Serum 25-hydroxyvitamin D levels and risk of colorectal cancer

Serum 25-hydroxyvitamin D levels and risk of colorectal cancer: an age-stratified analysis

Short title: 25(OH)D and early-onset colorectal cancer

Yejin Kim[1], Yoosoo Chang[1,2,3], Yoosun Cho[1,4], Jiwon Chang[4], Kyungeun Kim[5], Dong-Il Park[6], Soo-Kyung Park[6], Hee-Kyung Joh[7,8,9], Mi Kyung Kim[10], Chanmin Kim[11], Sarah H. Wild[12], Christopher D Byrne[13,14], Seungho Ryu[1,2,3]

[1]Center for Cohort Studies, Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[2]Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[3]Department of Clinical Research Design & Evaluation, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea

[4]Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[5]Department of Pathology, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[6]Division of Gastroenterology, Department of Medicine, School of Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University, Seoul, Republic of Korea

[7]Department of Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

[8]Department of Family Medicine, Seoul National University Hospital, Seoul, Republic of Korea

[9]Department of Family Medicine, Seoul National University Health Service Center, Seoul, Republic of Korea

[10] Department of Preventive Medicine, College of Medicine, Hanyang University, Seoul, Republic of Korea.

[11] Department of Statistics, Sungkyunkwan University, Seoul, Republic of Korea

[12] Usher Institute, University of Edinburgh, Edinburgh, UK

[13]Nutrition and Metabolism, Faculty of Medicine, University of Southampton, Southampton, UK

[14]National Institute for Health and Care Research Southampton Biomedical Research Centre, University Hospital Southampton, Southampton, UK

Address for correspondence: Prof. Seungho Ryu

Kangbuk Samsung Hospital, Samsung Main Building B2, 250, Taepyung-ro 2ga, Jung-gu, Seoul 04514, Republic of Korea

E-mail: sh703.yoo@gmail.com; Telephone: 82-2-2001-5137; Fax: 82-2-757-0436

Co-corresponding author: Prof. Yoosoo Chang

Kangbuk Samsung Hospital, Samsung Main Building B2, 250, Taepyung-ro 2ga, Jung-gu, Seoul 04514, Republic of Korea

E-mail: yoosoo.chang@gmail.com; Telephone: 82-2-2001-5139; Fax: 82-2-757-0436

Acknowledgements: This study was supported by the SKKU Excellence in Research Award Research Fund, Sungkyunkwan University, 2021. CDB was supported in part by the

Southampton National Institute for Health Research Biomedical Research Centre (grant code NIHR 203319), UK.

Financial support statement: None to declare

Conflict of interest statement: The authors have no conflicts of interest to disclose.

Data availability statement: The data are not available to be shared publicly as per the policy of institutional review board (IRB) regarding data distribution. However, data is available from the corresponding author on reasonable request.

Author Contributions:

Yejin Kim: study concept, interpretation of data, and drafting and critical revision of the manuscript

Yoosoo Chang: study concept and design, acquisition of data, interpretation of data, and drafting and critical revision of the manuscript

Yoosun Cho: interpretation of data and critical revision of the manuscript

Jiwon Chang: interpretation of data and critical revision of the manuscript

Kyungeun Kim: interpretation of data and critical revision of the manuscript

Dong-Il Park: interpretation of data and critical revision of the manuscript

Soo-Kyung Park: interpretation of data and critical revision of the manuscript

Hee-Kyung Joh: interpretation of data, and critical revision of the manuscript

Mi Kyung Kim: interpretation of data and critical revision of the manuscript

Chanmin Kim: interpretation of data and critical revision of the manuscript

Sarah H. Wild: interpretation of data and critical revision of the manuscript

Christopher D Byrne: interpretation of data and critical revision of the manuscript

Seungho Ryu: study concept and design, acquisition of data, analysis and interpretation of data, and critical revision of the manuscript

ABSTRACT

Background and aims: The role of circulating 25-hydroxyvitamin D (25(OH)D) in prevention of early-onset colorectal cancer (CRC) in young adults under 50 years is uncertain. We evaluated the age-stratified associations (<50 vs. ≥50 years) between circulating 25(OH)D levels and the risk of CRC in a large sample of Korean adults. **Methods:** Our cohort study included 236,382 participants (mean [standard deviation] age, 38.0 [9.0] years) who underwent a comprehensive health examination, including measurement of serum 25(OH)D levels. Serum 25(OH)D levels were categorized as follows: <10, 10–20, and ≥20 ng/mL. CRC, along with the histologic subtype, site, and invasiveness was ascertained through linkage with the national cancer registry. Cox proportional hazard models were used to estimate hazard ratios (HRs; 95% confidence intervals [CIs]) for incident CRC according to the serum 25(OH)D status, with adjustment for potential confounders. **Results:** During the 1,393,741 person-years of follow-up (median, 6.5 years; interquartile range, 4.5–7.5 years), 341 participants developed CRC (incidence rate, 19.2 per 10^5 person-years). Among young individuals aged <50 years, serum 25(OH)D levels were inversely associated with the risk of incident CRC with HRs (95% CIs) of 0.61 (0.43–0.86) and 0.41 (0.27–0.63) for 25(OH)D 10-19 and ≥20 ng/mL, respectively, with respect to the reference (<10 ng/mL) (*p* for trend <0.001, time-dependent model). Significant associations were evident for adenocarcinoma, colon cancer, and invasive cancers. For those aged ≥50 years, associations were similar, although slightly attenuated compared to younger individuals. **Conclusions:** Serum 25(OH)D levels may have beneficial associations with the risk of developing CRC for both early-onset and late-onset disease.

Keywords: early-onset colorectal cancer; cohort study; serum 25-hydroxyvitamin D; risk factor

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer worldwide [1]. Although the incidence and mortality of CRC among older adults are declining, the incidence of early-onset CRC in individuals aged <50 years has doubled since the 1990s [2] and is especially high in high-income countries [3]. Accumulating evidence suggests that the recent increase in the incidence of early-onset CRC is likely attributable to changes in environmental and lifestyle factors, such as elevated exposure to the Western diet, obesity, sedentary lifestyle, and physical inactivity [2, 4].

Vitamin D, 25-hydroxyvitamin D (25(OH)D), is a pro-hormone involved in various actions across different tissues and possesses anticancer properties [5, 6]. Vitamin D deficiency is increasing faster among young adults [7-10], potentially due to physical inactivity or decreased vitamin D intake [9, 11, 12], which are also linked to a rising incidence of CRC in this subpopulation. Previous epidemiological studies suggest that low vitamin D status may be implicated in the development of CRC, mostly later-onset CRC [13-20]. A recent meta-analysis indicates that vitamin D consumption may have protective effects on risks of CRC [14]. However, as CRC is typically considered a problem of older age, no studies have specifically addressed whether low vitamin D status may be responsible for the recent rise of CRC in young adults. There is only one recent study from the Nurses' Health Study II indicated potential protective associations between dietary vitamin D intake and risk of early-onset CRC [21]. To date, however, large cohort studies have yet to explore the association between circulating 25(OH)D levels and early-onset CRC in East Asian populations, characterized by a high prevalence of vitamin D deficiency. It is also unclear whether the associations differ by CRC tumor subtype.

Preventing and managing CRC in young adults is challenging due to limited understanding of the underlying pathophysiology and the absence of screening guidelines for

those under 50 years. Therefore, identifying predisposing risk factors for early-onset CRC in young adults is an essential step towards effective CRC prevention. In this study, we aimed to investigate the relationship between circulating 25(OH)D levels and incident CRC risks, as well as its anatomical and histological subtypes, in a large sample of Korean adults and examine whether the associations differ by age.

METHODS

Study population and design

The present study was conducted using data from the Kangbuk Samsung Health Study, a cohort study of men and women aged ≥18 years who underwent a comprehensive annual or biennial health examination at the clinics of the Kangbuk Samsung Hospital Total Healthcare Screening Center in Seoul and Suwon, South Korea, as previously described in detail [22]. The present study used de-identified retrospective data routinely collected during the health screening process of all study participants aged over 18 years who underwent a comprehensive health examination, provided informed consent for linkage to the national registries between January 2011 and December 2018 and were followed up until December 2019 (n = 242,187). In total, 5,805 participants were excluded because of missing data on serum 25(OH)D levels or body mass index (BMI) (n = 223); a self-reported history of cancer (n = 5,141), including a history of CRC (n = 370); or cancer (n = 2,088), including CRC (n = 381), registered before baseline based on the national cancer registry data. Some participants satisfied more than one exclusion criterion. Finally, 236,382 participants were included in the final analyses.

Ethics statement

This study was approved by the Institutional Review Board (IRB) of Kangbuk Samsung Hospital (IRB No. KBSMC 2021-11-057), which waived the requirement of informed consent specific to the current study due to the use of pre-existing, de-identified, retrospective data that were routinely collected during the health screening process and linked to the national cancer registry data in Korea. The names of the human participants and other HIPAA identifiers were not used during the study and have not been included in any section of the manuscript, including the supplementary information.

Data collection

Data on standardized self-administered questionnaires, physical measurements, and serum biochemical measurements were collected at each visit as part of the basic health check-up program [23]. The current average alcohol consumption per day was assessed using data on the frequency of alcohol consumption per week and amount of alcohol consumed per drinking day. Physical activity levels were assessed using the validated Korean version of the International Physical Activity Questionnaire Short Form [24, 25]. For further information on self-administered data collection, see **Supplementary Materials**.

Sitting blood pressure and anthropometric parameters were measured by trained nurses. Obesity was defined as a BMI of ≥25 kg/m^2, which is the proposed cut-off for the diagnosis of obesity among Asians [26]. Blood and urine specimens were collected after at least 10 h of fasting.

Serum 25(OH)D measurement

To assess serum 25(OH)D status, total 25(OH)D levels, including $25(OH)D_2$ and $25(OH)D_3$, were measured with a competitive immunoassay using an Elecsys Vitamin D Total assay on the Modular E170 (Roche Diagnostics, Tokyo, Japan) until April 2015 and

Cobas e801 (Roche Diagnostics) thereafter [27]. Total 25(OH)D measurement using the Elecsys Vitamin D Total assay demonstrated acceptable performance compared to that using liquid chromatography–tandem mass spectrometry, the reference standard for 25(OH)D measurement [27, 28]. When the analytical performance for precision was evaluated according to CLSI-EP15-A2 guidelines [29], the inter-assay coefficients of variation for quality control specimens with lower and higher levels of total 25(OH)D were 2.01%–5.94% and 2.69%–5.03%, respectively, during the study period, with a detection limit of <3 ng/mL, which was determined according to the CLSI EP17-A2 guidelines [30], as previously described [31].

Incident CRC cases

CRC or carcinoma in situ cases with anatomical site and histological type data were ascertained from the national cancer registry data that were available until December 2019. For further information on the cancer registry, see **Supplementary Material**. Based on the codes of the Classification of Diseases for Oncology, 3rd edition (ICD-O-3), CRC was defined as ICD-O-3 C18–C20 with a behaviour code of "/2 (carcinoma in situ)" or "/3 (malignant)" [32] and further categorized by anatomical site into colon cancer (C18) and cancer of the recto-sigmoid junction or rectum (C19 and C20). The ICD-O-3 morphology codes were used to define carcinoid tumors and adenocarcinomas [33]. Early-onset CRC was defined as CRC diagnosed before the age of 50 years; late-onset CRC was defined as CRC diagnosed after the age of 50 years.

Statistical analyses

1) Categorization of serum 25(OH)D levels

First, serum 25(OH)D levels were categorized as <10, 10–20, and ≥20 ng/mL (<25, 25–50, and ≥50 nmol/L, respectively) [34] as serum 25(OH)D levels ≥20 ng/mL indicate

vitamin D sufficiency according to the recommendation for the healthy general population [35-37]. We also evaluated a flexible estimate of the concentration–response relationship between 25(OH)D levels as a continuous variable and the risk of incident CRC by modelling 25(OH)D levels as restricted cubic splines with knots at the 5th, 27.5th, 50th, 72.5th, and 95th percentiles of the sample distribution. Furthermore, we assessed the risk of CRC for every 5 ng/mL increment of 25(OH)D level.

2) Baseline and main analyses

Descriptive statistics were used to summarise the participants' characteristics according to 25(OH)D categories. To describe potential linear trends in the incidence of CRC, the 25(OH)D categories were used as a continuous variable in the regression models.

The primary outcome was incident CRC, including both invasive cancer and carcinoma in situ. The secondary outcomes included separate assessments by histological subtype (adenocarcinoma vs. other types), anatomical location (colon vs. rectum), and invasiveness (invasive vs. carcinoma in situ).

Each participant was followed up from the baseline examination to the development of the endpoint or the end of follow-up (31 December 2019), whichever came first. Incidence rates were calculated as the number of incident cases divided by the person-years of follow-up. Cox proportional hazard models were used to estimate the hazard ratios (HRs) with 95% confidence intervals (CIs) for the development of incident CRC. Missing values were treated as the separate category of 'unknown' and were included in the models. We initially adjusted for age and sex. Multivariable model was additionally adjusted for the study center (Seoul and Suwon), year of screening examination, season (spring, summer, autumn, winter), alcohol consumption (0, <20, ≥20 g/d, or unknown), smoking status (never, past, current, or unknown), physical activity (inactive, minimally active, health enhancing physical activity, or

unknown), total energy intake (quintiles, or unknown), education level (<community college graduate, ≥community college graduate, or unknown), BMI (continuous), history of diabetes (yes, no, unknown), family history of CRC (yes, no, unknown), use of multi-vitamin (yes, no, unknown) and calcium supplements (yes, no, unknown) (presented in Supplementary Materials). To incorporate changes in serum 25(OH)D levels and other covariates between baseline and the end of the study, we performed analyses by introducing serum 25(OH)D levels and other variables, including smoking, alcohol consumption, physical activity, total energy intake, BMI, season, history of diabetes, and use of multi-vitamin and calcium supplements, as time-varying covariates in the models (time-dependent model, primary analysis). The proportional hazards assumption was assessed using estimated log (-log) survival curves, and no violation of this assumption was found.

3) Sensitivity analyses

i) Seasonal calibration

To further account for potential effect of season on serum 25(OH)D levels, we standardized vitamin D levels, in addition to adjustment for season of blood measure in the models, by adding the overall mean of vitamin D for all subjects to the residuals derived from a regression of vitamin D levels on the periodic function two variables ($R = \cos([2\pi/12]*X)$ and $S = \sin([2\pi/12]*X)$, where X is the month of blood measurement [38].

ii) Restricted analyses to early-onset CRC cases during follow-up

We restricted the analysis to early-onset cases by excluding incident CRC cases that occurred after the age of 50 years during follow-up.

iii) Restricted analyses to incident CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels

To address a potential issue of reverse causality, we conducted sensitivity analysis that was restricted to CRC cases that occurred at least 2-4 years after the measurement of 25(OH)D.

iv) Population Attributable Fraction (PAF)

To assess the fraction of CRC cases that can be potentially prevented by addressing vitamin D deficiency, PAFs of CRC risk, comparing vitamin D deficient individuals to non-deficient individuals, were calculated using the following formula [39]: PAF (%) = $P_e \times [RR-1] \times 100 / (1+ P_e \times [RR-1])$, where P_e is the exposed proportion of the population, and *RR* is the rate ratio.

v) Subgroup analysis

Subgroup analyses were performed according to sex (women vs men), BMI (<25 kg/m^2 vs ≥25 kg/m^2), and family history of CRC (no vs yes). The effect modifications by subgroup including age were tested using likelihood ratios to compare the models with and without multiplicative interaction terms.

Statistical analyses were performed using STATA version 17.0 (StataCorp LP, College Station, TX, USA). Statistical significance was set at $p < 0.05$.

RESULTS

Baseline characteristics

The mean (standard deviation) age of the 236,382 participants was 38.0 (9.0) years. The baseline characteristics of the study participants by age group (<50 years vs. ≥50 years) are presented according to 25(OH)D levels at baseline in **Table 1**. For participants aged <50 years, 25(OH)D levels were positively associated with age; male sex; alcohol intake; current smoking status; physical activity, educational level; and use of multi-vitamin, vitamin D, and calcium supplements and were inversely associated with total energy intake. For participants aged ≥50 years, 25(OH)D levels were positively associated with male sex; alcohol intake; physical activity; use of multivitamin and calcium supplement and were inversely associated with family history of CRC and total energy intake.

Risks of CRC according to serum 25(OH)D categories

During the 1,393,741 person-years of follow-up (median, 6.5 years; interquartile range, 4.5–7.5 years), 341 participants developed CRC (incidence rate, 19.2 per 10^5 person-years). Among participants aged <50 years, 229 individuals developed CRC during the 1,265,609 person-years of follow-up (incidence rate [95% CI], 18.1 [15.9–20.6]) per 10^5 person-years. Among participants aged ≥50 years, 112 developed CRC during 128,132 person-years of follow-up (incident rate [95% CI], 87.4 [72.6-105.2]). The median age at CRC diagnosis among the incident cases was 42.4 years (interquartile range, 37.8–46.7 years) for participants aged<50 years; 62.4 years (interquartile range, 57.1-68.2) for participants aged ≥50 years. The median (interquartile range) frequency of serum 25(OH)D measurement was 5 (3-7) for participants aged <50 years and 6 (3-8) for those aged ≥50 years. Further details on tumor characteristics can be found in **eTable 1**.

The analysis of clinical categories (as categorical variables) of 25(OH)D showed an inverse association between 25(OH)D levels and incident CRC in both age groups after adjusting for time-varying changes of potential confounders (**Table 2**; **eTable 2** for age-sex-

and multivariable-adjusted analyses). HRs (95% CIs) for incident CRC among participants aged <50 years were 0.61 (0.43–0.86) and 0.41 (0.27–0.63) for 25(OH)D 10-19 ng/mL and ≥20 ng/mL, respectively, compared to the reference level (<10 ng/mL) (*p* for trend = <0.001). Among those ≥50 years, comparable but slightly weaker associations were observed. The interaction of age and 25(OH)D levels with the risk of CRC based on data from the time-dependent regression model was not significant (*p*-value = 0.793). Additionally, the analysis considering every 5 ng/mL increment of 25(OH)D as a continuous variable consistently demonstrated an inverse association with incident CRC in both age groups (**eTable 3**).

Dose-response relationship between serum 25(OH)D and incident CRC risk

When we evaluated a dose-response relationship of 25(OH)D as a continuous variable with risks of incident CRC based on the spline regression models, the risk of CRC decreased in a dose-response manner across the range of 25(OH) levels in both age groups (**Figure 1**).

Serum 25(OH)D levels and risk of CRC by histological subtype

Table 3 shows the association between 25(OH)D levels and the development of CRC according to histological subtype, site, and invasiveness among individuals aged <50 years. In the analysis by histological subtype (adenocarcinoma vs other types), the risk of adenocarcinoma significantly decreased with increasing levels of 25(OH)D among individuals aged <50 years (HRs [95% CIs] of adenocarcinoma for 25(OH)D 10–20 and ≥20 ng/mL with respect to the reference, 0.59 [0.38–0.91] and 0.46 [0.27–0.77], respectively; *p* for trend = 0.005; **Table 3**; see **eTable 4** for age-sex- and multivariable-adjusted analyses**)**. However, no significant associations between the risk of other types of tumors and 25(OH)D levels were found.

For individuals aged ≥50 years (**Table 4**; **eTable 5** for age-sex- and multivariable-adjusted analyses), significant inverse associations with adenocarcinoma were observed only in the highest 25(OH)D level.

Serum 25(OH)D levels and risk of CRC by tumor location

The incidence of colon cancer (per 10^5 PY) was 8.9 and 59.3 for participants aged <50 years and ≥50 years, respectively; the incidence of rectal cancer was 9.4 and 29.6 for participants aged <50 years and ≥50 years, respectively.

For individuals aged <50 years, risk of colon cancer was significantly and inversely associated with increasing 25(OH)D levels (**Table 3**). The HRs (95% CI) estimated using time-dependent model comparing 25(OH)D 10-19 and ≥20 ng/mL to the reference were 0.55 (0.34-0.91) and 0.32 (0.17-0.59), respectively. For rectal cancer, however, the associations did not reach statistical significance. For those aged ≥50 years, similar, although slightly weaker, associations were observed compared to those in younger individuals (**Table 4**). These patterns were similarly observed in analysis considering every 5 ng/mL increment of 25(OH)D as a continuous variable (**eTable 6**).

Site-specific associations (colon vs rectum) between 25(OH)D levels and the risk of adenocarcinoma, excluding carcinoid tumors, are shown in **eTable 7.** In our study, we observed that 85.5% (52 cases out of 61) of carcinoid tumors in younger individuals and 80.0% (5 cases out of 6) in older individuals were located in the rectum. These findings accounted for 47.7% of rectal cancer cases among younger individuals and 17.9% among older adults (as shown in **eTable 1**). The associations between 25(OH)D levels and the risk of adenocarcinoma by different site were similar for both age groups, with significant associations with colon cancer found only in the younger age group after the adjustment for time-varying covariates.

Serum 25(OH)D levels and risk of CRC by invasiveness

Among younger individuals aged <50 years, significant inverse associations between 25(OH)D levels and invasive cancer were observed. HRs (95% CI) for serum 25(OH)D 10–20 and ≥20 ng/mL were 0.63 (0.42-0.95) and 0.42 (0.26–0.70), respectively (**Table 3**). For carcinoma in situ, a dose–response reduction in risk was found with increasing 25(OH)D levels, but the associations were not statistically significant. Similar pattern was observed among those aged ≥50 years (**Table 4**).

Sensitivity analysis

i) Seasonal calibration

In the analysis where season was calibrated, similar trends of associations to the main analysis were found (**eTable 8**), both for younger and older individuals.

ii) Restricted analyses to early-onset CRC cases during follow-up

When we limited our analyses strictly to early-onset cases by excluding diagnosis after 50 years of age, the associations between 25(OH) and risks of early-onset CRC were consistent and significant (**eTable 9**). In the analyses evaluating risks of tumor subtypes, significant associations with colon cancer, invasive cancer, and adenocarcinoma were still observed (**eTable 10**).

iii) Restricted analyses to incident CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels

In sensitivity analyses restricted to incident cases of CRC occurring from 2 to 4 years after the initial measurements (**eTable 11**), associations remained fairly similar to the original

observations among individuals aged <50 years. However, after further restriction to incident CRC cases that occurred after 4 years from sample collection, we did not detect significant associations. For those aged ≥50 years, associations were not significant (**eTable 12**).

iv) PAFs

PAF, comparing vitamin D deficient individuals to non-deficient individuals, is presented as a percentage of RR for individuals aged <50 years and for those aged ≥50 years (**eTable 13**). The estimated PAFs for these age groups are 5.4% and 3.0%, respectively.

v***) Subgroup analyses***

In the pre-specified subgroup analyses (**eTable 14**), the significant inverse association between 25(OH)D levels and the risk of incident CRC was found both in men and women; individuals with BMI <25 kg/m^2; and those without family history of CRC. However, there was no significant interaction between any of the subgroups.

Among participants aged <50 years, none of 695 participants who had a history of inflammatory bowel disease (IBD) developed incident CRC; for those aged ≥50 years, one out of 96 with a history of IBD developed incident CRC.

DISCUSSION

This large cohort study of 236,382 men and women with approximately 1.2 million person-years of follow-up demonstrated that serum 25(OH)D levels were significantly and inversely associated with the risk of CRC after adjustment for potential confounders including BMI, season, and physical activities as well as temporal changes in 25(OH)D levels and other covariates during follow-up; the associations were found both in younger and older individuals. In stratified analyses based on histologic subtype, subsite, and invasiveness, the

significant and inverse associations between 25(OH)D and CRC were evident in adenocarcinoma, colon cancer, and invasive cancer in young adults aged under 50 years; similar but slightly weaker inverse associations were observed among adults aged over 50 years.

We did not find statistically significant effect modification by age, although associations between 25(OH)D and incident CRC risks tended to be stronger in younger adults across the analyses compared to older adults. We observed that, while the relative risks associated with low 25 (OH)D were greater among young adults compared to older adults, the absolute risk of CRC associated with 25(OH)D levels was lower across all vitamin D categories. It is possible that the risk of CRC in older individuals could be influenced by cumulative exposure to age-associated risk factors, possibly leading to reduced magnitude of the risk attributable to serum concentration of 25(OH)D as compared with younger individuals. Evidence suggests that there are stronger associations between modifiable risk factors, CVD, and certain types of cancer in younger individuals owing to low absolute risks of a disease in this age group [40-42]. Additionally, the implications of vitamin D levels measured at a single time point may vary by age in terms of their value as a proxy measure of long-term 25(OH)D levels. Given the process of carcinogenesis may span up to several decades, vitamin D status at an earlier stage may be more relevant in the etiology of CRC development [43, 44]. However, given the lack of a statistically significant interaction between age groups, the beneficial associations between 25(OH)D and CRC seem to be applicable regardless of whether it is early-onset or late-onset disease, which warrants further investigation.

To our knowledge, there is only one recent study that evaluated the role of vitamin D exclusively in early-onset CRC [21]. This study reported significant inverse associations between total vitamin D intake and the risk of early-onset CRC in women (mean age, early

40s) [21]. However, the total vitamin D intake was derived from questionnaires in that study, which may not fully correlate with bioavailable serum 25(OH)D levels [21, 45]. One of the strengths of our study is that we directly measured serum 25(OH)D levels, which is considered the best indicator of vitamin D status [46]. Direct measurement, however, can also be subject to confounding by factors such as obesity, season, or outdoor activity and is not an ideal measure of long-term 25(OH)D status [47]. Although our study lacked information on outdoor activity or sunlight exposure, we accounted for other relevant factors such as BMI, physical activity, use of vitamin D or calcium supplements, season, and the time-varying changes in 25(OH)D and other covariates and revealed significant associations between serum 25(OH)D levels and the risk of CRC.

It is worth noting that prior randomized controlled trials (RCTs) have been equivocal regarding the role of vitamin D supplementation in CRC [19, 48, 49]. While many of the previous trials included vitamin D-replete individuals, recent studies, including few Mendelian randomisation analyses, have highlighted that benefit of vitamin D in reducing disease risks including mortality was not observed among individuals with sufficient vitamin D levels. These studies suggest that mortality decreased steeply as 25(OH)D levels increase up to 20 ng/mL, reaching a minimum of 25(OH)D-associated mortality at 30 ng/mL, [50, 51] but diminished beyond those levels [52-54]. One of the largest trials investigating health effects of vitamin D, the Vitamin D and Omega-3 Trial (VITAL) trial, has recently reported null effects of vitamin D supplementation in incident CRC as a secondary outcome. However, the population of the VITAL trial had comparatively high average serum 25(OH)D (approximately 30 ± 10.0 ng/mL), whereas our sample had a high prevalence of those with 25(OH)D levels <20 ng/mL (75% for younger group and ~60% for older group versus 12.7% in the VITAL study), consistent with the prevalence of suboptimal 25(OH)D levels in South Korea [7, 55]. Thus, in the trial, further increase in serum 25(OH)D may not have translated into

a stronger anti-cancer association. Another factor to consider is adiposity. Previous RCTs focused on cancer incidence including the VITAL trial used samples with a relatively high mean BMI over 27 kg/m^2 [19, 48, 49, 56, 57], which contrasts our sample with mean BMI of 23.3 kg/m^2 (SD, 3.4). Overweight and obesity are known to substantially reduce bioavailability of 25(OH)D and blunt responses to supplementation even in vitamin D-deficient individuals [58, 59]. Thus, whether vitamin D supplementation is effective in treating or preventing CRC remains inconclusive, partly since its effects can highly vary by individual health needs and responses to the intervention. Further RCTs should account for baseline vitamin D status or factors that modify vitamin D availability such as BMI to better elucidate benefits of vitamin D supplementation for CRC prevention.

Serum 25(OH)D level was significantly associated with a decreased risk of adenocarcinoma, but not with the risk of other cancer types, in which carcinoids comprise the majority of the cases. The differential effect of 25(OH)D likely relates to etiological differences between subtypes [60]. In our study, we tried to distinguish between site-specific effect versus histologic differences, as previous studies have suggested a higher prevalence of carcinoids in the rectum compared to the colon [61]. In the analyses limited to adenocarcinoma only, associations for both colon and rectal cancer were similarly observed among younger individuals, although associations with rectal adenocarcinoma were not clear in the time-dependent analysis. This finding suggests that the null associations between 25(OH)D levels and the risk of rectal cancer observed in our study may be partly attributable to selective effects of serum 25(OH)D levels on adenocarcinoma. While the incidence of carcinoid tumors is growing (accounting for 8%–34% of all rectal cancers) in young individuals with CRC [33], adenocarcinomas still account for the majority of CRC cases with a disproportionately high incidence among early-onset cases [60, 62]. In addition, adenocarcinomas, the primary target of CRC screening programs, are preventable by early

detection and removal of precancerous polyps [60]. Thus, in adults aged <50 years who are unlikely to undergo routine screening, the association of higher 25(OH)D with decreased risk of adenocarcinoma, if confirmed, may have an important clinical significance particularly in the prevention of early-onset CRC. The mechanistic involvement of serum 25(OH)D in histological subtypes of CRC should be further investigated.

Our study has several limitations. First, as mentioned earlier, although we included information on the use of vitamin D or multi-vitamin supplements, detailed information on dose, type, and frequency of supplementation; outdoor activities; or sunlight exposure was lacking. Therefore, the potential for residual confounding remains. Nevertheless, we directly measured serum 25(OH)D levels, which are considered to reflect the overall vitamin D status and cumulative effect of sunlight exposure and dietary intake of vitamin D. Second, because of the observational nature of our study, a causal association could not be established, and the possibility of residual confounding remains due to unmeasured factors such as genetic factors. Third, a self-administered structured questionnaire was employed to gather data on smoking, alcohol use, physical activity, and medical history in health checkup programs under the National Health Insurance plan in Korea. However, it is possible that measurement error in these variables may have caused some residual confounding, and thus cannot be entirely ruled out. Fourth, the median follow-up duration of 6.5 years may not be sufficient to capture an adequate number of early-onset CRC events. However, according to a previous pooled analysis of 17 cohorts, the reported median time from blood draw for the 25(OH)D assay to CRC (mostly late-onset) diagnosis was 5.5 years [13]. However, since early-onset CRC is a relatively rare condition compared to later-onset CRC, further large-scale studies with longer follow-up periods are necessary to ensure sufficient power. Lastly, we included a relatively young and healthy Korean working population; thus, the generalizability of our

findings to other populations with different racial/ethnic backgrounds or sociodemographic characteristics is limited.

In conclusion, in this large cohort study of Korean men and women, serum 25(OH)D levels were associated with a decreased risk of developing CRC for both early- and late-onset disease. Our findings provide support for the inverse correlation between vitamin D status and the risk of CRC. Additional investigations are needed to confirm whether maintaining adequate 25(OH)D levels can decrease the risk of developing CRC, particularly with a focus on differentiating between early-onset and later-onset cases, examining specific histologic types, and considering the baseline vitamin D deficiency status of individuals.

REFERENCES

1. Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. Lancet Gastroenterol Hepatol 2019;4:511-518.
2. Vuik FE, Nieuwenburg SA, Bardou M, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut 2019;68:1820-1826.
3. **Pan H, Zhao Z, Deng Y**, et al. The global, regional, and national early-onset colorectal cancer burden and trends from 1990 to 2019: results from the Global Burden of Disease Study 2019. BMC Public Health 2022;22:1896.
4. Burnett-Hartman AN, Lee JK, Demb J, et al. An Update on the Epidemiology, Molecular Characterization, Diagnosis, and Screening Strategies for Early-Onset Colorectal Cancer. Gastroenterology 2021;160:1041-1049.
5. Feldman D, Krishnan AV, Swami S, et al. The role of vitamin D in reducing cancer risk and progression. Nature Reviews Cancer 2014;14:342-357.
6. Jeon S-M, Shin E-A. Exploring vitamin D metabolism and function in cancer. Experimental & Molecular Medicine 2018;50:1-14.
7. Park JH, Hong IY, Chung JW, et al. Vitamin D status in South Korean population: Seven-year trend from the KNHANES. Medicine (Baltimore) 2018;97:e11032.
8. Ginde AA, Liu MC, Camargo CA, Jr. Demographic differences and trends of vitamin D insufficiency in the US population, 1988-2004. Arch Intern Med 2009;169:626-32.
9. Herrick KA, Storandt RJ, Afful J, et al. Vitamin D status in the United States, 2011–2014. The American Journal of Clinical Nutrition 2019;110:150-157.
10. Crowe FL, Jolly K, MacArthur C, et al. Trends in the incidence of testing for vitamin D deficiency in primary care in the UK: a retrospective analysis of The Health Improvement Network (THIN), 2005–2015. BMJ Open 2019;9:e028355.

11. Joh H-K, Lim CS, Cho B. Lifestyle and Dietary Factors Associated with Serum 25-Hydroxyvitamin D Levels in Korean Young Adults. jkms 2015;30:1110-1120.
12. Tangpricha V, Pearce EN, Chen TC, et al. Vitamin D insufficiency among free-living healthy young adults. Am J Med 2002;112:659-62.
13. **McCullough ML, Zoltick ES, Weinstein SJ**, et al. Circulating Vitamin D and Colorectal Cancer Risk: An International Pooling Project of 17 Cohorts. J Natl Cancer Inst 2019;111:158-169.
14. **Boughanem H, Canudas S**, Hernandez-Alonso P, et al. Vitamin D Intake and the Risk of Colorectal Cancer: An Updated Meta-Analysis and Systematic Review of Case-Control and Prospective Cohort Studies. Cancers (Basel) 2021;13.
15. Jenab M, Bueno-de-Mesquita HB, Ferrari P, et al. Association between pre-diagnostic circulating vitamin D concentration and risk of colorectal cancer in European populations:a nested case-control study. BMJ 2010;340:b5500.
16. Lee JE, Li H, Chan AT, et al. Circulating levels of vitamin D and colon and rectal cancer: the Physicians' Health Study and a meta-analysis of prospective studies. Cancer Prev Res (Phila) 2011;4:735-43.
17. Song M, Nishihara R, Wang M, et al. Plasma 25-hydroxyvitamin D and colorectal cancer risk according to tumour immunity status. Gut 2016;65:296-304.
18. Theodoratou E, Tzoulaki I, Zgaga L, et al. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. BMJ 2014;348:g2035.
19. Wactawski-Wende J, Kotchen JM, Anderson GL, et al. Calcium plus vitamin D supplementation and the risk of colorectal cancer. N Engl J Med 2006;354:684-96.
20. Yin L, Grandi N, Raum E, et al. Meta-analysis: longitudinal studies of serum vitamin D and colorectal cancer risk. Alimentary Pharmacology & Therapeutics 2009;30:113-125.
21. Kim H, Lipsyc-Sharf M, Zong X, et al. Total Vitamin D Intake and Risks of Early-Onset Colorectal Cancer and Precursors. Gastroenterology 2021;161:1208-1217 e9.
22. Chang Y, Ryu S, Choi Y, et al. Metabolically healthy obesity and development of chronic kidney disease: a cohort study. Annals of internal medicine 2016;164:305-312.
23. Chang Y, Ryu S, Sung KC, et al. Alcoholic and non-alcoholic fatty liver disease and associations with coronary artery calcification: evidence from the Kangbuk Samsung Health Study. Gut 2019;68:1667-1675.
24. Chun MY. Validity and reliability of korean version of international physical activity questionnaire short form in the elderly. Korean J Fam Med 2012;33:144-51.
25. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35:1381-95.
26. WHO Western Pacific Region, IASO and IOTF. The Asia-Pacific Perspective: Redefining Obesity and Its Treatment. Health Communications Australia Pty Limit: Sdyney, Australia. 2000.
27. Shin SY, Kwon MJ, Song J, et al. Measurement of serum total vitamin D (25-OH) using automated immunoassay in comparison [corrected] with liquid chromatography tandem-mass spectrometry. J Clin Lab Anal 2013;27:284-9.
28. Heijboer AC, Blankenstein MA, Kema IP, et al. Accuracy of 6 routine 25-hydroxyvitamin D assays: influence of vitamin D binding protein concentration. Clin Chem 2012;58:543-8.
29. Clinical and Laboratory Standards Institute. EP15-A2: User Verification of Performance for Precision and Trueness; Approved guideline. Wayne, PA: Clinical and Laboratory Standards Institute, 2012.
30. Clinical and Laboratory Standards Institute. EP17-A2: Evaluation of detection capability for clinical laboratory measurement procedures; Approved guideline. Wayne, PA: Clinical and Laboratory Standards Institute, 2012.

31. Kim Y, Chang Y, Ryu S, et al. Resolution of, and Risk of Incident Non-alcoholic Fatty Liver Disease With Changes in Serum 25-hydroxy Vitamin D Status. J Clin Endocrinol Metab 2022;107:e3437-e3447.
32. Hong S, Won YJ, Lee JJ, et al. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2018. Cancer Res Treat 2021;53:301-315.
33. Montminy EM, Zhou M, Maniscalco L, et al. Contributions of Adenocarcinoma and Carcinoid Tumors to Early-Onset Colorectal Cancer Incidence Rates in the United States. Ann Intern Med 2021;174:157-166.
34. . In: Ross AC, Taylor CL, Yaktine AL, Del Valle HB, eds. Dietary Reference Intakes for Calcium and Vitamin D. Washington (DC), 2011.
35. Giustina A, Adler RA, Binkley N, et al. Controversies in Vitamin D: Summary Statement From an International Conference. J Clin Endocrinol Metab 2019;104:234-240.
36. Slomski A. IOM endorses vitamin D, calcium only for bone health, dispels deficiency claims. JAMA 2011;305:453-4, 456.
37. Munns CF, Shaw N, Kiely M, et al. Global Consensus Recommendations on Prevention and Management of Nutritional Rickets. J Clin Endocrinol Metab 2016;101:394-415.
38. Sachs MC, Shoben A, Levin GP, et al. Estimating mean annual 25-hydroxyvitamin D concentrations from single measurements: the Multi-Ethnic Study of Atherosclerosis. Am J Clin Nutr 2013;97:1243-51.
39. Hanley JA. A heuristic approach to the formulas for population attributable fraction. J Epidemiol Community Health 2001;55:508-14.
40. **Tromp J, Paniagua SMA**, Lau ES, et al. Age dependent associations of risk factors with heart failure: pooled population based cohort study. Bmj 2021;372:n461.
41. Kaneko H, Yano Y, Okada A, et al. Age-Dependent Association Between Modifiable Risk Factors and Incident Cardiovascular Disease. J Am Heart Assoc 2023;12:e027684.
42. Yuan C, Kim J, Wang QL, et al. The age-dependent association of risk factors with pancreatic cancer. Ann Oncol 2022;33:693-701.
43. Joh HK, Lee DH, Hur J, et al. Simple Sugar and Sugar-Sweetened Beverage Intake During Adolescence and Risk of Colorectal Cancer Precursors. Gastroenterology 2021;161:128-142.e20.
44. Siegel RL, Jakubowski CD, Fedewa SA, et al. Colorectal Cancer in the Young: Epidemiology, Prevention, Management. Am Soc Clin Oncol Educ Book 2020;40:1-14.
45. Millen AE, Wactawski-Wende J, Pettinger M, et al. Predictors of serum 25-hydroxyvitamin D concentrations among postmenopausal women: the Women's Health Initiative Calcium plus Vitamin D clinical trial. Am J Clin Nutr 2010;91:1324-35.
46. Chandler PD, Buring JE, Manson JE, et al. Circulating Vitamin D Levels and Risk of Colorectal Cancer in Women. Cancer Prev Res (Phila) 2015;8:675-82.
47. Bertrand KA, Giovannucci E, Liu Y, et al. Determinants of plasma 25-hydroxyvitamin D and development of prediction models in three US cohorts. Br J Nutr 2012;108:1889-96.
48. Baron JA, Barry EL, Mott LA, et al. A Trial of Calcium and Vitamin D for the Prevention of Colorectal Adenomas. N Engl J Med 2015;373:1519-30.
49. Manson JE, Cook NR, Lee IM, et al. Vitamin D Supplements and Prevention of Cancer and Cardiovascular Disease. N Engl J Med 2019;380:33-44.
50. Deng X, Song Y, Manson JE, et al. Magnesium, vitamin D status and mortality: results from US National Health and Nutrition Examination Survey (NHANES) 2001 to 2006 and NHANES III. BMC Med 2013;11:187.
51. Gaksch M, Jorde R, Grimnes G, et al. Vitamin D and mortality: Individual participant data meta-analysis of standardized 25-hydroxyvitamin D in 26916 individuals from a European consortium. PLoS One 2017;12:e0170791.

52. Sofianopoulou E, Kaptoge SK, Afzal S, et al. Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses. The Lancet Diabetes & Endocrinology 2021;9:837-846.
53. Sutherland JP, Zhou A, Hyppönen E. Vitamin D Deficiency Increases Mortality Risk in the UK Biobank. Annals of Internal Medicine 2022.
54. Durup D, Jørgensen HL, Christensen J, et al. A reverse J-shaped association of all-cause mortality with serum 25-hydroxyvitamin D in general practice: the CopD study. J Clin Endocrinol Metab 2012;97:2644-52.
55. Park HY, Lim YH, Park JB, et al. Environmental and Occupation Factors Associated with Vitamin D Deficiency in Korean Adults: The Korea National Health and Nutrition Examination Survey (KNHANES) 2010-2014. Int J Environ Res Public Health 2020;17.
56. Chatterjee R, Fuss P, Vickery EM, et al. Vitamin D Supplementation for Prevention of Cancer: The D2d Cancer Outcomes (D2dCA) Ancillary Study. J Clin Endocrinol Metab 2021;106:2767-2778.
57. Virtanen JK, Nurmi T, Aro A, et al. Vitamin D supplementation and prevention of cardiovascular disease and cancer in the Finnish Vitamin D Trial: a randomized controlled trial. Am J Clin Nutr 2022;115:1300-1310.
58. Tobias DK, Luttmann-Gibson H, Mora S, et al. Association of Body Weight With Response to Vitamin D Supplementation and Metabolism. JAMA Network Open 2023;6:e2250681-e2250681.
59. Keum N, Chen QY, Lee DH, et al. Vitamin D supplementation and total cancer incidence and mortality by daily vs. infrequent large-bolus dosing strategies: a meta-analysis of randomised controlled trials. British Journal of Cancer 2022;127:872-878.
60. Fields PM, Jr., Butterly LF, Anderson JC. Inclusion of Carcinoids in Early Onset Colorectal Tumor Incidence Rates: Adenocarcinoma in Young Adults Still the Major Problem. Gastroenterology 2021;160:2613-2615.
61. Levy AD, Sobin LH. From the archives of the AFIP: Gastrointestinal carcinoids: imaging features with clinicopathologic comparison. Radiographics 2007;27:237-57.
62. Meester RGS, Mannalithara A, Lansdorp-Vogelaar I, et al. Trends in Incidence and Stage at Diagnosis of Colorectal Cancer in Adults Aged 40 Through 49 Years, 1975-2015.

Author names in bold designate shared co-first authorship.

Table 1. Estimated[a] mean values (95% confidence intervals) and adjusted[a] proportion (95% confidence intervals) of baseline characteristics by baseline 25(OH)D levels among young adults (under the age of 50) and older adults

Characteristics	Vitamin D levels (ng/mL)			*p*–trend
	<10	10-19	≥20	
Among participants aged under 50 (*n* = 212,885)				
Number of participants (%)	34,326 (16.1)	125,812 (59.1)	52,747 (24.8)	
Age (years)	35.1 (35.0-35.2)	35.6 (35.6-35.7)	36.6 (36.6-36.7)	<0.001
Male (%)	31.6 (31.1-32.1)	54.9 (54.6-55.2)	65.7 (65.3-66.1)	<0.001
Alcohol intake (%) [b]	17.4 (16.9-17.8)	21.8 (21.6-22)	26.3 (26-26.7)	<0.001
Current smoker (%)	18.5 (18.0-19.0)	19.8 (19.6-20)	22.3 (22-22.6)	<0.001
HEPA (%)	12.5 (12.1-12.8)	14.8 (14.6-15)	18.1 (17.8-18.5)	<0.001
Education level (%) [c]	77.8 (77.4-78.3)	81.6 (81.3-81.8)	83.1 (82.8-83.4)	<0.001
History of diabetes (%)	1.3 (1.1-1.4)	1.2 (1.2-1.3)	1.3 (1.2-1.3)	0.886
Family history of colorectal cancer (%)	2.5 (2.3-2.7)	2.6 (2.5-2.7)	2.6 (2.5-2.7)	0.375
Multi–vitamin supplement (%)	6.6 (6.3-6.9)	10.9 (10.7-11.0)	17.3 (17-17.6)	<0.001
Vitamin D supplement (%)	0.4 (0.3-0.4)	0.8 (0.8-0.9)	2.7 (2.6-2.9)	<0.001
Calcium supplement (%)	0.3 (0.2-0.3)	0.5 (0.5-0.6)	1.4 (1.3-1.5)	<0.001
Obesity (%) [d]	27.0 (26.5-27.5)	27.8 (27.5-28)	27.5 (27.2-27.9)	0.287
Body mass index (BMI) (kg/m^2)	23.2 (23.2-23.2)	23.2 (23.2-23.3)	23.2 (23.2-23.2)	0.331
Total energy intake (kcal/d)[e]	1567.5 (1559.3-1575.6)	1561 (1556.8-1565.1)	1525.2 (1518.7-1531.6)	<0.001
Among participants aged over 50 (*n* = 23,497)				
Number of participants (%)	2,594 (11.0)	11,414 (48.6)	9,489 (40.4)	
Age (years)	57.9 (57.7-58.1)	57.8 (57.7-58)	59.2 (59.1-59.3)	<0.001
Male (%)	29.2 (27.4-30.9)	53.6 (52.7-54.5)	62.1 (61.1-63.1)	<0.001
Alcohol intake (%) [b]	19.0 (17-21)	24.2 (23.4-25)	29 (28.1-29.8)	<0.001
Current smoker (%)	20.5 (18.4-22.5)	17.0 (16.3-17.8)	17.3 (16.6-18.1)	0.138
HEPA (%)	20.2 (18.5-21.8)	25.5 (24.7-26.3)	29.5 (28.6-30.4)	<0.001
Education level (%) [c]	43.5 (41.6-45.4)	44.5 (43.6-45.3)	44.3 (43.3-45.2)	0.727
History of diabetes (%)	13.2(11.8-14.6)	10.9 (10.4-11.5)	10.3 (9.8-10.9)	<0.001
Family history of colorectal cancer (%)	5.2 (4.3-6)	4.8 (4.4-5.2)	4.7 (4.3-5.1)	0.362
Multi–vitamin supplement (%)	9.2 (8.2-10.3)	15.2 (14.6-15.9)	29.8 (28.9-30.8)	<0.001
Vitamin D supplement (%)	0.8 (0.5-1.1)	1.7 (1.4-1.9)	7.1 (6.5-7.6)	<0.001

Calcium supplement (%)	0.8 (0.5-1.1)	1.7 (1.5-2)	5.7 (5.2-6.2)	<0.001
Obesity (%) [d]	33.4 (31.5-35.3)	35.6 (34.7-36.5)	32.7 (31.8-33.6)	0.010
Body mass index (BMI) (kg/m^2)	23.9 (23.8-24)	24.1 (24.1-24.2)	23.9 (23.8-24)	0.013
Total energy intake (kcal/d)[f]	1565.2 (1536.1-1594.2)	1539.5 (1525.6-1553.3)	1506.1 (1490.7-1521.6)	<0.001

[a]Adjusted for age and sex; [b] ≥20 g/day; [c] ≥ college graduate; [d] BMI ≥ 25 kg/m^2; [e] among 159,182 participants and [f] 12,767 participants with plausible estimated energy intake levels (within three standard deviations from the log-transformed mean energy intake); HEPA was defined as follows: (1) vigorous activity ≥ 3 days/week with ≥1,500 accumulated metabolic equivalent (MET)-min/week, or (2) a combination of walking, moderate, or vigorous intensity activities for 7 days accumulating to ≥3,000 MET-min/week.

Abbreviations: HEPA, health-enhancing physical activity

Table 2. Development of CRC by 25(OH)D levels among young adults (under the age of 50) and older adults

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	HR (95% CI)[a] in a model with time-dependent variables
Among participants aged under 50 (*n* = 212,885)				
<10	211,796.1	49	23.1	1.00 (reference)
10-19	754,687.6	130	17.2	0.61 (0.43-0.86)
≥20	299,126.1	50	16.7	0.41 (0.27-0.63)
p–trend				<0.001
Among participants aged over 50 (*n* = 23,497)				
<10	15,752.9	17	107.9	1.00 (reference)
10-19	63,732.6	54	84.7	0.75 (0.43-1.32)
≥20	48,646.8	41	84.3	0.52 (0.28-0.94)
p–trend				0.018

The *p*-value for the interaction of age and 25(OH)D levels with the risk of CRC was 0.793.
[a] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

Table 3. Development of CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness among participants under the age of 50 (*n* = 212,885)

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	HR (95% CI)[a] in a model with time-dependent variables
Histologic subtype				
Adenocarcinoma				
<10	211,796.1	34	16.1	1.00 (reference)
10-19	754,687.6	86	11.4	0.59 (0.38-0.91)
≥20	299,126.1	33	11	0.46 (0.27-0.77)
p–trend				0.005
Other types				
<10	211,796.1	13	6.1	1.00 (reference)
10-19	754,687.6	33	4.4	0.81 (0.40-1.63)
≥20	299,126.1	16	5.3	0.51 (0.22-1.23)
p–trend				0.113
Subsite (colon vs rectum)				
Colon cancer				
<10	211,796.1	26	12.3	1.00 (reference)
10-19	754,687.6	66	8.7	0.55 (0.34-0.91)
≥20	299,126.1	21	7.0	0.32 (0.17-0.59)
p–trend				<0.001
Rectal cancer				
<10	211,796.1	24	11.3	1.00 (reference)
10-19	754,687.6	65	8.6	0.72 (0.43-1.20)
≥20	299,126.1	30	10.0	0.61 (0.34-1.12)
p–trend				0.129
Invasiveness				
Invasive cancer				
<10	211,796.1	39	18.4	1.00 (reference)
10-19	754,687.6	94	12.5	0.63 (0.42-0.95)

≥20	299,126.1	36	12.0	0.42 (0.26-0.70)
p–trend				0.001
Carcinoma in situ, colorectal neoplasms				
<10	211,796.1	10	4.7	1.00 (reference)
10-19	754,687.6	36	4.8	0.62 (0.30-1.26)
≥20	299,126.1	14	4.7	0.46 (0.20-1.05)
p–trend				0.076

[a] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
There were three overlapping cases of both colon and rectal cancers: two cases with synchronous CRC at different site (C20, C18.7; C20, C18.9) and one case with metachronous case (C20.9, C18.2).
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

Table 4. Development of CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness among participants over the age of 50 (*n* = 23,497)

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	HR (95% CI)[a] in a model with time-dependent variables
Histologic subtype				
Adenocarcinoma				
<10	15,752.9	15	95.2	1.00 (reference)
10-19	63,732.6	52	81.6	0.73 (0.40-1.34)
≥20	48,646.8	35	71.9	0.43 (0.22-0.84)
p–trend				0.005
Other types				
<10	15,752.9	0	0	1.00 (reference)
10-19	63,732.6	2	3.1	-
≥20	48,646.8	4	8.2	-
p–trend				0.525
Subsite (colon vs rectum)				
Colon cancer				
<10	15,752.9	12	76.2	1.00 (reference)
10-19	63,732.6	38	59.6	0.71 (0.36-1.42)
≥20	48,646.8	26	53.4	0.44 (0.21-0.95)
p–trend				0.022
Rectal cancer				
<10	15,752.9	5	31.7	1.00 (reference)
10-19	63,732.6	16	25.1	0.61 (0.22-1.71)
≥20	48,646.8	17	34.9	0.48 (0.16-1.43)
p–trend				0.213
Invasiveness				

Invasive cancer				
<10	15,752.9	15	95.2	1.00 (reference)
10-19	63,732.6	35	54.9	0.55 (0.29-1.06)
≥20	48,646.8	32	65.8	0.40 (0.20-0.82)
p–trend				0.018
Carcinoma in situ, colorectal neoplasms				
<10	15,752.9	2	12.7	1.00 (reference)
10-19	63,732.6	19	29.8	1.24 (0.36-4.33)
≥20	48,646.8	9	18.5	0.60 (0.15-2.39)
p–trend				0.189

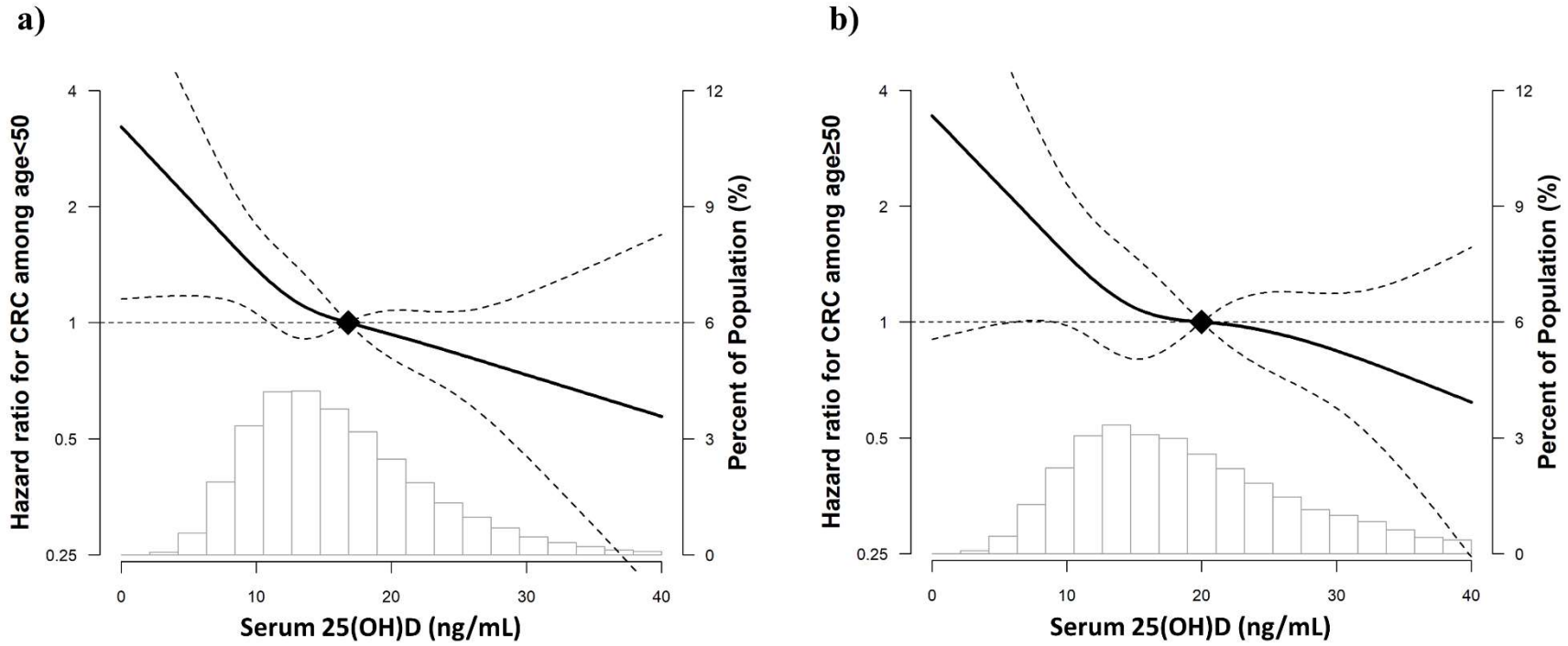
[a] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
There were two overlapping cases of both colon and rectal cancers, all synchronous CRC at different site (C20, C18.7; C20, C18.9)
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

FIGURE LEGENDS

Figure 1 Multivariable-adjusted hazard ratios for the development of colorectal cancer (CRC) with restricted cubic splines

a) Age under 50 years

b) Age over 50 years


a)
b)
Hazard ratio for CRC among age<50
Hazard ratio for CRC among age≥50
Percent of Population (%)
Serum 25(OH)D (ng/mL)
4
2
1
0.5
0.25
12
9
6
3
0
0
10
20
30
40

Serum 25-hydroxyvitamin D levels and risk of colorectal cancer: an age-stratified analysis

Design

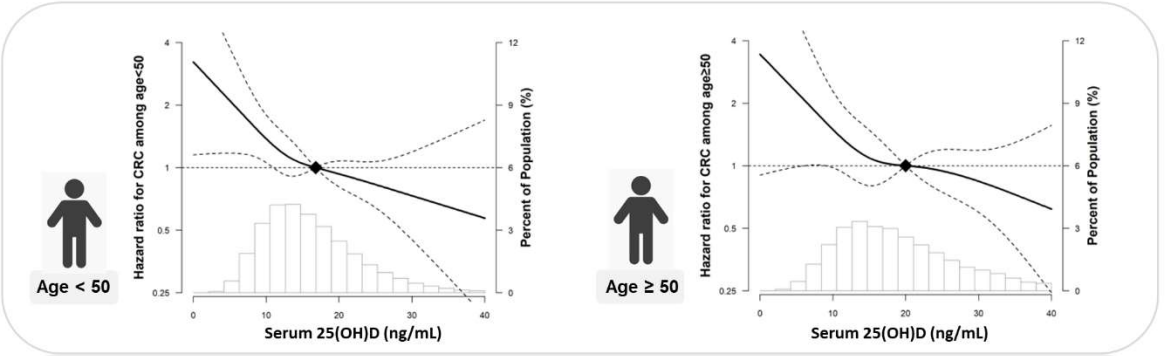
A retrospective cohort study

236,382 participants (mean age, 38.0 years) who underwent a comprehensive health examination

Median follow-up 6.5 years; 341 participants developed incident colorectal cancer (CRC) (incidence rate, 19.2 per 10^5 person-years).

Results

Risk of developing CRC according to serum 25(OH)D levels



- Serum 25(OH)D levels were inversely associated with the risk of incident CRC both in young and older adults. No significant age-associated differences were found (*p* for interaction = 0.793).
- Significant associations were evident for adenocarcinoma, colon cancer, and invasive cancers.

Gastroenterology

SUPPLEMENTARY MATERIALS

Title: Serum 25-hydroxyvitamin D levels and risk of colorectal cancer: an age-stratified analysis

Table of Contents

eTable 9. Development of early-onset CRC by 25(OH)D levels when patients diagnosed with CRC after the age of 50 were censored at the age of 50 (*n* = 212,885)

eTable 10. Development of early-onset CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness e when patients diagnosed with CRC after the age of 50 were censored at the age of 50 (*n* = 212,885)

eTable 11. Hazard ratios (95% CIs) for CRC by 25(OH)D levels when restricted to incident cases of CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels among participants under the age of 50

eTable 12. Hazard ratios (95% CIs) for CRC by 25(OH)D levels when restricted to incident cases of CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels among participants over the age of 50

eTable 13. Population attributable fraction of CRC due to 25(OH)D deficiency (<10 ng/mL) among young adults (under the age of 50) and older adults

eTable 14. Hazard ratios (95% CIs) of incident early-onset CRC according to serum 25(OH)D levels by clinically relevant subgroups (*n* = 236,382)

S1-1. Methods: Completeness of the cancer registry

The diagnosis of colorectal cancer (CRC) was ascertained by linkage to the Korea Central Cancer Registry (KCCR). The KCCR annually collects data on newly diagnosed cancer patients from hospitals across the country during the previous year. Since 2002, the KCCR has combined the KCCR mother database with nine other databases. These databases include site-specific cancer registries (such as breast, ovary, cervix, and liver cancer), data from additional medical record review surveys, and cancer mortality data from the National Statistical Office (NSO), to create the Korea National Cancer Incidence Data Bases (KNCIDB). The KCCR gathers information on primary tumor site, histology, date of diagnosis, stage at diagnosis, treatment within four months of diagnosis, patient age, and sex. Data from the previous year is combined with information collected by the central and 11 regional cancer registries, including data on cancer patients missed in hospital-based registrations, resulting in a two-year process to calculate the year's KNCIDB and cancer statistics. The 2019 KNCIDB was estimated to be 98.3% complete [1].

S1-2. Methods: Additional details on self-reported data collection

1) Physical activity: In our study, physical activity levels were assessed using the validated Korean version of the International Physical Activity Questionnaire (IPAQ) short form. The data were then converted to metabolic equivalents (METs; min/week) and divided into inactive, minimally active, or health-enhancing physical activity (HEPA) [2]. Health-enhancing physical activity (HEPA) was defined as physical activity that meets one of two following criteria: (1) vigorous activity on three or more days per week with ≥1,500 of accumulated metabolic equivalent (MET)-minutes/week (1 MET is energy expenditure at rest); or (2) seven days of any combination of walking, moderate intensity, or vigorous intensity activities achieving at least 3,000 MET min/week.[2].

IPAQ was first developed by the World Health Organization in 1998, and its questions have been adapted to cultural diversity and use country-specific examples to help define physical activity and sitting. The short form of the IPAQ consists of 7 items and the IPAQ long form consists of 27 items, both of which have been validated against accelerometer measurements as a gold standard in 12 countries. However, the short form is generally preferred because it is simpler and has equivalent psychometric properties to the long form. According to a previous study published in Korea reporting reliability and validity of IPAQ short form, the Kappa value was estimated to be 0.365~0.620 (median 0.471) and was above 0.4 in 5 out of 7 questionnaires [2, 3].

2) Smoking status: For smoking status, participants who had smoked < 100 cigarettes during their lifetime were classified as never-smokers. Participants who had smoked > 100 cigarettes in their lifetime were further categorized as (1) current smokers who smoked currently or (2) former smokers who no longer smoked at the time of their screening examination. This definition also has widely been used in major cohort studies such as Nurses' Health Study (NHS) and UK biobank [4, 5].

3) Alcohol consumption: The amount of alcohol consumed per drinking day was recorded in units of 'soju,' which is the most popular alcoholic beverage in Korea. The alcoholic content in soju was estimated to be 20%, and the traditional unit of jan of soju (1 jan of soju=50 mL) contained 8 g [6]. Current alcohol use was assessed as the frequency of alcohol consumption and the amount of alcohol consumed per drinking day using the following questions [7]: "How often do you drink alcohol in a week on average (please fill in the details that apply to your current situation)?" and "How much alcohol do you usually drink per drinking day?". Then, these data were used to estimate the average alcohol consumption per day.

REFERENCES

1. Kang MJ, Won YJ, Lee JJ, et al. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2019. Cancer Res Treat 2022;54:330-344.
2. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35:1381-95.
3. Chun MY. Validity and reliability of korean version of international physical activity questionnaire short form in the elderly. Korean J Fam Med 2012;33:144-51.
4. Carter BD, Abnet CC, Feskanich D, et al. Smoking and Mortality — Beyond Established Causes. New England Journal of Medicine 2015;372:631-640.
5. Kar D, El-Wazir A, Delanerolle G, et al. Predictors and determinants of albuminuria in people with prediabetes and diabetes based on smoking status: A cross-sectional study using the UK Biobank data. eClinicalMedicine 2022;51.
6. Chang Y, Ryu S, Kim Y, et al. Low Levels of Alcohol Consumption, Obesity, and Development of Fatty Liver With and Without Evidence of Advanced Fibrosis. Hepatology 2020;71.
7. Chang Y, Cho YK, Cho J, et al. Alcoholic and Nonalcoholic Fatty Liver Disease and Liver-Related Mortality: A Cohort Study. Am J Gastroenterol 2019;114:620-629.

eTable 1. Characteristics for 341 colorectal cancers by age groups

	n (%)	
	among subjects under the age of 50 (*n* = 229)	among subjects over the age of 50 (*n* = 112)
Cancer type*		
Carcinoma in situ	60 (26.2)	30 (26.8)
Invasive	169 (73.8)	82 (73.2)
Location (ICD-O-3 code)		
Cecum (C18.0)	9 (3.9)	5 (4.5)
Appendix (C18.1)	6 (2.6)	2 (1.8)
Ascending colon (C18.2)	16 (7.0)	23 (20.5)
Hepatic flexture (C18.3)	3 (1.3)	4 (3.6)
Transverse colon (18.4)	8 (3.5)	6 (5.4)
Splenic flexture (C18.5)	2 (0.9)	1 (0.9)
Descending colon (C18.6)	4 (1.8)	3 (2.7)
Sigmoid colon (C18.7)	52 (22.7)	25 (22.3)
Rectosigmoid colon (C19.9)	12 (5.2)	9 (3.4)
Rectum (C20.9)	109 (47.6)	28 (25.0)
Colon, unspecified (C18.9)	12 (5.2)	5 (4.5)
SEER		
Localized (including carcinoma in situ)	169 (73.8)	72 (64.3)
Regional	33 (14.4)	26 (23.2)
Distant	12 (5.2)	12 (10.7)
Unknown	15 (6.6)	2 (1.8)
Histology		
Adenocarcinoma in situ	55 (24.0)	29 (25.9)
Adenocarcinoma	98 (22.8)	75 (67.0)
Carcinoid tumor**	61 (26.6)	6 (5.4)
Stromal sarcoma	3 (1.3)	0 (0.0)
Others***	12 (5.2)	2 (1.8)

Unless otherwise specified, data are numbers of participants with percentages in parentheses. SEER = surveillance, epidemiology, and end results.
*ICD-O-3, behavioral code,"/2" for carcinoma in situ and "/3" for invasive malignant cancer

****Location of carcinoid tumor**: out of 61 cases among subjects under the age of 50-

52 cases in rectum, 9 cases in colon (1 in cecum, 2 in appendix, 1 in sigmoid colon and 5 in colon, unspecified); out of 6 cases among subjects over the age of 50- 5 cases in rectum and 1 case in splenic flexture.

For carcinoid, the ICD-O-3 morphology codes included 8013/3, 8240/3, 8241/3, 8243/3, 8244/3, 8245/3, 8246/3, and 8249/3; for adenocarcinoma, morphology codes included 8140/3, 8141/3, 8143/3, 8144/3, 8210/3, 8211/3, 8213/3, 8220/3, 8221/3, 8260-8265/3, 448 8255/3, 8260-8263/3, 8310/3, 8323/3, 8440/3, 8460/3, 8470/3, 8472/3, 8480-8482/3, 8490/3, 8570/3, 449 8574/3, 8576/3.

***out of 12 cases among subjects under the age of 50: 8000/3 (malignant neoplasm, others, n=2), 8010/2 (carcinoma in situ, NOS, n=4), 8148/2 (Glandular intraepithelial neoplasia, high grade, n=1), 9680/3 (diffuse large B-cell lymphoma, NOS, n=3), 9691/3 (follicular lymphoma, n=1), 9699/3 (Marginal zone B-cell lymphoma, NOS, n=1); out of 2 cases among subjects over the age of 50: 8000/2 (neoplasm in situ, others, n=1), 9680/3 (diffuse large B-cell lymphoma, NOS, n=1)

eTable 2. Development of CRC by 25(OH)D levels among young adults (under the age of 50) and older adults (age-sex- and multivariable-adjusted models)

25(OH)D levels (**ng/mL**)	Age & sex-adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)
Clinical categories		
Participants **aged < 50** (*n* = 212,885)		
<10	1.00 (reference)	1.00 (reference)
10-19	0.65 (0.47-0.91)	0.60 (0.43-0.84)
≥20	0.56 (0.37-0.84)	0.47 (0.31-0.72)
p–trend	0.008	0.001
Participants **aged ≥ 50** (n = 23,497)		
<10	1.00 (reference)	1.00 (reference)
10-19	0.76 (0.44-1.31)	0.70 (0.40-1.21)
≥20	0.66 (0.37-1.16)	0.56 (0.31-1.00)
p–trend	0.171	0.058

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.

Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 3. Development of CRC for every 5 ng/ml increase in 25(OH)D levels among young adults (under the age of 50) and older adults

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	HR (95% CI)[a] in a model with time-dependent variables
All CRC				
participants aged < 50	1,265,609.8	229	18.1	0.84 (0.75-0.93)
participants aged ≥ 50	128,132.3	112	87.4	0.87 (0.78-0.97)

[a] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.

eTable 4. Development of CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness among young adults (under the age of 50) (age-sex- and multivariable-adjusted models)

25(OH)D levels (**ng/mL**)	Participants aged < 50 (*n* = 212,885)	
	Age & sex-adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)
Histologic subtype		
Adenocarcinoma		
<10	1.00 (reference)	1.00 (reference)
10-19	0.62 (0.41-0.94)	0.62 (0.41-0.94)
≥20	0.53 (0.32-0.87)	0.52 (0.30-0.89)
p–trend	0.017	0.021
Other types		
<10	1.00 (reference)	1.00 (reference)
10-19	0.65 (0.34-1.26)	0.64 (0.32-1.25)
≥20	0.74 (0.35-1.60)	0.68 (0.30-1.54)
p–trend	0.543	0.432
Anatomic site-specific CRC		
Colon cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.61 (0.38-0.98)	0.57 (0.35-0.93)
≥20	0.43 (0.23-0.78)	0.37 (0.19-0.70)
p–trend	0.006	0.002
Rectal cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.57 (0.35-0.93)	0.69 (0.42-1.12)
≥20	0.37 (0.20-0.70)	0.71 (0.39-1.28)
p–trend	0.002	0.322
Invasiveness		
Invasive cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.61 (0.42-0.90)	0.58 (0.39-0.86)
≥20	0.53 (0.33-0.85)	0.49 (0.29-0.80)
p–trend	0.012	0.007
Carcinoma in situ		
<10	1.00 (reference)	1.00 (reference)
10-19	0.88 (0.43-1.80)	0.86 (0.41-1.80)
≥20	0.75 (0.32-1.74)	0.67 (0.27-1.65)
p–trend	0.487	0.363

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.

Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 5. Development of CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness among individuals over the age of 50 (age-sex- and multivariable-adjusted models)

25(OH)D levels (**ng/mL**)	Participants aged ≥ 50 (n = 23,497)	
	Age & sex-adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)
Histologic subtype		
Adenocarcinoma		
<10	1.00 (reference)	1.00 (reference)
10-19	0.78 (0.44-1.40)	0.66 (0.36-1.20)
≥20	0.60 (0.32-1.11)	0.43 (0.22-0.83)
p–trend	0.083	0.008
Other types		
<10	1.00 (reference)	1.00 (reference)
10-19	-	-
≥20	-	-
p–trend		
Anatomic site-specific CRC		
Colon cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.73 (0.38-1.41)	0.64 (0.33-1.26)
≥20	0.56 (0.28-1.12)	0.45 (0.21-0.96)
p–trend	0.096	0.040
Rectal cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.69 (0.25-1.92)	0.54 (0.19-1.52)
≥20	0.85 (0.31-2.37)	0.47 (0.16-1.41)
p–trend	0.960	0.271
Invasiveness		
Invasive cancer		
<10	1.00 (reference)	1.00 (reference)
10-19	0.52 (0.28-0.95)	0.43 (0.23-0.80)
≥20	0.54 (0.29-1.02)	0.37 (0.18-0.73)
p–trend	0.173	0.017
Carcinoma in situ		
<10	1.00 (reference)	1.00 (reference)
10-19	2.30 (0.53-10.02)	2.04 (0.46-8.95)
≥20	1.19 (0.25-5.63)	1.01 (0.21-4.91)
p–trend	0.513	0.492
p–trend	0.513	0.492

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.

Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 6. Development of CRC for every 5 ng/ml increase in 25(OH)D levels according to histologic subtype, subsite, and invasiveness among young adults (under the age of 50) and older adults

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	HR (95% CI)[a] in a model with time-dependent variables
Histologic subtype				
Adenocarcinoma				
participants aged < 50	1,265,609.8	153	12.1	0.83 (0.73-0.95)
participants aged ≥ 50	128,132.3	102	79.6	0.85 (0.75-0.96)
Other types				
participants aged < 50	1,265,609.8	62	4.9	0.92 (0.76-1.12)
participants aged ≥ 50	128,132.3	6	4.7	1.07 (0.69-1.67)
Anatomic site-specific CRC				
Colon cancer				
participants aged < 50	1,265,609.8	113	8.9	0.75 (0.63-0.88)
participants aged ≥ 50	128,132.3	76	59.3	0.85 (0.74-0.98)
Rectal cancer				
participants aged < 50	1,265,609.8	119	9.4	0.95 (0.83-1.08)
participants aged ≥ 50	128,132.3	38	29.7	0.84 (0.69-1.03)
Invasiveness				
Invasive cancer				
participants aged < 50	1,265,609.8	169	13.4	0.83 (0.73-0.94)
participants aged ≥ 50	128,132.3	82	64.0	0.88 (0.77-1.01)
Carcinoma in situ				
participants aged < 50	1,265,609.8	60	4.7	0.89 (0.74-1.08)
participants aged ≥ 50	128,132.3	30	23.4	0.77 (0.60-0.98)

[a] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.

For participants aged <50, there were three overlapping cases of both colon and rectal cancers: two cases with synchronous CRC at different site (C20, C18.7; C20, C18.9) and one case with metachronous case (C20.9, C18.2). For participants aged ≥ 50, there were two overlapping cases of both colon and rectal cancers, all synchronous CRC at different site (C20, C18.7; C20, C18.9)

Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 7. Development of **adenocarcinoma** by 25(OH)D levels according to subsite among young adults (under the age of 50) and older adults

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	Age sex adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)	HR (95% CI)[b] in a model with time-dependent variables
Adenocarcinoma at the colon						
Participants aged <50 (n = 212,885)						
<10	211,796.1	21	9.9	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	754,687.6	55	7.3	0.62 (0.37-1.04)	0.56 (0.33-0.95)	0.57 (0.33-0.97)
≥20	299,126.1	19	6.4	0.46 (0.25-0.88)	0.39 (0.20-0.75)	0.37 (0.19-0.70)
p–trend				0.021	0.006	0.003
Participants aged ≥50 (n = 23,497)						
<10	15,752.9	10	63.5	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	63,732.6	37	58.1	0.89 (0.44-1.80)	0.85 (0.42-1.74)	0.94 (0.45-1.98)
≥20	48,646.8	24	49.3	0.63 (0.30-1.33)	0.59 (0.27-1.27)	0.58 (0.26-1.28)
p–trend				0.146	0.112	0.069
Adenocarcinoma at the rectum						
Participants aged <50 (n = 212,885)						
<10	211,796.1	14	6.6	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	754,687.6	32	4.2	0.57 (0.30-1.08)	0.50 (0.26-0.96)	0.53 (0.27-1.05)
≥20	299,126.1	14	4.7	0.56 (0.26-1.21)	0.44 (0.20-0.97)	0.51 (0.24-1.11)
p–trend				0.174	0.059	0.162
Participants aged ≥50 (n = 23,497)						
<10	15,752.9	5	31.7	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	63,732.6	15	23.5	0.71 (0.26-1.98)	0.60 (0.22-1.69)	0.61 (0.22-1.70)
≥20	4,8646.8	13	26.7	0.72 (0.25-2.03)	0.53 (0.18-1.57)	0.50 (0.17-1.45)
p–trend				0.642	0.330	0.245

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.

[b] Estimated from Cox proportional hazard models with tertiles of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
Abbreviations: BMI, body mass index; CI, confidence interval; HR, hazard ratio

eTable 8. Development of CRC by season-calibrated 25(OH)D categories among young adults (under the age of 50) and older adults

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	Age sex adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)	HR (95% CI)[b] in a model with time-dependent variables
Season-calibrated 25(OH) levels, approach 1[a]						
Participants aged <50 (n = 211,105)						
<10	145,866.3	36	24.7	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	752559	137	18.2	0.65 (0.45-0.95)	0.61 (0.42-0.89)	0.6 (0.41-0.89)
≥20	367184.4	56	15.3	0.48 (0.31-0.74)	0.41 (0.27-0.65)	0.4 (0.25-0.61)
p–trend				0.001	<0.001	<0.001
Participants aged ≥50 (n = 22,726)						
<10	9,388.9	12	127.8	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	57,239	48	83.9	0.63 (0.33-1.19)	0.59 (0.31-1.12)	0.71 (0.35-1.40)
≥20	61,504.4	52	84.5	0.54 (0.29-1.03)	0.48 (0.25-0.91)	0.55 (0.28-1.10)
p–trend				0.101	0.040	0.068
Season-calibrated 25(OH) levels, approach 2[b]						
Participants aged <50 (n = 211,105)						
<10	211,796.1	49	23.1	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	754,687.6	130	17.2	0.78 (0.53-1.15)	0.73 (0.50-1.08)	0.53 (0.37-0.77)
≥20	299,126.1	50	16.7	0.53 (0.34-0.83)	0.46 (0.29-0.73)	0.39 (0.26-0.60)
p–trend				0.003	<0.001	<0.001
Participants aged ≥50 (n = 22,726)						
<10	15,752.9	17	107.9	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	63,732.6	54	84.7	0.88 (0.52-1.49)	0.83 (0.49-1.41)	0.87 (0.50-1.49)
≥20	48,646.8	41	84.3	0.61 (0.34-1.1)	0.53 (0.29-0.97)	0.54 (0.29-0.98)
p–trend				0.063	0.022	0.019

The *P*-value for the interaction of age and 25(OH)D levels with the risk of CRC was 0.794 in season calibrated model 1, 0.916 in season calibrated model 2, and 0.324 for season calibrated time-dependent model.

[a] Vitamin D levels were standardized by adding the overall mean of vitamin D for all subjects to the residuals derived from a regression of vitamin D levels on the periodic function two variables (R= cos([2π/12]*X) and S= sin([2π/12]*X), where X is the month of blood measurement.

[b] Vitamin D levels were standardized by adding the overall mean of vitamin D for all subjects to the residuals derived from a regression of vitamin D levels on the periodic function two variables (R= cos([2π/12]*X) and S= sin([2π/12]*X), where X is the month of blood measurement as well age and sex.
[c] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.
[d] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 9. Development of early-onset CRC by 25(OH)D levels when patients diagnosed with CRC after the age of 50 were censored at the age of 50 (*n* = 212,885)

25(OH)D levels (**ng/mL**)	Person year (PY)	Incident cases	Incidence density (/ 10^5 PY)	Age sex adjusted HR (95% CI)	Multivariable-adjusted HR[a] (95% CI)	HR (95% CI)[b] in a model with time-dependent variables
As a categorical variable						
<10	17,8455.2	42	23.5	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	63,3257.4	115	18.2	0.65 (0.45-0.93)	0.63 (0.44-0.92)	0.69 (0.46-1.01)
≥20	24,2105.7	43	17.8	0.52 (0.33-0.81)	0.49 (0.31-0.79)	0.49 (0.31-0.79)
p–trend				0.005	0.005	0.003
As a continuous variable[c]						
Per 5 ng/ml increase	1,053,818.3	200.0	19.0	0.83 (0.74-0.94)	0.82 (0.72-0.93)	0.86 (0.77-0.97)

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.
[b] Estimated from Cox proportional hazard models with categories of vitamin D levels, smoking, alcohol consumption, physical activity, season, total energy intake, BMI, history of diabetes, multi-vitamin supplements, and calcium supplements as time-dependent variables, baseline age, center, year of screening exam, education level and family history of CRC as time-fixed variables.
[c] Models were separately constructed; one using 25(OH) levels as a categorical variable and another using it as a continuous variable.

Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 10. Development of early-onset CRC by 25(OH)D levels according to histologic subtype, subsite, and invasiveness when patients diagnosed with CRC after the age of 50 were censored at the age of 50 (n = 212,885)

25(OH)D levels (**ng/mL**)	Multivariable-adjusted HR[a]					
	Histologic subtype		Subsite (colon vs rectum)[b]		Invasiveness	
	Adenocarcinoma	Other types	Colon cancer	Rectal cancer	Invasive cancer	Carcinoma in situ, colorectal neoplasms
As a categorical variable						
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.63 (0.40-1.01)	0.63 (0.31-1.26)	0.54 (0.32-0.91)	0.71 (0.42-1.21)	0.55 (0.36-0.84)	1.03 (0.44-2.42)
≥20	0.49 (0.27-0.89)	0.65 (0.28-1.53)	0.29 (0.14-0.59)	0.74 (0.39-1.40)	0.44 (0.26-0.76)	0.74 (0.27-2.08)
p–trend	0.023	0.403	0.001	0.451	0.004	0.466
As a continuous variable[c]						
Per 5 ng/ml increase	0.75 (0.63-0.89)	0.96 (0.77-1.19)	0.70 (0.57-0.85)	0.93 (0.79-1.09)	0.82 (0.71-0.95)	0.81 (0.63-1.04)

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.
[b] There were three overlapping cases of both colon and rectal cancers: two cases with synchronous CRC at different site (C20, C18.7; C20, C18.9) and one case with metachronous case (C20.9, C18.2).
[c] Models were separately constructed; one using 25(OH) levels as a categorical variable and another using it as a continuous variable.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 11. Hazard ratios (95% CIs) for CRC by 25(OH)D levels when restricted to incident cases of CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels among participants under the age of 50

25(OH)D levels (**ng/mL**)	Multivariable-adjusted HR[a] (95% CI)		
	Restricted to 178 incident cases of CRC occurring from 2 years after the baseline measurement	Restricted to 153 incident cases of CRC occurring from 3 years after the baseline measurement	Restricted to 127 incident cases of CRC occurring from 4 years after the baseline measurement
Colorectal cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.63 (0.42-0.93)	0.59 (0.39-0.90)	0.57 (0.36-0.91)
≥20	0.59 (0.37-0.95)	0.56 (0.34-0.94)	0.63 (0.36-1.10)
p–trend	0.046	0.045	0.188
Colon cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.58 (0.34-0.97)	0.57 (0.33-1.00)	0.57 (0.31-1.08)
≥20	0.43 (0.22-0.83)	0.46 (0.23-0.93)	0.61 (0.29-1.30)
p–trend	0.014	0.036	0.264
Rectal cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.65 (0.36-1.18)	0.57 (0.31-1.06)	0.50 (0.26-0.99)
≥20	0.78 (0.39-1.54)	0.67 (0.32-1.38)	0.61 (0.28-1.35)
p–trend	0.666	0.426	0.384
Invasive cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.59 (0.37-0.94)	0.52 (0.32-0.85)	0.50 (0.29-0.86)
≥20	0.61 (0.35-1.07)	0.53 (0.30-0.96)	0.58 (0.31-1.10)
p–trend	0.136	0.071	0.204
Carcinoma in situ			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.73 (0.35-1.54)	0.83 (0.36-1.89)	0.79 (0.32-1.94)
≥20	0.54 (0.21-1.35)	0.64 (0.23-1.78)	0.76 (0.25-2.28)

p–trend	0.181	0.377	0.639
Adenocarcinoma			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.59 (0.38-0.93)	0.52 (0.33-0.84)	0.51 (0.30-0.86)
≥20	0.52 (0.30-0.90)	0.47 (0.26-0.85)	0.57 (0.30-1.07)
p–trend	0.028	0.019	0.131
Other types			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.74 (0.29-1.93)	0.72 (0.25-2.07)	0.65 (0.20-2.07)
≥20	1.26 (0.44-3.59)	1.21 (0.38-3.82)	0.98 (0.27-3.48)
p–trend	0.425	0.494	0.734

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 12. Hazard ratios (95% CIs) for CRC by 25(OH)D levels when restricted to incident cases of CRC occurring from 2 to 4 years after the initial measurement of 25(OH)D levels among participants over the age of 50

25(OH)D levels (**ng/mL**)	Multivariable-adjusted HR[a] (95% CI)		
	Restricted to 70 incident cases of CRC occurring from 2 years after the baseline measurement	Restricted to 56 incident cases of CRC occurring from 3 years after the baseline measurement	Restricted to 45 incident cases of CRC occurring from 4 years after the baseline measurement
Colorectal cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.67 (0.34-1.30)	0.62 (0.29-1.30)	0.80 (0.34-1.89)
≥20	0.58 (0.28-1.17)	0.60 (0.27-1.30)	0.61 (0.24-1.54)
p–trend	0.167	0.295	0.272
Colon cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.76 (0.35-1.64)	0.71 (0.31-1.64)	0.85 (0.33-2.16)
≥20	0.53 (0.23-1.25)	0.52 (0.21-1.30)	0.54 (0.19-1.53)
p–trend	0.128	0.156	0.195
Rectal cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.50 (0.13-1.97)	0.42 (0.08-2.34)	0.82 (0.09-7.39)
≥20	0.82 (0.22-3.06)	1.08 (0.23-5.17)	1.40 (0.16-11.93)
p–trend	0.825	0.389	0.503
Invasive cancer			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.42 (0.19-0.92)	0.37 (0.15-0.90)	0.51 (0.18-1.47)
≥20	0.53 (0.24-1.16)	0.50 (0.21-1.22)	0.52 (0.17-1.54)
p–trend	0.320	0.377	0.382
Carcinoma in situ			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	2.04 (0.46-9.06)	0.95 (0.18-5.12)	1.79 (0.39-8.24)
≥20	0.79 (0.15-4.23)	0.78 (0.57-1.06)	0.89 (0.16-5.07)

p–trend	0.349	0.634	0.642
Adenocarcinoma			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	0.80 (0.39-1.66)	0.79 (0.35-1.78)	1.19 (0.45-3.18)
≥20	0.63 (0.29-1.37)	0.69 (0.29-1.63)	0.89 (0.31-2.53)
p–trend	0.22	0.41	0.611
Other types			
<10	1.00 (reference)	1.00 (reference)	1.00 (reference)
10-19	-	-	-
≥20	-	-	-
p–trend	0.242	-	-

[a] Estimated from Cox proportional hazards models. Multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements, and calcium supplements.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

eTable 13. Population attributable fraction of CRC due to 25(OH)D deficiency (<10 ng/mL) among young adults (under the age of 50) and older adults

	Among participants aged under 50		Among participants aged over 50
Population attributable fraction	5.4%		3.0%
Population attributable fraction =	$\frac{P_e(RR-1)}{1+P_e(RR-1)}$	x 100	where P_e is the exposed proportion of the population, and RR is the rate ratio

eTable 14. Hazard ratios[a] (95% CIs) of incident early-onset CRC according to serum 25(OH)D levels by clinically relevant subgroups (*n* = 236,382)

Subgroup	Vitamin D levels (ng/mL)			*p*–trend	*p* for interaction
	<10	10-19	≥20		
Sex					0.634
Women (n = 108,993)	Reference	0.66 (0.45-0.97)	0.43 (0.25-0.74)	0.002	
Men (n = 127,389)	Reference	0.60 (0.38-0.95)	0.52 (0.32-0.84)	0.025	
BMI					0.120
<25 kg/m^2 (n = 167,773)	Reference	0.55 (0.40-0.76)	0.41 (0.27-0.60)	<0.001	
≥25 kg/m^2 (n = 66,058)	Reference	1.07 (0.53-2.17)	0.96 (0.46-2.03)	0.750	
Family history of CRC					0.268
No (n = 227,264)	Reference	0.62 (0.46-0.84)	0.47 (0.33-0.68)	<0.001	
Yes (n = 6,567)	Reference	0.62 (0.13-3.10)	1.12 (0.23-5.44)	0.562	

[a]Estimated using Cox proportional hazard models. The multivariable model was adjusted for age, sex, center, year of screening examination, alcohol consumption, smoking, physical activity, season, total energy intake, education level, BMI, history of diabetes, family history of CRC, multi-vitamin supplements and calcium supplements except for the stratifying variable of each stratum.
Abbreviations: BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio

WHAT YOU NEED TO KNOW

Background and Context

The potential benefits of 25-hydroxyvitamin D (25(OH)D) in prevention of colorectal cancer, especially for early-onset disease, are not well understood.

New Findings

Serum 25(OH)D levels were significantly and inversely associated with risk of colorectal cancer development in younger individuals aged <50 years; the associations were particularly evident for adenocarcinoma, colon cancer, and invasive cancer. Among those aged ≥50 years, associations were similarly observed.

Limitations

Since our study population consisted of East Asian men and women, the generalizability to other ethnic groups may be limited.

Clinical Research Relevance

Maintaining an adequate 25(OH)D level needs to be tested as an alternative strategy for colorectal cancer prevention both in young and older adults.

LAY SUMMARY

Individuals with sufficient vitamin D levels may have a lower likelihood of developing colorectal cancer.