Biostatistics Research
https://ojs.wiserpub.com/index.php/BSR/

UNIVERSAL WISER PUBLISHER

Research Article

Nonoptimal Follow-up Times Make It Difficult to Detect the Epidemiological Inverse Relationship Between 25-Hydroxyvitamin D and Lung Cancer

Ari Voutilainen*, Jyrki K. Virtanen, Sari Hantunen, Tarja Nurmi, Petra Kokko, Tomi-Pekka Tuomainen

Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
E-mail: ari.voutilainen@uef.fi

Received: 19 May 2021; **Revised:** 5 August 2021; **Accepted:** 22 August 2021

Abstract: Background: Previous studies have reported controversial conclusions regarding the association between circulating 25-hydroxyvitamin D (25-HVD) and lung cancer risk. **Objectives:** To test the hypothesis that the controversial conclusions can be due to different follow-up times (FUT) and because of interpreting findings predominantly based on statistical significance. **Methods:** The Kuopio Ischaemic Heart Disease Risk Factor Study provided data. We used the Cox regression to study the association between 25-HVD and lung cancer risk in 2578 middle-aged Finnish men. Out of them, 808 were free of cancer and willing to participate in follow-up examinations 11 years after baseline. We repeated all analyses for them. **Results:** Higher circulating 25-HVD predicted lower lung cancer risk over the entire follow-up period of 33 years. The hazard ratio (HR) of the highest vs. the lowest 25-HVD tertile adjusted for age, smoking, alcohol consumption, body weight status, inflammatory status, physical activity, and diet was lowest, 0.39 (95% CI: 0.17-0.87), when the FUT was 15 years. The HR was statistically significant ($p < 0.05$) only when the FUT was 15-17 years. In the sub-cohort, Cohen's *d* denoted a large or medium effect during the first 11 years. **Conclusions:** The optimal FUT in this prospective cohort study investigating the association between circulating 25-HVD and lung cancer risk in a middle-aged male population, 50-55 years, was 15 years. When the population aged for 11 years, shorter FUTs became more pertinent. In general, interpreting results of prospective cohort studies with respect to FUTs and effect sizes may lead to more precise conclusions. Moreover, researchers should consider FUTs when they combine studies meta-analytically.

***Keywords*:** cohort study, follow-up time, lung cancer, vitamin D

1. Introduction

The steroid hormone calcitriol, 1,25-dihydroxyvitamin D, has many potential anticancer effects due to its key role in common biochemical mechanisms. These effects relate to the anti-proliferation, apoptosis, differentiation of cancer cells to less malignant phenotypes, anti-inflammation, and inhibition of invasion, metastasis, and angiogenesis [1]. Vitamin D is the precursor of calcitriol, and 25-hydroxyvitamin D (25-HVD) is a circulating form of vitamin D.

The potential anticancer effects of calcitriol suggest that maintaining sufficient vitamin D levels should reduce the risk of cancer. Regarding lung cancer, the following prospective cohort studies reported an inverse relationship

between circulating 25-HVD concentrations and disease risk in the general population: The Copenhagen City Heart Study [2], The Mini-Finland Health Survey [3], ESTHER [4], TROMSØ [5], Monica10 [6], Inter99 [6], and Health2006 [6]. However, only the Copenhagen City Heart Study [2] concluded that lower circulating 25-HVD concentrations are associated with a higher risk of lung cancer. All analyses [2-6] included sex, smoking, body mass index (BMI), and education as covariates.

A meta-analysis of 12 studies [7] concluded that the relative risk (RR) of lung cancer for the highest versus lowest quintile of serum 25-HVD concentration is 0.83 (95% CI: 0.77-0.90, $p < 0.001$). The result, however, is oversimplified and based on heterogeneous data, as the meta-analysis included only five prospective cohort studies [2-4,6,8] investigating the association between circulating 25-HVD concentrations and lung cancer risk in general population.

In general, follow-up intervals potentially affect the results of cohort studies. The inverse association between circulating 25-HVD concentrations and breast cancer risk is most evident for follow-up periods less than three years, whereas the same association for colorectal cancer appears to be significant up to 12 years or so [9].

Lung cancer is traditionally diagnosed at advanced stages mainly because it is symptomless when localized [10]. From the viewpoint of cohort studies, the long delay between the onset of lung cancer and its diagnosis denotes that short follow-up intervals for only a few years are nonoptimal. Among cohorts studying the association between circulating 25-HVD concentrations and lung cancer risk in the general population, the average follow-up time (FUT) ranged from 4 years in Health2006 [6] to 21 years in the Copenhagen City Heart Study [2]. In the latter [2], the inverse association between circulating 25-HVD concentrations and lung cancer risk strengthened together with the increasing FUT from 10 to 20 years.

In this study, we hypothesized that partly controversial conclusions regarding the epidemiological relationship between circulating 25-HVD concentrations and lung cancer risk can be due to differences in FUTs across studies as well as because of interpreting findings mainly from the perspective of statistical significance. To test our hypothesis, we investigated the association between circulating 25-HVD concentrations and lung cancer risk in an ongoing prospective cohort study and interpreted our results with respect to different FUTs. Furthermore, we demonstrated how statistical significance relates to effect sizes to emphasize the importance of other than *p*-value based conclusions.

2. Methods

2.1 *Study design and sampling*

The Kuopio Ischaemic Heart Disease Risk Factor Study (KIHD) is an ongoing population-based follow-up study [11]. Originally, the main purpose of KIHD was to investigate previously unestablished risk factors for acute myocardial infraction. Subsequently, the purpose has widened, and now the KIHD dataset includes all causes of hospitalization and death. The causes of hospitalization base on linkages to national health care registers maintained by the National Institute for Health and Welfare (permission number THL/93/5.05.00/2013) and the causes of death base on linkages to registers maintained by Statistics Finland (permission number TK-53-1770-16). Cancer cases are verified based on linkages to registers maintained by the Finnish Cancer Registry (permission number THL/93/5.05.00/2013). KIHD has received ethical approval from the Research Ethics Committee of the University of Kuopio on December 1, 1983. In the 1980s, the committee did not provide reference numbers, but it referred to studies by dates. All KIHD study participants have given written informed consent.

KIHD recruited all men who were 42-, 48-, 54-, or 60-year-old and lived in or near the city of Kuopio at the time of KIHD baseline between March 1984 and December 1989 ($n = 3235$). Of them, 2682 were willing to participate. For this study, we excluded participants with cancer ($n = 51$) as well as those with no 25-HVD measurements at baseline ($n = 53$). Consequently, the final number of participants in this study was 2578 (Figure 1). We repeated our analyses for 808 men who participated in follow-up examinations between March 1998 and February 2001 and were then still free of cancer. Comparisons between these primary ($n = 2578$) and secondary analyses ($n = 808$) enabled us to assess the plausibility of our results concerning the epidemiological relationship between 25-HVD and lung cancer as well as to estimate the reliability of 25-HVD measurements in KIHD.

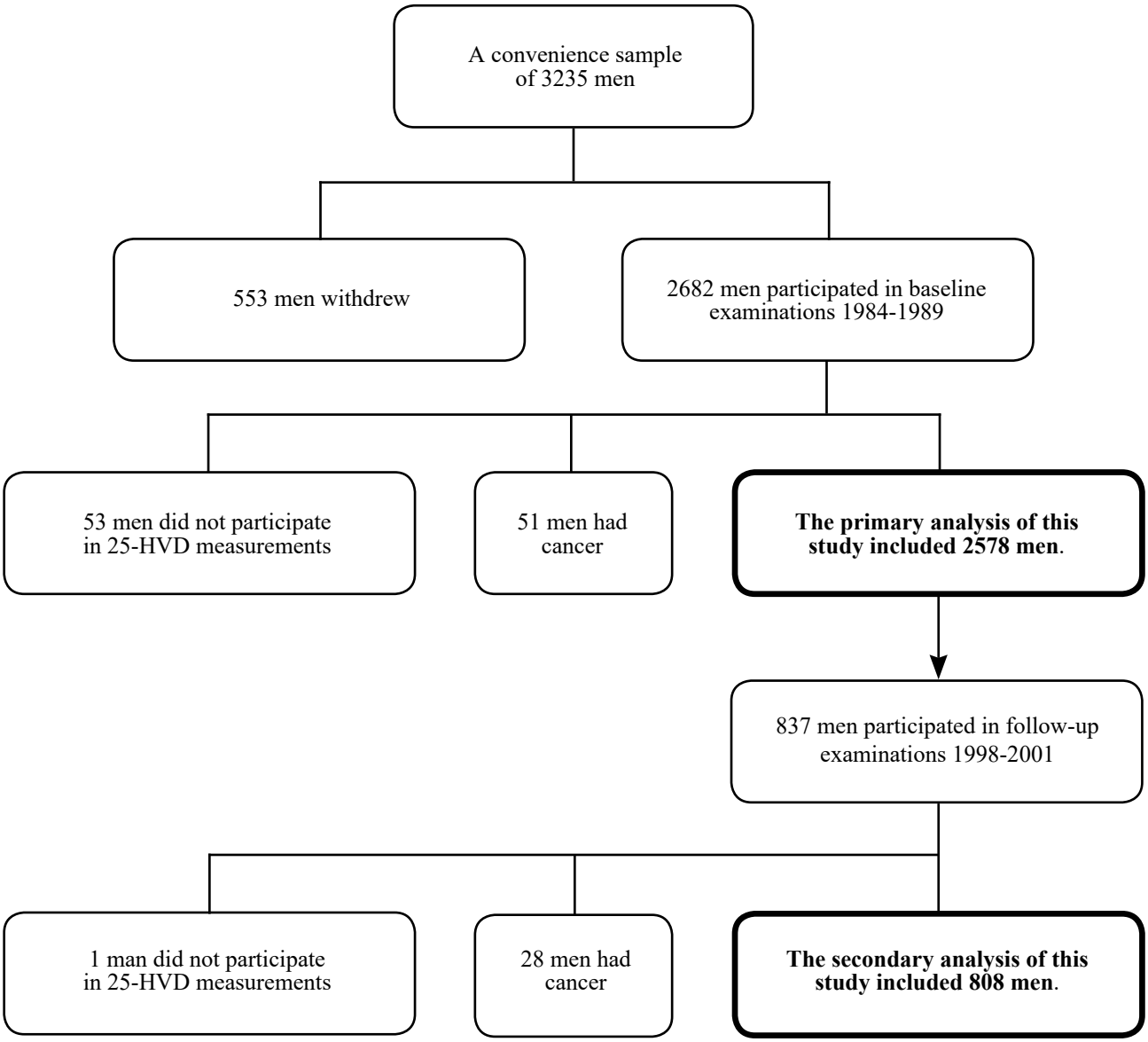

Figure 1. Study flow chart

2.2 *Determination of dependent and independent variables*

In KIHD, cancer refers to malign neoplasms and carcinoma in situ excluding basal cell carcinoma and myelofibrosis based on cancer types by ICD-10 [12] and tumor behavior by ICD-O-3 [13]. In this study, lung cancer referring to ICD-10 codes of C33-34 served as a dependent variable. We included all new lung cancer cases diagnosed between KIHD baseline visits and December 31, 2016.

The main independent variable was the serum 25-HVD concentration, precisely 25-HVD3. At baseline, a study nurse collected blood samples from study participants. The laboratory of our institute separated serum from blood and stored serum gel tubes mainly at -70 ˚C, but for a few years at -20 ˚C. Starting in 2012, the laboratory determined serum 25-HVD concentrations using a high-performance liquid chromatography device (Shimadzu, Kyoto, Japan) equipped with a coulometric electrode array detector (CEAD, ESA, Chelmsford, MA, USA) [14]. Coefficients of variation were less than 8% for controls with 25-HVD from 27.5 to 88.0 nmol/L. Participation in the Vitamin D External Quality Assessment Scheme four times per year served as the external quality control.

Other independent variables were age in years, smoking status (smoking within a month vs. smoking over a month ago or never smoked), alcohol consumption in standard drinks per week, BMI, inflammation status measured as high-sensitive C-reactive protein (hsCRP) in milligrams per liter, total physical activity (TPA) measured as metabolic equivalent hours (METh) per day, and diet described as fiber and meat intakes in grams per day together with a total daily energy intake in kilocalories. According to the Cancer Society of Finland [15], these covariates comprise the biggest lifestyle risk factors for cancer.

To describe study participants' nutritional status more thoroughly we reported the modified Baltic Sea Diet Score (BSDS) [16]. The score ranges from 0 to 25 and it indicates the adherence to a healthy Nordic diet. The higher the score, the higher the adherence. We did not use BSDS as an independent variable because the KIHD follow-up visit data do not include BSDS results.

2.3 *Statistical analyses*

In the analyses, we used season-specific concentrations of 25-HVD. First, we identified two seasons that represent the sunniest months in Finland from June to September and the less sunny months from October to May. Second, we distributed study participants into tertiles based on their season-specific levels of 25-HVD. Third, we performed Cox regressions to predict the hazard of lung cancer in different season-specific 25-HVD tertiles and adjusted the hazard for age, smoking, alcohol consumption, BMI, hsCRP, METh, fiber intake, meat intake, and total energy intake. We applied FUTs starting from 1989 when all 2578 men had participated in baseline examinations.

As an effect size measure, we selected Cohen's *d* [17]. To ease the demonstration of connections across hazards, risks, statistical significance, and Cohen's *d* we used a binary logistic regression to create odds ratios (OR) that are helpful in converting among effects sizes [18].

Moreover, we used the analysis of variance and Kruskal-Wallis test to detect differences in the independent variables across the 25-HVD tertiles. IBM® SPSS® Statistics Version 25 provided a platform for the analyses.

We repeated the analyses for 808 men who participated in follow-up examinations including 25-HVD measurements between 1998 and 2001. In these analyses, we applied FUTs starting from 2001, when all 808 men had participated in the examinations.

3. Results

3.1 *Baseline characteristics*

Table 1 presents baseline characteristics of 2578 men at different 25-HVD tertiles, and Table 2 presents the same follow-up characteristics for 808 men. Briefly, the groups based on 25-HVD levels differed from each other according to the number of smokers and average hsCRP concentration. At baseline, also the average total energy intake and BSDS differed across the groups. The average (SD, range) FUT from baseline visits to the lung cancer, death, or December 2016 was 23.3 (8.5, 0.02-32.76) years. The season of 25-HVD measurements, 'winter' or 'summer', did not affect FUTs (23.3 vs. 23.0 years, $p = 0.344$). The average (SD, range) FUT from follow-up visits was 14.9 (4.3, 0.45-18.81) years. Again, the season did not affect FUTs (14.9 vs. 14.9 years, $p = 0.946$). The correlation, Pearson's r, between baseline and follow-up 25-HVD measurements was 0.274 ($p < 0.001$).

3.2 *Risk of lung cancer from 1989 to 2016*

By the end of 2016, 103 men had lung cancer. Of them, 27 belonged to the highest 25-HVD tertile, and 43 belonged to the lowest tertile. The risk of getting lung cancer was lowest in the highest 25-HVD tertile and highest in the lowest tertile over the follow-up period (Figure 2). The RR of lung cancer in the highest 25-HVD tertile reached its minimum after 15 years of follow-up. Then the risk of getting lung cancer was 0.9% in the highest tertile and 2.2% in the lowest tertile, which yielded the RR of 0.41.

In the crude Cox model, the hazard of lung cancer in the highest 25-HVD tertile was statistically significantly lower compared to that in the lowest tertile from the average FUT of 12 years onwards (Figure 2). The statistically significant hazard ratio (HR) was lowest, 0.31 (95% CI: 0.14-0.70, $p = 0.005$, $d = 0.62$), when the average FUT was 15 years. The

Cohen's *d* effect size in relation to the statistically significant hazard of lung cancer in the highest vs. lowest 25-HVD tertile was above medium (> 0.5) only when the average FUT ranged from 12 to 17 years. This finding indicates a rather narrow follow-up interval for detecting the epidemiological relationship between circulating 25-HVD concentrations and lung cancer risk.

Table 1. Baseline characteristics of study participants (*n* = 2578) by season-specific 25-HVD tertiles

Variable	1st	2nd	3rd	*p*
n	859	860	859	n/a
Vitamin D (nmol/L), 'Winter' [a]	22.7 (4.7)	35.8 (3.9)	57.5 (12.9)	n/a
Vitamin D (nmol/L), 'Summer' [a]	38.1 (7.4)	56.8 (4.3)	78.4 (11.5)	n/a
Smoker [*n* (%)]	321 (37.4)	263 (30.6)	238 (27.7)	< 0.001
Age at baseline	52.9 (5.1)	53.0 (5.2)	53.2 (5.1)	0.422
Alcohol consumption (g/week) [b]	75.9 (129.8)	78.5 (165.9)	72.1 (109.6)	0.617
BMI	26.9 (3.8)	26.9 (3.5)	26.9 (3.4)	0.990
hsCRP (mg/L)	2.7 (5.3)	2.3 (3.4)	2.3 (3.4)	0.026
Physical activity (METh/day)	40.2 (11.9)	40.2 (10.8)	40.5 (10.8)	0.825
Fiber intake (g/day)	25.4 (9.6)	25.0 (8.6)	24.9 (8.0)	0.457
Meat intake (g/day)	158 (82)	165 (87)	158 (82)	0.116
Energy intake (kcal/day)	2485 (664)	2423 (603)	2385 (588)	0.004
Baltic Sea Diet Score (BSDS)	12.1 (3.9)	12.9 (4.0)	13.5 (3.9)	< 0.001
BSDS: missing data (*n*)	10	10	6	n/a

Notes. Numbers indicate mean (SD) unless otherwise informed. P-value is for the between-group difference.
[a] Winter: October to May, Summer: June to September.
[b] In Finland, one standard drink contains 12 grams of pure alcohol.

Table 2. Follow-up characteristics of study participants (*n* = 808) by season-specific 25-HVD tertiles

Variable	1st	2nd	3rd	*p*
n	268	270	270	n/a
Vitamin D (nmol/L), 'Winter' [a]	30.3 (4.9)	43.0 (3.8)	60.4 (9.5)	n/a
Vitamin D (nmol/L), 'Summer' [a]	40.9 (6.6)	56.6 (3.7)	74.0 (8.7)	n/a
Smoker [*n* (%)]	67 (25.8)	50 (19.1)	30 (11.3)	< 0.001
Smoker: missing data (*n*)	8	8	5	n/a
Age at baseline	62.4 (6.7)	62.2 (6.5)	62.6 (6.1)	0.754
Alcohol consumption (g/week) [b]	95.9 (175.1)	72.7 (97.8)	73.1 (108.3)	0.070
BMI	27.4 (3.6)	27.6 (3.7)	27.0 (3.4)	0.166
hsCRP (mg/L)	3.5 (6.2)	2.7 (3.7)	2.2 (4.4)	0.010
Physical activity (METh/day)	46.7 (8.8)	46.9 (8.4)	48.1 (7.6)	0.124
Fiber intake (g/day)	24.6 (10.9)	24.6 (9.6)	25.2 (9.2)	0.680
Meat intake (g/day)	158 (105)	155 (86)	152 (90)	0.794
Energy intake (kcal/day)	2092 (570)	2101 (599)	2125 (540)	0.795

Note. Numbers indicate mean (SD) unless otherwise informed. P-value is for the between-group difference.
[a] Winter: October to May, Summer: June to September.
[b] In Finland, one standard drink contains 12 grams of pure alcohol.

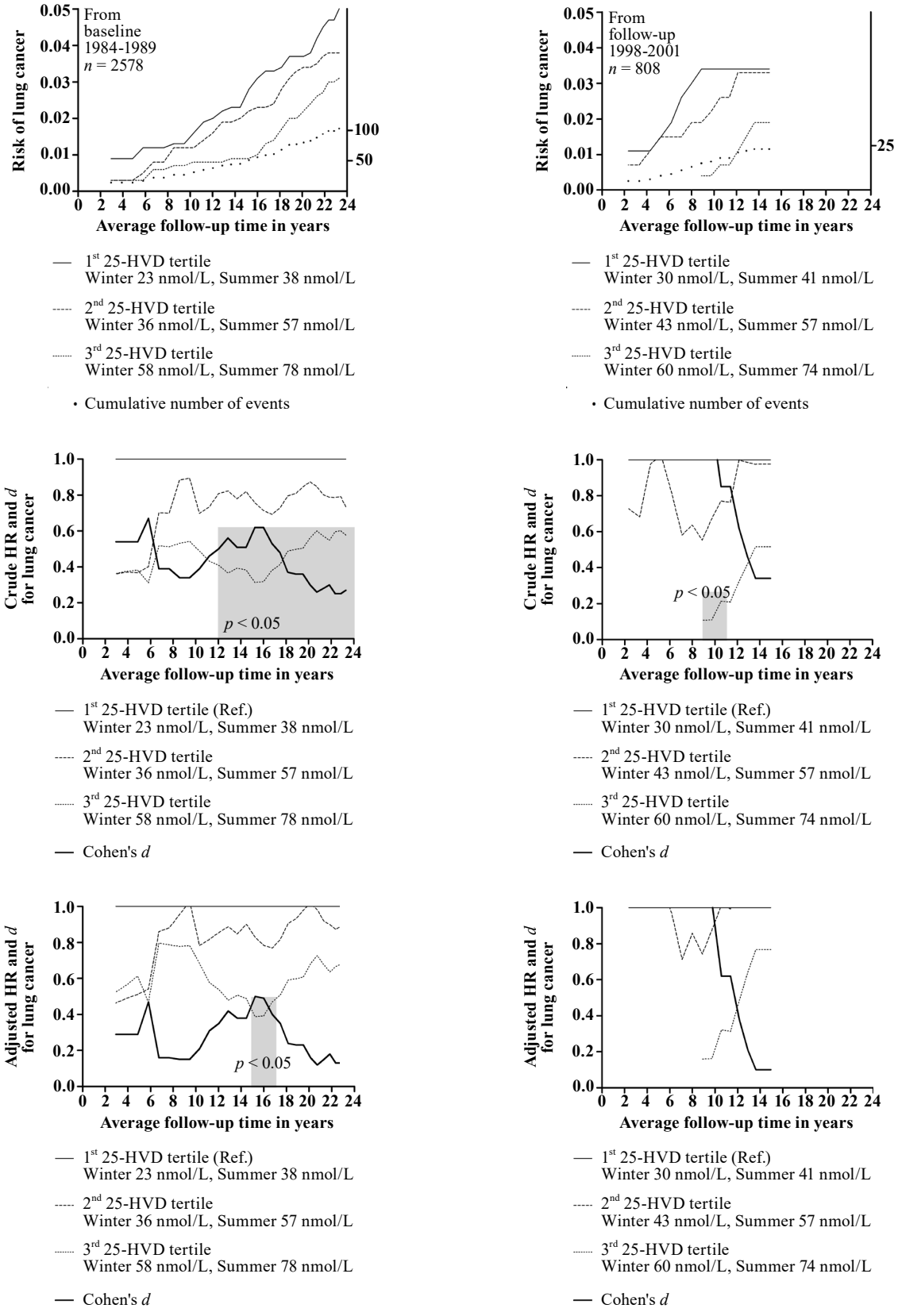

Figure 2. Risk, hazard ratio, and Cohen's *d* for lung cancer in 25-HVD tertiles starting from baseline (left panel) and 11-year follow-up visits (right panel)

In the Cox model adjusted for age, smoking, alcohol consumption, BMI, hsCRP, METh, fiber intake, meat intake, and total energy intake, the hazard of lung cancer was statistically significantly lower in the highest 25-HVD tertile only when the average FUT ranged from 15 to 17 years. Again, the HR was lowest, 0.39 (95% CI: 0.17-0.87, $p = 0.022$, $d = 0.50$), when the average FUT was 15 years. In addition to the serum 25-HVD concentration, the model restricted to 15 years of follow-up highlighted effects of smoking, age, and alcohol consumption on the hazard of lung cancer. The HR was 9.24 (95% CI: 4.38-19.49, $p < 0.001$, $d = 1.17$) for smokers vs. non-smokers, 1.17 (95% CI: 1.08-1.26, $p < 0.001$, $d = 0.08$) for one-year increase in age, and 1.02 (95% CI: 1.00-1.03, $p = 0.032$, $d = 0.01$) for one additional standard drink per week. The Cohen's d effect size in relation to the statistically significant hazard of lung cancer in the highest vs. lowest 25-HVD tertile peaked, being 0.5 when the average FUT was 15 years.

3.3 *Risk of lung cancer from 2001 to 2016*

Approximately 20% of lung cancer cases, 23 out of 103, were diagnosed after the KIHD follow-up visits. Again, the risk of getting lung cancer was lowest in the highest 25-HVD tertile and highest in the lowest tertile over the follow-up period (Figure 2). Until 2008 no lung cancer cases appeared in the highest 25-HVD tertile. After that, the RR of lung cancer in the highest vs. lowest 25-HVD tertile stayed at 0.12 for two years until it started to increase.

In the crude Cox model, the hazard of lung cancer in the highest 25-HVD tertile was statistically significantly lower compared to that in the lowest tertile, when the FUT ranged from nine to 11 years (Figure 2). The statistically significant HR was lowest, 0.11 (95% CI: 0.01-0.85, $p = 0.034$, $d = 1.23$), when the average FUT was nine years. The Cohen's d effect size in relation to the hazard of lung cancer in the highest vs. lowest 25-HVD tertile was large (> 0.8), when the average FUT was less than 12 years.

The adjusted Cox model did not result in statistically significant HRs regarding the serum 25-HVD concentration. Only statistically significant HRs related to smoking and age. The HR was 9.11 (95% CI: 2.79-29.7, $p < 0.001$, $d = 1.22$) for smokers vs. non-smokers and 1.21 (95% CI: 1.08-1.36, $p = 0.001$, $d = 0.11$) for one-year increase in age. The Cohen's d effect size in relation to the hazard of lung cancer in the highest vs. lowest 25-HVD tertile was large (> 0.8), when the average FUT was less than 10 years.

4. Discussion

The main finding of this study was the inverse association between circulating 25-HVD concentrations and lung cancer risk irrespective of the FUT up to 33 years. The FUT, however, greatly affected statistical significance that per se depends on sample size and the number of events. In KIHD, p-values regarding the HR of the highest vs. the lowest 25-HVD tertile reached the level of statistical significance (< 0.05), when the FUT was restricted to 15-17 years. After that, p-values increased, although the number of lung cancer cases in the cohort increased. In the KIHD sub-cohort, the sample size was approximately three times smaller compared to the main cohort, and p-values stayed above the level of statistical significance over the entire follow-up period. On the other hand, in the main cohort, Cohen's d indicated only a medium effect of 25-HVD on lung cancer risk at its peak, whereas, in the sub-cohort, Cohen's d expressed a large effect. This strengthens the general view that study results should not be interpreted solely based on statistical significance [19].

The level of circulating 25-HVD potentially modifies the strength and direction of the relationship between 25-HVD and lung cancer risk so that at lower levels the relationship is inverse, at moderate levels weak, and at higher levels even direct. A dose-response meta-analysis of 12 cohorts [20] concluded a nonlinear relationship between circulating 25-HVD concentrations and lung cancer risk with the greatest risk reduction at c. 50 nmol/L. The meta-analysis included the Copenhagen City Heart Study [2], The Mini-Finland Health Survey [3], ESTHER [4], TROMSØ [5], the Health in Men Study [8], Monica10 [6], Inter99 [6], and Health2006 [6] together with four other cohorts: the Health Professionals Follow-Up Study [21] composed of male health professionals, the ATBC Study [22] composed of male smokers, the Third National Health and Nutrition Examination Survey [23] reporting lung cancer mortality, and a multi-institution inflammatory bowel disease cohort [24]. Among the prospective cohort studies of the general population, the average level of 25-HVD ranged from 41 nmol/L in the Copenhagen City Heart Study [2] to 68 nmol/L in the Health in Men Study [25]. A more recent dose-response meta-analysis [26] including the same cohorts as the

meta-analysis by Chen and colleagues [20], except for the Third National Health and Nutrition Examination Survey [23], but additionally a hospital-based case-control study [27] concluded that each 10 nmol/L increase in the circulating 25-HVD concentration reduces the risk of lung cancer by 8%.

The age of the study population has effects on the relationship between 25-HVD and lung cancer risk. In general, only people at the age of 55 and above are, potentially, at high risk for lung cancer [28]. In cohorts studying the association between circulating 25-HVD concentrations and lung cancer risk in the general population, the average age of study subjects ranged from younger than 50 years in Inter99 [6] and Health2006 [6] to 77 years in the Health in Men Study [8]. With respect to average FUTs, this means that the study subjects were c. 50 years old in Health2006 [6] and c. 85 years old in the Health in Men Study [8] at the end of follow-up, which for its part may explain the somewhat controversial results across the studies.

The KIHD main cohort resembles the Copenhagen City Heart Study [2] regarding the FUT, the average circulating 25-HVD concentration, and age. In the Copenhagen City Heart Study [2], however, the RR of lung cancer was high especially among subjects with very low circulating 25-HVD concentrations, below 15 nmol/L. In KIHD, the RR was evidently increased among all subjects with 'winter' 25-HVD concentrations below 40 nmol/L and 'summer' 25-HVD concentrations below 60 nmol/L. Broadly, an insufficient serum level of 25-HVD refers to 25-75 nmol/L [29].

The KIHD main cohort also resembles the Mini-Finland Health Survey [3] with respect to the FUT and circulating 25-HVD concentrations. One reason why the inverse association between circulating 25-HVD concentrations and lung cancer risk in the Mini-Finland Health Survey did not reach statistical significance could be age. The cohort [3] represented ≥ 30-year-old people and, consequently, the youngest of them were not at high risk for lung cancer over the FUT of 24 years at most [28]. Inevitably, this has reduced the number of lung cancer cases, i.e., statistical power in the Mini-Finland Health Survey.

The KIHD sub-cohort of 808 men resembles ESTHER [4,5], except for the average FUT that is five years shorter in ESTHER. As in the Mini-Finland Health Survey [3], in ESTHER [5], the risk of lung cancer was highest in the middle tertile of circulating 25-HVD concentration, c. 30-50 nmol/L. In KIHD, the risk was highest in the lowest 25-HVD tertile. One explanation for the unexpected results in the Mini-Finland Health Survey and ESTHER could be that study subjects with the lowest circulating 25-HVD concentrations died for other reasons before getting lung cancer. In general, circulating 25-HVD concentrations relate to all-cause mortality [30].

The KIHD sub-cohort also resembles TROMSØ [5], but the TROMSØ data lack current smokers, which seriously weakens the generalizability of TROMSØ results.

The Health in Men Study [8] is the only prospective cohort study that reports a reduced risk of lung cancer due to low 25-HVD levels. In the Health in Men Study [8,25], the population was old, and the average circulating 25-HVD concentration was high.

4.1 *Strengths and limitations*

The main strength of our study is the KIHD dataset that enabled us to investigate the effects of different follow-up intervals on the epidemiological relationship between 25-HVD and lung cancer risk. The maximum follow-up time in KIHD is exceptionally long. Moreover, KIHD provided us with an opportunity to repeat our analyses owing to its follow-up visits 11 years after baseline visits. On the other hand, although the KIHD data as such show a significant amount of internal variation [31], the lack of heterogeneity, i.e., differences in KIHD participant characteristics including gender and origin, is the main limitation of our study in terms of generalizability. Moreover, as a limitation regarding our discussion, we acknowledge uncertainties related to comparisons across studies due to non-standardized methods to measure circulating 25-HVD concentrations in general.

5. Conclusion

Our findings suggest that the optimal FUT for a prospective cohort study investigating the association between circulating 25-HVD concentrations and lung cancer risk in a middle-aged population, 50-55 years old, is 15 years. If the population is older, 60-65 years old, a shorter FUT up to 10 years appears to be more pertinent. Moreover, our findings propose that study results regarding the epidemiological relationship between circulating 25-HVD concentrations and

lung cancer risk should be routinely interpreted with respect to FUTs. In addition, researchers should consider FUTs when they combine studies meta-analytically. Overall, this study strengthens the general view that scientific conclusions should be drawn by considering both statistical significance and effect sizes.

References

[1] Feldman D, Krishnan AV, Swami S, Giovannucci E, Feldman BJ. The role of vitamin D in reducing cancer risk and progression. *Nature Reviews Cancer.* 2014; 14(5): 342-357.

[2] Afzal S, Bojesen SE, Nordestgaard BG. Low plasma 25-hydroxyvitamin D and risk of tobacco-related cancer. *Clinical Chemistry.* 2013; 59(5): 771-780.

[3] Kilkkinen A, Knekt P, Heliövaara M, Rissanen H, Marniemi J, Hakulinen T, et al. Vitamin D status and the risk of lung cancer: A cohort study in Finland. *Cancer Epidemiology, Biomarkers & Prevention.* 2008; 17(11): 3274-3278.

[4] Ordóñez-Mena JM, Schöttker B, Haug U, Müller H, Köhrle J, Schomburg K, et al. Serum 25-hydroxyvitamin D and cancer risk in older adults: Results from a large German prospective cohort study. *Cancer Epidemiology, Biomarkers & Prevention*. 2013; 22(5): 905-916.

[5] Ordóñez-Mena JM, Schöttker B, Fedirko V, Jenab M, Olsen A, Halkjær J, et al. Pre-diagnostic vitamin D concentrations and cancer risks in older individuals: An analysis of cohorts participating in the CHANCES consortium. *European Journal of Epidemiology*. 2016; 31(3): 311-323.

[6] Skaaby T, Husemoen LLN, Thuesen BH, Pisinger C, Jørgensen T, Roswall N, et al. Prospective population-based study of the association between serum 25-hydroxyvitamin-D levels and the incidence of specific types of cancer. *Cancer Epidemiology, Biomarkers & Prevention*. 2014; 23(7): 1220-1229.

[7] Zhang L, Wang S, Che X, Li X. Vitamin D and lung cancer risk: A comprehensive review and meta-analysis. *Cellular Physiology and Biochemistry*. 2015; 36(1): 299-305.

[8] Wong YYE, Hyde Z, McCaul KA, Yeap BB, Golledge J, Hankey GJ, et al. In older men, lower plasma 25-hydroxyvitamin D is associated with reduced incidence of prostate, but not colorectal or lung cancer. *PLoS One*. 2014; 9(6): e99954.

[9] Grant WB. Effect of interval between serum draw and follow-up period on relative risk of cancer incidence with respect to 25-hydroxyvitamin D level: Implications for meta-analyses and setting vitamin D guidelines. *Dermato-Endocrinology*. 2011; 3(3): 199-204.

[10] Leiro-Fernández V, Mouronte-Roibás C, Ramos-Hernández C, Botana-Rial M, González-Piñeiro A, García-Rodríguez E, et al. Changes in clinical presentation and staging of lung cancer over two decades. *Archivos de Bronconeumología*. 2014; 50(10): 417-421.

[11] Salonen JT. Is there a continuing need for longitudinal epidemiologic research? The Kuopio Ischaemic Heart Disease Risk Factor Study. *Annals of Clinical Research*. 1988; 20(1-2): 46-50.

[12] THL. *The Finnish version of the International Statistical Classification of Diseases and Related Health Problems*. 3rd ed. Mikkeli: The National Institute for Health and Welfare (THL); 2011.

[13] WHO. *Internal Classification of Diseases for Oncology (ICD-O)*. 3rd ed. Geneva: World Health Organization (WHO); 2013.

[14] Nurmi T, Tuomainen TP, Virtanen J, Mursu J, Voutilainen S. High-performance liquid chromatography and coulometric electrode array detector in serum 25-hydroxyvitamin D(3) and 25-hydroxyvitamin D(2) analyses. *Analytical Biochemistry*. 2013; 435(1): 1-9.

[15] Pukkala E., Pitkäniemi J, Heikkinen S. *Cancer in Finland 2016: Risk factors of cancer*. Helsinki: Cancer Society of Finland; 2016. Available from: http://www.cancersociety.fi/publications/reports/cancer-in-finland-2016/risk-factors-of-cancer/ [Accessed 12th September 2019].

[16] Tertsunen HM, Hantunen S, Tuomainen TP, Virtanen JK. Healthy Nordic diet and risk of disease death among men: The Kuopio Ischaemic Heart Disease Risk Factor Study. *European Journal of Nutrition*. 2020; 59(8): 3545-3553.

[17] Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press; 1969.

[18] Hasselblad V, Hedge1s LV. Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*. 1995; 117(1): 167-178.

[19] Amrhein V, Greenland S, McShane B. Retire statistical significance. *Nature*. 2019; 567(7748): 305-307.

[20] Chen GC, Zhang ZL, Wan Z, Wang L, Weber P, Eggersdorfer M, et al. Circulating 25-hydroxyvitamin D and risk of lung cancer: A dose-response meta-analysis. *Cancer Causes & Control*. 2015; 26(12): 1719-1728.

[21] Giovannucci E, Liu Y, Rimm EB, Hollis BW, Fuchs CS, Stampfer MJ, et al. Prospective study of predictors of

vitamin D status and cancer incidence and mortality in men. *Journal of the National Cancer Institute*. 2006; 98(7): 451-459.
[22] Weinstein SJ, Yu K, Horst RL, Parisi D, Virtamo J, Albanes D. Serum 25-hydroxyvitamin D and risk of lung cancer in male smokers: A nested case-control study. *PLoS One*. 2011; 6(6): e20796.
[23] Cheng TY, Neuhouser ML. Serum 25-hydroxyvitamin D, vitamin A, and lung cancer mortality in the US population: A potential nutrient-nutrient interaction. *Cancer Causes & Control*. 2012; 23(9): 1557-1565.
[24] Ananthakrishnan AN, Cheng SC, Cai T, Cagan A, Gainer VS, Szolovits P, et al. Association between reduced plasma 25-hydroxy vitamin D and increased risk of cancer in patients with inflammatory bowel diseases. *Clinical Gastroenterology and Hepatology*. 2014; 12(5): 821-827.
[25] Wong YYE, McCaul KA, Yeap BB, Hankey GJ, Flicker L. Low vitamin D status is an independent predictor of increased frailty and all-cause mortality in older men: The Health in Men Study. *The Journal of Clinical Endocrinology & Metabolism*. 2013; 98(9): 3821-3828.
[26] Feng Q, Zhang H, Dong Z, Zhou Y, Ma JP. Circulating 25-hydroxyvitamin D and lung cancer risk and survival: A dose-response meta-analysis of prospective cohort studies. *Medicine (Baltimore)*. 2017; 96(45): e8613.
[27] Wang X, Cui J, Gu J, He H, Li B, Li W. Plasma 25-hydroxyvitamin D deficiency is associated with the risk of non-small cell lung cancer in a Chinese population. *Cancer Biomarkers*. 2015; 15(5): 663-668.
[28] Yang D, Liu Y, Bai C, Wang X, Powell CA. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. *Cancer Letters*. 2020; 468: 82-87.
[29] Rosen CJ. Clinical practice. Vitamin D insufficiency. *The New England Journal of Medicine*. 2011; 364(3): 248-254.
[30] Gaksch M, Jorde R, Grimnes G, Joakimsen R, Schirmer H, Wilsgaard T, 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(2): e0170791.
[31] Brester C, Stanovov V, Voutilainen A, Tuomainen TP, Semenkin E, Kolehmainen M. Evolutionary Fuzzy Logic-based model design in predicting coronary heart disease and its progression. In: Merelo JJ, Garibaldi J, Barranco AL, Madani K, Warwick K. (eds.) *Proceedings of the 11th International Joint Conference on Computational Intelligence*. Vienna: SciTePress Digital Library; 2019. p.360-366.