

Marion Denos

# Vitamin D status and its interaction with genetic profile on risk of type 2 diabetes in Norwegian adults

The HUNT Study

Master's thesis in Global Health

Supervisor: Xiao-Mei Mai & Yi-Qian Sun

June 2020



Marion Denos

# **Vitamin D status and its interaction with genetic profile on risk of type 2 diabetes in Norwegian adults**

The HUNT Study

Master's thesis in Global Health  
Supervisor: Xiao-Mei Mai & Yi-Qian Sun  
June 2020

Norwegian University of Science and Technology  
Faculty of Medicine and Health Sciences  
Department of Public Health and Nursing





## **ACKNOWLEDGEMENTS**

This research was supported by the Faculty of Medicine and Health Sciences at the Norwegian University of Science and Technology (NTNU). I thank all the individuals who participated in the Nord-Trøndelag Health Study (HUNT).

I would like to thank my supervisors, Dr. Xiao-Mei Mai and Dr. Yi-Qian Sun, for sharing their knowledge and insight, and for their invaluable guidance and encouragement throughout this process.

I would also like to thank my friends, Martine, Hanne and Birthe, with whom I was sharing the same reading room. They helped me to keep my motivation up.

I wish to thank my beloved partner, Romain, for his support and constructive advices, and to have been a nice co-worker at home office in this challenging time.

I am very grateful to my family: my mother, sister and brother, whose love and positive vibes are with me in whatever I pursue. My father would have been proud of me, and I thank him a lot to have been such an essential pillar in my education and professional path all along the years.

## ABSTRACT

**Objective:** The aim of the study was to investigate the relationship between serum 25-hydroxyvitamin D [25(OH)D] level and risk of type 2 diabetes mellitus (T2DM) in adults participating in the Nord-Trøndelag Health Study (HUNT) over 11 years of follow-up. We also evaluated if sex, body mass index (BMI) and genetic profile for T2DM had potential interactions with low vitamin D levels on the risk of T2DM.

**Introduction:** With more than 400 million people affected globally, diabetes has been identified by the United Nations as one of the four high priority non-communicable diseases targeted for action. Meanwhile, low vitamin D status remains a potential threat to health globally and has been associated with many diseases. However, its potential harmful effect on risk of diabetes is inconclusive.

**Methods:** This prospective cohort study included 3 586 diabetes-free adults at baseline. They were  $\geq 19$  years of age and participated in the second and third surveys of the Nord-Trøndelag Health Study [HUNT2 (1995-1997), defined as the baseline, and HUNT3 (2006-2008)] in Norway. Serum 25(OH)D levels were determined at baseline and classified as  $< 50$  nmol/L and  $\geq 50$  nmol/L. Comprehensive lifestyle and sociodemographic data were collected at baseline. A polygenic risk score (PRS) for T2DM based on 14 single-nucleotide polymorphisms (SNPs) was generated. Incident diabetes was defined by self-reporting of diabetes and/or non-fasting glucose levels greater than 11 mmol/L and serum glutamic acid decarboxylase antibodies (GADA) level lower than 0.08 ai at the follow-up in HUNT3. Odds ratios (ORs) with 95% confidence intervals (CI) for the association of baseline 25(OH)D level with risk of T2DM were evaluated using multivariable logistic regression models. We estimated effect modification by sex, BMI and PRS for diabetes for the risk of T2DM. Biological interaction of low serum 25(OH)D and high PRS for risk of T2DM was also evaluated. Multiple imputation was used to address possible bias due to missing data of the covariates.

**Results:** Over the follow-up period, a total of 104 (2.9%) participants developed T2DM. A higher risk of incident T2DM seemed to be present but not statistically significant in participants with serum 25(OH)D level under 50 nmol/L compared to those with serum 25(OH)D level above 50 nmol/L (OR=1.55, 95% CI=0.97-2.49). We used the top one third PRS as the cut-off value to study statistical interaction between serum 25(OH)D and PRS. Level of 25(OH)D under 50 nmol/L was associated with a significant increased risk of T2DM in the higher PRS group (OR=2.98, 95% CI=1.19-7.45). There was no effect modification by sex and BMI. We also found a biological interaction of low serum 25(OH)D and high PRS for the risk of T2DM.

**Conclusion:** Low serum 25(OH)D was associated with an increased risk of T2DM in Norwegian adults. The inverse association appeared particularly significant for people with high polygenic risk score for T2DM. Given the high prevalence of vitamin D insufficiency and T2DM worldwide, our study provided insights into new ways to prevent T2DM, particularly by incorporating the genetic factor of the disease with modifiable factors in prevention.

**Keywords:** 25-hydroxyvitamin D, cohort studies, diabetes, genetic profile, HUNT, polygenic risk score, vitamin D

# TABLE OF CONTENTS

|  |    |
|--|----|
| ACKNOWLEDGEMENTS .....   | 1  |
| ABSTRACT .....   | 2  |
| LIST OF FIGURES.....   | 5  |
| LIST OF TABLES .....   | 5  |
| ABBREVIATIONS.....   | 6  |
| 1. Background.....   | 7  |
| 1.1. Global prevalence and burden of type 2 diabetes.....                      | 7  |
| 1.2. Risk factors for T2DM .....   | 8  |
| 1.3. Vitamin D synthesis and metabolism .....                                  | 9  |
| 1.4. Prevalence and prevention of vitamin D insufficiency and deficiency ..... | 10 |
| 1.5. The biological role of vitamin D .....                                    | 11 |
| 1.6. Association between vitamin D status and T2DM .....                       | 12 |
| 2. Study aims.....   | 14 |
| 3. Methods.....  | 15 |
| 3.1. Study design and data collection .....                                    | 15 |
| 3.2. Serum 25(OH)D level as exposure variable .....                            | 16 |
| 3.3. Covariates .....  | 18 |
| 3.4. Polygenic risk score for T2DM .....                                       | 20 |
| 3.5. T2DM as outcome variable .....  | 21 |
| 3.6. Statistical analysis.....   | 21 |
| 3.7. Ethical reflections .....   | 22 |
| 4. Results.....  | 23 |
| 4.1. Baseline characteristics of the study population.....                     | 23 |
| 4.2. Serum 25(OH)D level and risk of T2DM.....                                 | 25 |
| 4.3. Effect modification by sex, BMI and genetic profile of T2DM .....         | 26 |
| 4.4. Biological interaction .....  | 28 |

|   |    |
|---|----|
| 5. Discussion .....   | 29 |
| 5.1. Main findings.....   | 29 |
| 5.2. Comparison with previous studies.....                              | 29 |
| 5.2.1. Vitamin D status as an independent risk factor for T2DM .....    | 29 |
| 5.2.2. Effect modification by sex, BMI and genetic profile of T2DM..... | 31 |
| 5.3. Potential biological mechanisms .....                              | 32 |
| 5.4. Strengths .....  | 33 |
| 5.5. Limitations.....   | 33 |
| 6. Conclusion .....   | 35 |
| 7. References.....  | 36 |
| APPENDIX: Supplementary tables .....                                    | 42 |



## LIST OF FIGURES

|  |    |
|--|----|
| <b>Figure 1.</b> Selection process of the analytical sample, Nord-Trøndelag Health Study (HUNT).<br>.....                      | 16 |
| <b>Figure 2.</b> Serum 25(OH)D in the Nord-Trøndelag Health Study (HUNT). .....  | 18 |
| <b>Figure 3.</b> Directed acyclic graph (DAG) underlining associations between exposure, outcome variable and covariates. .... | 19 |

## LIST OF TABLES

|  |    |
|--|----|
| <b>Table 1.</b> Baseline characteristics of subjects overall and by baseline serum 25(OH)D levels in the HUNT2 study, 1995-1997. ....  | 24 |
| <b>Table 2.</b> Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up .....   | 26 |
| <b>Table 3.</b> Different cut-off groups evaluated for effect modification of T2DM PRS on association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up. .... | 27 |
| <b>Table 4.</b> Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up, stratified by PRS of T2DM .....  | 28 |

## ABBREVIATIONS

|                         |  |
|-------------------------|--|
| 1,25(OH) <sub>2</sub> D | 1,25-dihydroxyvitamin D                                    |
| 25(OH)D                 | 25-hydroxyvitamin D  |
| BMI                     | Body Mass Index  |
| CI                      | Confidence interval  |
| DAG                     | Directed acyclic graph                                     |
| GADA                    | Glutamic Acid Decarboxylase Antibodies                     |
| GWAS                    | Genome-wide association study                              |
| HUNT                    | Nord-Trøndelag Health Study                                |
| MR                      | Mendelian Randomization                                    |
| OR                      | Odds ratio   |
| PRS                     | Polygenic risk score                                       |
| REK                     | Regional Committees for Medical and Health Research Ethics |
| RERI                    | Relative Excess Risk due to Interaction                    |
| SNP                     | Single-Nucleotide Polymorphism                             |
| T2DM                    | Type 2 Diabetes Mellitus                                   |
| UVB                     | Ultraviolet-B  |
| VDR                     | Vitamin D Receptor   |
| WHO                     | World Health Organization                                  |

# 1. Background

## 1.1. Global prevalence and burden of type 2 diabetes

According to the World Health organization (WHO) definition, diabetes is a serious, chronic disease that occurs either when the pancreas does not produce enough insulin (type 1), or when the body cannot effectively use the insulin it produces (type 2) (1). It is characterized by elevated levels of blood glucose that may lead, if not well controlled, to serious damage to the heart, blood vessels, eyes, kidneys and nerves (1). Therefore, individuals with diabetes have a shorter life expectancy, with up to 75% dying of macrovascular complications (2).

More than 400 million people currently live with diabetes. The global prevalence of diabetes has almost doubled since 1980, rising from 4.7% to 8.5% in the adult population (1). It has also been predicted to rise by 55% in the world and by 22% in European countries between 2013 and 2035 (3). In Norway, diabetes is one of the most common chronic diseases, and around 5% of the population have a diabetes diagnosis (4). Type 2 diabetes mellitus (T2DM) seems to be more prevalent among males than females (5-7). One possible explanation would be that men could develop T2DM with relatively lower weight gain compared to women (8). This is partly because men, due to their elevated visceral and hepatic adipose tissue, are generally more insulin resistant than women (9). In contrast to prevalence figures, the incidence of T2DM in high income countries might have been falling in the latest years (10). However, this trend could be different in low- and middle-income countries where data are limited (10).

Diabetes is an important public health problem, one of the four high priority non-communicable diseases targeted for action by the General Assembly of the United Nations (11). Besides coronary heart disease, stroke and lung cancer, T2DM has become one of the top ten leading causes of death globally in 2017 (7). The number of deaths due to T2DM has increased in the most populous countries, such as India, China and Indonesia (7). In addition, diabetes

also leads to a considerable increase in healthcare expenditures over the years, which is no longer due to the treatment of complications, but rather on the prevention of vascular complications and the ongoing management of the condition (12).

Several successful trials showed that prevention, or at least delaying the onset, of T2DM is possible (13). On a life-course perspective, promotion of healthy lifestyle for everyone and controlling blood pressure and lipids could have a significant impact on the disease and its complications. Additionally, the WHO encourages mass media campaigns and social marketing to influence positively changes and make healthy behaviours (1).

## **1.2. Risk factors for T2DM**

T2DM is a result of the interplay between lifestyle, environmental and genetic factors. There are well-defined biological and behavioural risk factors that are thought to act through increasing insulin resistance (14). Among them, overweight and obesity, particularly abdominal obesity, and physical inactivity are the most important (15, 16). According to the WHO, the prevalence of obesity has tripled over the past two decades, with 39% of adults being overweight (defined by a body mass index (BMI)  $\geq 25$  kg/m<sup>2</sup>) and 13 % being obese (BMI  $\geq 30$  kg/m<sup>2</sup>) (17). Other behavioural risk factors include smoking habits (18) and dietary patterns, such as consumption of sugar-sweetened soft drinks (19, 20).

In addition to lifestyle factors, there is compelling evidence that the individual risk of T2DM is strongly influenced by genetic factors (21, 22). The genome-wide association study (GWAS) investigates the associations between hundreds of thousands to millions of single-nucleotide polymorphisms (SNPs) and phenotype (23). GWAS have provided the possibility to identify specific risk alleles of SNPs for T2DM. A recent meta-analysis study, combining data from 32 GWAS, has provided new insights by identifying 243 loci, including 135 newly implicated in

T2DM predisposition (24). These loci are related to genes which mostly influence energy metabolism and regulation of insulin secretion.

Among them, the genetic variant rs7578326 near gene *IRS1* has been identified in several studies (25). *IRS1* is an insulin receptor, a necessary component of insulin action such as in insulin-secreting beta cells of the pancreas (26). The A allele of rs7578326 near gene *IRS1* would be associated with increased T2DM risk. Voight et al. suggested that the risk alleles at *CENTD2* modify T2DM susceptibility through a detrimental effect on beta-cell function. In contrast, the risk alleles at *KLF14* and at *IRS1* appear to have a primary effect on insulin action (by reducing insulin sensitivity) (25). They found that their data were consistent with the idea that common T2DM-risk alleles, while usually acting through beta-cell dysfunction, can also exert their primary effects on insulin action (25).

### **1.3. Vitamin D synthesis and metabolism**

Vitamin D is a vital vitamin mainly created by our skin after exposure to ultraviolet-B (UVB) radiation in sunlight. Few foods naturally contain vitamin D, including fatty fish and cod liver oil, and food fortification is often inadequate to satisfy human vitamin D requirement (27). Approximately 80% of vitamin D is produced in the skin, while the remaining 20% comes from dietary sources, but this varies depending on factors such as season, latitude, nutrition intake or ethnicity (28-30). Moreover, anything that diminishes the penetration of UVB radiation into our skin will affect the cutaneous synthesis of vitamin D. Therefore, uses of sunscreen or clothes covering too much of the skin, as well as diminution of outdoor activity, reduce the possibility of dermal synthesis of vitamin D (31).

Vitamin D is considered biologically inert until it undergoes two consecutive hydroxylation reactions: first in the liver to 25-hydroxyvitamin D [25(OH)D], which is used to

determine a patient's vitamin D status, and then in the kidney to its active form, 1,25-dihydroxyvitamin D [1,25(OH)<sub>2</sub>D] (32, 33). Serum 25(OH)D is the main circulating vitamin D metabolite that is considered to best indicate overall vitamin D status as it reflects vitamin D supply from diverse sources (34). Although there is no consensus on optimal levels of serum 25(OH)D, vitamin D insufficiency is commonly defined as serum 25(OH)D level < 50 nmol/L, and vitamin D deficiency as serum 25(OH)D level < 30 nmol/L (32, 33, 35).

#### **1.4. Prevalence and prevention of vitamin D insufficiency and deficiency**

Vitamin D insufficiency is very common worldwide in all ethnicities and age groups. Even in countries with low latitude, meaning adequate sun exposure all year round, and in developed countries with access to vitamin-rich foods, vitamin D deficiency is present (36). Although prevalence of vitamin D insufficiency is uncertain due to data lacking from many countries, it has been estimated that 1 billion people worldwide have low vitamin D levels (36). It may be linked to many issues, such as less vitamin D photosynthesis in response to UVB radiations in individuals with high skin melanin content or due to aging (37), use of extensive skin coverage and few exposure to sunlight, which has been shown in individuals from Africa, the Middle East, Central and South America (36). Additionally, the problem can be due to both a low vitamin D intake and high rates of obesity worldwide (36).

According to recent surveys, 40% of the general population in Europe have low levels of vitamin D (38), compared to 24% in the US (39). The prevalence seems to be even higher in low and middle-income countries (40). Among young and middle-aged Norwegian adults in HUNT, the prevalence of vitamin D insufficiency is about 40%, with a seasonal variation ranging from 20% in the summer to a peak at 64% in the winter (41). Larose et al. found that the high prevalence of vitamin D insufficiency in a Norwegian adult population was associated

with winter season, high BMI and current smoking (41). The sunlight being low in Norway between October and March due to the high latitude, almost no vitamin D is produced in the skin.

Despite the ongoing debates on impact of vitamin D status on many aspects of human health, there is a general agreement that prevention of vitamin D deficiency is a public health priority (27). The Institute of Medicine and the Nordic Nutrition Recommendations Committee recommend an intake of 10 µg/day in order to maintain serum 25(OH)D level around 50 nmol/L among the majority of the population, and of 20 µg/day for the elderly (32, 42). To prevent vitamin D deficiency, it has been recommended to increase intakes of naturally vitamin D containing food, food fortification, vitamin D supplements, to increase solar UVB exposure and to encourage weight loss (43-45). Supplementation can improve vitamin D intake on an individual basis, but adherence within the population as well as potential overdosing of vitamin D supplements are significant limitations (46). Overall, food fortification seems to represent the best opportunity to increase the vitamin D supply to the population, as highlighted by the European project ODIN (47). Well-designed sustainable fortification strategies, which use a range of various foods such as dairy products, meat, eggs and mushrooms, have potential to increase vitamin D intakes across the population distribution and minimize the prevalence of low serum 25(OH)D (47, 48).

### **1.5. The biological role of vitamin D**

Vitamin D is essential for musculoskeletal health. It promotes calcium absorption from the gut, enables mineralisation of newly formed osteoid tissue in bone and has an important role in muscle function (33, 49). The main manifestation of vitamin D deficiency is rickets in children and osteomalacia in adults, generally associated with increased risk at serum 25(OH)D

levels < 30 nmol/L (33, 35, 50). Vitamin D insufficiency may lead to bone loss, muscle weakness, falls and fragility fractures in older people (51). Moreover, as explained in the review of Dr. Holick, most tissues and cells in the body, including immune cells, have a vitamin D receptor (VDR) and respond to the active form of vitamin D: 1,25(OH)<sub>2</sub>D (33). This active form helps control the expression of more than 200 genes (52), including genes responsible for regulating cell growth and cellular differentiation via paracrine or autocrine regulatory mechanisms (33, 35, 53, 54). Therefore, vitamin D insufficiency may be related to the risk of many illnesses, including diabetes (55-57), cancers (58) and cardiovascular diseases, among others (59). These conditions are major public health problem worldwide.

#### **1.6. Association between vitamin D status and T2DM**

There is a debate about whether low vitamin D status increases the risk of T2DM or is just a marker of overall poor health. However, there is accumulating evidence to suggest that altered vitamin D may play a role in the development of T2DM. Defects in pancreatic  $\beta$ -cell function, insulin sensitivity and systematic inflammation appear to be related to development of T2DM (60, 61), and vitamin D may influence these mechanisms (62).

Most cross-sectional studies have reported inverse association between vitamin D status and T2DM (63-65), but others failed to show any association (66, 67). It is unclear if the association between serum 25(OH)D and metabolic syndrome may simply reflect the well-established inverse association between serum 25(OH)D and body weight (68). Contrary to cross-sectional studies, longitudinal observational studies may suggest the direction of the association between vitamin D and T2DM. Many have reported that higher 25(OH)D levels are significantly associated with lower risks of diabetes (57, 69), but overall the results are inconsistent (70, 71). The meta-analysis of Song at al. combined data from 21 prospective



studies, including 4 996 incident T2DM cases with average follow-up periods from 1 to 22 years (57). They showed a significant inverse association between baseline 25(OH)D levels and risk of T2DM (57). They also evaluated sex as potential effect modifier in stratified analyses, and the association tended to be stronger for men than for women (57). A Norwegian prospective study demonstrated that lower baseline serum 25(OH)D concentrations were associated with a higher risk of T2DM, and stratified analysis by BMI showed that the inverse association was present among people with a BMI  $\leq$  23 kg/m<sup>2</sup> (72). No studies were found on the interaction between vitamin D status and genetic profile on risk of T2DM.

## **2. Study aims**

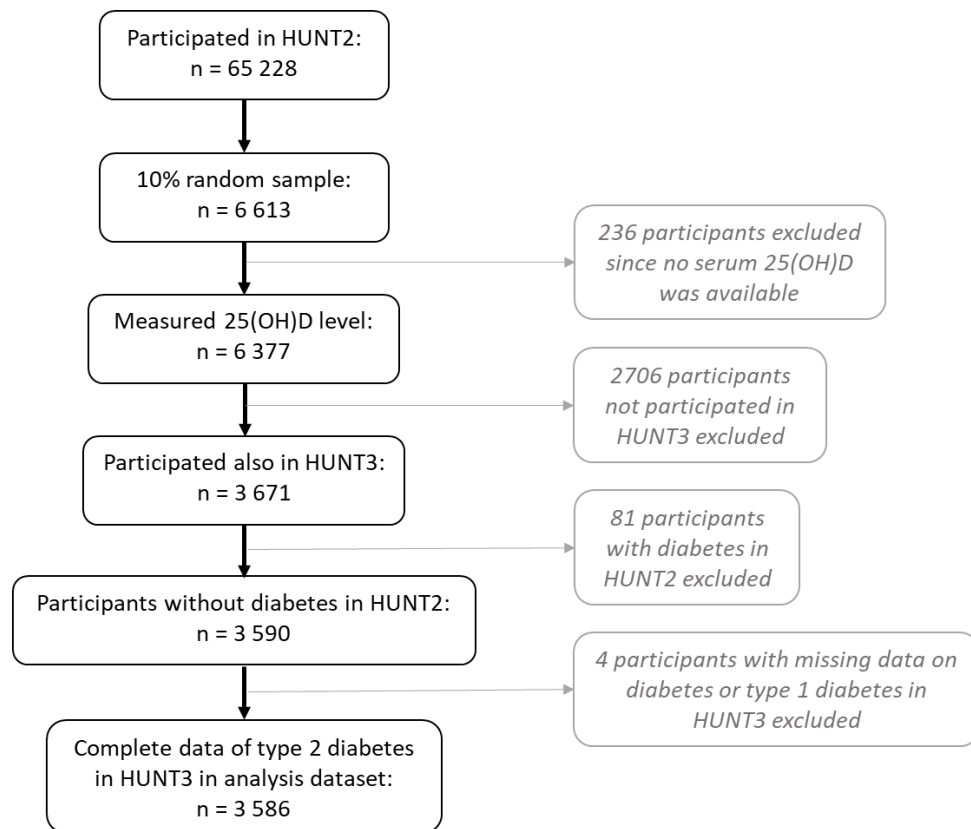
The aim of the study was to investigate the relationship between serum 25(OH)D level and risk of T2DM in adults participating in the Nord-Trøndelag Health Study (HUNT) over 11 years of follow-up. We also evaluated if sex, BMI and genetic profile for T2DM had potential interactions with low vitamin D levels on the risk of T2DM.

### **3. Methods**

#### **3.1. Study design and data collection**

The Nord-Trøndelag Health Study (HUNT) is one of the largest collections of health data in Norway. The study has enrolled about 150 000 participants aged over 19 years old in four different surveys: HUNT1 (1984-1986), HUNT2 (1995-1997), HUNT3 (2006-2008) and HUNT4 (2017-2019). The Nord-Trøndelag Study area is located in the middle of Norway, situated at latitude 64° North (73). The study population was mostly Caucasian (97%), with sociodemographic characteristics generally representative of Norway. Participants may be followed up longitudinally between the surveys and in several national health- and other registers covering the total population (73, 74). The HUNT study includes data from questionnaires, interviews, clinical measurements and biological samples (blood and urine) (73). Blood and serum samples collected in HUNT2-3-4 were stored in freezers at -80°C, in the HUNT Biobank for research use.

In this study, we linked data from HUNT2 (n=65 228) to HUNT3, in an average 11-year follow-up. Of the participants in HUNT2, a 10% random sample (n=6 613) was selected for measurement of serum 25(OH)D levels. Baseline levels were established for participants with sufficient blood sample volume (n=6 377). Among them, 3 671 also participated in HUNT3. Participants with diabetes in HUNT2, missing information on diabetes or type 1 diabetes in HUNT3 were excluded. The final analysis dataset included 3 586 participants. Figure 1 presents a flow diagram outlining the selection process of the analysis dataset.



**Figure 1.** Selection process of the analytical sample, Nord-Trøndelag Health Study (HUNT). Analytical sample comprised of participants with complete data on serum 25(OH)D, participating in HUNT2 and HUNT3, without diabetes at baseline and complete data on T2DM in HUNT3. *25(OH)D: 25-hydroxyvitamin D*

### 3.2. Serum 25(OH)D level as exposure variable

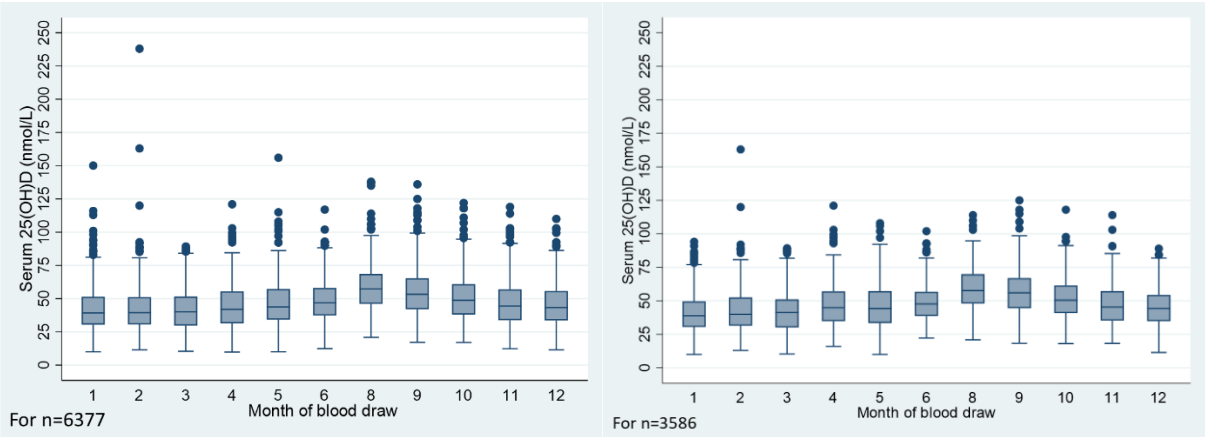
Measurement of serum 25(OH)D levels is known as the best approach to estimate body vitamin D status as the serum level integrates sun exposure, dietary intake, supplement use and storage (75, 76).

Blood samples collected in HUNT2 were stored at  $-80^{\circ}\text{C}$ . LIAISON 25-OH vitamin D TOTAL (DiaSorin, Saluggia, Italy) was used to measure serum 25(OH)D levels. It is a fully automated antibody-based chemiluminescence assay for the direct measurement of 25-OH vitamin D in human serum. The detection range of the assay is 10-375 nmol/L. The assay has an intra- and inter-assay coefficient of variation of 4% and 8% respectively. A cosinor model based on month of blood draw was used to calculate season-standardized 25(OH)D level

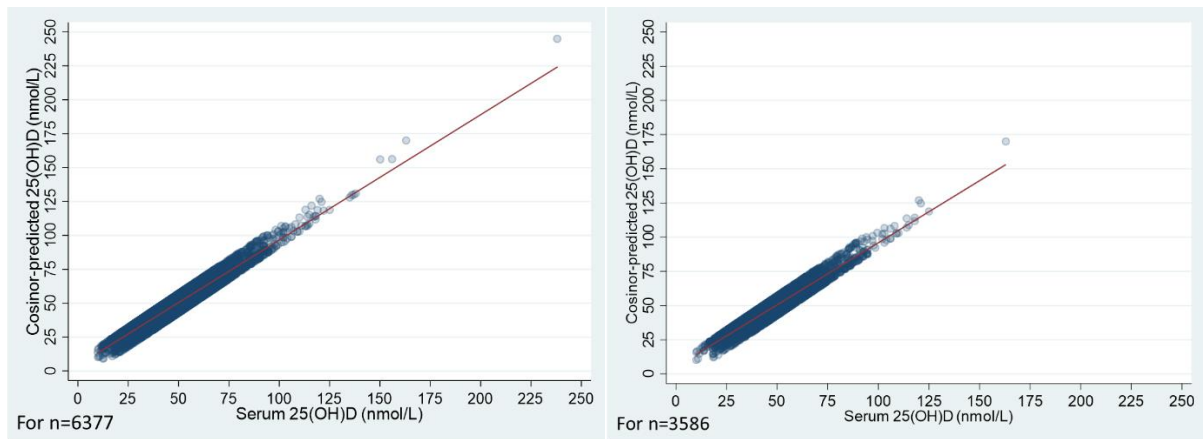
(nmol/L) in order to reduce the effect of seasonal fluctuations due to the high-latitude geographical position of Norway. This season-standardized 25(OH)D level represents the annual average value of serum 25(OH)D for each participant. It has been suggested that the cosinor model is a better approach than adjustment for seasons in regression model (77, 78). The cosinor model appears to minimize the mean squared error which includes both bias squared and variance. The season-standardized 25(OH)D levels were treated both as a continuous variable and as a categorical variable classified into two categories:  $<50$  and  $\geq 50$  nmol/L based on the widely accepted definition of vitamin D insufficiency (33).

Figure 2a) shows the distribution of measured serum 25(OH)D by month of blood draw in the 10% random sample of HUNT2 participants whose blood sample were available (n=6 377) and in the analysis sample (n=3 586). Levels of serum 25(OH)D are higher around summer and early autumn months (June-October) when sun exposure is greater, and lower in winter. Figure 2b) underlines the linear relationship between the measured and the season-standardized 25(OH)D levels. For each figure, the patterns are similar between the two samples, indicating no selection bias.

a)



b)



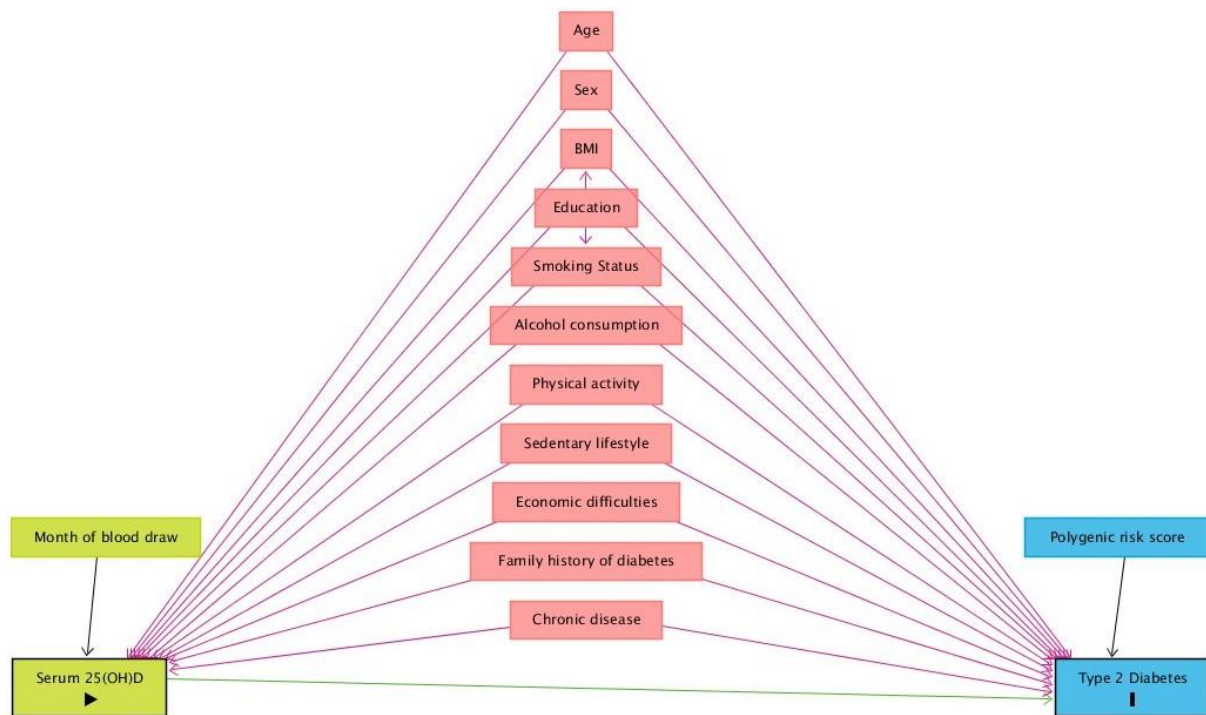
**Figure 2.** Serum 25(OH)D in the Nord-Trøndelag Health Study (HUNT).

**a)** Box plot of the distribution of measured serum 25(OH)D (nmol/L) by month of blood draw (January-December) in participants in HUNT2 (n=6377) and in the analytical sample (n=3586). Horizontal lines indicate 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles. **b)** Scatter plot comparing measured serum 25(OH)D (nmol/L) and cosinor-predicted 25(OH)D (nmol/L). The linear regression line is marked as red.

*25(OH)D*: 25-hydroxyvitamin D

### 3.3. Covariates

Several confounding factors that could bias the association between exposure and outcome variables were identified using a directed acyclic graph (DAG), which serves as a visual representation of causal assumptions by making underlying relations explicit (Figure 3) (79).



**Figure 3.** Directed acyclic graph (DAG) underlining associations between exposure, outcome variable and covariates.  
*25(OH)D: 25-hydroxyvitamin D*

Covariates data were collected in HUNT2. Sociodemographic variables included age (as a continuous variable), sex (female or male), years of education (<10, 10-12,  $\geq$ 13, or unknown (2%)), economic difficulties in the past year (yes, no, or unknown (19%)). Body weight and standing height of participants were measured in light clothing and without shoes by health professionals (73). Based on these measures, BMI was calculated as weight divided by the square value of height ( $\text{kg}/\text{m}^2$ , used as a continuous variable). Lifestyle factors included daily smoking (never, former, current, or unknown (1%)), frequency of alcohol consumption per month (0, 1-4,  $\geq$ 5, or unknown (7%)), level of physical activity (inactive or very low, low, moderate, high, or unknown (27%)), average hours of sitting time per day as a marker for sedentary life ( $\leq$ 4, 5-7,  $\geq$ 8, or unknown (18%)). Family history of diabetes was identified by asking the following question concerning the participant’s relatives (mother, father, sister, brother): “Do they have or have they had diabetes?” (yes, no, or unknown (14%)). Chronic diseases were determined by the question: “Do you suffer from any long-term illness or injury

of a physical or psychological nature that impairs your functioning in your everyday life?” (yes, no, or unknown (3%)). The unknown category for education, smoking, alcohol consumption, physical activity, sedentary life, economic difficulties, family history of diabetes, history of chronic diseases, was included in the primary analysis. This categorization of covariates has been used in previous HUNT publications (41, 59, 80).

### **3.4. Polygenic risk score for T2DM**

DNA was isolated from blood samples collected in HUNT2 and stored at the HUNT biobank. Genotyping was performed using HumanCoreExome (Illumina, San Diego, CA, USA) arrays as described elsewhere (81). Imputation was performed on samples of recent European ancestry using Minimac3 (82) from a merged reference panel constructed from the Haplotype Reference Consortium panel (83) and a local reference panel based on 2201 whole-genome sequenced HUNT participants (84).

We generated a polygenic risk score (PRS) for T2DM based on 14 SNPs identified in a large-scale genome-wide association study (GWAS) by Voight et al. (25). The 14 SNPs include rs243021 (nearby gene *BCL11A*), rs4457053 (*ZBED3*), rs972283 (*KLF14*), rs896854 (*TP53INP1*), rs13292136 (*CHCHD9*), rs231362 (*KCNQ1*), rs1552224 (*CENTD2*), rs1531343 (*HMGA2*), rs7957197 (*HNF1A*), rs11634397 (*ZFAND6*), rs8042680 (*PRC1*), rs5945326 (*DUSP9*), rs7578326 (*IRS1*) and rs1387153 (*MTNR1B*). The risk allele was coded 1 and the non-risk allele 0. The PRS for each participant was calculated by multiplying the number of risk allele by the weight and summing across the 14 SNPs (85). The weights are effect sizes derived from the GWAS results of Voight et al. (25).



### **3.5. T2DM as outcome variable**

Diabetes were defined using the following question: “Have you had or do you have diabetes?”, and/or by non-fasting blood glucose measurement. Participants were diagnosed as having diabetes if they answered yes to the question and/or had a non-fasting blood glucose level above 11 mmol/L, in accordance with the WHO guidelines (1). Incident diabetes cases were the participants who developed diabetes between HUNT2 and HUNT3 among individuals who were free of diabetes at baseline. The questionnaire-based diabetes diagnoses have been validated by comparison with medical records, and were verified in 96% of the cases (86).

Type 1 and type 2 diabetes were defined based on the values of serum glutamic acid decarboxylase antibodies (GADA). Antibody levels are expressed as an antibody index (ai) relative to a standard serum. The incident cases were classified as type 1 diabetes if the value of GADA  $\geq 0.08$  ai and as T2DM if the value of GADA  $< 0.08$  (87). One incident case of type 1 diabetes was excluded from the analysis.

### **3.6. Statistical analysis**

Baseline characteristics were presented for participants in the analysis dataset (n=3 586), and stratified by two categories of serum 25(OH)D level ( $< 50.0$  and  $\geq 50.0$  nmol/L). The baseline covariates were presented as percentages for categorical variables and mean  $\pm$  standard deviation for continuous variables. Comparisons between baseline serum 25(OH)D level categories were made for each covariate using Pearson chi-square tests for categorical covariates and t-tests for continuous covariates.

The relationship between serum 25(OH)D and risk of T2DM was evaluated using logistic regression models. Crude and adjusted odds ratios (ORs), and 95% confidence intervals (CIs) were estimated. The multivariable regression analysis was performed to control for potential

confounders: age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary life, economic difficulties, family history of diabetes and chronic disease. The higher serum 25(OH)D category ( $\geq 50.0$  nmol/L) was used as the reference group.

Potential effect modification by sex, BMI and PRS for diabetes for the risk of T2DM was assessed by the likelihood ratio test. A p-value  $< 0.05$  indicates a statistical interaction. BMI and PRS were categorized as a binary variable for stratification. The two BMI categories were defined as  $< 30$  and  $\geq 30$  kg/m<sup>2</sup>. The PRS was classified as the top one third and the other two thirds. Biological interaction of low serum 25(OH)D and high PRS for risk of T2DM was evaluated by the relative excess risk due to interaction (RERI). A RERI  $> 0$  and the lower limit of 95% CI  $> 0$  suggest a biological interaction (88).

Multiple imputation (mi command in Stata) was used to address possible biases due to missing data of the covariates, assuming missing at random (89). White et al. showed that the number of imputations performed should be at least higher than the percentage of missing values in the analysis (90). As missing data in the analysis dataset were around 10%, 10 imputations were performed to follow the recommendations. All statistical analyses were performed using Stata statistical software: release 16 (College Station, TX: StataCorp LLC).

### **3.7. Ethical reflections**

Participation in the HUNT study was voluntary, and written informed consent was obtained from all participants prior to participation. The data were already collected at the HUNT Research Center. All names and personal ID numbers were removed; therefore, all information was de-identified. Nobody was contacted for gathering of new data. This master project was conducted as a sub-study under the approval of the Regional Committees for Medical Research Ethics – REK (2015-1562 REK sør-øst C).

## **4. Results**

### **4.1. Baseline characteristics of the study population**

The study population was more represented by women than men (54.7% vs 45.3%), with approximately the same distribution within the two categories of serum 25(OH)D level (Table 1). More participants had baseline serum 25(OH)D levels under 50 nmol/L than above (57.6% vs 42.4%). Study participants were 46.7 years old on average; they were relatively older in the group with serum 25(OH)D level above 50 nmol/L compared to the other group (mean age 47.9 vs 45.7 years). The study participants with baseline serum 25(OH)D levels under 50 nmol/L had a BMI of 1.3 point higher compared to the group with serum 25(OH)D levels above 50 nmol/L (26.6 vs 25.3 kg/m<sup>2</sup>). More study participants with serum 25(OH)D levels under 50 nmol/L had a high polygenic risk score for T2DM compared to the other group with 25(OH)D levels above 50 nmol/L (33.6% vs 30.9%). Participants with baseline 25(OH)D levels under 50 nmol/L tended to report the following characteristics in greater proportions than participants of baseline 25(OH)D levels above 50 nmol/L: <10 years of education, current smoker, alcohol abstainer, inactive or very low physical activity, and economic difficulties. There was no clear pattern for chronic diseases and total sitting time.

**Table 1.** Baseline characteristics of subjects overall and by baseline serum 25(OH)D levels in the HUNT2 study, 1995-1997.

| Characteristic                                      | Overall<br>n=3586 | Baseline seasonal-standardized serum<br>25(OH)D level (nmol/L) |                 | p-value <sup>1</sup> |
|---|-------------------|--|-----------------|----------------------|
|   |                   | <50.0<br>n=2066  | ≥50.0<br>n=1520 |                      |
| <b>Age (years)</b>                                  | 46.7±13.4         | 45.7±12.8  | 47.9±14.1       | <0.001               |
| <b>Sex</b>  |                   |  |                 | 0.75                 |
| Female  | 1961 (54.7%)      | 1125 (54.5%)   | 836 (55.0%)     |                      |
| Male  | 1625 (45.3%)      | 941 (45.5%)  | 684 (45.0%)     |                      |
| <b>BMI (kg/m<sup>2</sup>)</b>                       | 26.1±3.8          | 26.6±4.1   | 25.3±3.3        | <0.001               |
| <b>Education (years)</b>                            |                   |  |                 | 0.001                |
| <10   | 1057 (29.5%)      | 632 (30.6%)  | 425 (28.0%)     |                      |
| 10-12   | 1324 (36.9%)      | 776 (37.6%)  | 548 (36.1%)     |                      |
| ≥13   | 1116 (31.1%)      | 596 (28.9%)  | 520 (34.2%)     |                      |
| Unknown   | 89 (2.5%)         | 62 (3.0%)  | 27 (1.8%)       |                      |
| <b>Smoking status</b>                               |                   |  |                 | 0.002                |
| Never   | 1611 (44.9%)      | 906 (43.9%)  | 705 (46.4%)     |                      |
| Former  | 999 (27.9%)       | 548 (26.5%)  | 451 (29.7%)     |                      |
| Current   | 941 (26.2%)       | 591 (28.6%)  | 350 (23.0%)     |                      |
| Unknown   | 35 (1.0%)         | 21 (1.0%)  | 14 (0.9%)       |                      |
| <b>Alcohol intake per month</b>                     |                   |  |                 | <0.001               |
| 0 (abstainer)                                       | 1054 (29.4%)      | 653 (31.6%)  | 401 (26.4%)     |                      |
| 1-4   | 1845 (51.5%)      | 1052 (50.9%)   | 793 (52.2%)     |                      |
| ≥5  | 430 (12.0%)       | 215 (10.4%)  | 215 (14.1%)     |                      |
| Unknown   | 257 (7.2%)        | 146 (7.1%)   | 111 (7.3%)      |                      |
| <b>Physical activity</b>                            |                   |  |                 | <0.001               |
| Inactive or very low                                | 744 (20.8%)       | 481 (23.3%)  | 263 (17.3%)     |                      |
| Low   | 672 (18.7%)       | 395 (19.1%)  | 277 (18.2%)     |                      |
| Moderate  | 895 (25.0%)       | 462 (22.4%)  | 433 (28.5%)     |                      |
| High  | 324 (9.0%)        | 152 (7.4%)   | 172 (11.3%)     |                      |
| Unknown   | 951 (26.5%)       | 576 (27.9%)  | 375 (24.7%)     |                      |
| <b>Sitting time, hours/day</b>                      |                   |  |                 | 0.001                |
| ≤4  | 959 (26.7%)       | 538 (26.0%)  | 421 (27.7%)     |                      |
| 5-7   | 880 (24.5%)       | 479 (23.2%)  | 401 (26.4%)     |                      |
| ≥8  | 1088 (30.3%)      | 626 (30.3%)  | 462 (30.4%)     |                      |
| Unknown   | 659 (18.4%)       | 423 (20.5%)  | 236 (15.5%)     |                      |
| <b>Economic difficulties</b>                        |                   |  |                 | <0.001               |
| No  | 2122 (59.2%)      | 1154 (55.9%)   | 968 (63.7%)     |                      |
| Yes   | 782 (21.8%)       | 495 (24.0%)  | 287 (18.9%)     |                      |
| Unknown   | 682 (19.0%)       | 417 (20.2%)  | 265 (17.4%)     |                      |
| <b>Family history of diabetes</b>                   |                   |  |                 | 0.002                |
| No  | 2505 (69.9%)      | 1412 (68.3%)   | 1093 (71.9%)    |                      |
| Yes   | 585 (16.3%)       | 332 (16.1%)  | 253 (16.6%)     |                      |
| Unknown   | 496 (13.8%)       | 322 (15.6%)  | 174 (11.5%)     |                      |
| <b>Polygenic risk score for<br/>type 2 diabetes</b> |                   |  |                 | 0.04                 |

|                        |              |              |              |      |
|------------------------|--------------|--------------|--------------|------|
| Low (bottom 2/3)       | 2330 (65.0%) | 1311 (63.5%) | 1019 (67.0%) | 0.44 |
| High (top 1/3)         | 1164 (32.5%) | 694 (33.6%)  | 470 (30.9%)  |      |
| Unknown                | 92 (2.6%)    | 61 (3.0%)    | 31 (2.0%)    |      |
| <b>Chronic disease</b> |              |              |              |      |
| No                     | 2532 (70.6%) | 1476 (71.4%) | 1056 (69.5%) |      |
| Yes                    | 960 (26.8%)  | 538 (26.0%)  | 422 (27.8%)  |      |
| Unknown                | 94 (2.6%)    | 52 (2.5%)    | 42 (2.8%)    |      |

25(OH)D: 25-hydroxyvitamin D; BMI: body mass index

Data are given as number of subjects (column percentage) or mean±standard deviation.

<sup>1</sup>Comparisons between baseline serum 25(OH)D level categories; p-values reported using Pearson chi-square tests for categorical covariates or t-tests for continuous covariates

#### 4.2. Serum 25(OH)D level and risk of T2DM

A total of 104 (2.9%) participants were identified with T2DM during the 11-year follow up period (Table 2). The risk of T2DM during the 11-year follow up was 3.6% in the <50.0 nmol/L group compared to 2.0% in the ≥50.0 nmol/L group.

An 85% (95% CI=1.20- 2.83) higher risk of incident T2DM was observed in people with serum 25(OH)D level under 50 nmol/L compared to those with serum 25(OH)D level above 50 nmol/L in the crude analysis. However, the relative risk was not statistically significant in the adjusted model (OR=1.55, 95% CI=0.97-2.49). Every 25 unit decrease in serum 25(OH)D was associated with a 52% higher risk of getting T2DM (95% CI=1.09-2.12), but this risk was attenuated to 31% and not statistically significant in the adjusted model (95% CI=0.90-1.89). The analyses after performing multiple imputation of missing data showed similar results (Supplementary table S1).

**Table 2.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up.

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%) | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|----------|--------------------|-----------------------|
| <i>Categorical</i>                                 |                     |              |          |                    |                       |
| <50.0  | 2066                | 74           | 3.6%     | 1.85 (1.20 - 2.83) | 1.55 (0.97 - 2.49)    |
| ≥50.0  | 1520                | 30           | 2.0%     | 1.00 (reference)   | 1.00 (reference)      |
| <i>Continuous<sup>1</sup></i>                      |                     |              |          |                    |                       |
|  | 3586                | 104          | 2.9%     | 1.52 (1.09 - 2.12) | 1.31 (0.90 - 1.89)    |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of  $\geq 11$  mmol/L and/or reported diabetes.

<sup>1</sup>per 25 nmol/L decrease in serum 25(OH)D

\* Adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

### 4.3. Effect modification by sex, BMI and genetic profile of T2DM

We also studied the association between serum 25(OH)D and risk of T2DM stratified by sex, BMI and PRS for diabetes. The p-value of likelihood ratio test for interaction between serum 25(OH)D and sex was 0.69, and between serum 25(OH)D and BMI 0.71, indicating no evidence of statistical interaction. The results are presented in supplementary tables S2 and S3.

The likelihood ratio test showed evidence of statistical interaction between serum 25(OH)D and PRS. Table 3 presents results using different cut-off values of PRS for stratification. The higher the PRS is, the more the person is genetically susceptible to get T2DM. The ORs for 25(OH)D-diabetes association showed a dose-response relationship in the group with higher PRS when the cut-off value changed from the top 1/2 to the top 1/5 of the PRS, whereas the ORs remained similar in the lower PRS group. Using the top one third PRS as the cut-off value seemed to show the largest statistical power; and p-value for interaction was 0.04.

**Table 3.** Different cut-off groups evaluated for effect modification of T2DM PRS on association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up.

| PRS cut-off groups | No. of participants | No. of cases | Risk (%) |  | OR*  | 95% CI       |
|--------------------|---------------------|--------------|----------|--|------|--------------|
| PRS top 1/2        | 1747                | 59           | 3.4      |  | 1.82 | (0.95-3.49)  |
| PRS bottom 1/2     | 1747                | 38           | 2.2      |  | 0.96 | (0.45-2.05)  |
| PRS top 1/3        | 1164                | 40           | 3.4      |  | 2.98 | (1.19-7.45)  |
| PRS bottom 2/3     | 2330                | 57           | 2.5      |  | 0.92 | (0.50-1.71)  |
| PRS top 1/4        | 873                 | 34           | 3.9      |  | 3.24 | (1.18-8.87)  |
| PRS bottom 3/4     | 2621                | 63           | 2.4      |  | 1.04 | (0.58-1.85)  |
| PRS top 1/5        | 698                 | 27           | 3.9      |  | 4.58 | (1.33-15.76) |
| PRS bottom 4/5     | 2796                | 70           | 2.5      |  | 1.11 | (0.64-1.92)  |

T2DM: type 2 diabetes mellitus; PRS: polygenic risk score; 25(OH)D: 25-hydroxyvitamin D; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of  $\geq 11$  mmol/L and/or reported diabetes.

\*OR and 95% CI represent the risk in the group of 25(OH)D level  $< 50$  nmol/L ( $> 50$  nmol/L as reference) and were adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

The detailed results presenting the association between serum 25(OH)D and risk of T2DM stratified by PRS (cut-off value at 1/3) are shown in Table 4. Participants with a high PRS were more likely to get T2DM than those with a low PRS (3.4% vs 2.4%). Serum 25(OH)D level below 50 nmol/L seemed to be associated with a higher risk of T2DM in both the PRS groups (2.7% vs 2.2% for low PRS group, and 4.8% vs 1.5% for high PRS group). However, after adjustment for confounders, level of 25(OH)D under 50 nmol/L was associated with a significantly increased risk of T2DM in the higher PRS group (OR=2.98, 95% CI=1.19-7.45), but there was no association between low 25(OH)D level and risk of T2DM in the lower PRS group (OR=0.92, 95% CI=0.50-1.71). The analyses after performing multiple imputation of missing data showed similar results (Supplementary table S4).

**Table 4.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up, stratified by PRS of T2DM.

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%)    | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|-------------|--------------------|-----------------------|
| <b>Polygenic risk scores</b>                       |                     |              |             |                    |                       |
| <b>Low score (2/3)</b>                             | <b>2330</b>         | <b>57</b>    | <b>2.4%</b> |                    |                       |
| <50.0  | 1311                | 35           | 2.7%        | 1.24 (0.72 - 2.13) | 0.92 (0.50 - 1.71)    |
| ≥50.0  | 1019                | 22           | 2.2%        | 1.00 (reference)   | 1.00 (reference)      |
| <b>High score (1/3 top)</b>                        | <b>1164</b>         | <b>40</b>    | <b>3.4%</b> |                    |                       |
| <50.0  | 694                 | 33           | 4.8%        | 3.30 (1.45 - 7.53) | 2.98 (1.19 - 7.45)    |
| ≥50.0  | 470                 | 7            | 1.5%        | 1.00 (reference)   | 1.00 (reference)      |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; PRS: polygenic risk score; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of  $\geq 11$  mmol/L and/or reported diabetes.

\* Adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

#### 4.4. Biological interaction

We evaluated the potential biological interaction of low serum 25(OH)D and high PRS (the top 1/3 score) for risk of T2DM using the RERI. We found the RERI equal to 1.28 (95% CI=0.31-2.25). Since the 95% CI of RERI did not include 0, there is biological interaction of low serum 25(OH)D and high PRS for the risk of T2DM.



## **5. Discussion**

### **5.1. Main findings**

We investigated the relationship between serum vitamin D level and the risk of T2DM in a sample of 3 586 adults over 11 years of follow-up, during which 104 incident T2DM were identified. In this study, we found that lower 25(OH)D levels seemed to be associated with an increased risk of T2DM in adults. The inverse association was more evident and stronger among people who were genetically predisposed to T2DM than those who were not. In the group with a top one third PRS for T2DM, the participants with serum 25(OH)D level under 50 nmol/L had almost three times increased risk to get T2DM compared to those with serum 25(OH)D level above 50 nmol/L. We also observed a biological interaction of low serum 25(OH)D and high PRS for the risk of T2DM. Effect modification by sex and BMI did not seem to be present.

### **5.2. Comparison with previous studies**

#### *5.2.1. Vitamin D status as an independent risk factor for T2DM*

Many observational longitudinal studies have investigated the association of serum 25(OH)D level with risk of T2DM (57, 69-71, 91). Results of the present study support previous reports of the potential negative impact of vitamin D insufficiency on risk of T2DM. However, the association was not statistically significant in our adjusted models. Indeed, serum 25(OH)D <50.0 nmol/L was associated with a 55% increased risk of T2DM (OR=1.55, 95% CI=0.97-2.49) and a 31% increased risk of T2DM was present per 25 nmol/L decrease in serum 25(OH)D (OR=1.31, 95% CI=0.90-1.89). The wide CIs imply an inadequate sample size in our study. Our results are consistent with those in a Danish study that found a 20% increased risk of T2DM per 25 nmol/L decrease in serum 25(OH)D (70). Contrastingly, another Danish study did not

find a significant association between low 25(OH)D status and incident diabetes after adjustment for confounders (71). This may be due to the short follow-up duration of 5 years of the study. Forouhi et al. designed a nested case-cohort from the EPIC-Norfolk study, including 621 incident T2DM cases (69). Compared to participants with 25(OH)D  $\leq$  49 nmol/L, the risk of T2DM was halved in those with 25(OH)D  $>$ 80 nmol/L (69). The variation in the estimate magnitude of the risk of T2DM between these previous studies and our study could be partially explained by difference in categorization of serum 25(OH)D.

Several recent meta-analyses have summarized results from the observational studies. The meta-analysis performed by Song et al. combined data from 21 prospective studies, including 4 996 incident T2DM cases in diverse population (57). They found a summary relative risk for T2DM of 0.62 (95% CI=0.54-0.70) comparing the highest to lowest category of 25(OH)D levels, indicating a significant inverse association between 25(OH)D levels and risk of T2DM (57). In a more recent meta-analysis, using data from 22 longitudinal studies of populations of European descent and including 8492 cases of T2DM, a 25 nmol/L decrease in 25(OH)D concentration was significantly associated with a 21% increased risk of T2DM (91).

Recent research has also applied Mendelian randomisation (MR) to study the potential causal relationship between low vitamin D and risk of T2DM. MR approach attempts to overcome issues of confounding and reverse causation usually present in observational studies (92, 93). A recent MR study, based on data from the China Kadoorie Biobank, found that higher vitamin D status was associated with lower risk of diabetes, and provided moderate evidence for a causal protective effect of higher 25(OH)D levels on prevention of T2DM (94). Similarly, another recent MR analysis, based on 898 130 individuals of European ancestry from 32 studies, revealed a significant inverse association of genetically predicted serum 25(OH)D levels with risk of T2DM (95). Results of both studies were robust in the analysis using genetic variants affecting the synthesis of vitamin D. In contrast, a MR study including 28 144 cases found no

association between genetically predicted serum 25(OH)D and incidence of T2DM (91). However, it is important to note that this last study was using SNPs within or near four genes involved in both vitamin D synthesis and metabolism, without making any distinction in the analysis, and the sample size for the last MR study was relatively small.

### *5.2.2. Effect modification by sex, BMI and genetic profile of T2DM*

We evaluated sex, BMI and PRS for diabetes as potential effect modifiers of the relationship between vitamin D insufficiency and risk of T2DM. To our knowledge, this is one of the first prospective cohort studies to investigate effect modification by genetic profile for T2DM for the association between serum 25(OH)D and risk of T2DM.

Sex did not appear to be an effect modifier in the current study. Similarly, the Melbourne Collaborative Cohort Study and a prospective study from the Tromsø Study did not find that the association varied by sex (72, 96). Association tended to be stronger for men than for women in a meta-analysis, but the difference did not reach statistical significance (57). Another analysis based on pooling two nested case-control studies of 412 incident T2DM cases and 986 control subjects in Finland found that higher serum 25(OH)D was related to a reduced incidence of T2DM in men but not in women (97). However, this result may be due to residual confounding from dietary factors and cannot be generalized due to its small sample size.

Our data did not indicate possible effect modification by BMI. A population-based cohort from the Tromsø Study found an inverse association among people with a BMI  $\leq 23$  kg/m<sup>2</sup>, while there was no association in the other BMI categories (72). However, these results must be interpreted cautiously as the number of cases was low in the leanest BMI category (72).

Our results were indicative of potential effect modification by genetic profile of T2DM. Among participants with top one third PRS, those with baseline serum 25(OH)D under 50 nmol/L had a three fold risk of getting T2DM. No significant association was found in the low PRS group. These findings indicate that low vitamin D status may be an important risk factor for T2DM for people genetically predisposed to T2DM. In current literature, the effect modification of T2DM PRS for vitamin D-T2DM association is unknown. Our study also found a biological interaction of serum 25(OH)D under 50 nmol/L and the high PRS for the risk of T2DM. The results suggest that there was a 1.28 relative excess risk of T2DM due to the synergistic interaction of serum 25(OH)D under 50 nmol/L and the high PRS compared to the additive effect of both factors.

### **5.3. Potential biological mechanisms**

Several reviews have highlighted potential mechanisms for a role of vitamin D in the pathogenesis of T2DM (62, 98, 99). VDRs are present in many cell types and organs, and 1,25(OH)<sub>2</sub>D is locally produced in several extrarenal organs, including  $\beta$ -pancreatic cells, supporting potential broad-ranging effects of vitamin D outside of skeletal health such as T2DM (62). Vitamin D may have both direct (via activation of the VDR) and indirect (via the regulation of calcium homeostasis) effects on various mechanisms linked to the pathophysiology of T2DM, including impaired pancreatic- $\beta$  cell function and insulin resistance (62). Vitamin D could also indirectly contribute to the pathogenesis of T2DM via regulation of inflammatory processes, such as production of cytokines, associated with insulin resistance and  $\beta$ -cell death (99).

The biological mechanism by which low serum 25(OH)D level in people genetically predisposed to T2DM might lead to substantially increased risk of T2DM is not elucidated.

GWAS have identified hundreds of loci containing common variants robustly associated with the risk of T2DM (24), and explanation of the mechanisms through which these act can provide novel pathophysiological insights. Our study detected a gene-environment interaction effect, whereby low vitamin D status and genetic variants for T2DM showed a synergistic effect on the risk of T2DM. More fundamental research is warranted to understand the underlying mechanism of this biological interaction.

#### **5.4. Strengths**

To our knowledge, this prospective cohort study is the first to provide an insight into the influence of genetic profile for T2DM on the relationship between serum 25(OH)D and risk of T2DM. Serum 25(OH)D was measured in a large random sample of Norwegian adults aged over 19 years in HUNT2. The pattern of 25(OH)D was similar in the analysis dataset and in the sub-cohort of HUNT2 with measured 25(OH)D levels, which suggests less selection bias. In addition, the study has a long follow-up duration. A large range of sociodemographic and lifestyle variables were available in the questionnaire data, giving the possibility to include important potential confounders in the analysis, thus strengthening the validity of our results. Use of the cosinor model to calculate season-standardized serum 25(OH)D rather than adjustment for season of blood draw in the models minimized the mean squared error, as argued previously.

#### **5.5. Limitations**

There are several limitations related to our study. Participation rate declined from HUNT2 (70%) to HUNT3 (54%), increasing the possibility of selection bias. Participants in the HUNT

studies were shown to be healthier than nonparticipants (100). Due to the small sample size in the analysis cohort, serum 25(OH)D could only be classified into two categories. Data on lifestyle factors could be subject to misclassification due to self-reported methods and thus resulting in residual confounding; the outcome variable about diabetes is especially of concern. However, the questionnaire-based diabetes diagnoses of the HUNT study were verified in 96% of the cases by comparing them with medical records (86). Data about serum 25(OH)D and lifestyle were collected at baseline, meaning that we could not evaluate any changes over the 11 years follow-up. We cannot exclude measurement error due to the single serum 25(OH)D measure used for the study. Nevertheless, a previous Norwegian study demonstrated that 25(OH)D concentrations were relatively stable up to 14 years of follow-up (101). Missing data on baseline characteristics were classified as “unknown” category which may have resulted in residual confounding. However, results before and after multiple imputation are similar. Participants are mainly Caucasian, reducing the generalizability to other ethnic populations. Indeed, the genetic susceptibility for T2DM may differ between populations. For instance, some SNPs associated to T2DM are more common in South Asian than in Europeans (102).

## **6. Conclusion**

Overall, this is the first prospective cohort study to evaluate vitamin D status and its interaction with genetic profile on the risk of T2DM. We have shown that low serum 25(OH)D was associated with an increased risk of T2DM in Norwegian adults, especially for people with high polygenic risk score for T2DM. Future epidemiological studies with larger sample size should seek to replicate our results, while fundamental research on the underlying mechanisms could support the biological plausibility of our findings. Given the high prevalence of vitamin D insufficiency and T2DM worldwide, our study provided insights into new ways to prevent T2DM, particularly by incorporating the genetic factor of the disease with modifiable factors in prevention.

## 7. References

1. WHO. Global report on diabetes.; 2016.
2. Davies MJ, Tringham JR, Troughton J, Khunti KK. Prevention of Type 2 diabetes mellitus. A review of the evidence and its application in a UK setting. *Diabet Med.* 2004;21(5):403-14.
3. Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes research and clinical practice.* 2014;103(2):137-49.
4. NIPH. Diabetes in Norway 2017 [Available from: <https://www.fhi.no/en/op/hin/health-disease/diabetes-in-norway---public-health/>].
5. Lipscombe LL, Hux JE. Trends in diabetes prevalence, incidence, and mortality in Ontario, Canada 1995-2005: a population-based study. *Lancet.* 2007;369(9563):750-6.
6. Choi YJ, Kim HC, Kim HM, Park SW, Kim J, Kim DJ. Prevalence and management of diabetes in Korean adults: Korea National Health and Nutrition Examination Surveys 1998-2005. *Diabetes care.* 2009;32(11):2016-20.
7. Williams J, Loeffler M. Global Trends in Type 2 Diabetes, 2007-2017. *Jama.* 2019;322(16):1542.
8. Logue J, Walker JJ, Colhoun HM, Leese GP, Lindsay RS, McKnight JA, et al. Do men develop type 2 diabetes at lower body mass indices than women? *Diabetologia.* 2011;54(12):3003-6.
9. Geer EB, Shen W. Gender differences in insulin resistance, body composition, and energy balance. *Gen Med.* 2009;6 Suppl 1(Suppl 1):60-75.
10. Magliano DJ, Islam RM, Barr ELM, Gregg EW, Pavkov ME, Harding JL, et al. Trends in incidence of total or type 2 diabetes: systematic review. *Bmj.* 2019;366:15003.
11. Alleyne G, Stuckler D, Alwan AJG, Health. The hope and the promise of the UN Resolution on non-communicable diseases. 2010;6(1):15.
12. Sørensen M, Arneberg F, Line TM, Berg TJ. Cost of diabetes in Norway 2011. *Diabetes Res Clin Pract.* 2016;122:124-32.
13. Gillies CL, Abrams KR, Lambert PC, Cooper NJ, Sutton AJ, Hsu RT, et al. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. *Bmj.* 2007;334(7588):299.
14. Whiting D, Unwin N, Roglic G. Diabetes: equity and social determinants. *Equity, social determinants and public health programmes.* 2010;77:94.
15. Vazquez G, Duval S, Jacobs Jr DR, Silventoinen KJEr. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. 2007;29(1):115-28.
16. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. *Diabetes care.* 2007;30(3):744-52.
17. WHO. Obesity and overweight 2020 [Available from: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight>].
18. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *Jama.* 2007;298(22):2654-64.
19. Romaguera D, Norat T, Wark PA, Vergnaud AC, Schulze MB, van Woudenberg GJ, et al. Consumption of sweet beverages and type 2 diabetes incidence in European adults: results from EPIC-InterAct. *Diabetologia.* 2013;56(7):1520-30.
20. Bellou V, Belbasis L, Tzoulaki I, Evangelou E. Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses. *PLoS One.* 2018;13(3):e0194127.



21. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. *Nature*. 2016;536(7614):41-7.
22. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*. 2005;365(9467):1333-46.
23. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467-84.
24. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. 2018;50(11):1505-13.
25. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*. 2010;42(7):579-89.
26. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proença C, et al. Genetic variant near *IRS1* is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nature genetics*. 2009;41(10):1110-5.
27. Spiro A, Buttriss JL. Vitamin D: An overview of vitamin D status and intake in Europe. *Nutrition bulletin*. 2014;39(4):322-50.
28. Holick MF. The vitamin D deficiency pandemic: Approaches for diagnosis, treatment and prevention. *Rev Endocr Metab Disord*. 2017;18(2):153-65.
29. Christakos S, Dhawan P, Verstuyf A, Verlinden L, Carmeliet G. Vitamin D: Metabolism, Molecular Mechanism of Action, and Pleiotropic Effects. *Physiol Rev*. 2016;96(1):365-408.
30. Macdonald HM, Mavroei A, Fraser WD, Darling AL, Black AJ, Aucott L, et al. Sunlight and dietary contributions to the seasonal vitamin D status of cohorts of healthy postmenopausal women living at northerly latitudes: a major cause for concern? *Osteoporos Int*. 2011;22(9):2461-72.
31. Holick MF, Chen TC. Vitamin D deficiency: a worldwide problem with health consequences. *Am J Clin Nutr*. 2008;87(4):1080s-6s.
32. Ross AC, Taylor CL, Yaktine AL, Del Valle HB. Institute of Medicine Committee to review dietary reference intakes for vitamin D and calcium. Food and Nutrition Board. 2011.
33. Holick MF. Vitamin D deficiency. *N Engl J Med*. 2007;357(3):266-81.
34. Pilz S, März W, Cashman KD, Kiely ME, Whiting SJ, Holick MF, et al. Rationale and Plan for Vitamin D Food Fortification: A Review and Guidance Paper. *Front Endocrinol (Lausanne)*. 2018;9:373-.
35. Holick MF. High prevalence of vitamin D inadequacy and implications for health. *Mayo Clin Proc*. 2006;81(3):353-73.
36. Palacios C, Gonzalez L. Is vitamin D deficiency a major global public health problem? *The Journal of steroid biochemistry and molecular biology*. 2014;144 Pt A:138-45.
37. MacLaughlin J, Holick MF. Aging decreases the capacity of human skin to produce vitamin D<sub>3</sub>. *J Clin Invest*. 1985;76(4):1536-8.
38. Cashman KD, Dowling KG, Škrabáková Z, Gonzalez-Gross M, Valtueña J, De Henauw S, et al. Vitamin D deficiency in Europe: pandemic? *The American journal of clinical nutrition*. 2016;103(4):1033-44.
39. Schleicher RL, Sternberg MR, Looker AC, Yetley EA, Lacher DA, Sempos CT, et al. National Estimates of Serum Total 25-Hydroxyvitamin D and Metabolite Concentrations Measured by Liquid Chromatography-Tandem Mass Spectrometry in the US Population during 2007-2010. *The Journal of nutrition*. 2016;146(5):1051-61.
40. Cashman KD, Sheehy T, O'Neill CM. Is vitamin D deficiency a public health concern for low middle income countries? A systematic literature review. *Eur J Nutr*. 2019;58(1):433-53.

41. Larose TL, Chen Y, Camargo CA, Jr., Langhammer A, Romundstad P, Mai XM. Factors associated with vitamin D deficiency in a Norwegian population: the HUNT Study. *J Epidemiol Community Health*. 2014;68(2):165-70.
42. Ministers NCo. Nordic Nutrition Recommendations 2012. Integrating Nutrition and Physical Activity. 2014.
43. Cashman KD, Kiely M. Recommended dietary intakes for vitamin D: Where do they come from, what do they achieve and how can we meet them? *J Hum Nutr Diet*. 2014;27(5):434-42.
44. Hoel DG, Berwick M, de Gruijl FR, Holick MF. The risks and benefits of sun exposure 2016. *Dermato-endocrinology*. 2016;8(1):e1248325.
45. Himbert C, Ose J, Delphan M, Ulrich CM. A systematic review of the interrelation between diet- and surgery-induced weight loss and vitamin D status. *Nutr Res*. 2017;38:13-26.
46. Black LJ, Walton J, Flynn A, Cashman KD, Kiely M. Small Increments in Vitamin D Intake by Irish Adults over a Decade Show That Strategic Initiatives to Fortify the Food Supply Are Needed. *The Journal of nutrition*. 2015;145(5):969-76.
47. Kiely M, Cashman K, Consortium O. The ODIN project: Development of food-based approaches for prevention of vitamin D deficiency throughout life. *Nutrition bulletin*. 2015;40(3):235-46.
48. Cashman KD, Kiely M. EURRECA—Estimating vitamin D requirements for deriving dietary reference values. *Critical reviews in food science and nutrition*. 2013;53(10):1097-109.
49. Francis RM, Aspray TJ, Fraser W, Macdonald H, Patel S, Mavroei A, et al. Vitamin D and Bone Health: A Practical Clinical Guideline for Patient Management. England: National Osteoporosis Society; 2018.
50. Holick MF. Resurrection of vitamin D deficiency and rickets. *J Clin Invest*. 2006;116(8):2062-72.
51. Bischoff-Ferrari HA, Giovannucci E, Willett WC, Dietrich T, Dawson-Hughes B. Estimation of optimal serum concentrations of 25-hydroxyvitamin D for multiple health outcomes. *The American journal of clinical nutrition*. 2006;84(1):18-28.
52. Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrell A, et al. A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. *Genome Res*. 2010;20(10):1352-60.
53. DeLuca HF. Overview of general physiologic features and functions of vitamin D. *The American journal of clinical nutrition*. 2004;80(6 Suppl):1689S-96S.
54. Morris HA, Anderson PH. Autocrine and paracrine actions of vitamin d. *Clin Biochem Rev*. 2010;31(4):129-38.
55. Zipitis CS, Akobeng AK. Vitamin D supplementation in early childhood and risk of type 1 diabetes: a systematic review and meta-analysis. *Arch Dis Child*. 2008;93(6):512-7.
56. Pittas AG, Dawson-Hughes B, Li T, Van Dam RM, Willett WC, Manson JE, et al. Vitamin D and calcium intake in relation to type 2 diabetes in women. *Diabetes care*. 2006;29(3):650-6.
57. Song Y, Wang L, Pittas AG, Del Gobbo LC, Zhang C, Manson JE, et al. Blood 25-hydroxy vitamin D levels and incident type 2 diabetes: a meta-analysis of prospective studies. *Diabetes care*. 2013;36(5):1422-8.
58. Garland CF, Garland FC, Gorham ED, Lipkin M, Newmark H, Mohr SB, et al. The role of vitamin D in cancer prevention. *Am J Public Health*. 2006;96(2):252-61.
59. Sun Y-Q, Langhammer A, Skorpen F, Chen Y, Mai X-M. Serum 25-hydroxyvitamin D level, chronic diseases and all-cause mortality in a population-based prospective cohort: the HUNT Study, Norway. *BMJ Open*. 2017;7(6):e017256.

60. Hu FB, Meigs JB, Li TY, Rifai N, Manson JE. Inflammatory markers and risk of developing type 2 diabetes in women. *Diabetes*. 2004;53(3):693-700.
61. Lillioja S, Mott DM, Spraul M, Ferraro R, Foley JE, Ravussin E, et al. Insulin resistance and insulin secretory dysfunction as precursors of non-insulin-dependent diabetes mellitus. Prospective studies of Pima Indians. *N Engl J Med*. 1993;329(27):1988-92.
62. Pittas AG, Lau J, Hu FB, Dawson-Hughes B. The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2007;92(6):2017-29.
63. Scragg R, Sowers M, Bell C. Serum 25-hydroxyvitamin D, diabetes, and ethnicity in the Third National Health and Nutrition Examination Survey. *Diabetes care*. 2004;27(12):2813-8.
64. Zhao G, Ford ES, Li C. Associations of serum concentrations of 25-hydroxyvitamin D and parathyroid hormone with surrogate markers of insulin resistance among U.S. adults without physician-diagnosed diabetes: NHANES, 2003-2006. *Diabetes care*. 2010;33(2):344-7.
65. Zhang FF, Al Hooti S, Al Zenki S, Alomirah H, Jamil KM, Rao A, et al. Vitamin D deficiency is associated with high prevalence of diabetes in Kuwaiti adults: results from a national survey. *BMC Public Health*. 2016;16:100.
66. Luo C, Wong J, Brown M, Hooper M, Molyneaux L, Yue DK. Hypovitaminosis D in Chinese type 2 diabetes: lack of impact on clinical metabolic status and biomarkers of cellular inflammation. *Diab Vasc Dis Res*. 2009;6(3):194-9.
67. Majumdar V, Nagaraja D, Christopher R. Vitamin D status and metabolic syndrome in Asian Indians. *Int J Obes (Lond)*. 2011;35(8):1131-4.
68. Angellotti E, Pittas AG. Chapter 77 - The Role of Vitamin D in Type 2 Diabetes and Hypertension. In: Feldman D, editor. *Vitamin D (Fourth Edition)*: Academic Press; 2018. p. 387-423.
69. Forouhi NG, Ye Z, Rickard AP, Khaw KT, Luben R, Langenberg C, et al. Circulating 25-hydroxyvitamin D concentration and the risk of type 2 diabetes: results from the European Prospective Investigation into Cancer (EPIC)-Norfolk cohort and updated meta-analysis of prospective studies. *Diabetologia*. 2012;55(8):2173-82.
70. Husemoen LL, Skaaby T, Thuesen BH, Jørgensen T, Fenger RV, Linneberg A. Serum 25(OH)D and incident type 2 diabetes: a cohort study. *Eur J Clin Nutr*. 2012;66(12):1309-14.
71. Husemoen LL, Thuesen BH, Fenger M, Jørgensen T, Glümer C, Svensson J, et al. Serum 25(OH)D and type 2 diabetes association in a general population: a prospective study. *Diabetes care*. 2012;35(8):1695-700.
72. Grimnes G, Emaus N, Joakimsen RM, Figenschau Y, Jenssen T, Njølstad I, et al. Baseline serum 25-hydroxyvitamin D concentrations in the Tromsø Study 1994-95 and risk of developing type 2 diabetes mellitus during 11 years of follow-up. *Diabet Med*. 2010;27(10):1107-15.
73. Holmen J, Midthjell K, Krüger Ø, Langhammer A, Holmen TL, Bratberg GH, et al. The Nord-Trøndelag Health Study 1995-97 (HUNT 2) ; objectives, contents, methods and participation. *Norsk epidemiologi*. 2003;13(1):19-32.
74. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, et al. Cohort Profile: the HUNT Study, Norway. *International journal of epidemiology*. 2013;42(4):968-77.
75. Hollis BW. Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for establishing a new effective dietary intake recommendation for vitamin D. *The Journal of nutrition*. 2005;135(2):317-22.
76. Institute of Medicine Committee to Review Dietary Reference Intakes for Vitamin D, Calcium. *The National Academies Collection: Reports funded by National Institutes of*

Health. In: Ross AC, Taylor CL, Yaktine AL, Del Valle HB, editors. Dietary Reference Intakes for Calcium and Vitamin D. Washington (DC): National Academies Press (US)

National Academy of Sciences.; 2011.

77. Degerud E, Hoff R, Nygard O, Strand E, Nilsen DW, Nordrehaug JE, et al. Cosinor modelling of seasonal variation in 25-hydroxyvitamin D concentrations in cardiovascular patients in Norway. *Eur J Clin Nutr.* 2016;70(4):517-22.
78. Sachs MC, Shoben A, Levin GP, Robinson-Cohen C, Hoofnagle AN, Swords-Jenny N, et al. Estimating mean annual 25-hydroxyvitamin D concentrations from single measurements: the Multi-Ethnic Study of Atherosclerosis. *Am J Clin Nutr.* 2013;97(6):1243-51.
79. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10(1):37-48.
80. Carlsson S, Midthjell K, Grill V. Influence of family history of diabetes on incidence and prevalence of latent autoimmune diabetes of the adult: results from the Nord-Trondelag Health Study. *Diabetes care.* 2007;30(12):3040-5.
81. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet.* 2017;49(12):1752-7.
82. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48(10):1284-7.
83. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics.* 2016;48(10):1279-83.
84. Zhou W, Fritsche LG, Das S, Zhang H, Nielsen JB, Holmen OL, et al. Improving power of association tests using multiple sets of imputed genotypes from distributed reference panels. *Genet Epidemiol.* 2017;41(8):744-55.
85. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *International journal of epidemiology.* 2013;42(4):1134-44.
86. Midthjell K, Holmen J, Bjorndal A, Lund-Larsen G. Is questionnaire information valid in the study of a chronic disease such as diabetes? The Nord-Trondelag diabetes study. *Journal of epidemiology and community health.* 1992;46(5):537-42.
87. Sorgjerd EP, Skorpen F, Kvaloy K, Midthjell K, Grill V. Time dynamics of autoantibodies are coupled to phenotypes and add to the heterogeneity of autoimmune diabetes in adults: the HUNT study, Norway. *Diabetologia.* 2012;55(5):1310-8.
88. Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. *European journal of epidemiology.* 2005;20(7):575-9.
89. Rubin DB. *Multiple imputation for nonresponse in surveys*: John Wiley & Sons; 2004.
90. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine.* 2011;30(4):377-99.
91. Ye Z, Sharp SJ, Burgess S, Scott RA, Imamura F, Langenberg C, et al. Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a mendelian randomisation study. *The lancet Diabetes & endocrinology.* 2015;3(1):35-42.
92. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj.* 2018;362:k601.
93. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014;23(R1):R89-98.
94. Lu L, Bennett DA, Millwood IY, Parish S, McCarthy MI, Mahajan A, et al. Association of vitamin D with risk of type 2 diabetes: A Mendelian randomisation study in European and Chinese adults. *PLoS Med.* 2018;15(5):e1002566.

95. Yuan S, Jiang X, Michaëlsson K, Larsson SC. Genetic Prediction of Serum 25-Hydroxyvitamin D, Calcium, and Parathyroid Hormone Levels in Relation to Development of Type 2 Diabetes: A Mendelian Randomization Study. *Diabetes care*. 2019;42(12):2197-203.
96. Heath AK, Williamson EJ, Hodge AM, Ebeling PR, Eyles DW, Kvaskoff D, et al. Vitamin D status and the risk of type 2 diabetes: The Melbourne Collaborative Cohort Study. *Diabetes Res Clin Pract*. 2019;149:179-87.
97. Knekt P, Laaksonen M, Mattila C, Härkänen T, Marniemi J, Heliövaara M, et al. Serum vitamin D and subsequent occurrence of type 2 diabetes. *Epidemiology*. 2008;19(5):666-71.
98. Palomer X, González-Clemente JM, Blanco-Vaca F, Mauricio D. Role of vitamin D in the pathogenesis of type 2 diabetes mellitus. *Diabetes Obes Metab*. 2008;10(3):185-97.
99. Mathieu C. Vitamin D and diabetes: Where do we stand? *Diabetes Res Clin Pract*. 2015;108(2):201-9.
100. Langhammer A, Krokstad S, Romundstad P, Heggland J, Holmen J. The HUNT study: participation is associated with survival and depends on socioeconomic status, diseases and symptoms. *BMC Med Res Methodol*. 2012;12:143-.
101. Jorde R, Sneve M, Hutchinson M, Emaus N, Figenschau Y, Grimnes G. Tracking of serum 25-hydroxyvitamin D levels during 14 years in a population-based study and during 12 months in an intervention study. *Am J Epidemiol*. 2010;171(8):903-8.
102. Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat Genet*. 2011;43(10):984-9.

## APPENDIX: Supplementary tables

**Table S1.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up

*Analyses after performing multiple imputation of missing data*

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%) | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|----------|--------------------|-----------------------|
| <i>Categorical</i>                                 |                     |              |          |                    |                       |
| <50.0  | 2066                | 74           | 3.6%     | 1.85 (1.20 - 2.83) | 1.49 (0.94 - 2.38)    |
| ≥50.0  | 1520                | 30           | 2.0%     | 1.00 (reference)   | 1.00 (reference)      |
| <i>Continuous<sup>1</sup></i>                      |                     |              |          |                    |                       |
|  | 3586                | 104          | 2.9%     | 1.52 (1.09 - 2.12) | 1.28 (0.88 - 1.85)    |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of ≥11 mmol/L and/or reported diabetes.

<sup>1</sup>per 25 nmol/L decrease in serum 25(OH)D

\* Adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

**Table S2.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up, stratified by Sex

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%) | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|----------|--------------------|-----------------------|
| <b>Sex</b>   |                     |              |          |                    |                       |
| <i>Female</i>                                      |                     |              |          |                    |                       |
| <50.0  | 1125                | 38           | 3.4%     | 2.40 (1.25 - 4.62) | 1.95 (0.94 - 4.05)    |
| ≥50.0  | 836                 | 12           | 1.4%     | 1.00 (reference)   | 1.00 (reference)      |
| <i>Male</i>  |                     |              |          |                    |                       |
| <50.0  | 941                 | 36           | 3.8%     | 1.47 (0.83 - 2.61) | 1.47 (0.77 - 2.81)    |
| ≥50.0  | 684                 | 18           | 2.6%     | 1.00 (reference)   | 1.00 (reference)      |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of ≥11 mmol/L and/or reported diabetes.

\* Adjusted for age, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

**Table S3.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up, stratified by BMI

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%)    | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|-------------|--------------------|-----------------------|
| <b>BMI</b>   |                     |              |             |                    |                       |
| <b>&lt;30 kg/m<sup>2</sup></b>                     | <b>3105</b>         | <b>58</b>    | <b>1.9%</b> |                    |                       |
| <50.0  | 1713                | 37           | 2.2%        | 1.44 (0.84 - 2.47) | 1.34 (0.75 - 2.39)    |
| ≥50.0  | 1392                | 21           | 1.5%        | 1.00 (reference)   | 1.00 (reference)      |
| <b>≥30 kg/m<sup>2</sup></b>                        | <b>472</b>          | <b>46</b>    | <b>9.7%</b> |                    |                       |
| <50.0  | 349                 | 37           | 10.6%       | 1.50 (0.70 - 3.21) | 1.93 (0.82 - 4.55)    |
| ≥50.0  | 123                 | 9            | 7.3%        | 1.00 (reference)   | 1.00 (reference)      |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; BMI: body mass index; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of ≥11 mmol/L and/or reported diabetes.

\* Adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

**Table S4.** Association between baseline seasonal-standardized serum 25(OH)D level and incidence of T2DM over an 11-year follow up, stratified by PRS

*Analyses after performing multiple imputation of missing data*

| Seasonal-standardized serum 25(OH)D level (nmol/L) | No. of participants | No. of cases | Risk (%)    | Crude OR (95% CI)  | Adjusted OR (95% CI)* |
|--|---------------------|--------------|-------------|--------------------|-----------------------|
| <b>Polygenic risk scores</b>                       |                     |              |             |                    |                       |
| <b>Low score (2/3)</b>                             | <b>2330</b>         | <b>57</b>    | <b>2.4%</b> |                    |                       |
| <50.0  | 1311                | 35           | 2.7%        | 1.24 (0.72 - 2.13) | 0.93 (0.51 - 1.72)    |
| ≥50.0  | 1019                | 22           | 2.2%        | 1.00 (reference)   | 1.00 (reference)      |
| <b>High score (1/3 top)</b>                        | <b>1164</b>         | <b>40</b>    | <b>3.4%</b> |                    |                       |
| <50.0  | 694                 | 33           | 4.8%        | 3.30 (1.45 - 7.53) | 2.76 (1.15 - 6.65)    |
| ≥50.0  | 470                 | 7            | 1.5%        | 1.00 (reference)   | 1.00 (reference)      |

25(OH)D: 25-hydroxyvitamin D; T2DM: type 2 diabetes mellitus; PRS: polygenic risk score; OR: odds ratio; CI: confidence interval

T2DM was defined as non-fasting serum glucose levels of ≥11 mmol/L and/or reported diabetes.

\* Adjusted for age, sex, BMI, education, smoking status, alcohol consumption, physical activity, sedentary lifestyle, economic difficulties, family history of diabetes, chronic disease

