrants, (iii) visible (non-white) minorities, (iv) those speaking languages other than English and French (official languages of Canada) at home, (v) those having postsecondary education, (vi) those currently employed, and the CDA (vii) median household income.

2.4 Outcomes

In-hospital discharge diagnosis and dates were obtained from the Discharge Abstract Database (DAD) and National Ambulatory Care Reporting System (NACRS) via Alberta Health Services. Incident cancers were defined as the first and most responsible (primary) diagnosis of any cancer (ICD10 codes: C00.x-C97.x) if the patient was discharged alive or died in hospital. We further subdivided 'any cancer' into total and individual 'major cancer' [17] (breast [C50.x], colorectal [C18.x-C21.x], lung [C33.x-C34.x], prostate [C61.x] and skin [C43.x, C44.x, C46.x] [18]) and 'other cancer' – defined as 'any cancer' minus 'major cancer'. At the time of analysis, outcome data was available until December 31 2016.

2.5 Data cleaning

We kept only the first measurement of 25-hydroxyvitamin D to capture historic vitamin D status – which is more likely associated with cancer risk than vitamin D status affected by supplementation in response to an earlier diagnosis of deficiency. Patients were removed if BMI was measured beyond +/- 3 months from 25-hydroxyvitamin D testing or was within the top and bottom percentiles. Patients with cancer occurring before or within 3 months of 25-hydroxyvitamin D testing were eliminated to establish

temporality and reduce impact of behaviour changes or treatment in response to subclinical or previous disease. The cohort design is shown in Figure 1.

2.6 Statistical analysis

Patient characteristics were tabulated according to quartiles of serum 25-hydroxyvitamin D concentration. Linear trends for individual-level data were evaluated using linear and logistic regression. Linear trends for CDA-level SES covariates were evaluated using Poisson regression accounting for clustering of patients by CDA, and variance was calculated using a sandwich estimator.

The relationship between serum 25-hydroxyvitamin D quartile and incident cancer was evaluated using Cox proportional hazard models, with time from 25-hydroxyvitamin D testing to date of cancer diagnosis or censoring (December 26 2016) as follow-up time. For analyses of major and specific cancers, including 'other' cancer, those without the outcome of interest also included those without a diagnosis of any other cancer. We adjusted for age, sex, BMI, mean hours of daylight during month of testing, and CDA-level SES covariates in different models. Models adjusted for CDA-level SES covariates accounted for clustering of patients by CDA, and variance was calculated using a sandwich estimator. Because we examined 8 separate cancer outcomes, a Bonferroni correction (0.05 ÷ 8) was applied to reduce the nominal significance threshold of $p < 0.05$ to $p < 0.00625$ in order to minimize type I error. We tested the proportional hazards assumption for each variable by inserting time dependent covariates (e.g. 25-hydroxyvitamin D quartile * log(time)) into models. If time

dependent covariates reached nominal significance, they were included in all outcome analyses for a given model.[19]

For 25-hydroxyvitamin D-cancer associations that reached the Bonferroni-corrected threshold of significance, we explored possible interactions with all covariates using multiplicative terms in Cox models, and evaluated them using the nominal significance threshold. For convenience, associations were stratified by the median value of these covariates.

All data analysis was performed using Statistical Analysis Software (SAS 9.4).

3. Results:

After exclusions (Figure 2), there were 72 171 patients for analysis and 3439 cancer diagnoses (Major = 1719; Breast = 518, Colorectal = 317, Lung = 192, Prostate = 330, and Skin = 362; Other = 1720) over a median of 5.9 years of follow-up. Cancer diagnoses occurred after a median of 3.0 years. There were 2849 CDAs.

Mean age and daylight hours during month of testing significantly increased with serum 25-hydroxyvitamin D quartile whereas proportion of men and mean BMI significantly decreased. Among CDA-level SES covariates, mean proportions of Canadian citizens, those with post-secondary level education, employed individuals, and the median total household income significantly increased with 25-hydroxyvitamin D quartile whereas the mean proportion of recent immigrants, visible minorities, and those using non-official languages at home significantly decreased. (Table 1) The proportion of all cancer cases increased significantly across 25-hydroxyvitamin D quartile (Table 1; p for trends < 0.01).

Three Cox proportional hazards models were used to further evaluate the association of 25-hydroxyvitamin D and cancer risk: model 1: adjusted for age, model 2: model 1 adjusted for sex, BMI, mean daylight hours during month of testing, and model 3: model 2 adjusted for CDA-level SES covariates. Vitamin D quartile met the assumption of proportional hazards (i.e. no significant interaction with time) in every model, however several covariates did not and were therefore modeled using time dependent covariates in each model.

After adjusting for age, associations with any cancer, major cancer, breast cancer and skin cancer exceeded the threshold for Bonferroni significance (p for trend < 0.00625). Further adjustment for sex, BMI, mean daily hours of daylight during month of testing resulted in the association with breast cancer and other cancer becoming non-significant at the Bonferroni threshold (Table 2). Additional adjustment for CDA-level SES covariates resulted in only any cancer, major cancer, and skin cancer remaining significant at the Bonferroni threshold. Importantly, the association with major cancer was no longer significant after removing cases of skin cancer (p for trend = 0.15), confirming that this association was being driven by the association with skin cancer, which was the strongest observed. Compared to the bottom quartile of serum 25-hydroxyvitamin D, participants in the top quartile had a 2.56X greater risk of skin cancer (including melanoma; n=58) after adjusting for covariates.

We observed four nominally significant interactions between 25-hydroxyvitamin D quartile and CDA-level SES covariates on skin cancer risk (Table 3). For an increase in the CDA-level proportion of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking official languages (English or French) at home, the

individual-level association between 25-hydroxyvitamin D and skin cancer risk was stronger.

4. Discussion:

In a community population of patients under investigation for cardiovascular disease, higher serum 25-hydroxyvitamin D was associated with an increased risk of developing skin cancer. This association became stronger as the CDA-level proportion of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking official (English or French) languages at home increased. Associations with prostate and other cancers were weak and may have been due to chance.

Vitamin D, which can be synthesized in the skin from 7-dehydrocholesterol or obtained through diet, undergoes two hydroxylations to the biologically active 1,25-dihydroxyvitamin D (calcitriol) [20]. 1,25-dihydroxyvitamin D binds to vitamin D receptors (VDRs) on target tissues, causing increased update of calcium and phosphate from the small intestine, and increased calcium mobilization from bone via enhanced osteoclast activity. [20] However, 25-hydroxyvitamin D reflects an individual's true vitamin D status from both endogenous and exogenous sources because of its long half-life (approximately 2-3 weeks vs 4-6 hours of 1,25-dihydroxyvitamin D), its high concentration (1000 X greater than 1,25-dihydroxyvitamin D) and its resistance to metabolic changes [20].

In animal experiments, 1,25 dihydroxyvitamin D has important cellular effects that may decrease the risk of cancer or slow its progression. [10] These have been cited as evidence that associations between serum 25-hydroxyvitamin D and cancer risk identified in epidemiologic studies represent causal relationships. However many of

these studies are susceptible to unmeasured or residual confounding by factors associated with vitamin D deficiency and increased cancer risk (e.g. adiposity, low SES). They are also susceptible to reverse causality, particularly because low serum 25-hydroxyvitamin D may in part be a marker of ill health.[21, 22] This could result in individuals with subclinical cancer or other conditions being vitamin D deficient. As genetic variants that modestly reduce 25-hydroxyvitamin D are, for the most part, not associated with an increased risk of cancer in Mendelian randomization studies [23-25], this tends to support this hypothesis. However in randomized controlled trials, vitamin D supplementation slightly reduces cancer mortality. [26, 27] For example in the VITAL trial, 2000 IU/day supplementation of vitamin D3 significantly reduced the risk of metastatic or fatal cancer compared to placebo, and this effect was stronger among individuals who had a normal (< 25 kg/m^2) BMI [28]. Taken together, while the relationship between vitamin D status and cancer incidence may in part be due to confounding and reverse causality, vitamin D status may also be causally related to cancer mortality.

There is a well-established relationship between vitamin D status and sun exposure [29], and sun exposure is the most important risk factor for melanoma and non-melanoma skin cancer – particularly among individuals with a light skin tone. [30] As expected, serum 25-hydroxyvitamin D concentration was associated with a higher risk of skin cancer in a recent meta analyses of prospective cohort studies. [31] And while Mendelian randomization studies suggest that this likely does not represent a causal relationship [32, 33], serum 25-hydroxyvitamin D may still be useful as a skin

cancer risk marker because its concentration is related to sun exposure. However its concentration is also related to supplement use.

In our study, we found a positive association between serum 25-hydroxyvitamin D and risk of skin cancer which was consistent over time but stronger among individuals who resided in CDAs with a higher proportion of non-citizens, recent immigrants, visible minorities, and those who did not speak an official language at home. This may be because individuals living in these communities are less likely to take vitamin D supplements [34, 35], which would make their serum 25-hydroxyvitamin D concentration more representative of sun exposure than supplementation – resulting in a stronger overall association with skin cancer risk.

This study has some strengths. First, we used available secondary data to assemble a retrospective cohort study of a large community population while making several restrictions and exclusions to minimize bias. Second, while this population included only patients who received an ECG, any patients that had a 25-hydroxyvitamin D measured were eligible for inclusion. During the testing period, our laboratories experienced a very high volume of serum 25-hydroxyvitamin D testing – likely because of substantial interest in vitamin D t at the time. Third, we adjusted for several potentially important confounders, including mean daylight hours during month of testing, BMI, and community-level measures of SES.

This study also has some limitations. First, because we used secondary data, we had limited information on potential confounders and could not capture them at the same time as measurement of 25-hydroxyvitamin D. However for BMI, we dealt with this limitation by setting a +/- 3 month window for inclusion around measurement of 25-

hydroxyvitamin D. Second, as our study was observational, we could not determine whether the 25-hydroxyvitamin D-cancer relationship was causal. However this was not an objective of our study nor was it even achievable. Third, while we did not include participants without a BMI and ECG, we felt obtaining BMI was critical for reducing bias – even if it was based on self-report. However because ECGs are used to identify the presence of cardiovascular disease, our population may be at an elevated risk for both cardiovascular disease and cancer because many of the risk factors for cardiovascular disease are shared risk factors for cancer [36].

5. Conclusion:

Higher vitamin D status was associated with a greater risk of skin cancer in a large community population under investigation for cardiovascular disease. This association is likely due to sun exposure and may be modified by community variation in supplementation rates.

Table 1: Patient characteristics by quartile of serum 25-hydroxyvitamin D.

	Serum 25-hydroxyvitamin D quartile				
	Q1 (10 – 44 nmol/L)	Q2 (45 – 64 nmol/L)	Q3 (65 – 87 nmol/L)	Q4 (88 – 658 nmol/L)	P for linear trend
N	18053	18022	18056	18040	
Age, y, mean (sd)	48 (15)	51 (15)	54 (15)	58 (15)	< 0.01
% Male, (n)	56.3 (10166)	51.4 (9261)	45.4 (8199)	40.1 (7218)	< 0.01
BMI, kg/m^2, mean (sd)	27.4 (5.3)	27.3 (5.1)	26.9 (5.1)	26.1 (4.8)	< 0.01
Daylight hours during month of testing, mean (sd)	11.9 (2.9)	12.2 (2.9)	12.4 (2.9)	12.5 (2.9)	< 0.01
Census dissemination area-level measures of socioeconomic status (SES)					
Canadian citizens, mean % (sd)	88.1 (10.1)	89.8 (8.9)	91.0 (8.1)	91.9 (7.5)	< 0.01
Recent immigrants, mean % (sd)	33.4 (16.8)	30.0 (16.2)	27.4 (15.1)	25.2 (13.8)	< 0.01
Visible minorities, mean % (sd)	38.5 (26.0)	32.4 (24.9)	27.8 (22.6)	23.6 (19.9)	< 0.01
Do not speak official language (English or French) at home, mean % (sd)	39.4 (22.1)	34.5 (21.1)	31.0 (19.6)	27.7 (17.3)	< 0.01
Aboriginal identity, mean % (sd)	2.5 (5.7)	2.2 (4.4)	2.1 (4.2)	2.0 (4.1)	< 0.01
Postsecondary education, mean % (sd)	54.8 (14.0)	56.9 (13.6)	58.7 (13.5)	60.2 (12.9)	< 0.01
Employed, mean % (sd)	93.3 (5.7)	93.5 (5.8)	93.7 (5.5)	93.9 (5.2)	< 0.01
Household total income, $, median (sd)	85434 (32435)	90779 (35727)	944486 (38827)	96986 (41439)	< 0.01
					< 0.01
All cancer (n)	615	737	928	1159	< 0.01
Major cancer (n)	264	381	470	604	< 0.01
Breast cancer (n)	83	125	127	183	< 0.01
Colorectal cancer (n)	57	68	91	101	< 0.01
Lung cancer (n)	30	44	57	61	< 0.01
Prostate cancer (n)	63	80	77	110	< 0.01
Skin cancer (n)	31	64	118	149	< 0.01
Other cancer (n)	351	356	458	555	< 0.01

Caption: There were 2851 census dissemination areas.

Table 2 – Cox proportional hazard regression of serum 25-hydroxyvitamin D quartile and risk of cancer

	Q1	Q2	Q3	Q4	P for linear trend
Any cancer					
Model 1	1.00	1.04 (0.93-1.16)	1.13 (1.02-1.26)	1.21 (1.09-1.33)	< 0.01
Model 2	1.00	1.04 (0.94-1.16)	1.14 (1.03-1.27)	1.24 (1.12-1.37)	< 0.01
Model 3	1.00	1.04 (0.93-1.16)	1.12 (1.01-1.26)	1.21 (1.08-1.34)	< 0.01
Major cancer					
Model 1	1.00	1.24 (1.06-1.45)	1.32 (1.13-1.54)	1.43 (1.24-1.66)	< 0.01
Model 2	1.00	1.24 (1.06-1.45)	1.32 (1.14-1.54)	1.46 (1.26-1.69)	< 0.01
Model 3	1.00	1.21 (1.03-1.43)	1.24 (1.05-1.47)	1.34 (1.14-1.58)	< 0.01
Breast cancer					
Model 1	1.00	1.36 (1.03-1.80)	1.26 (0.95-1.66)	1.63 (1.25--2.12)	< 0.01
Model 2	1.00	1.23 (0.93-1.62)	1.02 (0.77-1.35)	1.23 (0.94-1.61)	0.30
Model 3	1.00	1.22 (0.92-1.63)	0.95 (0.71-1.27)	1.13 (0.85-1.51)	0.83
Colorectal cancer					
Model 1	1.00	1.01 (0.71-1.44)	1.14 (0.81-1.59)	1.04 (0.75-1.45)	0.72
Model 2	1.00	1.03 (0.72-1.46)	1.18 (0.84-1.65)	1.12 (0.80-1.56)	0.41
Model 3	1.00	1.03 (0.72-1.48)	1.20 (0.84-1.71)	1.16 (0.78-1.59)	0.44
Lung cancer					
Model 1	1.00	1.20 (0.75-1.91)	1.26 (0.81-1.96)	1.06 (0.68--1.65)	0.93
Model 2	1.00	1.20 (0.75-1.90)	1.25 (0.80-1.96)	1.06 (0.68-1.66)	0.91
Model 3	1.00	1.23 (0.76-1.98)	1.27 (0.80-2.00)	1.02 (0.65-1.61)	0.92
Prostate cancer					
Model 1	1.00	1.10 (0.79-1.53)	0.92 (0.65-1.28)	1.11 (0.81-1.52)	0.74
Model 2	1.00	1.20 (0.86-1.67)	1.13 (0.81-1.58)	1.57 (1.14-2.16)	0.01
Model 3	1.00	1.13 (0.80-1.59)	1.08 (0.76-1.52)	1.42 (1.02-1.97)	0.05
Skin cancer					
Model 1	1.00	1.75 (1.14-2.68)	2.72 (1.83-4.05)	2.84 (1.92-4.21)	< 0.01
Model 2	1.00	1.78 (1.16-2.73)	2.82 (1.89-4.20)	3.04 (2.05-4.51)	< 0.01
Model 3	1.00	1.66 (1.08-2.58)	2.42 (1.61-3.65)	2.56 (1.70-3.86)	< 0.01
Other cancer					
Model 1	1.00	0.89 (0.77-1.03)	1.00 (0.87-1.15)	1.04 (0.91-1.20)	0.20
Model 2	1.00	0.89 (0.77-1.03)	1.01 (0.88-1.16)	1.08 (0.94-1.24)	0.07
Model 3	1.00	0.91 (0.78-1.06)	1.04 (0.89-1.20)	1.11 (0.96-1.28)	0.04

Caption: Hazard ratios (HR) are indicated with 95% Confidence intervals in parenthesis. Model 1: adjusted for age and age*log(time) interaction. Model 2: model 1 additionally adjusted for sex, BMI, mean daylight hours during month of testing and log(time) interactions with age, sex, BMI and mean daylight hours during month of testing. Model 3: model 2 additionally adjusted for CDA-level proportion of Canadian citizens, recent immigrants, visible (non-white) minorities, those indicating Aboriginal identity, those not speaking official languages (English or French) at home, those with postsecondary education, those currently employed, the CDA median household income and log(time) interactions with CDA-level proportion of Canadian citizens, those indicating Aboriginal identity, and those with postsecondary education.

Table 3: Association of 25-hydroxyvitamin D with skin cancer risk stratified by median covariate values

	Q1	Q2	Q3	Q4	P for linear trend	P for interaction
CDA-level proportion of Canadian citizens						
≥ 92.7%	1.00	1.86 (1.05-3.31)	2.33 (1.37-3.97)	2.21 (1.28-3.81)	< 0.01	0.03
< 92.7%	1.00	1.36 (0.69-2.67)	2.55 (1.34-4.85)	3.16 (1.71-5.84)	< 0.01	
CDA-level proportion of recent immigrants						
≥ 26.7%	1.00	1.55 (0.72-3.31)	2.99 (1.51-5.95)	3.24 (1.65-6.34)	< 0.01	0.04
< 26.7%	1.00	1.66 (0.97-2.85)	2.08 (1.25-3.45)	2.18 (1.31-3.64)	< 0.01	
CDA-level proportion of visible (non-white) minorities						
≥ 24.3%	1.00	1.99 (0.95-4.15)	2.83 (1.39-5.73)	3.06 (1.53-6.09)	< 0.01	0.03
< 24.3%	1.00	1.47 (0.86-2.54)	2.16 (1.31-3.56)	2.26 (1.37-3.73)	< 0.01	
CDA-level proportion of non-official language (English or French) speakers at home						
≥ 29.2%	1.00	1.39 (0.67-2.90)	2.58 (1.31-5.07)	2.75 (1.41-5.35)	< 0.01	0.02
< 29.2%	1.00	1.79 (1.03-3.10)	2.33 (1.40-3.90)	2.47 (1.47-4.14)	< 0.01	

Caption: Hazard ratios (HR) are indicated with 95% Confidence intervals in parenthesis. All models are adjusted for age, sex, BMI, mean daylight hours during month of testing, CDA-level proportions of Canadian citizens, recent immigrants, visible (non-white) minorities, those not speaking official languages (English or French) at home, those indicating Aboriginal identity, those with postsecondary education, those currently employed, and the CDA median household income unless stratified by that variable. They were also adjusted for log(time) interactions with CDA-level proportion of Canadian citizens, those indicating Aboriginal identity, and those with postsecondary education. Only nominally significant interactions are shown.

Figure 1 – Cohort design

Caption: Note – Serum 25-hydroxyvitamin D measurements were made between 2009 and 2013. CDA-level SES covariates were measured in 2011.

Figure 2 – Patient exclusions leading to analysis cohort

References

1 Medicine). IIo. *Dietary Reference Intakes for Calcium and Vitamin D* . . Washington, DC, USA.: The National Academies Press. .

2 Wacker M, Holick MF. Sunlight and Vitamin D: A global perspective for health. *Dermatoendocrinol* 2013;**5**:51-108.

3 Janz T, Pearson C. Vitamin D Blood Levels of Canadians. Ottawa, Canada.: Statistics Canada 2013.

4 Yin L, Ordóñez-Mena JM, Chen T*, et al.* Circulating 25-hydroxyvitamin D serum concentration and total cancer incidence and mortality: a systematic review and meta-analysis. *Prev Med* 2013;**57**:753-64.

5 Ma Y, Zhang P, Wang F*, et al.* Association between vitamin D and risk of colorectal cancer: a systematic review of prospective studies. *J Clin Oncol* 2011;**29**:3775-82.

6 Dunn JA, Jefferson K, MacDonald D*, et al.* Low serum 25-hydroxyvitamin D is associated with increased bladder cancer risk: A systematic review and evidence of a potential mechanism. *J Steroid Biochem Mol Biol* 2019;**188**:134-40.

7 Mäkitie A, Tuokkola I, Laurell G*, et al.* Vitamin D in Head and Neck Cancer: a Systematic Review. *Curr Oncol Rep* 2020;**23**:5.

8 Zhang Y, Jiang X, Li X*, et al.* Serum Vitamin D Levels and Risk of Liver Cancer: A Systematic Review and Dose-Response Meta-Analysis of Cohort Studies. *Nutr Cancer* 2020:1-9.

9 Chowdhury R, Kunutsor S, Vitezova A*, et al.* Vitamin D and risk of cause specific death: systematic review and meta-analysis of observational cohort and randomised intervention studies. *BMJ* 2014;**348**:g1903.

10 Bandera Merchan B, Morcillo S, Martin-Nuñez G*, et al.* The role of vitamin D and VDR in carcinogenesis: Through epidemiology and basic sciences. *J Steroid Biochem Mol Biol* 2017;**167**:203-18.

11 Vanlint S. Vitamin D and obesity. *Nutrients* 2013;**5**:949-56.

12 Bhaskaran K, Douglas I, Forbes H*, et al.* Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet* 2014;**384**:755-65.

13 de Koning L, Henne D, Woods P*, et al.* Sociodemographic correlates of 25-hydroxyvitamin D test utilization in Calgary, Alberta. *BMC Health Serv Res* 2014;**14**:339.

14 Merletti F, Galassi C, Spadea T. The socioeconomic determinants of cancer. *Environ Health* 2011;**10 Suppl 1**:S7.

15 Bikos K. Calgary, Alberta, Canada — Sunrise, Sunset, and Daylength.

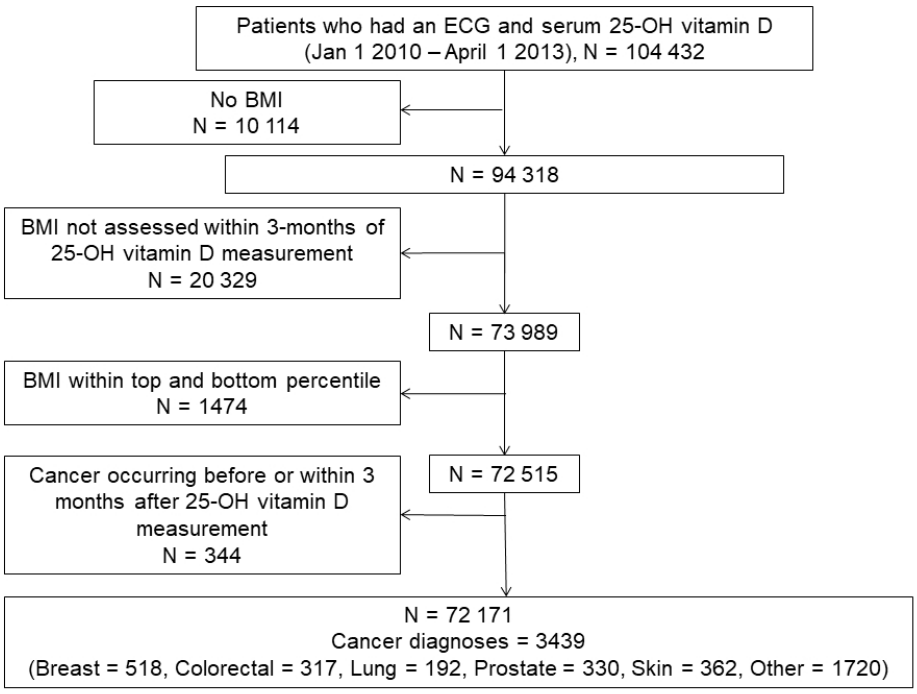
16 de Koning L, Al-Turkmani MR, Berg AH*, et al.* Variation in clinical vitamin D status by DiaSorin Liaison and LC-MS/MS in the presence of elevated 25-OH vitamin D2. *Clin Chim Acta* 2013;**415**:54-8.

17 Society CC. Cancer statistics at a glance. 2021.

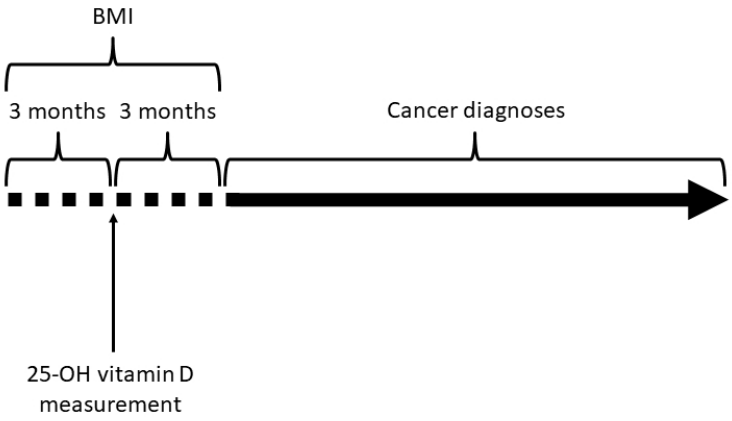
18 Cancer IAfRo. The Cancer Dictionary. *Cancer Mortality Database*. Lyon, France.: World Health Organizatio 2010.

19 Bradburn MJ, Clark TG, Love SB*, et al.* Survival analysis Part III: multivariate data analysis -- choosing a model and assessing its adequacy and fit. *Br J Cancer* 2003;**89**:605-11.
20 Fraser WD. Bone and Mineral Metabolism. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textboook of Clinical Chemistry and Molecular Diagnostics*. St. Louis, MO, USA.: Elsevier 2018:1422-91.
21 Autier P, Boniol M, Pizot C*, et al.* Vitamin D status and ill health: a systematic review. *Lancet Diabetes Endocrinol* 2014;**2**:76-89.
22 Autier P, Mullie P, Macacu A*, et al.* Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials. *Lancet Diabetes Endocrinol* 2017;**5**:986-1004.
23 Dimitrakopoulou VI, Tsilidis KK, Haycock PC*, et al.* Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. *BMJ* 2017;**359**:j4761.
24 Aspelund T, Grübler MR, Smith AV*, et al.* Effect of Genetically Low 25-Hydroxyvitamin D on Mortality Risk: Mendelian Randomization Analysis in 3 Large European Cohorts. *Nutrients* 2019;**11**.
25 Ong JS, Gharahkhani P, An J*, et al.* Vitamin D and overall cancer risk and cancer mortality: a Mendelian randomization study. *Hum Mol Genet* 2018;**27**:4315-22.
26 Keum N, Lee DH, Greenwood DC*, et al.* Vitamin D supplementation and total cancer incidence and mortality: a meta-analysis of randomized controlled trials. *Ann Oncol* 2019;**30**:733-43.
27 Association between vitamin D supplementation and mortality: systematic review and meta-analysis. *BMJ* 2020;**370**:m2329.
28 Chandler PD, Chen WY, Ajala ON*, et al.* Effect of Vitamin D3 Supplements on Development of Advanced Cancer: A Secondary Analysis of the VITAL Randomized Clinical Trial. *JAMA Netw Open* 2020;**3**:e2025850.
29 Jager N, Schöpe J, Wagenpfeil S*, et al.* The Impact of UV-dose, Body Surface Area Exposed and Other Factors on Cutaneous Vitamin D Synthesis Measured as Serum 25(OH)D Concentration: Systematic Review and Meta-analysis. *Anticancer Res* 2018;**38**:1165-71.
30 American Cancer Society. Cancer Facts & Figures 2021. . Atlanta, Georgia, USA. 2021.
31 Mahamat-Saleh Y, Aune D, Schlesinger S. 25-Hydroxyvitamin D status, vitamin D intake, and skin cancer risk: a systematic review and dose-response meta-analysis of prospective studies. *Sci Rep* 2020;**10**:13151.
32 Winsløw UC, Nordestgaard BG, Afzal S. High plasma 25-hydroxyvitamin D and high risk of nonmelanoma skin cancer: a Mendelian randomization study of 97 849 individuals. *Br J Dermatol* 2018;**178**:1388-95.
33 Liyanage UE, Law MH, Barrett JH*, et al.* Is there a causal relationship between vitamin D and melanoma risk? A Mendelian randomization study. *Br J Dermatol* 2020;**182**:97-103.
34 Gordon NP, Caan BJ, Asgari MM. Variation in vitamin D supplementation among adults in a multi-race/ethnic health plan population, 2008. *Nutr J* 2012;**11**:104.

35 Moffat T, Sellen D, Wilson W*, et al.* Comparison of infant vitamin D supplement use among Canadian-born, immigrant, and refugee mothers. *J Transcult Nurs* 2015;**26**:261-9.
36 Lau ES, Paniagua SM, Liu E*, et al.* Cardiovascular Risk Factors are Associated with Future Cancer. *JACC CardioOncol* 2021;**3**:48-58.



254x190mm (96 x 96 DPI)



254x190mm (96 x 96 DPI)

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and time frame within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Page 4 " "
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported			Page 5-6
Objectives	3	State specific objectives, including any prespecified hypotheses			Page 6
Methods					
Study Design	4	Present key elements of study design early in the paper			Page 6-7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection			Page 6-7

Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Page 6-8 “ Figure 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Page 6-8
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group			Page 6-8

Bias	9	Describe any efforts to address potential sources of bias			Page 8-9
Study size	10	Explain how the study size was arrived at			Page 8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why			Page 9
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses			Page 9-10
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	Page 7-8

				RECORD 12.2: Authors should provide information on the data cleaning methods used in this study.	
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Page 6-10
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Page 10
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)			Page 10, 16
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure			Table 1, Page 10

		category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			Table 1, Table 2, Table 3
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			Page 11, Table 3
Discussion					
Key results	18	Summarise key results with reference to study objectives			Page 11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibilities over time, as they pertain to the study being reported.	Page 14-15
Interpretation	20	Give a cautious overall interpretation of results considering objectives,			Page 15

		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results			Page 14
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based			Page 2
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Page 2

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

BMJ Open

Serum 25-hydroxyvitamin D and risk of cancer in a large community population under investigation for cardiovascular disease - retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056543.R1
Article Type:	Original research
Date Submitted by the Author:	03-Dec-2021
Complete List of Authors:	Yang, Jaeun; University of Alberta Faculty of Medicine & Dentistry, Department of Anesthesiology & Pain Medicine Dong, Yuan; University of Calgary Cumming School of Medicine, Community Health Sciences Naugler, Christopher; University of Calgary Cumming School of Medicine, Pathology and Laboratory Medicine de Koning, Lawrence; University of Calgary Cumming School of Medicine, Pathology and Laboratory Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Oncology
Keywords:	Epidemiology < ONCOLOGY, Clinical chemistry < PATHOLOGY, STATISTICS & RESEARCH METHODS

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

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

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

Serum 25-hydroxyvitamin D and risk of cancer in a large community population under investigation for cardiovascular disease – a retrospective cohort study

Jaeun Yang[1], Yuan Dong[2], Christopher T Naugler[3,4] and Lawrence de Koning[3,4]*

[1] Department of Anesthesiology & Pain Medicine, University of Alberta

2-150 Clinical Sciences Building (CSB) 8440 112 St NW, Edmonton, Alberta, Canada T6G 2G3.

[2] Department of Community Health Sciences, Cumming School of Medicine, Teaching, Research and Wellness Building, 5E25-3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6.

[3] Department of Pathology and Laboratory Medicine, Cumming School of Medicine, Health Sciences Centre, Foothills Campus, University of Calgary, 3030 Hospital Drive NW, Calgary, AB, Canada, T2N 4N1

[4] Alberta Precision Laboratories, Diagnostic and Scientific Research Centre, #9 3535 Research Way NW, Calgary, AB, Canada, T2L 2K8

*Address for correspondence:

Lawrence de Koning, PhD DABCC FAACC FCACB

Alberta Children's Hospital, Room B3-724, 28 Oki Drive NW

Calgary, Alberta, Canada

T3B6A8

Tel: 403-955-2277

Fax: 403-955-2321

Email: abldekon@ucalgary.ca

Twitter: **@ldekoning**

Word count: 3444

Tables/Figures: 5

Keywords: Vitamin D, cancer, epidemiology, risk factors

Abstract:

Objectives: It remains unclear whether vitamin D status is related to cancer risk. We examined this relationship using laboratory, administrative and survey data.

Design: Retrospective cohort study

Setting: All care settings within Calgary, Alberta, Canada and surrounding rural communities.

Participants: Patients tested for serum 25-hydroxyvitamin D from 2009-2013 without a past cancer diagnosis but with an electrocardiogram and body mass index (BMI) +/- 3 months from testing were included. Age, sex, mean hours of daylight during month of testing were linked to census dissemination area-level indicators of socioeconomic status (SES) measured in 2011.

Primary and secondary outcome measures: Hospital discharge diagnoses for any cancer, major cancer [colorectal, breast, lung, prostate, skin], and other cancers >3 months from testing from 2009-2016. Cox proportional hazard models were used to examine associations with incident cancer after adjusting for potential confounders. Interactions were tested using multiplicative terms.

Results: Among 72 171 patients, there were 3439 cancer diagnoses over a median of 5.9 years. After adjustment, increasing quartile of serum 25-OH vitamin D was significantly associated with an increased risk of any cancer and major cancer however this was completely driven by an increased risk of skin cancer. (Q4 vs Q1: HR = 2.56, 95% CI: 1.69-3.83, p for linear trend < 0.01). This association was strengthened among individuals residing in communities with higher proportions of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking an official Canadian language (English or French) at home.

Conclusions: Higher vitamin D status was associated with a greater risk of skin cancer in a large community population under investigation for cardiovascular disease. This association was likely due to sun exposure and may be modified by community variation in vitamin D supplementation.

Article summary:

Strengths and limitations of this study:

- We assembled a large retrospective cohort study of community patients by linking patient laboratory, national survey, and hospital administrative data during a time of high public and medical interest in vitamin D deficiency, and a commensurately high testing rate at our laboratory for serum 25-hydroxyvitamin D.
- Linkage allowed us to control for body mass index, a confounder of 25-hydroxyvitamin D-chronic disease relationships, and to explore variation in associations according to community-level socioeconomic factors correlated with vitamin D supplement use.
- As our study was restricted to patients who had received an electrocardiogram, it may not be generalizable to all patients.

1. Introduction:

Vitamin D deficiency is defined as a serum 25-OH (hydroxy) vitamin D concentration < 50 nmol/L [1 2] which causes reduced absorption of dietary calcium and phosphate and increases the risk of rickets in children and osteomalacia in adults. [1] Conversely, vitamin D toxicity is generally regarded to occur above 250 nmol/L– when symptoms of hypercalcemia begin to occur. [3]

Vitamin D deficiency is relatively common – especially in Northern latitudes where people experience less intense solar radiation, spend more time indoors due to cold, and experience prolonged periods of darkness during winter [4]. In Canada, 33% of residents may be vitamin D deficient.[5] As such, there remains significant interest in whether vitamin D deficiency is related to many common diseases – particularly cancer.

Meta-analysis of prospective cohort studies suggest that vitamin D deficiency is associated with an increased risk of multiple types of cancer, including all cancers [6], colorectal cancer [7], bladder cancer [8] head and neck cancer [9], liver cancer [10], and also death due to cancer [11]. These associations have been explained by *in vitro* and *in vivo* by effects of the active form of vitamin D (1,25 OH_2 [dihydroxy] vitamin D), which promotes cellular differentiation, decreases cancer cell growth, stimulates cell death (apoptosis), and reduces angiogenesis.[12] However associations may also be explained by the presence of confounding factors that are associated with vitamin D deficiency but also a higher risk of cancer.

For example, adiposity is a sink for and diluent of serum 25-hydroxyvitamin D [13] as well as a risk factor for several types of cancer [14]. As such, adjustment for some measure of adiposity (e.g. body mass index [BMI]) is generally recognized as essential

to control for bias in epidemiologic studies of serum 25-hydroxyvitamin D concentration and cancer risk. [1] Interestingly, low socioeconomic status (SES), while also a strong a risk factor for vitamin D deficiency [15] and cancer [16], is infrequently controlled for – probably because it is uncommonly measured in epidemiologic studies.

Historic uncertainty in the validity of epidemiologic findings have thus lead the Institute of Medicine (IOM) in the United States to indicate that evidence of a relationship between vitamin D status and non-skeletal chronic diseases does not meet criteria for establishing cause-and-effect. [1] However, vitamin D status could still be a useful and convenient cancer risk marker if its association with cancer risk is independent of other commonly measured factors and is observed in a large population of free-living individuals.

Our objectives were therefore to (i) examine the relationship between serum 25-hydroxyvitamin D (the major circulating form of vitamin D) and risk of cancer in a large community-based population, (ii) adjust for important confounders such as adiposity and socioeconomic status (SES), and (iii) test whether associations are modified by these and other factors.

2. Materials and Methods:

2.1 Ethics statement

This study was approved by the University of Calgary Conjoint Health Review Ethics Board (Ethics ID 25065). Research in this article was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2 *Patient and public involvement statement*

Due to the design of the study and because we did not collect the primary data, we did not involve patients or the public in the design, conduct or reporting of our research.

2.3 Population, primary exposure variable and covariates

We used the Cerner (Kansas City, MO, USA) Millennium laboratory information system (LIS) to identify those who had a serum 25-hydroxyvitamin D result with a test date from December 8 2009 to April 1 2013. This database contained all laboratory results on patients tested in Calgary, Alberta, Canada (population 1.4 million) as well as surrounding rural communities. During the time of this study, 25-hydroxyvitamin D testing was available to any ordering physician for any reason – and high test volumes (~16k / month) reflected a strong public and medical interest in vitamin D deficiency. For these patients, we retained data only for those who had an ECG because these patients had a self-reported height and weight entered into the LIS to calculate body mass index (BMI; weight in kg / (height in meters)^2). We then extracted age, sex, and the provincial health care number (PHN) to link to hospital administrative outcome data and postal codes. As vitamin D status is related to sun exposure, we also calculated mean hours of daylight during month of 25-hydroxyvitamin D testing for each person based on publicly available data [17]. This was done so that we could adjust for short-term variation in 25-hydroxyvitamin D related to seasonal changes in sun exposure at the time of testing. All serum 25-hydroxyvitamin D tests were performed on the DiaSorin (Saluggia, Italy)

Liaison total 25-hydroxyvitamin D automated immunoassay platform, which predominantly detects 25-hydroxyvitamin D3.[18] The performance of this assay was validated using guidelines from the Clinical Laboratory Standards Institute (CLSI; Annapolis Junction, MD, USA). Total imprecision was approximately 7%, and results from external proficiency survey samples from the Vitamin D External Quality Assessment Scheme (DEQAS) all fell within total allowable error intervals from peer group means (< 20 nmol/L: +/-5 nmol/L; ≥ 20 nmol/L: +/- 15 nmol/L).

Census-dissemination area (CDA)-level SES covariates were obtained from the 2011 Canadian National Household Survey (NHS) after postal code to CDA conversion. We extracted proportion of CDA residents who were (i) Canadian citizens, (ii) recent immigrants, (iii) visible (non-white) minorities, (iv) those speaking languages other than English and French (official languages of Canada) at home, (v) those having postsecondary education, (vi) those currently employed, and the CDA (vii) median household income.

2.4 Outcomes

In-hospital discharge diagnosis and dates were obtained from the Discharge Abstract Database (DAD) and National Ambulatory Care Reporting System (NACRS) via Alberta Health Services. Incident cancers were defined as the first and most responsible (primary) diagnosis of any cancer (ICD10 codes: C00.x-C97.x) if the patient was discharged alive or died in hospital. We further subdivided 'any cancer' into 'major cancer' [19] (breast [C50.x], colorectal [C18.x-C21.x], lung [C33.x-C34.x], prostate [C61.x] and skin [melanoma: C43.x, non-melanoma: C44.x, C46.x] [20]), specific cancers (breast,

colorectal, lung, prostate and skin) and 'other cancer' – defined as 'any cancer' other than 'major cancer'. At the time of analysis, outcome data was available until December 31 2016.

2.5 Data cleaning

We kept only the first measurement of 25-hydroxyvitamin D to capture historic vitamin D status – which is more likely associated with cancer risk than vitamin D status after supplementation in response to an earlier diagnosis of deficiency. Patients were removed if BMI was measured beyond +/- 3 months from 25-hydroxyvitamin D testing or was within the top and bottom percentiles. Patients with cancer occurring before or within 3 months of 25-hydroxyvitamin D testing were eliminated to establish temporality and to reduce impact of behaviour changes or treatment (including supplementation) in response to subclinical or previous disease. The cohort design is shown in Figure 1.

2.6 Statistical analysis

Patient characteristics were tabulated according to quartiles of serum 25-hydroxyvitamin D concentration. Linear trends for individual-level data were evaluated using linear and logistic regression. Linear trends for CDA-level SES covariates were evaluated using Poisson regression accounting for clustering of patients by CDA, and variance was calculated using a sandwich estimator.

The relationship between serum 25-hydroxyvitamin D quartile and incident cancer was evaluated using Cox proportional hazard models, with time from 25-hydroxyvitamin D testing to date of cancer diagnosis or censoring (December 26 2016)

as follow-up time. For analyses of major and specific cancers, including 'other' cancer, those without the outcome of interest also included those without a diagnosis of any other cancer. We adjusted for age, sex, BMI, mean hours of daylight during month of testing, and CDA-level SES covariates in different models. Models adjusted for CDA-level SES covariates accounted for clustering of patients by CDA, and variance was calculated using a sandwich estimator. Because we examined 8 separate cancer outcomes, a Bonferroni correction (0.05 ÷ 8) was applied to reduce the nominal significance threshold of $p < 0.05$ to $p < 0.00625$ in order to minimize type I error. We tested the proportional hazards assumption for each variable by inserting time dependent covariates (e.g. 25-hydroxyvitamin D quartile * log(time)) into models. If time dependent covariates reached nominal significance, they were included in all outcome analyses for a given model.[21]

For 25-hydroxyvitamin D-cancer associations that reached the Bonferroni-corrected threshold of significance, we explored possible interactions with all covariates using multiplicative terms in Cox models, and evaluated them using the nominal significance threshold. For convenience, associations were stratified by the median value of these covariates. Finally, we performed a sensitivity analyses where we excluded participants with a 25-OH vitamin D concentration of 100 nmol/L or greater, as these individuals may be more likely to be taking vitamin D supplements. [22]

All data analysis was performed using Statistical Analysis Software (SAS 9.4).

3. Results:

After exclusions (Figure 2), there were 72 171 patients for analysis and 3439 cancer diagnoses (Major = 1719; Breast = 518, Colorectal = 317, Lung = 192, Prostate = 330, and Skin = 362 [melanoma = 58, non-melanoma =304]; Other = 1720) over a median of 5.9 years of follow-up. Cancer diagnoses occurred after a median of 3.0 years. There were 2849 CDAs. Approximately 31% of patients were vitamin D deficient (i.e. serum 25-hydroxyvitamin D <50 nmol/L).

Mean age and daylight hours during month of testing significantly increased with serum 25-hydroxyvitamin D quartile whereas proportion of men and mean BMI significantly decreased. We also found that serum 25-hydroxyvitamin D was lowest when tested in the winter (median = 61 nmol/L) vs the summer (median = 69 nmol/L). Among CDA-level SES covariates, mean proportions of Canadian citizens, those with post-secondary level education, employed individuals, and the median total household income significantly increased with 25-hydroxyvitamin D quartile whereas the mean proportion of recent immigrants, visible minorities, and those using non-official languages at home significantly decreased. (Table 1) The proportion of all cancer cases increased significantly across 25-hydroxyvitamin D quartile (Table 1; p for trends < 0.001).

Three Cox proportional hazards models were used to further evaluate the association of 25-hydroxyvitamin D and cancer risk: model 1: adjusted for age, model 2: model 1 adjusted for sex, BMI, mean daylight hours during month of testing, and model 3: model 2 adjusted for CDA-level SES covariates. Vitamin D quartile met the assumption of proportional hazards (i.e. no significant interaction with time) in every

model, however several covariates did not and were therefore modeled using time dependent covariates in each model.

After adjusting for age, associations with any cancer, major cancer, breast cancer and skin cancer exceeded the threshold for Bonferroni significance (p for trend < 0.00625). Further adjustment for sex, BMI, mean daily hours of daylight during month of testing resulted in the association with breast cancer and other cancer becoming non-significant at the Bonferroni threshold (Table 2). Additional adjustment for CDA-level SES covariates resulted in only any cancer, major cancer, and skin cancer remaining significant at the Bonferroni threshold. Importantly, the association with major cancer was no longer significant after removing cases of skin cancer (p for trend = 0.15), confirming that this association was being driven by the association with skin cancer, which was the strongest observed. Compared to the bottom quartile of serum 25-hydroxyvitamin D, participants in the top quartile had a 2.56X greater risk of skin cancer after adjusting for covariates. Analysis by type of skin cancer yielded a similar association for non-melanoma, but the association for melanoma was not significant at either threshold of significance – perhaps due to a small number of melanomas in our study (n=58; results not shown).

We observed four nominally significant interactions between 25-hydroxyvitamin D quartile and CDA-level SES covariates on skin cancer risk (Table 3). For an increase in the CDA-level proportion of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking official languages (English or French) at home, the individual-level association between 25-hydroxyvitamin D and skin cancer risk was stronger.

Removal of 11 154 participants with 25-hydroxyvitamin D concentrations of at least 100 nmol/L deleted large numbers of cancer cases from the 4th quartile of serum 25-hydroxyvitamin vitamin D (any cancer: -66%, major cancer: -66%, breast cancer: -59%, colorectal cancer: -70%, lung cancer: -74%, prostate cancer: -67%, skin cancer: -66%, other cancer: -66%). While this did not change our overall findings, associations with any and major cancer were no longer significant at the Bonferroni threshold, and associations with prostate and other cancer were no longer borderline-significant or significant at the nominal threshold. Only the association with skin cancer remained significant at the Bonferroni threshold, and was in fact strengthened (HR per quartile change 1.33; 95% confidence interval: 1.17 to 1.50; p for trend < 0.001). Repeating the sensitivity analysis using serum 25-hydroxyvitamin D as a continuous variable yielded identical results.

4. Discussion:

In a community population of patients under investigation for cardiovascular disease, higher serum 25-hydroxyvitamin D was associated with an increased risk of developing skin cancer. This association became stronger as the CDA-level proportion of non-citizens, recent immigrants, visible (non-white) minorities, and those not speaking official (English or French) languages at home increased. Associations with prostate and other cancers were weak and may have been due to chance.

Vitamin D (vitamin D2 + vitamin D3), which can be synthesized in the skin (vitamin D3) from 7-dehydrocholesterol and UV radiation or obtained through diet (vitamin D2 or D3), undergoes two hydroxylations to the biologically active 1,25-dihydroxyvitamin D (D2 + D3; calcitriol) [23]. 1,25-dihydroxyvitamin D binds to vitamin D

receptors (VDRs) on target tissues, causing increased update of calcium and phosphate from the small intestine, and increased calcium mobilization from bone via enhanced osteoclast activity. [23] However, 25-hydroxyvitamin D (D2 + D3) reflects an individual's true vitamin D status from both endogenous and exogenous sources because of its long half-life (approximately 2-3 weeks vs 4-6 hours for 1,25-dihydroxyvitamin D), its high concentration (1000 X greater than 1,25-dihydroxyvitamin D) and its resistance to metabolic changes [23].

In animal experiments, 1,25 dihydroxyvitamin D has important cellular effects that may decrease the risk of cancer or slow its progression. [12] These have been cited as evidence that associations between serum 25-hydroxyvitamin D and cancer risk identified in epidemiologic studies represent causal relationships. However epidemiologic studies are susceptible to unmeasured or residual confounding by factors associated with vitamin D deficiency and increased cancer risk (e.g. adiposity, low SES). They are also susceptible to reverse causality, particularly because low serum 25-hydroxyvitamin D may in part be a marker of ill health. [24 25] This could cause individuals with subclinical cancer or other conditions to become vitamin D deficient. As genetic variants that modestly reduce 25-hydroxyvitamin D are, for the most part, not associated with an increased risk of cancer in Mendelian randomization studies [26-28], this tends to support this hypothesis. However in randomized controlled trials, vitamin D supplementation slightly reduces cancer mortality. [29 30] For example in the VITAL trial, 2000 IU/day supplementation of vitamin D3 significantly reduced the risk of metastatic or fatal cancer compared to placebo, and this effect was stronger among individuals who had a normal BMI (i.e. < 25 kg/m^2) [31]. Taken together, while the relationship

between vitamin D status and cancer incidence may in part be due to confounding and reverse causality, vitamin D status may be causally related to mortality risk from cancer.

There is a well-established relationship between vitamin D status and sun exposure [32]. For example, serum 25-hydroxyvitamin D concentration may rise above 100 nmol/L among individuals who perform extended outdoor activity in the central United States. [33] However this concentration is difficult to achieve without supplementation in locations where daylight hours are shorter (e.g. at higher latitude) and sunlight is weaker (e.g. at lower elevation). [22] Sun exposure is also the most important risk factor for melanoma and non-melanoma skin cancer – particularly among individuals with a light skin tone. [34] This is because ultraviolet radiation in sunlight not only induces the synthesis of vitamin D in skin, but damages its DNA without adequate protection by melanin .[35] As expected, higher serum 25-hydroxyvitamin D concentration was associated with a higher risk of skin cancer in a recent meta-analyses of prospective cohort studies. [36] And while results from Mendelian randomization studies suggest that this is not a causal relationship [37 38], serum 25-hydroxyvitamin D concentration may still be useful as a skin cancer risk marker because its concentration is related to sun exposure.

In our study, we found a positive association between serum 25-hydroxyvitamin D and risk of skin cancer which was consistent over time but stronger among individuals who resided in CDAs with a higher proportion of non-citizens, recent immigrants, visible minorities, and those who did not speak an official language at home. This may be because individuals living in these communities are less likely to take vitamin D supplements [39 40], which would make their serum 25-hydroxyvitamin D concentration

more representative of sun exposure than supplementation – resulting in a stronger overall association with skin cancer risk. Interestingly, removal of individuals with 25-OH vitamin D concentrations of 100 nmol/L or greater strengthened the association with skin cancer – which suggests we may have indeed removed individuals who were more likely to be taking vitamin D supplements. In a study of non-lactating women, daily oral supplementation of 5000 IU / day for 1 month raised serum 25-hydroxyvitamin D above 100 nmol/L. [41]

This study has some strengths. First, we used available secondary data to assemble a large retrospective cohort of community patients while making several restrictions and exclusions to minimize bias. Second, while this population included only patients who received an ECG, any patients that had a 25-hydroxyvitamin D measurement were eligible for inclusion. During the testing period, our laboratories experienced a very high volume of serum 25-hydroxyvitamin D testing – likely because of substantial interest in vitamin D at the time. Third, we adjusted for several potentially important confounders, including mean daylight hours during month of testing, BMI, and community-level measures of SES and examined variation in the association between serum 25-hydroxyvitamin D and cancer risk according to them.

This study also has some limitations. First, as it was based on secondary data, we had a limited number of variables and no control over when they were measured. Importantly, we could not tell if patients had taken vitamin D supplements. Supplementation elevates serum 25-hydroxyvitamin D in individuals who are vitamin D deficient, including those who are deficient because of low sun exposure and who are therefore at low risk for skin cancer. Including these individuals in our study would

weaken the strong and biologically plausible relationship we and others have observed between sun exposure, serum 25-hydroxyvitamin D concentration and skin cancer risk. As such, we may have underestimated the true association in our study – especially since we observed variation in this association according to community-level factors related to supplement use, and a strengthening of this association after removing patients with a serum 25-hydroxyvitamin D concentration of 100 nmol/L or greater. We also included only a single measure of serum 25-hydroxyvitamin D for patients in our study. However even a single measure may be useful in representing usual status due to its moderate intra-individual variation (Spearmen R, ICC = ~0.6) over similar time periods as our study. [42] Using the first measure may also better represent historic vitamin D status which is more likely to be associated with cancer risk than recent changes from supplementation in response to a diagnosis of deficiency. Interestingly, we found no change in our associations according to elapsed time between 25-hydroxyvitamin D testing and cancer diagnosis – which suggests our single measure may have adequately estimated usual vitamin D status. Second, as our study was observational, we could not determine whether the 25-hydroxyvitamin D-cancer relationship was causal. However this was not an objective of our study nor was it even achievable. Third, while we did not include participants without a BMI and ECG, we felt obtaining BMI was critical for reducing bias – even if it was based on self-report. We keep only participants who had BMI measured within a short period of time (+/- 3 month) from 25-hydroxyvitamin D measurement to maximize its relevance to 25-hydroxyvitamin D concentration. However because ECGs are used to identify the presence of cardiovascular disease, our population may be at an elevated risk for both

cardiovascular disease and cancer because many of the risk factors for cardiovascular disease are also risk factors for cancer (e.g. poor diet) [43].

5. Conclusion:

Higher vitamin D status was associated with a greater risk of skin cancer in a large community population under investigation for cardiovascular disease. This association is likely due to sun exposure and may be modified by community variation in supplementation rates.

Table 1: Patient characteristics by quartile of serum 25-hydroxyvitamin D

	Serum 25-hydroxyvitamin D quartile				
	Q1 (10 – 44 nmol/L)	Q2 (45 – 64 nmol/L)	Q3 (65 – 87 nmol/L)	Q4 (88 – 658 nmol/L)	P for linear trend
N	18053	18022	18056	18040	
Serum 25-hydroxyvitamin D, median, mean (sd)	33, 31 (8.5)	55, 55, (5.8)	75, 75 (6.4)	105, 114 (29.6)	
Age, y, mean (sd)	48 (15)	51 (15)	54 (15)	58 (15)	< 0.001
% Male, (n)	56.3 (10166)	51.4 (9261)	45.4 (8199)	40.1 (7218)	< 0.001
BMI, kg/m^2, mean (sd)	27.4 (5.3)	27.3 (5.1)	26.9 (5.1)	26.1 (4.8)	< 0.001
Daylight hours during month of testing, mean (sd)	11.9 (2.9)	12.2 (2.9)	12.4 (2.9)	12.5 (2.9)	< 0.001
Census dissemination area-level measures of socioeconomic status (SES)					
Canadian citizens, mean % (sd)	88.1 (10.1)	89.8 (8.9)	91.0 (8.1)	91.9 (7.5)	< 0.001
Recent immigrants, mean % (sd)	33.4 (16.8)	30.0 (16.2)	27.4 (15.1)	25.2 (13.8)	< 0.001
Visible minorities, mean % (sd)	38.5 (26.0)	32.4 (24.9)	27.8 (22.6)	23.6 (19.9)	< 0.001
Do not speak official language (English or French) at home, mean % (sd)	39.4 (22.1)	34.5 (21.1)	31.0 (19.6)	27.7 (17.3)	< 0.001
Aboriginal identity, mean % (sd)	2.5 (5.7)	2.2 (4.4)	2.1 (4.2)	2.0 (4.1)	< 0.001
Postsecondary education, mean % (sd)	54.8 (14.0)	56.9 (13.6)	58.7 (13.5)	60.2 (12.9)	< 0.001
Employed, mean % (sd)	93.3 (5.7)	93.5 (5.8)	93.7 (5.5)	93.9 (5.2)	< 0.001
Household total income, $, median (sd)	85434 (32435)	90779 (35727)	944486 (38827)	96986 (41439)	< 0.001
All cancer, n	615	737	928	1159	< 0.001
Major cancer, n	264	381	470	604	< 0.001
Breast cancer, n	83	125	127	183	< 0.001
Colorectal cancer, n	57	68	91	101	< 0.001
Lung cancer, n	30	44	57	61	< 0.001
Prostate cancer, n	63	80	77	110	< 0.001
Skin cancer, n (non-melanoma/melanoma)	31 (5/26)	64 (15/49)	118 (17/101)	149 (21/128)	< 0.001

Other cancer, n	351	356	458	555	< 0.001

Caption: There were 2851 census dissemination areas.

Table 2 – Cox proportional hazard regression of serum 25-hydroxyvitamin D quartile and risk of cancer

	Q1	Q2	Q3	Q4	Change in risk per quartile	P for linear trend
Any cancer						
Model 1	1.00	1.04 (0.93-1.16)	1.13 (1.02-1.26)	1.21 (1.09-1.33)	1.07 (1.04-1.10)	< 0.001
Model 2	1.00	1.04 (0.94-1.16)	1.14 (1.03-1.27)	1.24 (1.12-1.37)	1.08 (1.04-1.11)	< 0.001
Model 3	1.00	1.04 (0.93-1.16)	1.12 (1.01-1.26)	1.21 (1.08-1.34)	1.07 (1.03-1.10)	< 0.001
Major cancer						
Model 1	1.00	1.24 (1.06-1.45)	1.32 (1.13-1.54)	1.43 (1.24-1.66)	1.11 (1.06-1.16)	< 0.001
Model 2	1.00	1.24 (1.06-1.45)	1.32 (1.14-1.54)	1.46 (1.26-1.69)	1.12 (1.07-1.17)	< 0.001
Model 3	1.00	1.21 (1.03-1.43)	1.24 (1.05-1.47)	1.34 (1.14-1.58)	1.09 (1.04-1.14)	< 0.001
Breast cancer						
Model 1	1.00	1.36 (1.03-1.80)	1.26 (0.95-1.66)	1.63 (1.25--2.12)	1.14 (1.06-1.24)	< 0.001
Model 2	1.00	1.23 (0.93-1.62)	1.02 (0.77-1.35)	1.23 (0.94-1.61)	1.04 (0.96-1.13)	0.30
Model 3	1.00	1.22 (0.92-1.63)	0.95 (0.71-1.27)	1.13 (0.85-1.51)	1.01 (0.93-1.10)	0.83
Colorectal cancer						
Model 1	1.00	1.01 (0.71-1.44)	1.14 (0.81-1.59)	1.04 (0.75-1.45)	1.02 (0.93-1.13)	0.72
Model 2	1.00	1.03 (0.72-1.46)	1.18 (0.84-1.65)	1.12 (0.80-1.56)	1.04 (0.94-1.16)	0.41
Model 3	1.00	1.03 (0.72-1.48)	1.20 (0.84-1.71)	1.16 (0.78-1.59)	1.04 (0.91-1.17)	0.44
Lung cancer						
Model 1	1.00	1.20 (0.75-1.91)	1.26 (0.81-1.96)	1.06 (0.68--1.65)	1.01 (0.88-1.15)	0.93
Model 2	1.00	1.20 (0.75-1.90)	1.25 (0.80-1.96)	1.06 (0.68-1.66)	1.01 (0.88-1.15)	0.91
Model 3	1.00	1.23 (0.76-1.98)	1.27 (0.80-2.00)	1.02 (0.65-1.61)	0.99 (0.88-1.13)	0.92
Prostate cancer						
Model 1	1.00	1.10 (0.79-1.53)	0.92 (0.65-1.28)	1.11 (0.81-1.52)	1.02 (0.92-1.12)	0.74
Model 2	1.00	1.20 (0.86-1.67)	1.13 (0.81-1.58)	1.57 (1.14-2.16)	1.14 (1.04-1.27)	0.01
Model 3	1.00	1.13 (0.80-1.59)	1.08 (0.76-1.52)	1.42 (1.02-1.97)	1.11 (1.00-1.24)	0.05
Skin cancer						
Model 1	1.00	1.75 (1.14-2.68)	2.72 (1.83-4.05)	2.84 (1.92-4.21)	1.35 (1.23-1.50)	< 0.001
Model 2	1.00	1.78 (1.16-2.73)	2.82 (1.89-4.20)	3.04 (2.05-4.51)	1.39 (1.25-1.54)	< 0.001
Model 3	1.00	1.66 (1.08-2.58)	2.42 (1.61-3.65)	2.56 (1.70-3.86)	1.31 (1.18-1.46)	< 0.001
Other cancer						
Model 1	1.00	0.89 (0.77-1.03)	1.00 (0.87-1.15)	1.04 (0.91-1.20)	1.03 (0.98-1.08)	0.20
Model 2	1.00	0.89 (0.77-1.03)	1.01 (0.88-1.16)	1.08 (0.94-1.24)	1.04 (1.00-1.09)	0.07

Model 3	1.00	0.91 (0.78-1.06)	1.04 (0.89-1.20)	1.11 (0.96-1.28)	1.05 (1.00-1.10)	0.04

Caption: Hazard ratios (HR) are indicated with 95% Confidence intervals in parenthesis. Model 1: adjusted for age and log(time) interaction with age. Model 2: model 1 additionally adjusted for sex, BMI, mean daylight hours during month of testing and log(time) interactions with age, sex, BMI and mean daylight hours during month of testing. Model 3: model 2 additionally adjusted for CDA-level proportion of Canadian citizens, recent immigrants, visible (non-white) minorities, those indicating Aboriginal identity, those not speaking official languages (English or French) at home, those with postsecondary education, those currently employed, the CDA median household income and log(time) interactions with CDA-level proportion of Canadian citizens, those indicating Aboriginal identity, and those with postsecondary education.

Table 3: Association of serum 25-hydroxyvitamin D quartile with risk of skin cancer stratified by median values of Census Dissemination Area (CDA) covariates

	Q1	Q2	Q3	Q4	P for linear trend	P for interaction
Proportion of Canadian citizens						
≥ 92.7%	1.00	1.86 (1.05-3.31)	2.33 (1.37-3.97)	2.21 (1.28-3.81)	< 0.001	0.03
< 92.7%	1.00	1.36 (0.69-2.67)	2.55 (1.34-4.85)	3.16 (1.71-5.84)	< 0.001	
Proportion of recent immigrants						
≥ 26.7%	1.00	1.55 (0.72-3.31)	2.99 (1.51-5.95)	3.24 (1.65-6.34)	< 0.001	0.04
< 26.7%	1.00	1.66 (0.97-2.85)	2.08 (1.25-3.45)	2.18 (1.31-3.64)	< 0.001	
Proportion of visible (non-white) minorities						
≥ 24.3%	1.00	1.99 (0.95-4.15)	2.83 (1.39-5.73)	3.06 (1.53-6.09)	< 0.001	0.03
< 24.3%	1.00	1.47 (0.86-2.54)	2.16 (1.31-3.56)	2.26 (1.37-3.73)	< 0.001	
Proportion of non-official language (English or French) speakers at home						
≥ 29.2%	1.00	1.39 (0.67-2.90)	2.58 (1.31-5.07)	2.75 (1.41-5.35)	< 0.001	0.02
< 29.2%	1.00	1.79 (1.03-3.10)	2.33 (1.40-3.90)	2.47 (1.47-4.14)	< 0.001	

Caption: Hazard ratios (HR) are indicated with 95% Confidence intervals in parenthesis. All models are adjusted for age, sex, BMI, mean daylight hours during month of testing, Census Dissemination Area (CDA)-level proportions of Canadian citizens, recent immigrants, visible (non-white) minorities, those not speaking official languages (English or French) at home, those indicating Aboriginal identity, those with postsecondary education, those currently employed, and the CDA median household income unless stratified by that variable. They were also adjusted for log(time) interactions with CDA-level proportion of Canadian citizens, those indicating Aboriginal identity, and those with postsecondary education. Only nominally significant interactions are shown.

Figure 1 – Cohort design

Caption: Note – Serum 25-hydroxyvitamin D measurements were made between 2009 and 2013. CDA-level SES covariates were measured in 2011.

Figure 2 – Patient exclusions leading to analysis cohort

Contributor: LdK, JY and CN designed the study. LdK acquired the data. JY, YD and LdK conducted the statistical analysis. LdK, JY, YD, and CN interpreted the data. LdK and JY drafted the manuscript. JY, YD, CN and LdK critically revised and approved the final version of the manuscript. LdK supervised the project, obtained funding and acts as the guarantor.

Competing interests: The authors declare no potential competing interest.

Funding: This work was partially supported by funding from the MSI foundation of Alberta (grant #871) to LdK, and summer studentship awards to JY from Calgary Laboratory Services and the University of Calgary. All funding sources had no role in this work.

Data sharing statement: Please send any requests for data to the corresponding author (abldekon@ucalgary.ca).

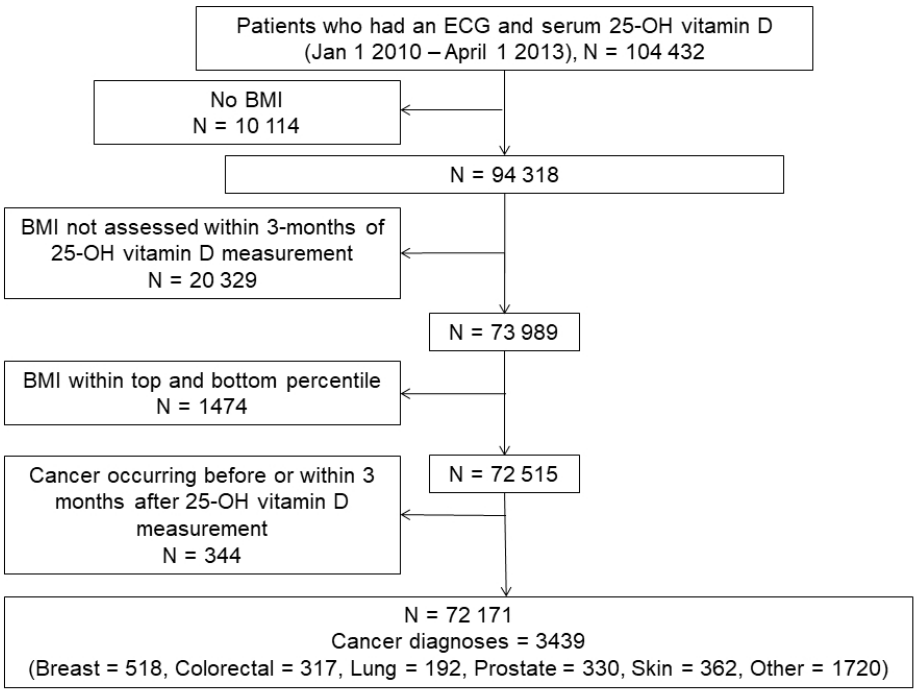
References

1. Medicine). IIo. *Dietary Reference Intakes for Calcium and Vitamin D* . . Washington, DC, USA.: The National Academies Press. .
2. Holick MF, Binkley NC, Bischoff-Ferrari HA, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011;96(7):1911-30. doi: 10.1210/jc.2011-0385 [published Online First: 20110606]
3. Fraser WD. Chapter 64: Bone and Mineral Metabolism. In: Rifai N, ed. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics. 6th ed. St. Louis, MO.: Elsevier 2018:1422-91.
4. Wacker M, Holick MF. Sunlight and Vitamin D: A global perspective for health. *Dermatoendocrinol* 2013;5(1):51-108. doi: 10.4161/derm.24494
5. Janz T, Pearson C. Vitamin D Blood Levels of Canadians. Ottawa, Canada.: Statistics Canada, 2013.
6. Yin L, Ordóñez-Mena JM, Chen T, et al. Circulating 25-hydroxyvitamin D serum concentration and total cancer incidence and mortality: a systematic review and meta-analysis. *Prev Med* 2013;57(6):753-64. doi: 10.1016/j.ypmed.2013.08.026 [published Online First: 2013/09/10]
7. Ma Y, Zhang P, Wang F, et al. Association between vitamin D and risk of colorectal cancer: a systematic review of prospective studies. *J Clin Oncol* 2011;29(28):3775-82. doi: 10.1200/JCO.2011.35.7566 [published Online First: 2011/08/29]
8. Dunn JA, Jefferson K, MacDonald D, et al. Low serum 25-hydroxyvitamin D is associated with increased bladder cancer risk: A systematic review and evidence of a potential mechanism. *J Steroid Biochem Mol Biol* 2019;188:134-40. doi: 10.1016/j.jsbmb.2019.01.002 [published Online First: 2019/01/14]
9. Mäkitie A, Tuokkola I, Laurell G, et al. Vitamin D in Head and Neck Cancer: a Systematic Review. *Curr Oncol Rep* 2020;23(1):5. doi: 10.1007/s11912-020-00996-7 [published Online First: 2020/11/20]
10. Zhang Y, Jiang X, Li X, et al. Serum Vitamin D Levels and Risk of Liver Cancer: A Systematic Review and Dose-Response Meta-Analysis of Cohort Studies. *Nutr Cancer* 2020:1-9. doi: 10.1080/01635581.2020.1797127 [published Online First: 2020/07/24]
11. Chowdhury R, Kunutsor S, Vitezova A, et al. Vitamin D and risk of cause specific death: systematic review and meta-analysis of observational cohort and randomised intervention studies. *BMJ* 2014;348:g1903. doi: 10.1136/bmj.g1903 [published Online First: 2014/04/01]
12. Bandera Merchan B, Morcillo S, Martin-Nuñez G, et al. The role of vitamin D and VDR in carcinogenesis: Through epidemiology and basic sciences. *J Steroid Biochem Mol Biol* 2017;167:203-18. doi: 10.1016/j.jsbmb.2016.11.020 [published Online First: 2016/11/30]
13. Vanlint S. Vitamin D and obesity. *Nutrients* 2013;5(3):949-56. doi: 10.3390/nu5030949 [published Online First: 2013/03/20]

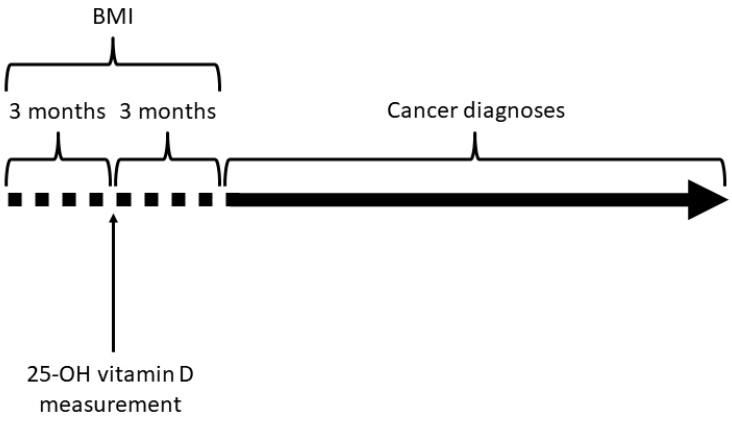
14. Bhaskaran K, Douglas I, Forbes H, et al. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet* 2014;384(9945):755-65. doi: 10.1016/S0140-6736(14)60892-8 [published Online First: 2014/08/13]
15. de Koning L, Henne D, Woods P, et al. Sociodemographic correlates of 25-hydroxyvitamin D test utilization in Calgary, Alberta. *BMC Health Serv Res* 2014;14:339. doi: 10.1186/1472-6963-14-339 [published Online First: 2014/08/09]
16. Merletti F, Galassi C, Spadea T. The socioeconomic determinants of cancer. *Environ Health* 2011;10 Suppl 1:S7. doi: 10.1186/1476-069X-10-S1-S7 [published Online First: 2011/04/05]
17. Bikos K. Calgary, Alberta, Canada — Sunrise, Sunset, and Daylength [Available from: https://www.timeanddate.com/sun/canada/calgary accessed July 23 2021 2021.
18. de Koning L, Al-Turkmani MR, Berg AH, et al. Variation in clinical vitamin D status by DiaSorin Liaison and LC-MS/MS in the presence of elevated 25-OH vitamin D2. *Clin Chim Acta* 2013;415:54-8. doi: 10.1016/j.cca.2012.09.002 [published Online First: 2012/09/10]
19. Society CC. Cancer statistics at a glance 2021 [Available from: https://www.cancer.ca/en/cancer-information/cancer-101/cancer-statistics-at-a-glance/?region=on.
20. Cancer IAfRo. The Cancer Dictionary Lyon, France.: World Health Organizatio; 2010 [Available from: https://www-dep.iarc.fr/WHOdb/WHOdb.htm accessed March 5 2021 2021.
21. Bradburn MJ, Clark TG, Love SB, et al. Survival analysis Part III: multivariate data analysis -- choosing a model and assessing its adequacy and fit. *Br J Cancer* 2003;89(4):605-11. doi: 10.1038/sj.bjc.6601120
22. Sowah D, Fan X, Dennett L, et al. Vitamin D levels and deficiency with different occupations: a systematic review. *BMC Public Health* 2017;17(1):519. doi: 10.1186/s12889-017-4436-z [published Online First: 20170622]
23. Fraser WD. Bone and Mineral Metabolism. In: Rifai N, Horvath AR, Wittwer CT, eds. Tietz Textboook of Clinical Chemistry and Molecular Diagnostics. 6th ed. St. Louis, MO, USA.: Elsevier 2018:1422-91.
24. Autier P, Boniol M, Pizot C, et al. Vitamin D status and ill health: a systematic review. *Lancet Diabetes Endocrinol* 2014;2(1):76-89. doi: 10.1016/S2213-8587(13)70165-7 [published Online First: 2013/12/06]
25. 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. doi: 10.1016/S2213-8587(17)30357-1 [published Online First: 2017/11/05]
26. Dimitrakopoulou VI, Tsilidis KK, Haycock PC, et al. Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. *BMJ* 2017;359:j4761. doi: 10.1136/bmj.j4761 [published Online First: 2017/10/31]
27. Aspelund T, Grübler MR, Smith AV, et al. Effect of Genetically Low 25-Hydroxyvitamin D on Mortality Risk: Mendelian Randomization Analysis in 3

Large European Cohorts. *Nutrients* 2019;11(1) doi: 10.3390/nu11010074 [published Online First: 2019/01/02]
28. Ong JS, Gharahkhani P, An J, et al. Vitamin D and overall cancer risk and cancer mortality: a Mendelian randomization study. *Hum Mol Genet* 2018;27(24):4315-22. doi: 10.1093/hmg/ddy307
29. Keum N, Lee DH, Greenwood DC, et al. Vitamin D supplementation and total cancer incidence and mortality: a meta-analysis of randomized controlled trials. *Ann Oncol* 2019;30(5):733-43. doi: 10.1093/annonc/mdz059
30. Association between vitamin D supplementation and mortality: systematic review and meta-analysis. *BMJ* 2020;370:m2329. doi: 10.1136/bmj.m2329 [published Online First: 2020/09/22]
31. Chandler PD, Chen WY, Ajala ON, et al. Effect of Vitamin D3 Supplements on Development of Advanced Cancer: A Secondary Analysis of the VITAL Randomized Clinical Trial. *JAMA Netw Open* 2020;3(11):e2025850. doi: 10.1001/jamanetworkopen.2020.25850 [published Online First: 2020/11/02]
32. Jager N, Schöpe J, Wagenpfeil S, et al. The Impact of UV-dose, Body Surface Area Exposed and Other Factors on Cutaneous Vitamin D Synthesis Measured as Serum 25(OH)D Concentration: Systematic Review and Meta-analysis. *Anticancer Res* 2018;38(2):1165-71. doi: 10.21873/anticanres.12336
33. Barger-Lux MJ, Heaney RP. Effects of above average summer sun exposure on serum 25-hydroxyvitamin D and calcium absorption. *J Clin Endocrinol Metab* 2002;87(11):4952-6. doi: 10.1210/jc.2002-020636
34. American Cancer Society. Cancer Facts & Figures 2021. . Atlanta, Georgia, USA., 2021.
35. Institute of Medicine. *Dietary Reference Intakes for Calcium and Vitamin D* . . Washington, DC, USA.: The National Academies Press. .
36. Mahamat-Saleh Y, Aune D, Schlesinger S. 25-Hydroxyvitamin D status, vitamin D intake, and skin cancer risk: a systematic review and dose-response meta-analysis of prospective studies. *Sci Rep* 2020;10(1):13151. doi: 10.1038/s41598-020-70078-y [published Online First: 2020/08/04]
37. Winsløw UC, Nordestgaard BG, Afzal S. High plasma 25-hydroxyvitamin D and high risk of nonmelanoma skin cancer: a Mendelian randomization study of 97 849 individuals. *Br J Dermatol* 2018;178(6):1388-95. doi: 10.1111/bjd.16127 [published Online First: 2018/04/14]
38. Liyanage UE, Law MH, Barrett JH, et al. Is there a causal relationship between vitamin D and melanoma risk? A Mendelian randomization study. *Br J Dermatol* 2020;182(1):97-103. doi: 10.1111/bjd.18238 [published Online First: 2019/09/11]
39. Gordon NP, Caan BJ, Asgari MM. Variation in vitamin D supplementation among adults in a multi-race/ethnic health plan population, 2008. *Nutr J* 2012;11:104. doi: 10.1186/1475-2891-11-104 [published Online First: 2012/12/11]
40. Moffat T, Sellen D, Wilson W, et al. Comparison of infant vitamin D supplement use among Canadian-born, immigrant, and refugee mothers. *J Transcult Nurs* 2015;26(3):261-9. doi: 10.1177/1043659614531793 [published Online First: 2014/05/05]

41. Meekins ME, Oberhelman SS, Lee BR, et al. Pharmacokinetics of daily versus monthly vitamin D3 supplementation in non-lactating women. *Eur J Clin Nutr* 2014;68(5):632-4. doi: 10.1038/ejcn.2013.278 [published Online First: 20140115]
42. Meng JE, Hovey KM, Wactawski-Wende J, et al. Intraindividual variation in plasma 25-hydroxyvitamin D measures 5 years apart among postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2012;21(6):916-24. doi: 10.1158/1055-9965.EPI-12-0026 [published Online First: 20120420]
43. Lau ES, Paniagua SM, Liu E, et al. Cardiovascular Risk Factors are Associated with Future Cancer. *JACC CardioOncol* 2021;3(1):48-58. doi: 10.1016/j.jaccao.2020.12.003 [published Online First: 2021/03/16]



254x190mm (96 x 96 DPI)



254x190mm (96 x 96 DPI)

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and time frame within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Page 4 " "
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported			Page 5-6
Objectives	3	State specific objectives, including any prespecified hypotheses			Page 6
Methods					
Study Design	4	Present key elements of study design early in the paper			Page 6-7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection			Page 6-7

Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Page 6-8 “ Figure 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Page 6-8
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group			Page 6-8

Bias	9	Describe any efforts to address potential sources of bias			Page 8-9
Study size	10	Explain how the study size was arrived at			Page 8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why			Page 9
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses			Page 9-10
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	Page 7-8

				RECORD 12.2: Authors should provide information on the data cleaning methods used in this study.	
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Page 6-10
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Page 10
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)			Page 10, 16
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure			Table 1, Page 10

		category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			Table 1, Table 2, Table 3
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			Page 11, Table 3
Discussion					
Key results	18	Summarise key results with reference to study objectives			Page 11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Page 14-15
Interpretation	20	Give a cautious overall interpretation of results considering objectives,			Page 15

		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results			Page 14
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based			Page 2
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Page 2

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.