Vitamin D and degenerative health: a large-scale study exploring deficiency, muscle traits, mortality, and fall risk using data sourced from the UK Biobank

Joshua Paul Sutherland

Bachelor of Health Science—Nutritional Medicine (Honours)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Nutritional and Genetic Epidemiology Research Group

Clinical and Health Sciences

University of South Australia



June 2024

© Joshua Paul Sutherland, 2024

Series of Dissertations Submitted to the School of Clinical Health Sciences, University of South Australia

All rights reserved. No part of this publication may be reproduced to transmitted, in any form by any means, without permission.

Navigating this thesis and contextual statement:

This thesis contains four individual studies. The first three studies are published, and they are included in their manuscript form, as constructed for journal publication. The fourth and final study is included as an unpublished manuscript.

- Each study has its own chapter containing the study manuscript, accompanying supplementary materials, and, for studies 3 and 4, an addendum.
- Referencing for each study (including intext citations and reference lists) are self-contained to within that study's chapter.
- The distinct thesis chapters (introduction, literature review, methods, and discussion) collectively constitute their own separate referenced body of work, with their reference list at the end of the thesis document.
- The thesis list of tables/figures denote those elements appearing in the literature review and method chapters.
- Each study contains its own manuscript table/figure list at the beginning of its respective chapter.
- The original manuscript text for studies 2-4 make specific references to supplementary document page numbers. Therefore, these supplementary sections have retained their original page numbering system (in addition to the continuous thesis numbering), thus allowing navigation from the manuscript text to the supplementary material to remain intact.

Each study contains a section contextualising author contribution. For any given study, Joshua Sutherland is the majority contributor.

Table of Contents

TABLE OF CONTENTS IV

LIST OF FIGURES VII

LIST OF TABLES VII

LIST OF SUPPLEMENTARY TABLES VII

GLOSSARY VIII

ABBREVIATIONS X

DECLARATION XIII

VITAMIN D AND DEGENERATIVE HEALTH: A LARGE-SCALE STUDY EXPLORING DEFICIENCY, MUSCLE TRAITS, MORTALITY, AND FALL RISK USING DATA SOURCED FROM THE UK BIOBANK. XIV

SUMMARY XIV
ACKNOWLEDGMENTS XVII
DEDICATION XVIII
LIST OF PUBLICATIONS GENERATED FROM THIS THESIS XIX
RELATED WORKS DURING THE CANDIDATURE XX
CONFERENCE PRESENTATIONS XX

CHAPTER ONE: INTRODUCTION 1

CHAPTER TWO: LITERATURE REVIEW 4

2.1 VITAMIN D: INTRODUCTION 4
2.2 VITAMIN D: ACQUISITION AND SKIN SYNTHESIS 5
2.3 VITAMIN D: HYDROXYLATION 7
2.4 VITAMIN D: TRANSPORTATION, CELLULAR INTEGRATION, GENETIC INTERACTION, AND CATABOLISM 9
2.5 VITAMIN D: A FOCUS ON METABOLIC INTERPLAY 12
2.6 VITAMIN D: SKELETAL EFFECTS 14
2.7 VITAMIN D: EXTRA-SKELETAL EFFECTS 14

2.8 VITAMIN D: ENVIRONMENTAL AND PHYSIOLOGICAL FACTORS, AND NUTRITIONAL STATES IN HUMANS 17
2.9 VITAMIN D: EXCESS AND ONGOING TOXICITY RISK—CAUTION REQUIRED. 19
2.10 VITAMIN D: MOTOR FUNCTION 20
2.11 VITAMIN D: MORTALITY 21
2.12 DEGENERATIVE HEALTH CONDITIONS 24
2.13 VITAMIN D: RESEARCH—2000S AND BEYOND 25
2.14 SUMMARY 26

CHAPTER THREE: THESIS AIMS AND OBJECTIVES 27

3.1 THESIS PRIMARY AIM 27
3.2 THESIS OBJECTIVES 27

CHAPTER FOUR: GENERAL METHODS 28

4.1 STUDY DATA—PARTICIPANTS 28
4.2 STUDY DATA—PHENOTYPIC DATA 28
4.3 STUDY DATA—PHENOTYPIC DATA FOR STUDY I EXPOSURE VARIABLES AND ETHNICITY/SEASON STRATIFICATION 30
4.4 STUDY DATA—PHENOTYPIC DATA FOR STUDY II OUTCOME VARIABLES (MUSCLE MASS/GRIP STRENGTH/PROBABLE SARCOPENIA) 31
4.5 STUDY DATA—PHENOTYPIC DATA FOR STUDY III OUTCOME VARIABLE (MORTALITY) 32
4.6 STUDY DATA—PHENOTYPIC DATA FOR STUDY IV OUTCOME VARIABLE (FALLS) 32
4.7 STUDY DATA—25(OH)D MEASUREMENT 33
4.8 STUDY DATA—OVERVIEW OF GENETIC DATA 33
4.9 STUDY DATA—QUALITY CONTROL OF GENETIC DATA 34
4.10 STUDY DATA—GENETIC INSTRUMENT, USED IN STUDIES II, III AND IV (FOR MENDELIAN RANDOMIZATION ANALYSES) 35
4.11 STUDY DATA—RETINOIC ACID RELATED GENETIC DATA FOR STUDY IV. 36
4.12 STUDY DESIGNS - OVERVIEW 38
4.13 STUDY DESIGNS—EPIDEMIOLOGY 38
4.14 STUDY DESIGNS—GENETIC EPIDEMIOLOGY 38
4.15 STUDY DESIGNS—MENDELIAN RANDOMISATION 39
4.16 STUDY DESIGNS—GENE-ENVIRONMENT INTERACTION. 40
4.17 STATISTICAL ANALYSES - LINEAR AND LOGISTIC REGRESSION, STUDIES I–IV (INCLUDING USE OF INTERACTION ANALYSES) 40
4.18 STATISTICAL ANALYSES—LINEAR AND LOGISTICAL REGRESSION USING THE FRACTIONAL POLYNOMIAL TECHNIQUE FOR PRESENTING PHENOTYPIC FINDINGS AS NONLINEAR (INCLUDING THE USE OF QUADRATIC TERMS IN REGRESSION, TO TEST FOR NONLINEARITY) 41
4.19 STATISTICAL ANALYSES—MENDELIAN RANDOMIZATION OVERVIEW, STUDIES II–IV 42

4.20 Statistical analyses—Linear Mendelian randomization using individual level data 42
4.21 Statistical analyses—linear Mendelian randomization using summary level data 43
4.22 Statistical analyses—non-linear Mendelian randomization, residual and doubly ranked method 45

CHAPTER FIVE: STUDY ONE 47
Study 1 Manuscript 48
Study 1 Supplementary 76

CHAPTER SIX: STUDY TWO 78
Study 2 Manuscript 79
Study 2 Supplementary 110

CHAPTER SEVEN: STUDY THREE 135
Study 3 Manuscript 136
Study 4 Supplementary 162
Study 3 Addendum 192

CHAPTER EIGHT: STUDY FOUR 206
Study 4 Manuscript 207
Study 4 Supplementary 238
Study 4 Addendum 261

CHAPTER NINE: DISCUSSION 279
9.1 The interconnectedness of vitamin D in degenerative health pathology. 282
9.2 The complexity of vitamin D in degenerative healthcare settings 283
9.3 Considerations for clinical practice and public health 284
9.4 Concerns and health equity 286
9.5 Considerations for future research 287
9.6 Strengths and limitations 290

CHAPTER TEN: CONCLUSIONS 293

REFERENCES: 295

SUPPLEMENTARY MATERIALS 321

APPENDICES – STATEMENTS OF AUTHORSHIP 325

List of Figures

Figure 1: Vitamin D isomer categorization 4
Figure 2: Vitamin D acquisition and skin synthesis, including skin synthesis regulation..... 5
Figure 3: Vitamin D hydroxylation. 8
Figure 4: Vitamin D transportation and cellular integration (A), genetic interaction (B and C), and catabolism (D). 10
Figure 5: Steroid hormone competition visually depicted via gene expression of a retinoic acid inducible gene in response to (A) an activating heterodimer species, versus (B) a repressive heterodimer species. 13
Figure 6: Latitudinal zoning of the UK Biobank Assessment Centres. 30
Figure 7: Mendelian Randomization. 39

List of Tables

Table 1: UK Biobank data and their source/collection point. 29
Table 2: Quality control measures for genetic data in the UK Biobank. 34
Table 3: 25(OH)D related SNPs used to construct the 35 SNP genetic instrument. 36
Table 4: Motor function related SNPs from retinoic responsive genes. 37

List of supplementary Tables

Supplementary Table 1: Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations. 321

Glossary

Allele: Alternative forms of a gene or DNA sequence.

Bolus dosing: Administration of a single large dose of a substance.

Bonferroni testing threshold: Statistical correction for multiple testing in research studies.

Calcitriol: The active form of vitamin D.

C-reactive protein: A marker of inflammation used to assess risk of cardiovascular disease.

Functional retinol deficiency: Insufficient active vitamin A levels in tissues despite adequate dietary intake.

Gene: Basic unit of heredity containing information for specific functions.

Genome: Sum total of genetic material of a cell or organism.

Genotype: Specific set of two alleles inherited at a genetic locus.

GWAS (Genome-wide association study): Study analyzing genetic variations across genomes.

GWAS catalogue: Repository summarizing genetic variant-trait associations.

Lean skeletal muscle mass: Muscle tissue free from fat and other non-muscular tissues.

Linkage: Two loci close enough that they rarely recombine independently.

Linkage disequilibrium: the non-random tendency of genetic variations located near each other on a chromosome to be inherited together more frequently than expected by chance.

Locus (plural loci): Physical site on a chromosome occupied by a gene.

Mendelian randomization: A method that uses genetic variants as unconfounded proxy measures to investigate causal relationships between and an outcome.

Minor allele: Less common allele in a population for a biallelic polymorphism.

Minor allele frequency (MAF): Frequency of the less common allele in a specific population.

Motor function: Ability to control and coordinate muscle movements for voluntary actions.

Non-linear Mendelian randomization (non-linear MR): Analysis allowing for non-linear relationships between genetic variants and exposure/outcome.

Phenotype: Observable characteristics resulting from genotype-environment interaction.

Phenotypic analyses: In the context of the thesis, these refer to analyses that examine observable characteristics (phenotypes) without directly incorporating genetic variables. They explore how environmental factors and other non-genetic influences contribute to the expression of traits or diseases in the observed population, distinct from analyses that primarily focus on genetic factors.

Pleiotropic effect: When a single gene influences multiple traits or diseases simultaneously.

Retinoic acid: Metabolite of vitamin A that influences gene transcription and cell differentiation.

Retinoic Acid Receptor: Nuclear receptor that binds retinoic acid, regulating gene transcription.

Retinoic Acid Response Element: DNA sequence where RAR binds in response to retinoic acid.

Retinoid X receptor: Nuclear receptor that forms heterodimers with other receptors.

Sarcopenia: Age-related loss of skeletal muscle mass and function.

Single nucleotide polymorphisms (SNPs): Variations in DNA sequence at a single nucleotide level.

Steroid hormone competition: Interactions where steroid hormones compete for binding sites.

Striatum: Brain region involved in motor and reward functions.

Summary statistics: Outcome measures from GWAS, including SNP data.

Threshold effects: Points at which a substance has an impact on health outcomes above a certain level.

Vitamin D receptor: Nuclear receptor protein that binds calcitriol, regulating biological processes.

Vitamin D response elements: DNA sequences where the VDR binds, influencing gene expression

Abbreviations

24(R),25(OH)2D: 24(R),25-Dihydroxyvitamin D

24,25(OH)2D: 24,25-Dihydroxyvitamin D

25(OH)D: 25-Hydroxyvitamin D

25(OH)D2: 25-Hydroxyvitamin D2

25(OH)D3: 25-Hydroxyvitamin D3

ADHD: Attention-Deficit/Hyperactivity Disorder

A-levels: Advanced Levels

AMPK: AMP-activated Protein Kinase

BMI: Body Mass Index

BRINP: BMP/retinoic acid-inducible neural-specific protein

CI: Confidence Interval

CLSI: Clinical Laboratory and Standards Institute

COVID-19: Coronavirus Disease 2019

CPY24A1: Cytochrome P450 Family 24 Subfamily A Member 1 gene

CRP: C-Reactive Protein

CSE: Certificate of Secondary Education

CV: Coefficient of Variation

CVD: Cardiovascular Disease

CYP2R1: Cytochrome P450 Family 2 Subfamily R Member 1

D2: Ergocalciferol

DHC: 7-Dehydrocholesterol

DHCR7: 7-Dehydrocholesterol Reductase gene

FGF23: Fibroblast Growth Factor-23

G x E: Gene-Environment Interaction

g: Gram

GABA: Gamma-Aminobutyric Acid

GBP: Great British Pound

GRS: Genetic Risk Score

GWAS: Genome-Wide Association Study

IL-6: Interleukin 6

IU: International Units

IVW: Inverse-Variance Weighted method

kg: Kilograms

LACE: Localized Average Causal Effect

LD: Linkage Disequilibrium

MAD: Major Depressive Disorder

MAF: Minor Allele Frequency

mg/L: Milligrams per Liter

mmol/L: Millimoles per Liter

MR: Mendelian Randomization

MR-Egger: Mendelian Randomization-Egger

mRNA: Messenger Ribonucleic Acid

MR-PRESSO: Mendelian Randomization Pleiotropy RESidual Sum and Outlier

mTOR: Mechanistic Target of Rapamycin

NAM: National Academy of Medicine

NHS: National Health Service

NHSCR: NHS Central Register

NRG (Gene): Neuregulin

NVQ: National Vocational Qualification

OR: Odds Ratio

PC: Personal Computer

PCA: Principal Component Analysis

PhenoScanner: Database of Genetic Association Studies

PTH: Parathyroid Hormone

Q: Quartiles

RA: Retinoic Acid or Rheumatoid Arthritis (context-dependent)

RAR: Retinoic Acid Receptor

RARE: Retinoic Acid Response Element

RCT: Randomized Controlled Trial

RSS: Residual Sum of Squares

RXR: Retinoid X Receptor

SD: Standard Deviation

SNP: Single Nucleotide Polymorphism

TH17: T Helper 17 Cells

UK Biobank: United Kingdom Biobank

UK: United Kingdom

USA: United States of America

UVB: Ultraviolet B

VDBP: Vitamin D Binding Protein

VDR: Vitamin D Receptor

Declaration

I declare that:

this thesis presents work carried out by myself and does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; to the best of my knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text; and all substantive contributions by others to the work presented, including jointly authored publications, is clearly acknowledged.

Signed

Joshua Paul Sutherland



June 2024

Vitamin D and degenerative health: a large-scale study exploring deficiency, muscle traits, mortality, and fall risk using data sourced from the UK Biobank.

Summary

Vitamin D's rich public health narrative spans centuries, now evolving with genetic insights aiding research beyond the scope of conventional clinical trials. Interest in vitamin D's role in degenerative health and aging stems from numerous observational studies linking deficiency [typically <25 nmol/L 25-hydroxyvitamin D (25(OH)D)] to conditions like accelerated muscle loss, respiratory issues, and increased fall risk. However, recent clinical trials—especially those using monthly/annual bolus dosing—have raised safety concerns, notably an elevated fall risk.

This thesis explores the contemporary dichotomy surrounding vitamin D, exploring its role in degenerative health while placing focus on warranted safety considerations. Pragmatically, focus is placed on motor function, given that motor function is a critical aspect where vitamin D is seen as supportive, while also being the research space to reveal unexpected risks from supplementation. It combines observational and genetic approaches, leveraging the extensive UK Biobank cohort data [approximately 500,000 individuals]. The specific aims are to:

- Explore 25(OH)D deficiency determinants in the UK Biobank, contextualizing 25(OH)D data for this project [Study I, Chapter 5].
- Test observational and genetic evidence for associations between 25(OH)D and muscle traits [Study II, Chapter 6].
- Test observational and genetic evidence for associations between 25(OH)D and sarcopenia [Study II, Chapter 6].
- Test observational and genetic evidence for associations between 25(OH)D and mortality [Study III, Chapter 7].
- Explore for the possibility of whether increased fall risk from higher vitamin D varies according to inflammatory status, and if the effect is dependent on a competitive steroid hormone relationship between retinol and vitamin D [Study IV, Chapter 8].

Study I: Logistic regression was used to explore vitamin D deficiency determinants across 440,581 participants. Deficiency prevalence was highest among Asians [57.2% in winter/spring,

50.8% in summer/autumn]. Socioeconomic deprivation increased deficiency odds [P = <0.0001]. Regular oily fish consumption reduced deficiency odds universally, but outdoor time in summer was less effective for Black Africans [OR = 0.89, 95% CI, 0.70-1.12] compared to White Europeans [OR = 0.40, 95% CI, 0.38-0.42]. High rates of vitamin D deficiency in the UK disproportionately affect specific ethnic and socioeconomic groups.

Study II: Mendelian randomization (MR) analysis tested the relationship between 25(OH)D concentrations and muscle traits, including grip strength and muscle mass, extending to sarcopenia and sarcopenic obesity. Genetic analyses supported a relationship between higher 25(OH)D and muscle traits, with MR analyses showing 0.11 kg greater grip strength per 10 unit higher 25(OH)D, and modest association with muscle mass (0.01 kg greater mass). There was evidence for lower odds of sarcopenia in non-obese sarcopenia cases [OR = 0.92, 95% CI, 0.86-0.98], but not in the obese.

Study III: To assess the causal role of low vitamin D in mortality, nonlinear and linear MR analyses were conducted. During 14 years of follow-up, 18,700 deaths occurred. Observational analyses showed a relationship between 25(OH)D deficiency and increase mortality risk. Genetically predicted 25(OH)D was L-shaped with all-cause mortality [P for nonlinearity < 0.001], decreasing steeply until 50 nmol/L. Associations were seen with cancer, cardiovascular, and for respiratory diseases using the linear MR method [$P \leq 0.033$ for all outcomes]. In addendum analysis, nonlinear associations could not be replicated using the rank nonlinear method, likely due to insufficient power.

Study IV: Study IV: We investigated the association between inflammation, vitamin D, and fall risk in 307,082 UK Biobank participants using a genetic epidemiology approach. The association between 25(OH)D and fall risk varied by CRP concentrations ($P_{Interaction}$ 0.009). In participants with little to no inflammation (CRP <5 mg/L), we observed an inverse association (OR = 0.9987, 95% CI, 0.998-0.999). However, among participants with CRP > 5 mg/L, the association was U-shaped, with elevated fall risk at both < 25 nmol/L and > 100 nmol/L 25(OH)D (P <0.004). The association with falls for both low and high 25(OH)D levels was particularly notable at higher CRP thresholds (CRP > 20 mg/L, ≥ 100 nmol/L vs. 50–74.99 nmol/L: OR = 2.40, 95% CI 1.46–3.77). Genetically predicted higher 25(OH)D was not linked to fall risk in the general population but showed an association with increased fall risk in participants with CRP > 20 mg/L (926

cases; OR = 1.20, 95% CI 1.00–1.44). Findings suggest that both low and high 25(OH)D levels increase fall risk in individuals with chronic inflammation.

In conclusion, this thesis presents observational and genetic evidence supporting a relationship between 25(OH)D and degenerative health factors, including grip strength, muscle composition, sarcopenia, and mortality. It highlights alarming deficiency rates in the UK and explores concerns about high-dose vitamin D supplementation and fall risk, with indirect evidence supporting a steroid hormone competition hypothesis. Future studies should investigate how steroid hormone competition affects genetic expression, reinforcing the importance of striving for universal vitamin D sufficiency for overall health and healthy aging.

Acknowledgments

Foremost, I acknowledge the years of unwavering support I have received from my primary supervisor, Professor Elina Hyppönen. You have both supported my learning and encouraged creativity and autonomy, and for that, I am eternally grateful. I will carry that encouragement and inspiration throughout my career. I also thank my co-supervisor, Dr. Ang Zhou, whose kindness is only exceeded by his patience, making all the difference in my Ph.D. journey. Furthermore, I thank my honours degree primary supervisor and current Ph.D. co-supervisor, Dr. Matthew Leach. Your kindness and mentorship were instrumental in my decision to pursue a Ph.D. Special thanks must be given to my friend, colleague, and undergraduate nutrition lecturer, Natasha Radcliffe, who has been an ever-present advocate for my professional and academic success, and who made the initial introduction between myself and the Australian Centre for Precision Health.

I importantly acknowledge and thank my nearest and dearest friends and family, who have supported me throughout the best and toughest of times. I also thank my academic and clinician colleagues, who each, in your own ways, provided support and sparks of inspiration that helped me along the way. To my Ph.D. colleagues, Muktar Ahmed and Kitty Pham, I thank you for us *always being in it together*. To Anwar Mulugeta Gebremichael and Amanda Lumsden, your supporting expertise and friendly chats made all the difference. I also thank Judith Burgemeister for her proofreading of this thesis. I thank the University of South Australia and the SAHMRI community for supporting my Ph.D.

Each study chapter contains an individual acknowledgement section specific to that piece of work. This research has been supported by the Commonwealth Government of Australia through the Research Training Program (RTP) International Scholarship. The complete funding support and author contributions for each manuscript are included in their respective chapters.

Dedication

To my grandfather,

I dedicate this thesis to the memory of my grandfather, Glen Malcom, who—as an ambulance officer—imparted to me my first interests in the health and medical sciences. He faced many hardships from an early age in rural Australia and he sacrificed greatly for his family. He would have likely ventured further into the medical sciences, should life have afforded him that opportunity. I am grateful to have completed a journey that I know in my heart is an extension of his passion and influence on my life, and one he would be very proud of.

List of publications generated from this thesis

Differences and determinants of vitamin D deficiency among UK biobank participants: A cross-ethnic and socioeconomic study

Sutherland JP, Zhou A, Leach MJ, Hyppönen E. Differences and determinants of vitamin D deficiency among UK biobank participants: A cross-ethnic and socioeconomic study. Clin Nutr. 2021 May;40(5):3436-3447. doi:10.1016/j.clnu.2020.11.019. Epub 2020 Nov 25. PMID: 33309415.

https://pubmed.ncbi.nlm.nih.gov/33309415/

Muscle Traits, Sarcopenia, and Sarcopenic Obesity: A Vitamin D Mendelian Randomization Study

Sutherland JP, Zhou A, Hyppönen E. Muscle Traits, Sarcopenia, and Sarcopenic Obesity: A Vitamin D Mendelian Randomization Study. Nutrients. 2023 Jun 9;15(12):2703. doi:10.3390/nu15122703. PMID: 37375607; PMCID: PMC10301137.

https://pubmed.ncbi.nlm.nih.gov/37375607/

Vitamin D Deficiency Increases Mortality Risk in the UK Biobank: A Nonlinear Mendelian Randomization Study

Sutherland JP, Zhou A, Hyppönen E. Vitamin D Deficiency Increases Mortality Risk in the UK Biobank: A Nonlinear Mendelian Randomization Study. Ann Intern Med. 2022 Nov;175(11):1552-1559. doi:10.7326/M21-3324. Epub 2022 Oct 25. PMID: 36279545.

https://pubmed.ncbi.nlm.nih.gov/36279545/

Related works during the candidature

Vitamin D and Type 1 Diabetes Risk: A Systematic Review and Meta-Analysis of Genetic Evidence

Najjar L, Sutherland J, Zhou A, Hyppönen E. Vitamin D and Type 1 Diabetes Risk: A Systematic Review and Meta-Analysis of Genetic Evidence. Nutrients. 2021 Nov 26;13(12):4260. doi:10.3390/nu13124260. PMID: 34959812; PMCID: PMC8707565.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8707565/

Conference presentations

Oral presentations

- **Title: Vitamin D: A large scale study exploring motor function risk.** University annual research week in 2019, School of Health Sciences, University of South Australia (UniSA), Adelaide
- **Title: Vitamin D: A large scale study exploring motor function risk.** University annual research week in 2020, Clinical & Health Sciences, UniSA, Adelaide
- **Title: Ethnic differences and determinants of vitamin D deficiency in the UK Biobank.** World Congress of Epidemiology 2021, Online.

Poster presentation

- **Title: Differences and Determinants of Vitamin D Deficiency among UK Biobank Participants: A Cross-Ethnic and Socio-Economic Study.** 44th Annual Scientific Meeting of the Nutrition Society of Australia. Online.

Chapter One: Introduction

Vitamin D constitutes a rich public health story spanning centuries, with current times using genetic insights to help deepen understanding, and to facilitate types of research that are challenging or unethical to conduct via conventional clinical trial methods (Cummings, Kiel & Black 2016; Nair & Maseeh 2012). Using genetic epidemiological approaches, this thesis focuses on vitamin D and how it intersects with degenerative health, namely motor function related matters [e.g., skeletal muscle health, sarcopenia, and fall risk] and then extends this consideration to mortality itself. Any exploration of vitamin D inevitably intersects with the consideration and discussion of vitamin D supplementation, given that supplementation—historically speaking—is a crucial part of the overall vitamin D story (Duguid 1930; Tobey 1936), and with supplementation generating ample and varied discourse in the public health and broader scientific communities (Cummings, Kiel & Black 2016; Kilbane et al. 2014; Stokes & Lammert 2016; Taylor & Davies 2018).

While not directly focusing on supplementation related analysis, the overall focus is constructed to help inform contemporary interests and concerns around vitamin D supplementation, namely concerns that have emerged regarding high dosage supplementation methods, which have—in part—arisen from the increased fall risk seen in some primary and subgroup analyses from clinical trials that used annual, monthly or high daily doses [$\geq$ 3200 IU] of the vitamin (Bischoff-Ferrari et al. 2016; Ginde et al. 2017; Smith, Gallagher & Suiter 2017; Sanders et al. 2010; Waterhouse et al. 2021; Zittermann et al. 2023). Some notable trials, such as the ViDA trial, did not observe an increased fall risk from monthly bolus dosing (Khaw, et al. 2017), and meta-analyses on vitamin D supplementation have shown conflicting results (Zittermann et al. 2023; Bolland, Grey & Avenell 2018a). These inconsistencies, however, may partly stem from variations in the vitamin D repletion status of trial cohorts (Sanders et al. 2010; Latham et al. 2003), raising the important question of whether other, as yet undetected physiological conditions also contribute to these divergent outcomes. In seeking to reflect this contemporary dichotomy of concern and confidence, a focus on risk is balanced in this thesis with a focus on optimism for vitamin D in the degenerative health and healthy aging space.

Vitamin D has long been used in the prevention of rickets, with vitamin D rich cod-liver oil used even prior to the vitamin's official discovery (Rajakumar 2003). At the turn of the twentieth

century, the connection between ultraviolet B [UVB] radiation exposure and vitamin D synthesis in humans was an emerging field (Riemerschmid 1921), which also included breakthroughs in using artificial UVB technology to both stimulate vitamin D synthesis in humans (Hess & Unger 1921) and in generating supplemental forms of vitamin D that were no longer reliant on cod-liver oil (Hess, Weinstock & Helman 1925). The consideration of vitamin D deficiency has dominated discourse in research, in no small part due to the public health crisis of rickets (Rajakumar 2003), and between 1920 and early 1930s, bona fide vitamin D deficiency studies were first undertaken (Grant 1930; Harris & Innes 1931). This soon led to the widespread utilisation of vitamin D fortified milks (Tobey 1936) and the continued efforts, seen today, to ensure appropriate supplementation (Tobey 1936; Weiler 2017).

Moving forward a century, supplemental vitamin D is now commonly prescribed by healthcare professionals, and self-administered by the health-conscious public, to treat vitamin D deficiency, manage varying health conditions [e.g. osteoporosis, depression, respiratory immune support (Avenell, Mak & O'Connell 2014; Jorde et al. 2008; Martineau et al. 2017)], to optimise performance [such as muscle strength (Halfon, Phan & Teta 2015)] and to support critical life-stages [such as infancy and old age (Jorde et al. 2008; Kweder & Eidi 2018; Martineau et al. 2011)]. Public health proponents around the world also specify guidelines for supplementation (Del Valle et al. 2010; Holick et al. 2011), given that some populations are at risk for deficiency, based on either latitudinal location [i.e., those living closer to the planetary magnetic poles] (Leary et al. 2017) or other UVB radiation exposure related factors [i.e., night shift workers, institutionalised individuals etc.] (Daugaard et al. 2018; Sambrook et al. 2002).

The enthusiasm for vitamin D in the fields of degenerative health and healthy ageing has emerged from a wealth of observational studies repeatedly associating vitamin D deficiency—typically characterised as 25(OH)D at <30 nmol/L (Bouillon et al. 2022)—with degenerative conditions, such as accelerated age-related loss of skeletal muscle integrity [such as sarcopenia (Remelli et al. 2019)], degenerative respiratory conditions [such as chronic obstructive pulmonary disease - COPD (Zhu et al. 2016)] and increased falling risk (Cheema & Chaudhry 2016). These conditions particularly comprise distinct-but-interlinked phenotypes of concern for older aged and frail people (Bone et al. 2017). These insights have put onus on the scientific community to explore these observed relationships more deeply, which has included the conduct

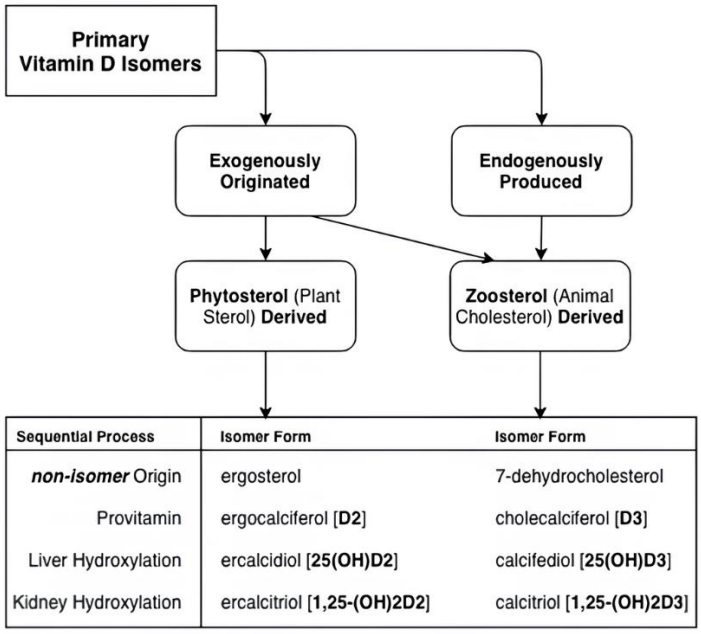
of clinical trials using bolus dosing (Camargo Jr et al. 2020; Martineau et al. 2017; Sanders et al. 2010). In the case of acute respiratory infections, bolus dosing has generally been shown to have net neutral effects [i.e., therapeutic nullity], while some improvements have been seen from smaller daily doses that better reflect the natural acquisition of vitamin D (Jolliffe et al. 2021). Meanwhile, primary and subgroup analyses from some clinical trials focused on motor function appear to have captured adverse effects from the use of bolus dosing, in way of increased fall risk (Bischoff-Ferrari et al. 2016; Ginde et al. 2017; Smith, Gallagher & Suiter 2017; Sanders et al. 2010; Waterhouse et al. 2021; Zittermann et al. 2023). Collectively, these studies demonstrate the need to diversify efforts of exploring vitamin D in the field of degenerative health, specifically, using safer and more expeditious research approaches, including the use of big-data, such as the UK Biobank (Sudlow et al. 2015).

Seeking to address this need [i.e., exploring relationships between vitamin D and degenerative health factors in a safe and ethical manner], using the UK Biobank [UKB] (Sudlow et al. 2015), most focus is pragmatically placed on motor function, given that motor function is an area where vitamin D is increasingly seen as a supportive factor for maintaining health, while also being an area to provide overt, and unexpected, evidence of risk from supplementation. Other factors of interest, such as mortality research, further serve to contribute context to the area of vitamin D in the field of degenerative health.

Chapter Two: Literature Review

2.1 Vitamin D: introduction

This literature review was undertaken by exploring published journal articles, which were (in most instances) found via Google Scholar searches with date range restriction set to 2019 onwards. Journal article reference lists were also explored to identify additional published research sources. The term vitamin D refers to a grouping of steroidal hormones that primarily mediate calcium absorption and regulation, as well as many other biological functions (Alia & Kerr 2021). Vitamin D contains a range of metabolites, all of which differ in their origin and metabolism (Holick & Slominski 2024). From a human-nutrition perspective, vitamin D isomers can be broadly differentiated into endogenously derived isomers [i.e., those synthesised via cutaneous-cholesterol exposure to UVB radiation], and those that originate exogenously from diet and supplementation (Benedik 2021). Exogenous isomers can be further subcategorized into those derived from sterols [plant based] and those derived from cholesterol [primarily animal-based, but can include some algae, select plant leaves, and tissues of the alfalfa plant] (Carlberg 2022a; Wang et al. 2001).



Sequential Process	Isomer Form	Isomer Form
non-isomer Origin	ergosterol	7-dehydrocholesterol
Provitamin	ergocalciferol [**D2**]	cholecalciferol [**D3**]
Liver Hydroxylation	ercalcidiol [**25(OH)D2**]	calcifediol [**25(OH)D3**]
Kidney Hydroxylation	ercalcitriol [**1,25-(OH)2D2**]	calcitriol [**1,25-(OH)2D3**]

Figure 1: Vitamin D isomer categorization

Sterols and cholesterol produce two different forms of vitamin D—provitamin D2 and D3, respectively. **Figure 1** demonstrates the broad categorization of phytosterol and zoosterol origins, along with their subsequent *liver hydroxylated* forms, **25(OH)D2** and **25(OH)D3** (Carlberg 2022b; Vieth 2020)—noting that algae/alfalfa dietary D3 is an exception, being plant-based, but cholesterol derived.

The two liver-derived forms of vitamin D are used as the primary metabolites for measuring vitamin D status of an individual (Amrein et al. 2020). However, the distinction—or

lack thereof—between D2 and D3 when measuring circulating concentrations is ultimately dependent on the type of assay being used (Fraser et al. 2020). For context, 25(OH)D concentrations in the UKB were measured by the Liaison XL assay, which does not differentiate between 25(OH)D2 and 25(OH)D3 (Binkley et al. 2010; Wallace et al. 2010). Consequently, 25(OH)D often refers to both sterol and cholesterol derived liver hydroxylated forms. The source of 25(OH)D, however, has been shown to have physiological significance, with 25(OH)D3 demonstrating greater vitamin D binding protein [VDBP] capacity (Chun et al. 2016), and is more likely to raise serum 25(OH)D levels than that derived from sterol source (Macdonald et al. 2012).

2.2 Vitamin D: acquisition and skin synthesis

The human body can generate D3 and downstream metabolites from a combination of endogenous cholesterols and subsequent exposure to sunlight, however, various forms of external acquisition of the vitamin serve to compliment—or compensate for—reduced internal synthesis of pre/pro D3 (Benedik 2021). As demonstrated in **Figure 2**, both provitamin forms of vitamin D [D2 and D3] are available in supplemental and food fortified sources, while whole food sources will usually contain a mixture of provitamin D and hydroxylated metabolites (Jakobsen & Christensen 2021; van den Heuvel et al. 2024).

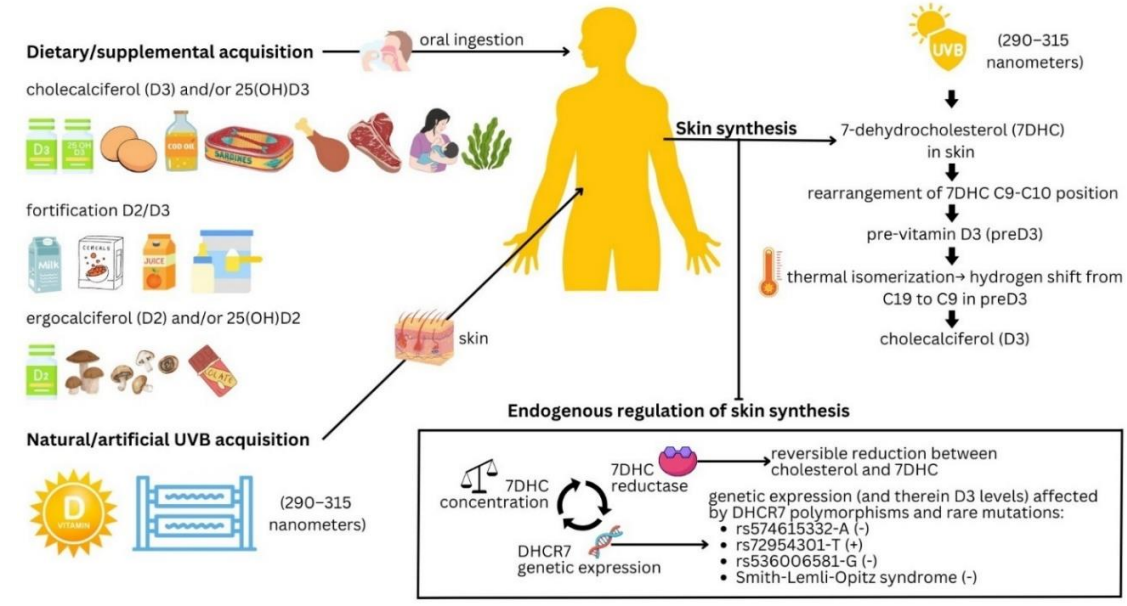


Figure 2: Vitamin D acquisition and skin synthesis, including skin synthesis regulation.
D2 = ergocalciferol. D3 = cholecalciferol. +/- = increase/decrease in enzyme expression. C = carbon. 7DHC = 7-dehydrocholesterol. DHCR7 = 7-dehydrocholesterol reductase gene.

Whole food sources may contain a mixture of metabolites, however, they primarily contain those forms consistent with their respective phytosterol or zoosterol origins [i.e., D2 from mushrooms and chocolate vs. D3 from egg yolks, oily fish, beef, lamb, chicken, and mammalian milks—including some algae] (Cashman et al. 2020; Dunlop et al. 2021). D2-based isomer content in beef, however, can account for anywhere up to one third of total vitamin D activity, and can vary in cattle depending on time of year and grazing conditions [(Cashman et al. 2020; Johnson et al. 2023)], and therein—depending on dietary choices—animal-based content could feasibly constitute an individual's primary source of dietary D2. Noteworthy is the recent work by Dunlop and colleagues that has ascertained canned salmon and fortified malted drink powder are the most concentrated sources of D3 available to the Australian diet, while the highest concentration of hydroxylated D3 metabolites included chicken eggs and chicken leg meat (Dunlop et al. 2021). D3 has emerged as the preferred provitamin form for supplementation and food fortification, and it is usually the most common form used in critical fortified foods, such as infant formulas (Gill & Indyk 2018). Access to pharmaceutical calcitriol [fully biologically activated D3] is restricted in most countries to medical-personnel-only [such as primary care providers, nurse practitioners, and endocrinologists] (Lung, Mowery & Komatsu 2022), however, in recent years the advancement of a hydroxylated supplement to the public marketplace, in the form of a calcifediol monohydrate supplement [partially biologically activated D3] has been seen, marking a significant shift in terms of what the public have access to, and what paramedical/allied health professionals can now prescribe (Turck et al. 2023).

Fundamentally, acquisition of vitamin D first depends on cholesterol, and secondly on sunshine. Specifically, synthesis of pre-vitamin D3 [preD3] is dependent on sufficient presence of 7-dehydrocholesterol [7DHC] within skin, with subsequent exposure to UVB radiation, within the 290-315 nanometre range, catalysing a hydrogen atom shift from the 9th position carbon molecule to that of the 10th position (Saponaro, Saba & Zucchi 2020) [Figure 2]. The next step from preD3 to provitamin D3 [wherein hydrogen shifts from carbon 19 position to carbon 9] is a thermal isomerization reaction and is dependent on body heat, with this reaction demonstrated—within skin at 37 degrees Celsius—to undergo maximal conversion within approximately 20 hours (Begum & Mirghani 2024; Wacker & Holick 2013).

The regulation of endogenous vitamin D3 synthesis is governed by a cyclical feedback loop involving the concentration of 7DHC, which in turn informs the synthesis of 7DHC reductase from the 7-dehydrocholesterol reductase gene [DHCR7], which converts cholesterol to 7DHC (Saponaro, Saba & Zucchi 2020). Smith-Lemli-Opitz syndrome is an example of a rare DHCR7 mutation, which results in accumulation of cholesterol in the 7DHC form, and therein can promote pathological hypocholesterolemia—with such affected individuals having multiple developmental issues but typically having higher than average levels of vitamin D metabolites (Latham & Cunniff 2021; Movassaghi et al. 2017). DHCR7 SNPs that have been shown, in recent investigation, to affect vitamin D metabolism [ascertained via lower/increased 25(OH)D concentration] include: rs574615332-A, rs536006581-G (Manousaki et al. 2020) / rs72954301-T (Sinnott-Armstrong et al. 2021), respectively.

2.3 Vitamin D: hydroxylation

Provitamin Ds ergocalciferol and cholecalciferol experience two hydroxylation steps in becoming fully biologically active (Acar & Özkan 2021), with the steps largely analogous, irrespective of the sterol/cholesterol origin. As outlined in **Figure 3**, the first hydroxylation step involves adding a hydroxyl group to the provitamin D carbon 25 position, and this primarily occurs via enzyme cytochrome P450 family 2 subfamily R member 1 [CYP2R1] in the liver (Makris, Sempos & Cavalier 2020), with research by Cheng and colleagues identifying this to occur in microsomal P450 fraction of hepatocytes (Cheng et al. 2003). The first provitamin D hydroxylation process is also highly expressed in males via the testes (Jensen 2021), with recent investigation showing transient differences in 25(OH)D levels for up to twelve months in men following an orchiectomy (Dieckmann et al. 2021). While CYP2R1 is not the only enzyme capable of catalysing 25(OH)D synthesis, it is the primary enzyme to do so, and research by Zhu and colleagues has demonstrated CYP2R1 gene knockout mice to have approximately 50% less 25(OH)D3 (Zhu et al. 2013). The CYP2R1 gene also has notable mutations, which are primarily responsible for vitamin D dependent rickets type 1B (Menon et al. 2024), while CYP2R1 SNPs rs117913124-A, rs117576073-T, rs117576073-G, and rs12794714-G are shown to lower (Manousaki et al. 2020; Revez et al. 2020) and rs12794714-G shown to increase 25(OH)D (Revez et al. 2020). Emerging research using mouse models suggests that CYP2R1 activity is

particularly affected by aging, with enzymatic activity diminishing by approximately thirty percent at the end of lifespan (Roizen et al. 2018).

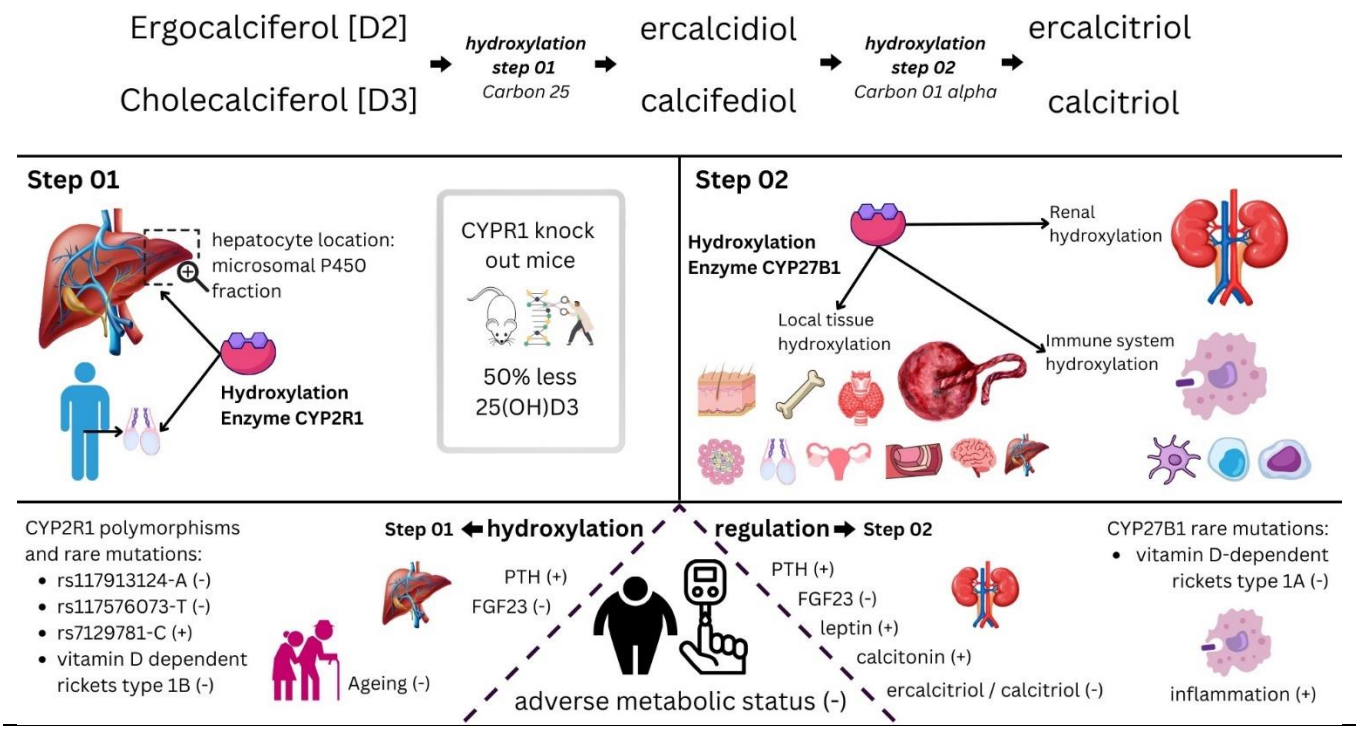


Figure 3: Vitamin D hydroxylation.
+/- = increase/decrease in enzyme expression or up/downregulation of hydroxylation. CYP2R1 = cytochrome P450 family 2 subfamily R member 1. CYP27B1 = cytochrome P450 family 27 subfamily B member 1.

Meanwhile, the second hydroxylation step is exclusively catalysed by enzyme 1α-hydroxylase from cytochrome P450 family 27 subfamily B member 1 [CYP27B1] and occurs by adding of a hydroxyl group to the carbon 1 position of 25(OH)D (Courbebaisse & Cavalier 2020), resulting in the formation of either ercalcitriol or calcitriol, depending on sterol origin [Figure 3]—*with here on in this review using cholesterol derived metabolite terminology*. Unlike the first hydroxylation step, the second can occur in a broader range of locations, with the process largely occurring via renal-based hydroxylation (Bikle 2020a), but with notable CYP27B1 activity also occurring within immune cells (Khashim Alswailmi et al. 2021; Lopez et al. 2021) and across a range of organs, glands and tissues, preforming various autocrine/paracrine functions (Hewison 2012) [Figure 3]. As discovered via investigating anephric [kidney function compromised] patients, macrophages [in response to inflammation, and independently of PTH regulation] were found to contribute the most extra-renal calcitriol synthesis, with placenta tissue also notably contributing during pregnancy (Hewison & Adams

2011). For other immune cells, including dendritic and T/B lymphocytes, their expression of CYP27B1 is dependent on their activation (Saponaro, Saba & Zucchi 2020).

Regarding regulation of renal calcitriol synthesis, factors such as parathyroid hormone [PTH] and calcitonin play crucial roles in modulating the activity of CYP27B1 (Young et al. 2022). PTH stimulates CYP27B1 activity, promoting calcitriol synthesis and thereby enhancing intestinal calcium absorption and contributing to calcium homeostasis (Fleet 2022). Conversely, calcitonin, released in response to elevated blood calcium levels, acts to lower calcium levels through slowing down the breakdown of bone [via the suppression of osteoclast activity], as well as via reducing renal reabsorption of calcium back into circulation (Srinivasan, Wong & Karponis 2020). Additionally, calcitonin stimulates CYP27B1 activity (Mady et al. 2023; Zhong, Armbrecht & Christakos 2009), likely as part of a compensatory mechanism in maintaining calcium balance. Leptin, a hormone produced by adipose tissue, also stimulates CYP27B1 activity [and VDR expression], potentially in a way to protect bone from obesogenic effects (Coomes & Haghbayan 2020). Conversely, calcitriol itself exerts negative feedback on CYP27B1, helping to maintain vitamin D homeostasis (Young et al. 2022). Furthermore, fibroblast growth factor-23 [FGF23], a hormone primarily produced by osteocytes in bone tissue, indirectly influences vitamin D metabolism by suppressing calcitriol production in the kidney. FGF23's main role is to prevent excess calcium absorption and maintain mineral balance in the body, contributing to overall mineral homeostasis alongside other factors involved in vitamin D metabolism (Nakatani et al. 2022). While there is great complexity regarding adverse metabolic status and vitamin D metabolism—fundamentally, evidence would suggest that insulin resistance has a direct negative impact on 1-alpha-hydroxylase activity (Trimarco et al. 2022), while obesity represses liver hydroxylation (Elkhwanky et al. 2020).

2.4 Vitamin D: transportation, cellular integration, genetic interaction, and catabolism

Calcitriol constitutes the fully biologically activated vitamin D metabolite that can impose functional changes at a genetic level (Lung, Mowery & Komatsu 2022). While calcitriol can circulate throughout the cardiovascular system in an unbound state, it is most likely found bound to VDBP, but can also be transported by albumin, at a respective ratio of approximately 85% to 15% (Palmer et al. 2022) [**Figure 4**]. VDBP has the strongest binding affinity to 25(OH)D, and it is in this sense that VDPB also acts as a reservoir for the metabolite (Alonso et al. 2023; Rivera-

Paredez et al. 2021). As per 'the free hormone hypothesis' [which suggests that only the un-protein-bound hormone can penetrate cells] free calcitriol/calcifediol and that liberated from albumin can diffuse into cells, while VDBP-bound metabolites require interaction with cellular membrane receptors, such as megalin receptors, to facilitate their cellular endocytic uptake (Bikle 2020b; Bikle & Schwartz 2019) [Figure 4]. Intracellular calcitriol can then interact with nuclear receptors, specifically the vitamin D receptor [VDR] and retinoid X receptor [RXR] (Meza-Meza, Ruiz-Ballesteros & de la Cruz-Mosso 2022) [Figure 4]. This interaction leads to the formation of VDR/RXR heterodimers, which are transcription factors that regulate gene expression (Nguyen et al. 2021). These heterodimers bind to specific DNA sequences known as vitamin D response elements [VDREs] which are typically located in the promoter regions of target genes, thereby modulating their expression (DeLuca & Clagett-Dame 2013).

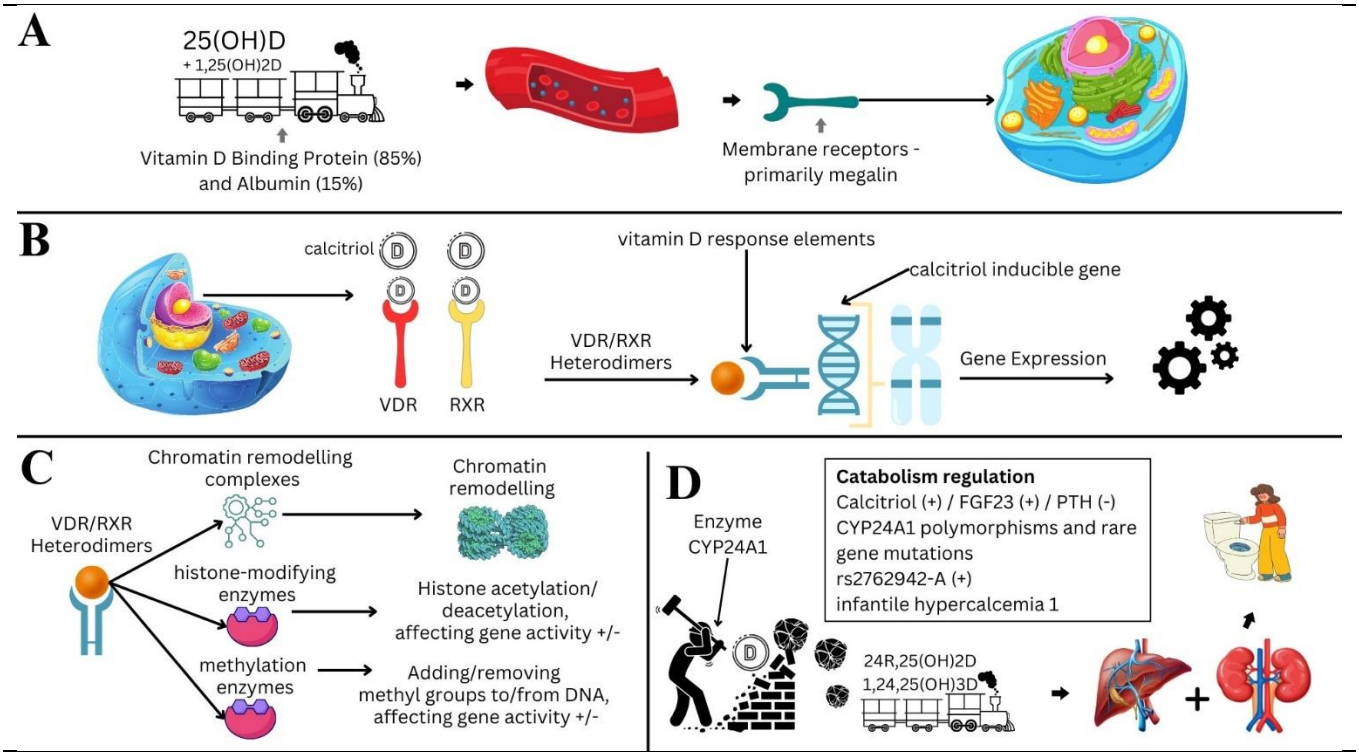


Figure 4: Vitamin D transportation and cellular integration (A), genetic interaction (B and C), and catabolism (D).
+/- = increase/decrease in enzyme expression. C = carbon. CYP24A1 = Cytochrome P450 family 24 subfamily A member 1 gene.

Calcitriol also promotes epigenetic influence via VDR/RXR heterodimers, which play a pivotal role in regulating gene expression through various epigenetic mechanisms (Snegarova & Naydenova 2020). These mechanisms include chromatin remodelling [i.e., chromatin being the

histone protein complex around which DNA is wound], with modification to chromatin altering the accessibility of DNA to transcription factors [i.e., including heterodimers] (Carlberg 2018). Chromatin remodelling can occur via the generation of chromatin remodelling complexes, which use energy to modify structure (Carlberg 2024; Tyagi et al. 2016), while acetylation/deacetylation of histone (Snegarova & Naydenova 2020) can serve to expand and contract histone, modifying accessibility [Figure 4]. VDR/RXR-mediated epigenetic changes also include the methylation/demethylation of DNA itself, with Català-Moll and colleagues having recently identified that VDR activation, via downstream interaction with Ten-Eleven Translocation 2 [a demethylation enzyme], initiates DNA demethylation in dendritic cells (Carlberg 2022; Català-Moll et al. 2022).

The catabolism of calcitriol marks the final stages of the vitamin D journey [Figure 4]. In this process, the Cytochrome P450 Family 24 Subfamily A Member 1 [CYP24A1] enzyme plays a central role as the primary catabolic factor for calcitriol, converting it into the inactive metabolite 24(R),25$(OH)_2$D (Kamiya et al. 2023). Additionally, CYP24A1 contributes to vitamin D regulation by metabolizing excess 25(OH)D to 24,25$(OH)_2$D. Recent work by Cavalier et al. promotes a 25(OH)D to 24,25$(OH)_2$D ratio measure as an evolving tool for predicting vitamin D deficiency, and functional pathologies involving CYP24A1 activity (Cavalier et al. 2020). Regulation of CYP24A1 is finely tuned to maintain vitamin D homeostasis, with calcitriol itself acting as a regulator by promoting the upregulation of CYP24A1 expression when calcitriol levels are elevated; this upregulation serves to increase the catabolism of calcitriol, preventing excessive accumulation of the active hormone (Shaurova, Seshadri & Hershberger 2018). PTH exerts an inhibitory effect on CYP24A1 expression, aiming to conserve calcitriol levels in situations where its synthesis or availability may be limited (Meyer & Pike 2020). Infantile hypercalcemia 1 constitutes a rare CPY24A1 mutation that results in elevated levels of calcitriol (Lenherr-Taube et al. 2021), while CYP24A1 SNP rs276942-A is associated with increased 25(OH)D measurement (Manousaki et al. 2020). Inactivated calcifediol and calcitriol are ultimately transported by VDBP for excretion, which primarily occurs via bile, and a little via urine (Del Valle et al. 2010).

2.5 Vitamin D: a focus on metabolic interplay

As discussed, additional factors such as calcium can interact with vitamin D, and vitamin D with calcium in return. Concerns have emerged around soft tissue calcification and the relationship therein with dysregulated vitamin D metabolism (Tintut & Demer 2021), with focus on both excess and deficiency mediated processes potentially driving such pathology (Wang et al. 2018). Meanwhile, vitamin A [specifically the retinoid metabolites] constitute another important factor that interacts with vitamin D (Ramanarayanan, Heine & Worm 2023), with the aforementioned RXR playing an integral role, alongside other factors, in facilitating the genetic effects of both vitamins (Osz et al. 2020; Rochel 2022). The retinoid metabolites constitute the animal-derived forms of vitamin A, with retinol being the primary circulating form, and then downstream metabolites, including retinoic acid, exerting function at a genetic level (Petkovich & Chambon 2022). In addition to RXR, the primary receptor that retinoic acid binds to—and from thence produce heterodimers—is the retinoic acid receptor [RAR] (Petkovich & Chambon 2022).

At the level of heterodimer synthesis, retinoic acid and calcitriol have a competitive antagonism relationship, with this dynamic being elucidated in seminal animal model studies, where large doses of one vitamin were used to mitigate the effects of supplementation induced by the other (Metz, Walser & Olson 1985; Rohde & DeLuca 2005; Rohde et al. 1999), while a high retinol status in humans has been demonstrated to aggravate the effect of vitamin D deficiency on conditions such as osteoporosis (Caire-Juvera et al. 2009; Mata-Granados et al. 2010; Navarro-Valverde et al. 2018). Heterodimer competition can have tangible significance for health, with breast cancer being a pertinent example of how heterodimer competition may act in balancing variations of possibly pro-tumorigenic retinoic acid mediated gene expression (Mezquita & Mezquita 2019). In this sense, part of the established role of calcitriol in protecting against cancer (Chakraborti 2011) likely includes a VDR-RXR mediated balancing of retinoic-acid response element [RARE] gene expression of cancer related cells (Jiménez-Lara & Aranda 1999). Contemporary research by Vyas and Goswami has demonstrated calcitriol and retinoic acid antagonism as an important mechanism behind the suppression of T helper 9 cell (Vyas & Goswami 2021), which has potentially protective implications for melanoma and lung adenocarcinomas (Kaplan, Hufford & Olson 2015).

Available literature shows that where steroid hormone balance is working against retinol [i.e., high vitamin D levels], the potentiality for calcitriol to bind to RXR—in addition to VDR—can produce excessive VDR-RXR heterodimers, which are capable of binding to RARE in retinoic acid responsive genes (Carlberg & Saurat 1996; Jiménez-Lara & Aranda 1999; Rehó et al. 2023). In this context, the VDR-RXR produces a repressive heterodimer species that can suppress gene expression that would otherwise occur from RAR-RXR stimulation (Carlberg & Saurat 1996; Jiménez-Lara & Aranda 1999; Rehó et al. 2023) [**Figure 5**].

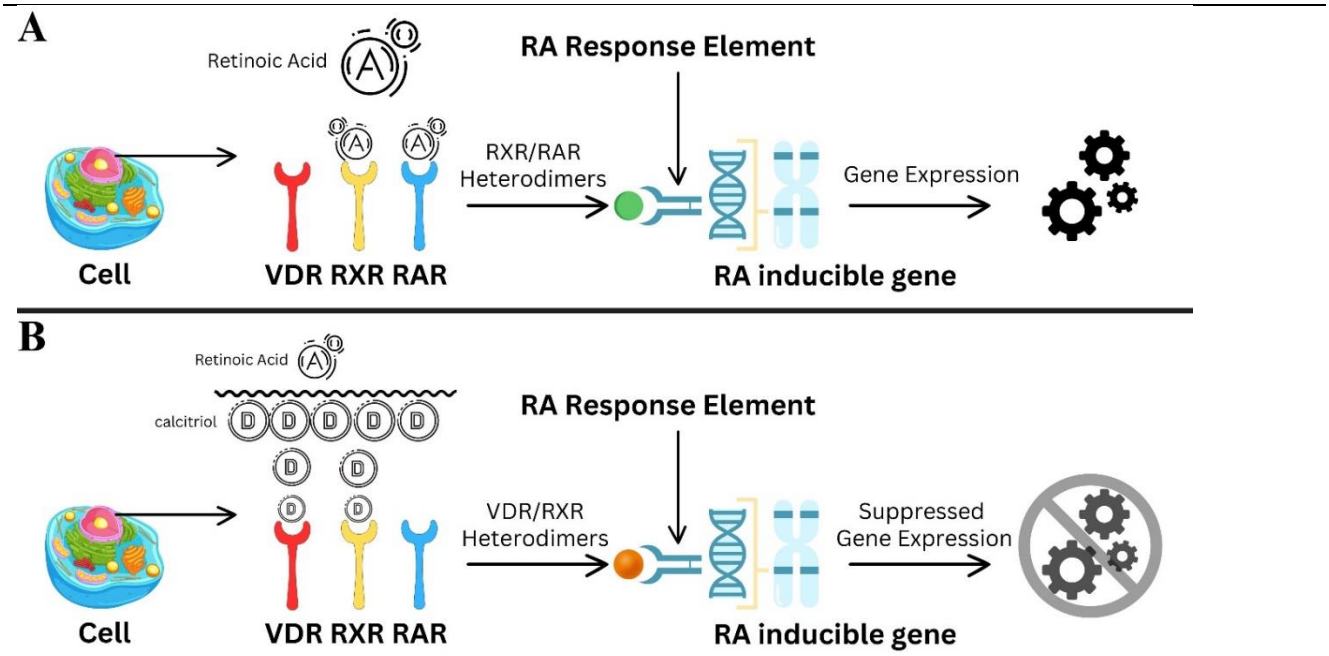


Figure 5: Steroid hormone competition visually depicted via gene expression of a retinoic acid inducible gene in response to (A) an activating heterodimer species, versus (B) a repressive heterodimer species.
A = depicting the uninhibited gene expression of a retinoic acid inducible gene in response to RA stimulated RXR/RAR heterodimers, and B = depicting inhibited gene expression of a retinoic acid inducible gene in response to calcitriol stimulated VDR/RXR heterodimers. RA = retinoic acid. VDR = vitamin D receptor. RXR = retinoid X receptor. RAR = retinoic acid receptor.

In this sense, a *functional*, as opposed to nutritional, retinol deficiency is being induced. For several decades, literature has offered examples of functional retinol deficiencies, such as foetal alcohol syndrome, wherein in 1994 it was put forth that inhibition of retinoic acid synthesis from ethanol could contribute to the pathogenesis of the condition (Duester 1994).

2.6 Vitamin D: skeletal effects

The effects of vitamin D on bone are well understood and do not constitute the specific focus of this thesis—with a greater attention rather placed on extra-skeletal effects. To briefly recap however, calcitriol enhances calcium absorption in the intestines by promoting the synthesis of calcium-binding proteins, facilitating the translocation of calcium for crucial use in deposition onto the bone matrix (Christakos et al. 2020). Calcitriol, via osteoblast-located VDR, regulates this deposition through stimulating osteocalcin and alkaline phosphatase (van Driel & van Leeuwen 2023). Calcitriol also suppresses the production of receptor activator of nuclear factor-kappa B ligand [RANKL] by osteoblasts and other cells, thereby reducing the differentiation and activity of bone-resorbing osteoclasts, with this inhibition of osteoclast function helping to prevent excessive bone breakdown while contributing to the maintenance of bone density (Bi et al. 2020).

2.7 Vitamin D: extra-skeletal effects

Immune System: Vitamin D exerts effects on the immune system, influencing both innate and acquired immunity through complex mechanisms that are still being elucidated, but which are fundamentally understood to occur via VDR expression in almost all cells of the immune system (Martens et al. 2020). For innate immune function, calcitriol plays a role in modulating inflammatory responses [via supressing proinflammatory cytokines like Tumour Necrosis Factor-α, RANKL, interleukin-6, cyclooxygenase-2, or by facilitating the release of anti-inflammatory cytokines, such as interleukin-10 or Transforming Growth Factor-beta from macrophages (Ao, Kikuta & Ishii 2021; Currò et al. 2020; Leal et al. 2020)], as well as stimulating antimicrobial peptide production [β-defensin 2 and cathelicidin antimicrobial peptide (White 2022)]. Moreover, calcitriol supports adaptive immunity by inhibiting B-cell differentiation and T-helper lymphocyte proliferation, promoting a more tolerogenic immune response through additionally increased regulatory T cell production and anti-inflammatory cytokine balance (Sîrbe et al. 2022). This helps to prevent/ameliorate the effects of autoimmune diseases like systemic lupus erythematosus, via the curbing of excessive antibody production (Irfan et al. 2022).

Cancer: Emerging research suggests that vitamin D plays a role in the prevention and progression of cancer, with VDR expressed in various cancer cell lines, and calcitriol shown to

exert anti-proliferative effects on cancer cells (Carlberg & Muñoz 2020). Notably, it has been demonstrated that some tumours express less VDR as they advance (Brożyna et al. 2011), while CYP24A1A [calcitriol's primary catabolism enzyme] appears to be overexpressed in human malignancy (Kamiya et al. 2023). While VDR knockout mice do not spontaneously develop neoplasia, they do demonstrate increased susceptibility to malignancies when exposed to cancer-causing agents or chemicals (Allen et al. 2020; Bouillon et al. 2008). It is therefore important to highlight that most major forms of cancer demonstrate an association between lower circulating 25(OH)D and increased risk, including: bladder cancer, breast cancer, colorectal cancer, glioblastoma, head and neck squamous cell carcinomas, liver cancer, lung cancer, melanoma, osteosarcoma, ovarian cancer, prostate cancer, squamous cell carcinoma (Feng et al. 2017; Francés et al. 2017; Horas et al. 2020; Na et al. 2022; Seraphin et al. 2023; Song et al. 2018; Yi, Wang & Tu 2021; Zhang et al. 2015) Literature already presented in this review has used breast cancer as an example to discuss VDR/RXR heterodimer activity and the role therein for regulating cellular expression. Additionally, presented literature on immunomodulation and inflammatory modulation is mechanistically relevant for cancer because chronic inflammation is linked to the heightened risk of many cancer types, while tumour progression is governed by the interplay of inflammatory cells and mediators [like cytokines and chemokines] (El-Sharkawy & Malki 2020). Literature presented for mortality will also cover the cancer topic.

Cardiovascular: In the cardiovascular system, VDR is expressed in major cardiac tissues (Kim, Norris & Artaza 2016) [including endothelial cells, vascular smooth muscle cells, and cardiomyocytes (Zittermann et al. 2021)] where continuing research elucidates a likely important role for VDR in modulating cardiac hypertrophy, among other functions (Liu et al. 2020). Mechanistically, calcitriol is implicated in calcium-dependent cellular processes, including the stimulation of calcium-binding protein synthesis [notably Calmodulin, for cardiomyocytes (Elíes et al. 2020)], and ionized calcium regulation. This calcium regulation is integral for normal cardiac contractility (Small 2006), normal cardiac electrical conduction, (Wei, Yohannan & Richards 2019) and therein necessary for heart rate regulation and arrhythmia prevention (Shah et al. 2022)]. Impaired heart rate measures can be seen in both human and animal vitamin D deficiency (Canpolat et al. 2015; Luo et al. 2022). Furthermore, calcitriol acts as a negative regulator of the renin-angiotensin system, contributing to blood pressure regulation and vascular function (Li 2024). Notably, VDR knockout mice exhibit signs of heart failure, including cardiac

hypertrophy and hypertension (Simpson 2011). A 2019 meta-analysis concluded that 25(OH)D deficiency [at <30 nmol/L] was associated with increases of combined incidence-mortality cardiovascular disease [CVD] (Mba et al. 2022), while a 2022 meta-analysis [using the higher insufficiency threshold of <50 nmol/L 25(OH)D], found an association with increased major adverse cardiovascular events, but no difference in mortality (Jaiswal et al. 2022). Benefit from vitamin D supplementation for reducing cardiovascular events continue to prove inconclusive, according to two recent meta-analyses (Pei et al. 2022; Ruiz-García et al. 2023).

Metabolic health: In the metabolic domain, vitamin D plays a multifaceted role in glucose and lipid metabolism (Danese et al. 2023). Elevated PTH is associated with altered glucose and lipid profiles, highlighting the calcitriol-PTH relationship as a fundamental mechanism in metabolic health (Modica et al. 2023). Calcitriol has also been implicated in pancreatic β-cell function and glucose-stimulated insulin secretion (Bornstedt et al. 2019), and low 25(OH)D status has been linked to components of metabolic syndrome, including type 2 diabetes mellitus (Taderegew et al. 2023) and obesity (Haghighat et al. 2023). In general, calcitriol enhances insulin sensitivity via several means, including the promotion of glucose transporter type 4 and their translocation to the cell surface (Manna & Jain 2012); this promotion increases glucose uptake and utilization. Calcitriol also activates insulin sensitivity pathways in the liver (Leung 2019), in part via mediating the Calcium/Calmodulin-dependent Protein Kinase Kinase-β and AMP-activated Protein Kinase [AMPK] signalling system (Cheng et al. 2016), wherein AMPK activation leads to increased glucose uptake, enhanced fatty acid oxidation, and increased mitochondrial biogenesis (Spaulding & Yan 2022). For adipose tissue, calcitriol similarly helps support healthy insulin signalling (Ding et al. 2012), with insulin resistance in adipocytes increasing free fatty acids into the bloodstream, adversely reducing glucose metabolism and ultimately promoting fat accumulation and hyperglycaemia (Sears & Perry 2015). Calcitriol/D3 supplementation in animal and human studies, respectively, has been shown to lower free fatty acids (Chehsmazar et al. 2021; Kumar et al. 2023).

Additionally, calcitriol contributes to healthy adipose tissue function via influence on inflammation and adipokine secretion (Park & Han 2021). Calcitriol effect on skeletal muscle tissue—and the effect therein on metabolic health—is addressed below.

Muscle: the motor function portion of this review outlines vitamin D and its integral nature in muscle health. Mechanistically, however, at the cellular level, VDR is present in skeletal muscle tissue, where they regulate the transcription of genes essential for muscle growth and repair (Girgis & Brennan-Speranza 2021). Calcitriol promotes myogenesis, and in doing so, inhibits myoblast proliferation—and rather promotes myoblast differentiation into myotubes, which are the precursor cells for muscle fibres (Romeu Montenegro et al. 2019). Calcitriol also enhances the synthesis of contractile proteins [myosin and actin-associated proteins], contributing to muscle strength and function (Gogulothu et al. 2020). Non-genomic effects of calcitriol include triggering a transduction signal that activates Mitogen-Activated Protein kinase and phospholipase C pathways, which causes a quick increase in calcium levels inside the cell, which is integral for downstream contractile action (Halfon, Phan & Teta 2015). Contemporary discussion also posits a potential role of calcitriol on mechanistic target of rapamycin [mTOR], which is important to note given mTOR's effect on enhancing muscle mass (Bollen & Atherton 2021). Calcitriol also improves insulin signalling in skeletal muscle tissue, likely via various mechanisms affecting how lipids are stored and processed within myocytes (Jefferson et al. 2017).

2.8 Vitamin D: environmental and physiological factors, and nutritional states in humans

Seasonal variability, geographical location, and indoor time affect sun exposure and therefore constitute factors that can influence vitamin D metabolism (Hauta-Alus et al. 2018). Infants and the elderly are particularly at risk for deficiency as both groups have relatively less sun exposure, with the elderly also having reduced 7DHC (Shakoor et al. 2021; Tung et al. 2021). Additionally, individuals with increased levels of UVB absorbing melanin [i.e. darker skin] will experience relatively lower dermal-based D3 synthesis (Hartono et al. 2023). UVB radiation will typically be reduced during winter months, and the UBV reduction is proportionate to the extent one lives closer to the planetary magnetic poles (Nair & Maseeh 2012). Vitamin D related pathology is seen to have seasonal variability, with seminal work by Scragg (1981) hypothesising a connection between CVD, vitamin D and seasonality. Numerous works have followed, expanding on these investigations with, for instance, Grant and colleagues in 2017 positing that normalising 25(OH)D throughout the year would be likely to reduce the pronounced seasonality of death rates (Berry, Hyppönen & Cortina-Borja 2013; Grant, Bhattoa & Boucher 2017).

Authorities governing populations that occupy higher northern latitude locations—such as Canada and Finland [northern]—have implemented vitamin D fortification measures aimed at normalising yearly 25(OH)D status and eradicating vitamin D deficiency at a population level (McCourt & O'Sullivan 2022), while in New Zealand [southern], some voluntary programs are in operation (Cairncross et al. 2017), but fortification is not mandated. In Canada, a 2021 ruling has doubled the allocation of D3 fortification, up to 80 IU/100 mL of cow/goat milk and 1040 IU/100 g for margarine, however, up to 76% of Canadians are still reported to be below a 400IU/day intake (Vatanparast & Longworth 2023). In Finland, proactive implementation of fortification policies [applicable to milks, yogurt, margarine, and spreads] since 2003, and doubling recommended levels in 2010 [40 IU/100 g for milk and 800 IU/100 g], has resulted in average daily intakes above 400 IU/day and deficiency prevalence has dropped from 12% to less than 1% between 2000 and 2011 (Cashman & Kiely 2024). New Zealand's fortification policies remain voluntary, however, assessment of children who consumed fortified milk products revealed virtually zero risk of 25(OH)D <25 nmol/L, standing as a testament to those measures where voluntarily implemented (Cairncross et al. 2017).

Definitive consensus on vitamin D nutritional states [e.g. deficiency, sufficiency and excess] differ among expert opinion, as well as between two dominating institutional voices in this space—that being, the National Academy of Medicine [NAM] *formally the Institute of Medicine* [advocating a 50 nmol/L 25(OH)D sufficiency cutoff (Del Valle et al. 2010)] and the Endocrine Society [who has advocated for a 75 nmol/L sufficiency cutoff (Holick et al. 2011), until recent guideline update has shifted focus from distinct 25(OH)D thresholds to population specific indications for vitamin D supplementation (Demay et al. 2024)]. The NAM recommendations for 50 nmol/L are shared by the European Food Safety Authority and Nordic guidelines (Darling & Lanham-New 2024). There continues to be a large amount of debate in this space, with perhaps a notable ongoing discordance being the UK Scientific Advisory Committee on Nutrition having 25 nmol/L as their sufficiency cut off, basing this on ensuring musculoskeletal health only (Darling & Lanham-New 2024). Contemporarily across multiple sources, a 25(OH)D at <30 - 25 nmol/L is readily considered as moderate/severe deficiency, and in need of correction (Amrein et al. 2020; Bouillon et al. 2022; Sizar et al. 2024). Meanwhile, 25(OH)D at 220 nmol/L—the highest level demonstrated to not cause acute serum calcium dysregulation—has generally been considered the upper limit (Balvers et al. 2015).

2.9 Vitamin D: excess and ongoing toxicity risk—caution required.

Cases of acute vitamin D toxicity have been referenced throughout the history of vitamin D, however, it is only in recent years that an increasing discussion on vitamin D excess [i.e., the consideration of high but not acutely toxic serum concentration levels] as a potential driver of various subclinical pathology [e.g. calcification, motor function issues] has emerged in the literature (Jiang et al. 2018; Mazess, Bischoff-Ferrari & Dawson-Hughes 2021). Contemporary discussions highlight ≥125 nmol/L 25(OH)D as potentially indicative of excess vitamin D (Rizzoli 2021). However, it remains uncertain whether the bolus dosing method itself is in part responsible for potential vitamin D excess associated pathology, as opposed to exclusively a higher 25(OH)D status.

Although still occupying a relatively grey area within scientific discourse, the topic of vitamin D excess is increasingly growing in relevance due to a '*more vitamin D is better*' culture prevailing in sub-populations of the public, which has also been embraced by some clinicians, resulting in both likely increases of excess and overt toxicity cases (Kilbane et al. 2014; Taylor & Davies 2018). A 2017 cohort study demonstrated that people using over-the-counter vitamin D serum-testing and supplementation are up to 50 times more likely to exceed 220 nmol/l 25(OH)D levels, versus those under medical supervision (Shea & Berg 2017). However, literature demonstrates that even within the medical setting, issues can still arise, with a review of deficiency overcorrection cases noting prescriber error among toxicity causes (Galior, Grebe & Singh 2018), while another review notes 12 toxicity case studies [post 2010] resulting from prescriber/dispensing error (Taylor & Davies 2018). These insights serve to highlight that although rare, it is important for all clinicians to be cognizant of excess/toxicity risk, especially when engaging bolus dosing practices. With concern for vitamin D related soft tissue calcification having been voiced as early as the 1930s (Dent 1964; Harris & Innes 1931) cardiovascular related concerns have been at the forefront of vitamin D safety analyses and discussion for many years (Batman et al. 2021; Department of Health and Aged Care 2022; Wang et al. 2018; Medicines and Healthcare products Regulatory Agency (MHRA) 2016). Meanwhile, contemporary concern over what effects higher vitamin D may have on steroid hormone competition—and possible relationship therein to essential retinol dependant nervous system function—appear to have garnered relatively less consideration.

2.10 Vitamin D: Motor function

Motor function broadly refers to those functions that include the voluntary and goal-directed actions of the body (Hallemans, Verbeque & Van de Walle 2020). Fundamentally, a deficiency of steroid hormone, be that vitamin D or vitamin A, has been associated with childhood maldevelopment issues, with vitamin A deficiency particularly affecting growth and developmental milestones (Sommer 2008), and vitamin D deficiency affecting both muscle health (Remelli et al. 2019; Sahni et al. 2021) [i.e., myopathy, sarcopenia] and skeletal maldevelopment (Wheeler et al. 2019) [i.e., rickets]. Indeed, retinoid receptors are found readily throughout the striatum (Carta et al. 2006), which in mammalian brains has been shown to be imperative for governing motor development and function (Cataldi et al. 2022; Suzuki & Nishimura 2022). The concept of steroid hormone competition in the context of motor function remains largely unexplored. However, it is nonetheless interesting that the following motor function research has provided tangible evidence that calls into question the notion that 'more vitamin D is better'.

Insight on motor function development in infants can be garnered from a 2016 randomized controlled trial [RCT] that administered D3 supplementation to a cohort of 55 healthy Canadian infants, who were separated into lower [400 IU], moderate [800 IU] and higher [1,200 IU] supplementation groups (Wicklow et al. 2016). In contrast to the researcher's hypothesis, gross motor achievements were significantly higher in infants receiving the lower dosing in comparison to the higher supplementation group (Wicklow et al. 2016). Prior to the Canadian infant study, however, a 2010 RCT (Sanders et al. 2010) using high dose D3 supplementation with ≥70-year women, had already reported overt adverse effects, specifically increased fall risk, using a single orally administered annual dose of 500,000 IU (Sanders et al. 2010). Subsequent primary and subgroup analyses from RCTs, comprising both male and female participants [≥60 years], showed similar adverse effects for monthly dosing of D3, which ranged between 60,000 IU and 100,000 IU (Bischoff-Ferrari et al. 2016; Ginde et al. 2017; Waterhouse et al. 2021). Similar associations were also reported in one study of females [mean age 66 years] that used daily dosing of D3 at 4,000 to 4,800 IU (Smith, Gallagher & Suiter 2017), and in a meta-analysis of longterm daily supplementation, 3200 to 400 IU, which also largely comprised older adults (Zittermann et al. 2023). Some notable trials that also used higher dosing methods, however, like

the ViDA trial—100,000 IU monthly—did not report increased fall risk (Khaw, et al. 2017). Importantly, the Sanders study remains unique, as no other motor function related clinical trial has administered a comparably high single oral dose to an almost entirely vitamin D-replete population (Sanders et al. 2010). Such research has unintentionally provided the scientific community with valuable insights into the risks of supplemental vitamin D, highlighting the need for further research. These risks are particularly relevant as the elderly are at an elevated risk of vitamin D deficiency (Giustina et al. 2023; Kweder & Eidi 2018), as well as having a higher risk of falls (Britton et al. 2019), and a recent consensus statement on vitamin D in the older population [arising from the 5th International Conference "Controversies in Vitamin D"] concluding that supplementation remains the most effective and practical measure for maintaining $\geq$50 nmol/L 25(OH)D for this population group (Giustina et al. 2023).

Despite the increasing evidence of potential risks associated with high-dose vitamin D supplementation, there nonetheless remains a crucial need to thoroughly explore the implications of vitamin D deficiency in the realm of motor function from a public health perspective. Numerous observational studies have consistently supported a link between vitamin D deficiency and compromised muscle strength (Abiri & Vafa 2020). This association is supported by evidence from multiple RCTs showing support for vitamin D supplementation on improving muscle strength, enhanced physical performance, with these effects seen to arise most effectively from 1000 IU D3 daily dosing, in older people, and in the vitamin D deficient (Halfon, Phan & Teta 2015; Rejnmark 2011). Looking forward in this space, insights on sarcopenic obesity (Di Filippo et al. 2022) are now specifically warranting further investigation, with this area constituting both an important emerging public health focus (Ji, Li & Ma 2022), and one needing innovative and safe research approaches.

2.11 Vitamin D: mortality

Notable challenges are encountered while exploring the impact of vitamin D on mortality —as, despite a wealth of observational research associating increased risk of premature death with vitamin D deficiency (Buniello et al. 2019), researchers managing RCTs have inevitably grappled with issues such as the recruitment of individuals with vitamin D deficiency, in addition to facing ethical restrictions that limit such inclusion (Bolland, Grey & Avenell 2018b; Cooper et al. 2016). For instance, a 2020 meta-analysis, encompassing nine trials of vitamin D

supplementation in critically ill patients (Lan et al. 2020), yielded inconclusive evidence, but the analysis included only one RCT that explicitly examined vitamin D-deficient participants (Hasanloei et al. 2020), while another study included a subgroup characterised by severe deficiency (Amrein et al. 2014). Interestingly, these two studies, independent of the meta-analysis, exhibited positive effects of vitamin D supplementation on mortality—indicating that consideration of threshold effects is needed when evaluating available literature. Indeed, this has been corroborated by some observational research that has demonstrated threshold effects of 25(OH)D on mortality outcome (Durup et al. 2015; Sempos et al. 2013). A contemporary example that demonstrates the consideration of both threshold effects and dosing methods include the D-Health Trial 2022 (Neale et al. 2022), which—although not showing an association between monthly vitamin D bolus supplementation and all-cause mortality—acknowledged "*the precautionary principle would suggest that this dosing regimen might not be appropriate in people who are vitamin D-replete*" (p.1).

According to the latest RCT meta-analysis, investigating vitamin D supplementation and mortality (Cao et al. 2023), over 116 distinct studies have now been undertaken in this field. Cao and colleagues' review constitutes the largest number of included studies in one analysis to date, with this work having built upon seminal reviews and meta-analyses preceding it (Bjelakovic et al. 2014; Bolland et al. 2014; Pei et al. 2022; Ruiz-García et al. 2023; Zhang et al. 2014; Zhang et al. 2019). The work of Cao and colleagues is an umbrella review [encompassing mortality across the following categorisations: all-cause mortality, CVD, cancer, chronic kidney disease, respiratory disease, critically ill populations, musculoskeletal disease, sepsis population, type 2 diabetes, and liver disease] and has provided an evidence-map outlining notable benefit of supplementation supporting survival for lung cancer patients, patients with liver disorders, and reduction of all-cause mortality in persons with COVID-19 infection (Cao et al. 2023). The second latest meta-analysis (Ruiz-García et al. 2023) [encompassing 80 RCTs] found evidence for a decreased risk following supplementation for all-cause mortality among the 'fair and good-quality' included RCTs. Other work shows modest associations between supplementation and improved survival among elderly people (Bjelakovic et al. 2014), while some studies have demonstrated survival benefits following supplementation in specific subgroups, such as those battling cancer (Zhang et al. 2019), with the most recent work from Zhang and colleagues also highlighting a relationship to mortality reduction in lung cancer patients (Zhang, R et al. 2022).

For cancer patients, vitamin D benefits might appear relatively stronger in meta-analysis due to action of supplementation on cancer related pathology, where acute medicinal effects could be impacting cellular processes—such as proliferation, differentiation, and apoptosis—potentially extending beyond nutritional sufficiency thresholds (Carlberg & Muñoz 2020; Chakraborti 2011).

The COVID-19 pandemic has provided numerous vitamin D related RCTs and observational analyses that have included measured 25(OH)D status, —with these works in a continued process of being meta-analysed. While reports of favourable effect from vitamin D supplementation on COVID-19 mortality has already been featured above (Cao et al. 2023), this topic warrants its own inspection. The following systematic reviews and meta-analyses [analysing RCTs and/or prospective/retrospective observational studies] have investigated vitamin D and the COVID-19 pandemic, and appear [in the following] in the order of their original publication date: December 2021 (Ebrahimzadeh et al. 2023), May 2022 (Hosseini, El Abd & Ducharme 2022), June 2022 (D'Ecclesiis et al. 2022), January 2023 (Argano et al. 2023), January 2023 (Petrelli et al. 2023), March 2023 (Zhang et al. 2023), July 2023 (Meng et al. 2023), August 2023 (Sîrbu et al. 2023), April 2024 (Begum & Mirghani 2024).

Among these papers, there is a variety of rigour, in way of journal positioning. Taking two papers from higher positioned journals, we can review both an observational meta-analysis [25(OH)D status and COVID-19 mortality—Critical Reviews in Food Science and Nutrition (Ebrahimzadeh et al. 2023)] and a RCT meta-analysis [D3/calcifediol supplementation in COVID-19 patients—Clinical Nutrition (Meng et al. 2023)]. Effectiveness of vitamin D within deficiency groups was a concordant feature across both papers, with vitamin D deficiency having a direct association with elevated risk for COVID-19 mortality (Ebrahimzadeh et al. 2023), and supplementation reducing morality in the patients with vitamin D deficiency subgroup (Meng et al. 2023). Meng and colleagues did not show a mortality reduction from supplementation except within this deficient subgroup. The most recent meta-analysis shows an overall effectiveness for vitamin D supplementation in reducing mortality among patient with COVID-19, featuring the inclusion of 186,500 events from observational studies (Begum & Mirghani 2024). Petrelli and colleagues' umbrella review of systematic reviews and meta-analyses (Petrelli et al. 2023) determined a benefit on mortality outcome from supplementation occurring after a COVID-19

diagnosis, with similar outcomes reached by Hosseini and colleagues (2022), and D'Ecclesiis and colleagues (2022). (D'Ecclesiis et al. 2022; Hosseini, El Abd & Ducharme 2022). Meanwhile, Zhang and colleagues (2023) determined no effect from supplementation but noted the limited number of quality studies included at the time (Zhang et al. 2023). Similarly in early 2023, Argano and colleagues indicated more studies are required in order to determine the likely positive benefit from supplementation on mortality being observed (Argano et al. 2023), with a similar conclusion reached by Sirbu and colleagues later in August (Sîrbu et al. 2023).

In addition to Ebrahimzadeh and colleagues and Meng and colleagues having presented benefits in deficient patient subgroups, D'Ecclesiis and colleagues, also specified that benefits from supplementation notably occur in older individuals and in those at higher latitudes [i.e., those most likely to be deficient] (D'Ecclesiis et al. 2022), and they further specified that in such individuals all dosing methods showed benefit—suggesting fundamental benefit from at least first serving to replete a likely nutritional deficiency. Overall, these emerging COVID-19 era studies support the pre-COVID-19 meta-analysis of vitamin D supplementation in critically ill patients (Lan et al. 2020), wherein benefit was seen only in those with established vitamin D deficiency.

With at least 13,000 known VDR binding sites in the human genome (Carlberg 2014), and most tissues and cell types being responsive to calcitriol (Wang, Zhu & Deluca 2012), the literature ultimately supports a myriad of biologically plausible scenarios where vitamin D would serve to maintain the minimal reserves necessary to sustain life, as suggested by the epidemiological evidence (Buniello et al. 2019). Furthermore, to the extent that clinical trial data can/has examined the matter, the emerging body of evidence supports this association, although contextualization of the findings is required [i.e., replete vs. deplete populations; the potential for anticarcinogenic medicinal action, even beyond the point of nutritional repletion].

2.12 Degenerative health conditions

The field of degenerative health is characterised by diverse nomenclature, such as "chronic degenerative diseases" (Di Renzo, Gualtieri & De Lorenzo 2021), "age-related diseases" (Acar & Özkan 2021), "chronic inflammatory diseases" (Schett & Neurath 2018)—as well as a range of discipline-specific terms, such as "neurodegenerative diseases" (Dugger & Dickson 2017),

and “degenerative joint diseases” (Rezuș et al. 2019). Fundamentally, these terms largely encompass chronically evolving conditions that progressively deteriorate an individual’s health—spanning cardiovascular, musculoskeletal, neurological, respiratory, endocrine, gastrointestinal, and renal systems, and the sensory organs. These conditions are often characterized by underlying chronic inflammatory or dysregulated metabolic phenomena, with risk increasing as aging progresses (Sanada et al. 2018). This thesis specifically explores motor function related degenerative muscular related conditions [i.e., sarcopenia and falling (Delmonico & Beck 2017)], though it also gives focus to mortality outcomes, and importantly includes analysis of respiratory related deaths—with these two areas [i.e., muscular and respiratory degeneration] demonstrating common co-morbidity patterns (Benz et al. 2019; Bone et al. 2017), and common underlying pathologized biochemistry that both intersect with vitamin D (Dzik & Kaczor 2019; Jolliffe et al. 2019; Remelli et al. 2019).

2.13 Vitamin D: research—2000s and beyond

Genetic discoveries of the late twentieth century (Jiang, Kiel & Kraft 2019) have enabled contemporary genetic epidemiological research to explore vitamin D related pathologies (Afzal et al. 2014; Aspelund et al. 2019; Berry et al. 2012; Hyppönen, Vimaleswaran & Zhou 2022; Jiang et al. 2018; Wang et al. 2010). A fundamental advancement in the genetic epidemiological field has been the steep advancement of the Mendelian randomization [MR] method for the explicit use of exposure-health-outcome associations around the early 2000s, with its use increasing sharply from 2010 onward (Bowden, Davey Smith & Burgess 2015). Leveraging genetic variants as instrumental or "proxy indicators," MR has emerged as a unique avenue for establishing causality in instances where conducting RCTs is either unethical or impractical (Davies, Holmes & Davey Smith 2018). Overcoming inherent challenges like reversed causation and confounding, MR has proven invaluable in providing causal evidence for various exposures, with the Lancet published body mass index [BMI]-Cardiometabolic disease association (Lyall et al. 2017) and the BMJ published alcohol consumption-CVD association (Holmes et al. 2014) being prominent [non vitamin D related] MR examples. With calls to investigate concerns with higher vitamin D status having emerged (Jiang et al. 2018; Mazess, Bischoff-Ferrari & Dawson-Hughes 2021), MR stands as a safe and innovative approach for such endeavours.

Gene-environment interaction analyses [G x E] is another helpful approach to examine the role of vitamin D on degenerative health (Berman et al. 2018). While there are various approaches, G x E commonly explores the combined influence that a candidate gene and a given environmental factor may have in determining a particular outcome (Hunter 2005). In the context of vitamin D, G x E analysis can help to assess concerns related to higher vitamin D status by exploring the relationship between relevant steroid hormone gene SNPs and phenotypic environmental factors, and therein safely providing proof of principle evidence (Butera et al. 2023) that might then inform other forms of research. Fundamentally, a project geared towards the genetic epidemiological approach is well suited to address the associations between vitamin D and degenerative health—given that the collective methods can explore the relationship between vitamin D and human health [i.e., skeletal muscle traits], while also exploring the relationship with overt disease states [i.e., sarcopenia and mortality], and also safely and ethically exploring the excess vitamin D phenomena [i.e., fall risk].

2.14 Summary

This literature review outlines key areas for further exploration. It acknowledges various factors affecting vitamin D status, and importantly explores the mechanisms underlying vitamin D's extra skeletal effects. It has also discussed discourse on vitamin D deficiency and excess, particularly in relation to sarcopenic issues and fall risk. Clinical insights on vitamin D and mortality are highlighted, paving the way for contextual insights from the thesis. Emerging concerns about vitamin D supplementation safety are addressed, aiming to fill a gap in understanding identified in the literature.

Chapter Three: Thesis aims and objectives

3.1 Thesis primary aim

- The primary purpose of this thesis is to explore the relationship between 25(OH)D and motor function related degenerative health conditions, including mortality outcomes.

3.2 Thesis objectives

- To undertake a thorough exploration of 25(OH)D in the UK Biobank, and present findings pertaining to determinants of deficiency—serving to orientate and contextualise UK Biobank 25(OH)D data for use in this project [Chapter 5],
- To test for observational and genetic evidence to support the association between 25(OH)D and muscle traits [Chapter 6],
- To test for observational and genetic evidence to support the association between 25(OH)D and the risk of sarcopenia [Chapter 6],
- To test for observational and genetic evidence to support the association between 25(OH)D and the risk of Mortality. [Chapter 7],
- To explore for the possibility of whether increased fall risk from higher vitamin D varies by inflammatory status, and if the effect is dependent on a competitive steroid hormone relationship between retinol and vitamin D [Chapter 8]

Chapter Four: General Methods

In this section, an overview of the methods that have been applied in this thesis is provided, with each subsequent study chapter [chapters 5–8] providing a more detailed description of the methodology and expanding upon each method/data source and the relationships therein with the study analyses. Additionally, the UK Biobank, which constitutes the main dataset used across the four included studies (Sudlow et al. 2015), is described. Other than the first study [Study I]—which included phenotypic analysis only—the remaining studies [Studies II–IV] include both phenotypic and genetic-based analyses, utilising both conventional regression techniques and genetic epidemiological approaches, such as MR. The health outcomes assessed within this thesis include the following: vitamin D deficiency [Study I], skeletal muscle mass and grip strength/probable sarcopenia [Study II], mortality [Study III], and falling [Study IV].

4.1 Study data—participants

The expansive UK Biobank is a comprehensive prospective cohort with an initial assessment conducted on 502,316 individuals, aged 37 to 73 years (Sudlow et al. 2015). The baseline assessment was conducted for the first participants in 2006 and continued until 2010, encompassing the territories of England, Scotland, and Wales, —with follow up assessment and data collection continuing until the present day. Volunteers contributed by responding to a series of in-depth surveys, subjecting themselves to thorough physical examinations, and offering biological specimens to facilitate the analysis of biomarkers and genetic characteristics. The participant inclusion/exclusion criteria for this thesis were individualized for each study and are outlined in each individual study chapter. For instance, Study I includes multi-ancestry analyses, while the remaining studies are primarily genetics-based and therein restricted to participants with European ancestry.

4.2 Study data—phenotypic data

Most of the phenotypic exposure and covariate information used in the studies included in this thesis were obtained at the baseline assessment and **Table 1** details their data source and method of data collection. Covariates are pivotal in safeguarding the accuracy and reliability of effect estimates in a study, enhancing the validity of findings by accounting for additional variables alongside exposure variables in statistical models (Morrow, Duff & Mayberry 2022). In the

context of the UK Biobank, a wealth of participant information has been carefully collated, providing a rich dataset from which variables can be derived for potential use as covariates in statistical models (Fry et al. 2019; Sudlow et al. 2015). These covariates span a spectrum of domains, ranging from demographic characteristics, lifestyle factors, and biological measurements (Sudlow et al. 2015; UK Biobank 2007, UK Biobank 2015a). Each individual study chapter in this thesis contains a detailed overview of covariates used, the rationale for their inclusion, and how they were used by the varying statistical methodologies.

Table 1: UK Biobank data and their source/collection point.

Variable	**Data Source**	**Collection Point**
Age, sex, physical activity, smoking, alcohol, Sun protection practices, outdoor time (winter/summer), employment/shiftwork, vitamin D supplementation, oily fish consumption, leisure time TV/PC use, ethnicity, birth location, fall	Self-reported touchscreen questionnaires	Baseline
Townsend deprivation index	Residential address data	Baseline
Serum 25(OH)D	LIAISON XL 25(OH)D assay	Baseline
CRP	Beckman Coulter AU5800	Baseline
Bioimpedance Spectroscopy	Tanita BC418MA body composition analyser	Baseline
Height	Physical assessment using Seca 202 stadiometer	Baseline
BMI	Height and weight measures	Baseline
Vitamin D sample season	Blood sampling month	Baseline
Weight	Tanita BC418MA, or body weight scales if body composition not analysed	Baseline
Assessment centre	Location of baseline assessment	Baseline
Longitudinal location	Assessment centre by latitude	Based on assessment centre location
25(OH)D, Month of blood sample, fasting time before blood sample	Biological sample acquisition data	Baseline
Sample aliquots for 25(OH)D measurement and SNP array data	Laboratory data	After baseline collection
Top 40 genetic principal components	Statistical analysis preformed on genetic data	After baseline collection
Mortality data	Data linkage	Follow up

Given that longitudinal location played an important role in this thesis, detailed information is included on how the variable based on the UK Biobank assessment centres was constructed, as visually demonstrated below in **Figure 06.**



Categorical latitudinal zoning, in ascending order of light to dark: <=51° / 52°-53° / 54°-≥55°. <=51° (Croydon, Barts, Hounslow, Reading, Oxford, Bristol, Cardiff, Swansea), 52°-53° (Birmingham, Stoke, Nottingham, Wrexham, Liverpool, Bury, Manchester, Sheffield, Leeds, Stockport), and 54°-≥55° (Middlesborough, Newcastle, Glasgow, Edinburgh).

Figure 6: Latitudinal zoning of the UK Biobank Assessment Centres.

4.3 Study data—phenotypic data for study I exposure variables and ethnicity/season stratification

For Study I, exposure variables were selected based on their relevance to various 25(OH)D related factors, including outdoor exposure, vitamin D intake, and independent potential to influence 25(OH)D concentrations. All exposure data were self-reported at baseline, except for latitude and socioeconomic status, which were derived from assessment centre location and residential data, respectively [Table 01].

The following exposure variables were included in analysis:

- BMI
- Latitudinal location
- Vitamin D supplementation
- Time spent outdoors in winter
- Time spent outdoors in summer
- Oily fish consumption
- Leisure time TV usage
- Leisure time computer usage
- Use of sun protection
- Smoking

Ethnicity was self-reported at baseline and subsequently collapsed into five primary ethnic groups:

- White European [encompassing White, British, Irish, and any other White backgrounds]
- Chinese
- Asian [covering Asian British, Indian, Pakistani, Bangladeshi, and any other Asian backgrounds]

- Black African [including Black British, Caribbean, African, and any other Black backgrounds]
- Mixed [comprising the "Other" ethnic group, White and Black Caribbean, White and Black African, White and Asian, and any other mixed backgrounds].

Season was determined based on the month of blood sampling, with categorization into:

- Summer [July to September]
- Autumn [October to November]
- Winter [January to March]
- Spring [April to June]

Body mass index [BMI] was calculated at baseline from height and weight measurements—calculated as body weight [kilograms] divided by height squared [meters]) (Nuttall 2015).

4.4 Study data—phenotypic data for Study II outcome variables (muscle mass/grip strength/probable sarcopenia)

At the initial baseline assessment centre visit, The UK Biobank conducted grip strength assessments and bioimpedance analyses of skeletal muscle mass and whole-body fat mass (Franssen et al. 2014). Grip strength measurements were obtained using a Jamar J00105 hydraulic hand dynamometer (UK Biobank 2011a). Tanita BC418MA body composition analyser, in conjunction with inputted height data from a Seca 202 stadiometer, was used to provide the bioimpedance body composition data (UK Biobank 2011b). For statistical analyses on grip strength, the average contractile force for both the left and right hands was calculated, measured in kilograms. A previously established definition for probable sarcopenia was applied, categorizing participants as cases if their grip strength was less than 27 kg for males and less than 16 kg for females (Cruz-Jentoft et al. 2019; Dodds et al. 2020). To calculate body fat percentage, the whole-body fat mass was divided by body weight and the result was multiplied by 100, and the conventional cutoff points of ≥25% for males and ≥35% for females were used to identify obesity (Deurenberg, Yap & Van Staveren 1998). Analysis then focused on probable sarcopenic obesity and non-obesity probable sarcopenia, with participants without sarcopenia serving as the control group. Additionally, for the assessment of lean skeletal muscle mass, data from the left and right arms was combined.

4.5 Study data—phenotypic data for Study III outcome variable (mortality)

Data on mortality were procured by the UK Biobank from public health sources; for participants residing in England and Wales, data were sourced from NHS England, while for participants in Scotland, data were obtained from the NHS Central Register (NHSCR), which is a component of the National Records of Scotland (UK Biobank 2021). Both sources provide their respective morality data to the UK Biobank on a quarterly basis. The latest recorded death occurred in June of 2020, as per the August 2020 showcase update, which included Death Register records up to June 2020. To explore the underlying factors, primary causes of death were categorised according to the delineations set forth by the International Classification of Diseases, 10th Revision (Steindel 2010). This classification framework maps out distinct codes attributed to various conditions, facilitating a comprehensive understanding of mortal events, specifically related to—in the context of this thesis—cancer [indicated by codes C00 to D48], CVD [marked by codes I00 to I89], and respiratory disorders [encompassing codes J09 to J18, J20 to J22, and J40 to J47] (Steindel 2010). By employing this approach, the mortality analyses afforded insight into the landscape of mortality patterns in the study cohort.

4.6 Study data—phenotypic data for Study IV outcome variable (falls)

Information pertaining to fall incidents was gathered at the baseline through a self-reported questionnaire. Participants were asked, "In the past year, have you experienced any falls?" For the study IV analysis, this was dichotomized into a 'yes' or 'no' responses, excluding those who chose not to answer. Then, to explore the outcome of falling within a state of likely depressed retinol functionality, separate outcome variables were established to identify falls coinciding with high baseline CRP levels, starting at $\geq$5 mg/L. In the absence of retinol concentration data from the UK Biobank, CRP levels $\geq$5 mg/L was used as a surrogate marker to reflect retinol reduction of at least ~25% (Thurnham et al. 2003). Serum CRP concentrations were assessed using a high sensitivity immunoturbidimetric assay conducted on a Beckman Coulter AU5800 instrument. The assay had an analytical range spanning from 0.08 to 79.96 mg/L (Fry et al. 2019).

4.7 Study data—25(OH)D measurement

Serum 25(OH)D reflects both average dietary intake and endogenous production stimulated by UVB exposure, serving as a marker of vitamin D nutritional status (Balvers et al. 2015). This information was used both as an outcome [Study 1] and as an exposure [Studies 2 to 4]. Measured 25(OH)D serum concentrations in the UK Biobank were acquired using the LIAISON XL 25(OH)D assay, developed by DiaSorin in Stillwater, USA. The assay employs magnetic microparticle separation and operates as a fully automated chemiluminescence immunoassay system (Diasorin 2019). To gauge precision, the assay underwent evaluation in line with CLSI EP5-A2 standards, resulting in a coefficient of variation of 2.3% for intra-run measurements and a coefficient of variation of 7.8% for total measurements, based on data derived from six samples and two controls. Notably, the LIAISON XL adheres to Clinical Laboratory and Standards Institute (CLSI) EP17-A guidelines, having a functional sensitivity threshold of ≤4.0 ng/ml. The assay captures total concentrations of 25(OH)D, encompassing both 25(OH)D2 and 25(OH)D3 forms (Diasorin 2019). Analysis of baseline serum 25(OH)D concentrations from peripheral venous blood samples involved meticulous steps, as detailed by Allen and colleagues, to address potential batch effects and variability across multiple batches and aliquots; these steps included important standardizations applied to batches to mitigate dilution of sample aliquots (Allen et al. 2020). Each sample underwent rigorous processing and was assigned a unique aliquot identifier, which is used in this thesis as a covariate to account for batch variability. All studies employed categorical analysis of 25(OH)D levels, conforming to the consistent categorization thresholds of: <25.0, 25.0 - 49.9, 50.0 - 74.9, 75.0—99.99, and ≥100 nmol/L, albeit with slight variations; study 1 capped the upper category at ≥75 nmol/L, while study 2 reported findings only for <25 nmol/L. Across all studies, 25(OH)D levels <25 nmol/L were indicative of vitamin D deficiency.

4.8 Study data—overview of genetic data

Genetic data plays a pivotal role across Studies II, III, and IV. All three studies leverage SNPs associated with 25(OH)D levels, forming a genetic instrument for use as the instrumental variable in MR analyses (Davies, Holmes & Davey Smith 2018). Study IV extended beyond the 25(OH)D-related SNPs to also include functional variant SNPs from retinoic acid-inducible genes involved in the regulation of healthy motor function. In addition to utilizing these 25(OH)D and retinoic acid related SNPs, genetic-related covariate information was also

integrated into genetic analyses. This includes top 40 genetic principal components [to account for population stratification and ancestry differences], and birth location information [enabling exploration of potential geographical influences on genetic predispositions] and accounting for any potential variation of genotyping between SNP arrays (Bycroft et al. 2018). Each study individually outlines the use of genetic-related covariate information. The genetic data obtained from the UK Biobank comprises genotype details for 488,377 individuals, with two nearly identical genotyping array platforms being employed for this purpose. Initially, the Applied Biosystems™ UK BiLEVE Axiom™ Array by Affymetrix was utilized to genotype 49,950 participants, covering 807,411 markers. Subsequently, the remaining participants were genotyped using the Applied Biosystems™ UK Biobank Axiom™ Array, which shares approximately 95% of markers with the UK BiLEVE Axiom® Array(Bycroft et al. 2017).

4.9 Study data—quality control of genetic data

The genetic data in this study originated from the UK Biobank, where researchers employed the Axiom array, a specialized genetic marker array custom-designed for the UK Biobank project (UK Biobank 2015a). This array encompasses a vast collection of genetic markers, comprising approximately 820,000 variants, including both common and rare genetic variations. To maintain data accuracy and reliability, the UK Biobank implemented a rigorous quality assurance regimen encompassing sample handling, genotyping procedures, and imputation quality assessments. The fundamental steps included quality control of SNPs and then quality control of the samples. An account of these quality control processes can be found in the genotyping and quality control report (UK Biobank 2015b) and are summarised below in **Table 02**.

Table 2: Quality control measures for genetic data in the UK Biobank.

QC Step	Description	Functional Result
SNP QC: Identification of SNPs with less reliable genotyping results	SNP-based metrics used to identify unreliable SNPs. SNPs failing QC filters flagged and set to missing in affected batches.	Removal of unreliable SNPs from the dataset.
Sample QC: population structure	PCA performed on 150,000 UK Biobank samples using 100,000 SNPs. PCA used to capture population structure. Individuals with European ancestry were selected based on PCA results from HapMap populations. UK Biobank-specific PCA results were used for QC metrics.	Identification of samples with similar ancestry for downstream analyses.
Sample QC: Heterozygosity and Missing Rates	Extreme heterozygosity and low call rates were identified as indicators of poor sample quality. A linear regression	Removal of poor-quality samples from the dataset.

QC Step	Description	Functional Result
	model was used to correct for population structure effects. Outliers with high missingness or heterozygosity were flagged. Samples with extreme heterozygosity were distinguished from poor-quality samples using ROH analysis. 480 outlier samples were identified and provided in a list to researchers.	
PCA = Principal Component Analysis. QC = Quality Control. HapMap = HapMap population refers to a group of individuals from a specific geographic region or ethnic background whose genetic data were included in the HapMap project. ROH = Runs of Homozygosity.		

4.10 Study data—genetic instrument, used in Studies II, III and IV (for Mendelian randomization analyses)

The genetic score for measured 25(OH)D serum concentration was constructed using a weighted approach. Common autosomal genome-wide significant SNPs [with a minor allele frequency >5%] were compiled, identified from a genome wide association study [GWAS] analysis conducted on the UK Biobank dataset, which revealed 134 independent loci for serum 25(OH)D concentration (Revez et al. 2020). To ensure the robustness of GWAS signals, the instrument was restricted to 35 SNP (Hyppönen, Vimaleswaran & Zhou 2022) which can be replicated (with a consistent direction and a p-value <0.05) in the earlier GWAS for 25(OH)D by the SUNLIGHT consortium (Jiang et al. 2018). To construct the genetic score for 25(OH)D serum concentration, the weighted average of the number of measured 25(OH)D-increasing alleles was calculated for each individual and subsequently multiplied this by the count of available variants. The weight for each SNP was the effect estimate of the association of the SNP with serum 25(OH)D in the SUNLIGHT Consortium (Jiang et al. 2018). **Table 03** presents the 25(OH)D related SNPs used in the 35 SNP instrument and is a condensed version of a larger table included in the supplementary. The expanded supplementary table (**Thesis Supplementary Table 1**) additionally includes base pair position information, allele frequency for the 25(OH)D increasing alleles, the beta coefficients, standard error, and P values from both the UK Biobank and the sunlight consortium, as well as the SNP data for the 122 SNP instrument used in sensitivity analyses.

Table 3: 25(OH)D related SNPs used to construct the 35 SNP genetic instrument.

35 SNP Instrument						35 SNP Instrument, *continued*					
No	**SNP**	**CHR**	**Gene**	**A1**	**A2**	**No**	**SNP**	**CHR**	**Gene**	**A1**	**A2**
1	rs6671730[a]	1	*PEX10*	G	A	**19**	rs61891388	11	*RP11-867G23.13*	G	T
2	rs35408430	1	*PADI1*	C	T	**20**	rs1660839	11	*AP002387.1*	A	G
3	rs7522116	1	*FOXO6*	C	T	**21**	rs12803256	11	*AP002387.1*	G	A
4	rs7528419	1	*CELSR2*	G	A	**22**	rs12798050[b]	11	*S100A11P3*	T	C
5	rs1933064	1	*FLG-AS1*	A	G	**23**	rs72997623	11	*DGAT2*	A	C
6	rs76798800	1	*DCST2*	G	T	**24**	rs1149605	11	*RP11-21L23.4*	C	T
7	rs6672758	1	*GALNT2*	T	C	**25**	rs10859995	12	*HAL*	T	C
8	rs727857	2	*LINC01122*	G	A	**26**	rs8018720	14	*SEC23A*	G	C
9	rs1047891	2	*CPS1*	C	A	**27**	rs261291	15	*ALDH1A2*	T	C
10	rs2012736	2	*UGT1A5 to A10*	C	A	**28**	rs77924615	16	*PDILT*	G	A
11	rs6782190	3	*CADM2*	G	A	**29**	rs212100	19	*SULT2A1*	T	C
12	rs705117	4	*GC*	C	T	**30**	rs10426	19	*KLK10*	A	G
13	rs1352846	4	*GC*	A	G	**31**	rs6123359	20	*BCAS1*	G	A
14	rs78151190	6	*CARMIL1 (LRRC16A)*	A	C	**32**	rs17216707	20	*CYP24A1*	T	C
15	rs75741381	7	*VGF*	C	G	**33**	rs2585442	20	*CYP24A1*	G	C
16	rs12056768	8	*LINC00536*	T	G	**34**	rs2762943	20	*CYP24A1*	G	T
17	rs77532868	10	*GRID1*	T	C	**35**	rs2074735	22	*PLA2G3*	C	G
18	rs12794714	11	*CYP2R1*								

A1: serum-25(OH)D-increasing allele; A2: alternative allele; SNP: single nucleotide polymorphism; CHR: chromosome number. Modified from the expanded Supplementary Table ##. See page ## for the expanded table and additional table attributions, which include: (Hyppönen, Vimaleswaran & Zhou 2022; Revez et al. 2020).

4.11 Study data—retinoic acid related genetic data for study IV.

Identifying genetic variants responsive to retinoic acid with potential impacts on motor function, began via a thorough exploration of the GWAS Catalogue (Buniello et al. 2019). This approach was twofold:

- *Genes Known for Retinol Responsiveness*: initially, a list was compiled of genes recognized for their responsiveness to retinoic acid, and which—supported by evidence from experimental animal models or human observational research—were also acknowledged contributors to the regulation of healthy motor function.
- *SNPs Associated with fall risk related pathology*: building on this gene list, exploration of the GWAS Catalogue uncovered SNPs [from the aforementioned genes] that demonstrated trait associations related to acquired neurological conditions linked to the risk of falls.

The gene *BRINP1* [BMP/Retinoic Acid Inducible Neural Specific Protein 1] stood out as a prominent candidate. In knock-out mouse models, the gene's absence was observed to induce dysregulated motor function (Berkowicz et al. 2016 a,b; Kobayashi et al. 2014). In humans, functional variants from the gene have been associated with Parkinson's disease and attention-deficit/hyperactivity disorder (Chairta et al. 2021; Pol-Fuster et al. 2021). Notably, these

conditions are characterized by motor regulation issues and an elevated risk of fall-related injuries (Amiri et al. 2017; Mazzoni, Shabbott & Cortes 2012; Pelicioni et al. 2019; Stray et al. 2013).

While other genes with possible connections to retinoic acid responsiveness and motor function were identified, the experimental evidence supporting these links was comparatively weaker, and therefore the relationships were more hypothetical in nature:

- *Nurr1/NR42A2*: due to its role in developing dopamine neurons, this gene is associated with Parkinson's disease and Alzheimer's disease, both of which entail motor function dysregulation (Al-Nusaif et al. 2022; Moon et al. 2019; Pan et al. 2019; Sacchetti et al. 2006; Stark et al. 2013; Zetterstrom et al. 1997).
- *GABA Subunit Genes*: these genes, linked to GABA insufficiency, are associated with cervical dystonia (Addae et al. 2012; Berman et al. 2018; Neelands, Zhang & Macdonald 1999; Song et al. 2005).

Beyond these, additional genes associated with various neurological disorders that may hold relevance:

- *BRINP2*: Associated with schizophrenia (Smeland et al. 2017).
- *NRG1*: Linked to bipolar disorder, schizophrenia, and depression—all of which elevate the risk of falls (Chapman et al. 2015; Hausdorff et al. 2004; Levchenko et al. 2020; Prata et al. 2009; Stefansson et al. 2002).

See **Table 04** for the list of motor function related SNPs from retinoic responsive genes, used in study IV. This data from the GWAS Catalogue was accessed in April 2023.

Table 4: Motor function related SNPs from retinoic responsive genes.

Mapped Genes	SNP	Trait	Risk Allele	MAF	Trait GWAS P Value
BRINP1	rs4837631	ADHD	C	0.4347	2×10^{-11}
	rs4837628	Parkinson's disease	T	0.4449	$^{*}1 \times 10^{-6}$
	rs1331621	Bipolar disorder	T	0.4299	$^{*}4 \times 10^{-7}$
	rs10818399	Insomnia	C	0.4066	1×10^{-11}
	rs942516	Bipolar disorder	A	0.1026	$^{*}5 \times 10^{-6}$
	rs7847153	Schizophrenia	C	0.2127	$^{*}9 \times 10^{-7}$
BRINP 2	rs17659437	Schizophrenia	T	0.1607	2×10^{-8}
	rs16851048	Schizophrenia	T	0.2666	5×10^{-14}
	rs6670165	Schizophrenia	T	0.1885	4×10^{-11}
NRG1	rs1383890	Insomnia	T	0.3385	4×10^{-8}

rs17673138	Schizophrenia, Bipolar, depression	C	0.0711	*6 x 10^{-8}
rs117715236	MAD	A	0.0072	*4 x 10^{-6}

*Trait association above 5 x 10^{-8}. MAF = minor allele frequency, ADHD = attention-deficit/hyper-activity disorder, MAD = major depressive disorder. Data obtained using the GWAS catalogue April 2023 (Buniello et al. 2019)

4.12 Study designs - overview

The analyses in this study are based on a prospective cohort study [UKB], with analyses conducted using cross-sectional data from the baseline assessment [Studies I—IV], while Study III includes prospective analyses [wherein the mortality outcome data was progressively captured over the course of nearly 15 years, up to the point of inclusion in the analysis] (Mann 2003). While study I used a classic epidemiological approach, studies II to IV used genetic epidemiological approaches and the Mendelian randomisation method, with Study IV additionally using G x E for additional investigations included in the addendum.

4.13 Study designs—epidemiology

While epidemiology is a continually evolving field, fundamentally, it is focused on investigating the determinants, incidence, and dispersion of health and disease within a population (Frérot et al. 2018). At its core, epidemiology employs observational studies—such as cross-sectional, case-control, and cohort studies (Mann 2003)—to investigate associations between exposures and outcomes. These approaches allow for the gathering of data on population characteristics, risk factors, and disease occurrence over time, facilitating the identification of patterns and trends. Through data analysis, epidemiology seeks to quantify the strength and nature of associations between exposures and outcomes, often utilizing measures of disease frequency [such as incidence and prevalence (Noordzij et al. 2010)] and measures of association [such as relative risk and odds ratio (Tenny & Hoffman 2017)]. The examination of associations between exposures and outcomes often involves statistical techniques, including regression analysis (Turner & Deng 2020), to assist in exploring potential confounding variables, which supports discerning differences between correlations and likely causation.

4.14 Study designs—genetic epidemiology

Genetic epidemiology—while sharing fundamental principles and statistical analysis techniques with traditional epidemiology—is a specialized discipline that explores the role of genetic factors in disease, or uses genetic factors as tools in elucidating deeper understanding of diseases (Fallin,

Duggal & Beaty 2016). This discipline employs diverse study designs, including family-based studies (Nudel et al. 2022), GWAS (Visscher et al. 2017), and statistical methods like polygenic risk scoring to uncover genetic determinants of disease risk (Choi, Mak & O'Reilly 2020). Moreover, the discipline includes methods like MR and G x E interaction analyses, with MR [leveraging genetic variants as instrumental variables (Davey Smith & Hemani 2014)] and G x E methods [exploring how genetic predisposition interacts with environmental exposures to influence disease susceptibility (Lovely et al. 2017)], helping to advance understanding of causal relationships in epidemiological research.

4.15 Study designs—Mendelian randomisation

Mendelian randomisation (MR) is a genetic methodology that uses specific genetic variants closely associated with the exposure of interest as instrumental or 'proxy' indicators (Davies, Holmes & Davey Smith 2018) [**Figure 07**]. MR is based on Mendel's law of segregation and independence, which asserts that children inherit alleles in a random manner—relative to their parents, and also random in respect to other genes in the genome (Gagliano Taliun & Evans 2021). This approach serves as a valuable tool for establishing causal relationships in situations where the conducting of RCTs is ethically or logistically challenging (Davies, Holmes & Davey Smith 2018). MR effectively addresses significant challenges inherent in other observational methodologies, including issues related to reverse causation and confounding factors (Burgess et al. 2020). Causal inference in MR analysis relies on three fundamental assumptions (Burgess et al. 2020):

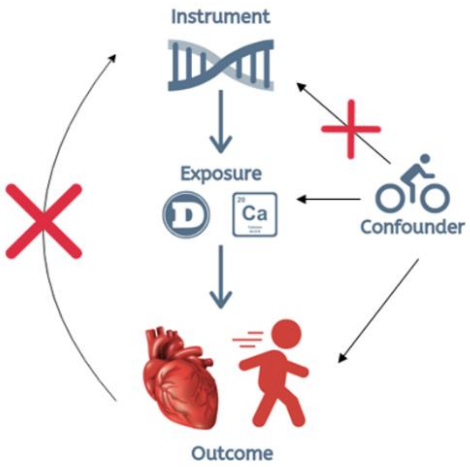


Figure 7: Mendelian Randomization.

Depicting the direct pathway from genes to genetic-related exposure, and from exposure to outcome. However—as depicted by the crosses—inherent constraint is placed on outcome and confounding variables from influencing the genetic makeup determined at conception.

- The genetic score for the exposure is associated with the measured exposure concentrations.
- The genetic score for the exposure does not directly impact the health outcome under investigation.
- The genetic score for the exposure is not associated with confounding factors affecting both the exposure and the outcome.
- The studies contained in this thesis have addressed these important assumptions in various ways, outlined below in MR sections 4.19-22 and Studies II–IV.

4.16 Study designs—Gene-environment interaction.

G x E analyses can be used to examine whether the effect of exposure on the outcome is modified by one or more genetic factors (Hunter 2005). Related analyses can be applied across different types of outcomes, but in any one analysis, typically only one outcome is addressed (Moonesinghe et al. 2012; Murcray, Lewinger & Gauderman 2008). A typical G x E analysis [such as that utilised by this project] may involve using one or more SNPs, and the assessment of whether the exposure—here, 25(OH)D—has a similar association with the outcome between individuals with differing genetic makeup (Han & Chatterjee 2018). Such analysis may provide insight on possible mechanistic considerations. An example G x E question [pertinent to this thesis] is as follows: *is the effect between an environmental exposure and disease trait modified [i.e., decreased or increased] due to genetic factors that affect/influence/interact with or are affected by the environmental exposure*?—and for instance, in this G X E question, the environmental exposure could hypothetically be *vitamin D*, the genetic factor could be *impaired retinoic acid responsive heterodimer signalling*, and the disease trait could be *increased fall risk*.

4.17 Statistical analyses - Linear and logistic regression, Studies I–IV (including use of interaction analyses)

In this section, an overview is presented of statistical methods used across Studies I to IV– linear/logistic regression, interaction analyses and use of quadratic terms. In Study IV, these methods were used for analysing some genetic data, however, in all other studies, these methods were used in the analysis of phenotypic data (see next section for where regression was used in the context of MR, using genetic data). For phenotypic analyses, the analytical approach predominantly revolved around linear and logistic regression models, implemented in STATA (StataCorp. 2023. Stata Statistical Software: Release 16-18. College Station, TX: StataCorp

LLC.) These models served as the foundation for investigating the associations between 25(OH)D concentrations and a range of health outcomes (Castro & Ferreira 2023)—these being, vitamin D deficiency as <25 nmol/L 25(OH)D, skeletal muscle mass, grip strength, probable sarcopenia, mortality, and falls.

Linear regression models are employed to examine the relationship between an exposure [i.e., 25(OH)D] and a continuous health outcome of interest [i.e., grip strength, skeletal muscle mass]. The association estimate is expressed in terms of a regression coefficient representing the change in the health outcome for each unit increase in the exposure (Castro & Ferreira 2023). On the other hand, logistic regression is used when the health outcome is binary, such as the presence or absence of a specific condition [i.e., falling or not falling, or having or not having vitamin D deficiency]; it models the probability of the binary outcome occurring based on independent variable. The association estimate is expressed in terms of an odds ratio describing the odds of having the outcome per unit increase in the exposure. In both instances, additional independent variables serve as the adjustment covariates in the model (Turner & Deng 2020).

Regarding the use of interaction terms [within the regression models, undertaken in STATA], these terms were introduced to account for the complexity of the associations. Interaction terms are formed by multiplying independent variables and incorporating the resulting products as extra independent variables within the regression equation (Brunner 2016). In the context of this thesis, interaction terms allowed us to assess if the relationship between 25(OH)D concentrations and a given health outcome varied according to the levels of a specific independent variable—and therein justify presenting such analyses stratified according to those independent variables of interest. For instance, work undertaken in this thesis examined how the associations between 25(OH)D and risk of sarcopenia varied according to obesity, as well as how the association between 25(OH)D and risk of falls varied according to CRP levels—and how 25(OH)D and multiple health outcomes varied according to seasonality.

4.18 Statistical analyses—linear and logistical regression using the fractional polynomial technique for presenting phenotypic findings as nonlinear (including the use of quadratic terms in regression, to test for nonlinearity)

In Studies II and III, fractional polynomials were used to model non-linearity between exposure-outcome association (Royston, Ambler & Sauerbrei 1999), and employed a likelihood ratio test

to select the best-fitting fractional polynomial function for the exposure of interest. In another instance, where fractional polynomials were not being used, the evidence of non-linearity or curvature was assessed by including a quadratic term of the exposure in the regression model.

Also investigated was whether the association between fall prevalence and 25(OH)D contained a statistically meaningful nonlinear distribution. To accomplish this, a quadratic term was introduced into the regression model. Quadratic terms assist in identifying nonlinearity by allowing the effect of an independent variable on the health outcome of interest to vary according to the level of the independent variable itself.

4.19 Statistical analyses—Mendelian randomization overview, Studies II–IV

All studies in this thesis using MR (Studies II—IV) include the linear form of MR. Study III (and some sensitivities in Study II) utilized nonlinear MR. The following description outlined the basic methodological approach for MR, and the supplementary methods sections for Studies II–III contain highly detailed information pertaining to the methodological differences between the linear and nonlinear approaches, as well as an addendum to the published study III, specifically addressing nonlinear MR, included at the end of Chapter 7. In this thesis, both linear and non-linear MR approaches are utilized to examine the genetic evidence for exposure-outcome relationships, with the latter capable of assessing evidence for curvature. For the linear MR analysis, MR estimates are computed using either individual level data or summary level data where appropriate. The non-linear MR can be viewed as an extension to the linear MR, which allows the effect of the exposure on the outcome to vary by the level of exposure.

4.20 Statistical analyses—Linear Mendelian randomization using individual level data

In this approach, MR analysis is performed using individual level data from a single sample in which genetic variants, exposure and outcome are measured in the same individuals. The MR estimate is computed - *using the ratio of coefficients method* - where the **Genetic Score—Measured Exposure** (γ_1, equation 1) and the **Genetic Score—Outcome** association estimates (Γ_1, equation 2) are computed and then used to calculate the ratio estimator (Palmer et al. 2011),

$$\hat{\beta} = \widehat{\Gamma_1} \Big/ \widehat{\gamma_1}. \qquad (1)$$

Using 25(OH)D and Fall as an example, Genetic Score—Measured Exposure and the Genetic

Score—Outcome association estimates are obtained from the models below:

$$\text{Measured 25(OH)D}_i = \gamma_0 + \gamma_1 \text{ Vitamin D Genetic Score}_i + \gamma_{2j} \text{ covariate}_j + \epsilon_i,\ \epsilon_i \sim N(0, \delta^2)$$

$$\text{Logit(FALL}_i) = \Gamma_0 + \Gamma_1 \text{ Vitamin D Genetic Score}_i + \Gamma_{2j} \text{ covariate}_j \quad (2)$$

In these equations:

- The subscript i denotes an individual.
- The subscript j signifies a covariate.
- The model encompasses various covariates such as age, sex, assessment centre, birth locations, SNP array, top genetic principal components, and other factors influencing the exposure variable, including sample collection timing, fasting duration before sample collection, and sample aliquots for measurement.

4.21 Statistical analyses—linear Mendelian randomization using summary level data

Linear MR estimates can also be computed using summary level data-based approaches, which leverages SNP-level information, utilizing SNP-exposure and SNP-outcome association estimates from two independent samples as inputs. In this thesis, five distinct summary level data-based approaches are performed. These methods encompass the inverse-variance weighted method (IVW), the Mendelian randomization Egger method, the weighted median Mendelian randomization method, the weighted mode method, and the Mendelian randomization pleiotropy residual sum and outlier (Mendelian randomization Presso) test. These methodologies offer complementary approaches, each resilient to different patterns of horizontal pleiotropy. Consequently, a convergence of results across these methods signifies robust evidence supporting a causal association.

- **IVW**: IVW estimator (β, equation 3) is obtained by the weighted regression of the effect estimates of variant-outcome association ($\hat{\Gamma}_j$, equation 3) on the effect estimates of variant-exposure association ($\hat{\gamma}_j$, equation 3) (Burgess, Butterworth & Thompson 2013) with the weight being the inverse variance of variant's ratio estimate. IVW assumes that there is no pleiotropy of variants, or the net pleiotropy zero.

  $$\hat{\Gamma}_j = \beta\, \hat{\gamma}_j + \epsilon_j,\ \epsilon_j \sim N\left(0, \delta_j^2\right) \quad (3)$$

  where, the subscript j denotes a variant.

- **Mendelian randomization Egger:** Mendelian randomization Egger is similar to the IVW method, except that the regression model additionally includes an intercept term, β_0 (equation 4), which represents the average pleiotropic effect (Bowden, Davey Smith & Burgess 2015) The Mendelian randomization Egger method can account for directional pleiotropy (when $\beta_0 \neq 0$), however, it introduces an additional untestable assumption, which states that pleiotropic effect must be independent of variant-exposure association.

  $$\hat{\Gamma}_j = \beta_0 + \beta\,\hat{\gamma}_j + \epsilon_j,\ \epsilon_j \sim N\left(0, \delta_j^2\right) \quad (4)$$

  where, the subscript j denotes a variant.

- **Mendelian randomization Presso:** Mendelian randomization Presso (Verbanck et al. 2018) operates by iteratively executing the IVW method, with each variant sequentially excluded from the analysis. During each iteration, the residual sum of squares (RSS) is computed and utilized as a metric for assessing heterogeneity in the ratio estimate. If the observed RSS significantly exceeds the null distribution of expected RSS values, the respective variant is flagged as an outlier. The causal estimate is derived using the IVW method after eliminating all identified outlier variants.

- **Weighted mode Mendelian Randomization method:** In the weighted mode method (Hartwig, Davey Smith & Bowden 2017), variants are grouped into clusters based on the similarity of their ratio estimates, with weighting determined by the inverse variance of these estimates. Clustering is executed using a normal kernel density function, with the spread determined by a bandwidth parameter. The causal effect is then estimated from the cluster with the highest weight. This method provides an unbiased causal effect if the variants within the most heavily weighted cluster are valid instruments.

- **Weighted median Mendelian randomization method:** In the weighted median-based approach (Bowden et al. 2016), variants undergo ranking based on their ratio estimates, and the causal estimate is determined as the median of this distribution. In the weighted version of the median method, the median is derived from a distribution of ratio estimates where

variants with greater precision receive higher weighting. This method yields a reliable estimate if at least 50% of the weight is derived from valid instruments.

4.22 Statistical analyses—non-linear Mendelian randomization, residual and doubly ranked method

In the non-linear MR analysis, the cohort is first stratified into subgroups with different average levels of the exposure. Localized average cause effect (LACE) estimates are then computed for each stratum of the cohort, and subsequently used to piece together the exposure-outcome association curve.

In the stratification step, two stratification methods were used to divide the cohort into strata, including the residual method (Staley & Burgess 2017) and doubly-ranked method (Tian et al. 2023). For the residual method, strata are formed using quantiles of residual exposure, which is defined as exposure minus the variation induced by the genetic score. For the doubly ranked method, strata are formed by first ranking participants into pre-strata based on their level of genetic score and then by ranking participants within each pre-stratum according to their level of exposure. The residual stratification method assumes the effect of the genetic instrument be linear and constant for all individuals in the population (Staley & Burgess 2017). In contrast to the residual method, the doubly-ranked stratification method does not make any parametric assumption on the effect of the genetic instrument; instead, it assumes that an individual's counterfactual exposure ranking is the same at all levels of the genetic instrument, which is a much weaker assumption than the linear and homogeneous assumption made by the residual method (Tian et al. 2023).

Following the stratification step, stratum-specific MR estimate is computed using the ratio of coefficients method, and then meta-regressed on the mean stratum exposure level. The meta-regression was performed by fitting a range of fractional polynomial exposure-outcome models of degree 1 (equation 5) and degree 2 (equation 6), with the best-fitting model being selected based on the likelihood ratio test. A non-linear effect is evident, if the best-fitting fractional polynomial model of degree 1 has a better fit than the linear model (i.e. $P_{non\text{-}linear} < 0.05$).

Fractional polynomials of degree 1 are defined as:

$$f(x) = \beta_0 + \beta_1 x^p \quad (5)$$

where, p ∈ {-2, -1, -0.5, 0, 0.5, 1, 2, 3}, and power of 0 refers to the natural-log function.

Fractional polynomials of degree 2 are defined as:

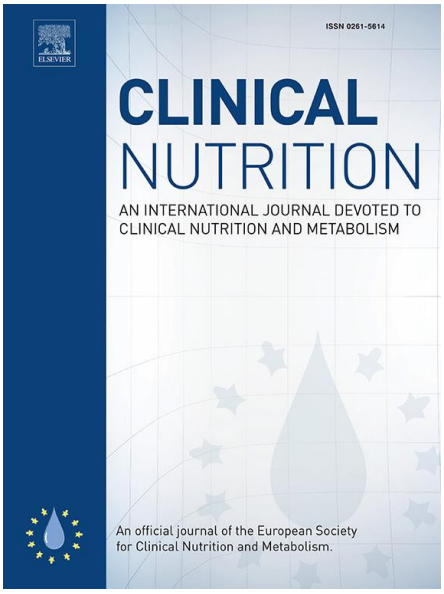
$$f(x) = \begin{cases} \beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2} \; if \; p_1 \neq p_2 \\ \beta_0 + \beta_1 x^{p} + \beta_2 x^{p} \log(x) \; if \; p_1 = p_2 = p \end{cases} \quad (6)$$

where, p_1 and p_2 ∈ {-2, -1, -0.5, 0, 0.5, 1, 2, 3}, and power of 0 refers to the natural-log function.

Conceptually, these MR methodologies form an interconnected framework for causal inference, each addressing specific challenges and nuances in causal estimation. Linear MR serves as the foundational approach, while non-linear methods MR offer extensions to capture complex relationships and heterogeneity in causal effects. Summary level data-based methods complement individual level data-based approach by leveraging SNP-level information and offering resilience against pleiotropy. Together, these methodologies contribute to a comprehensive toolkit for causal inference in epidemiological research.

Chapter Five: Study One

Differences and determinants of vitamin D deficiency among UK Biobank participants: A cross-ethnic and socioeconomic study



Clinical Nutrition (Elsevier) has published this manuscript and is the original source of the published article.

Impact factor: 7.3 (2020)

Rank in field: 4th out of 34 Quartile 1 journals in Nutrition & Dietetics.

Manuscript

Supplementary Material (supplementary tables 1 and 2)

Elsevier: *as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source.*

Study 1 Manuscript

List of study 1 manuscript figures:

Study 1 - Figure 1: Study population and participant inclusion criteria. 53

Study 1 - Figure 2: Distribution of categorical 25(OH)D concentration (stacked bar chart) and geometric mean of 25(OH)D concentration (line chart) by months. 59

Study 1 - Figure 3: Distribution of 25(OH)D concentration by seasons and socioeconomic, dietary, or lifestyle indicators. 62

Study 1 - Figure 4: Distribution of categorical 25(OH)D concentration (A) <25 nmol/L, (B) <50 nmol/L, across seasons and UK latitudinal zoning. 63

Study 1 - Figure 5: Distribution of 25(OH)D deficiency (<25nmol/L) by ethnic groupings and income (A) and by ethnic groupings and education (B). Educational level terms: National Vocational Qualification (NVQ), Certificate of Secondary Education (CSE), and A-level 64

List of study 1 manuscript tables:

Study 1 - Table 1: Demographic characteristics of UK Biobank participants. 57

Study 1 - Table 2: 25(OH)D concentration in UK Biobank participants, by season and demographic characteristics. 60

Study 1 - Table 3: Association of 25(OH)D deficiency (<25nmol/L) with selected risk factors by ethnic groups 65

Summary

Background: The public health relevance of true vitamin D deficiency is undisputed, although controversy remains regarding optimal vitamin D status. Few contemporary cross-ethnic studies have investigated the prevalence and determinants of very low 25-hydroxyvitamin D [25(OH)D] concentrations.

Methods: We conducted cross-ethnic analyses on the prevalence and determinants of vitamin D deficiency (25(OH)D ≤ 25 nmol/L) using data from 440,581 UK Biobank participants, of which 415,903 identified as White European, 7,880 Asian, 7,602 Black African, 1,383 Chinese, and 6,473 of mixed ancestry. Determinants of vitamin D deficiency were examined by logistic regression.

Results: The prevalence of vitamin D deficiency was highest among participants of Asian ancestry (57.2% in winter/spring and 50.8% in summer/autumn) followed by those of Black African ancestry (38.5% and 30.8%, respectively), mixed (36.5%, 22.5%), Chinese (33.1%, 20.7%) and White European ancestry (17.5%, 5.9%). Participants with higher socioeconomic deprivation were more likely to have 25(OH)D deficiency compared to less deprived participants (P = < 1 x 10-300); this pattern was more apparent among those of White European ancestry and in summer (Pinteraction ≤ 6.4 x 10-5 for both). In fully-adjusted analyses, regular consumption of oily fish was associated with reduced odds of vitamin D deficiency across all ethnicities, while outdoor-time in summer was less effective for Black Africans (OR = 0.89, 95% CI, 0.70-1.12) than White Europeans (OR = 0.40, 95% CI, 0.38-0.42).

Conclusions: Severe vitamin D deficiency remains an issue throughout the UK, particularly in lower socioeconomic areas. In some groups, levels of deficiency are alarmingly high with one-half of Asian and one-third of Black African ancestry populations affected across seasons.

Key messages: The prevalence of vitamin D deficiency in the UK is alarming, with certain ethnic and socioeconomic groups considered particularly vulnerable.

Key words: Vitamin D, Vitamin D Deficiency, Nutritional Status, Public Health, Socioeconomic Factors, Ethnic Groups.

Introduction

Vitamin D is a hormone precursor synthesized in the skin during exposure to UVB radiation. The active hormone regulates widespread biological functions, with receptors for vitamin D located throughout the body [1, 2]. A primary role of the hormonal vitamin D system is the mediation of osteogenic-related mineral absorption and utilisation [2], with skeletal development and maintenance constituting principal foci of public health strategies [3].

25-hydroxyvitamin D [25(OH)D] is the primary circulating vitamin D metabolite. As 25(OH)D reflects both dietary intake and sunlight-induced synthesis, it is often used as an indicator for nutritional vitamin D status [4]. Severe skeletal diseases, such as rickets, arise when 25(OH)D concentrations fall too low [3], with symptoms of skeletal maldevelopment and degradation becoming evident when serum concentrations fall below 25 nmol/L [5]. To this end, the scientific community broadly accepts < 25 nmol/L as the cutoff point for overt vitamin D deficiency [6].

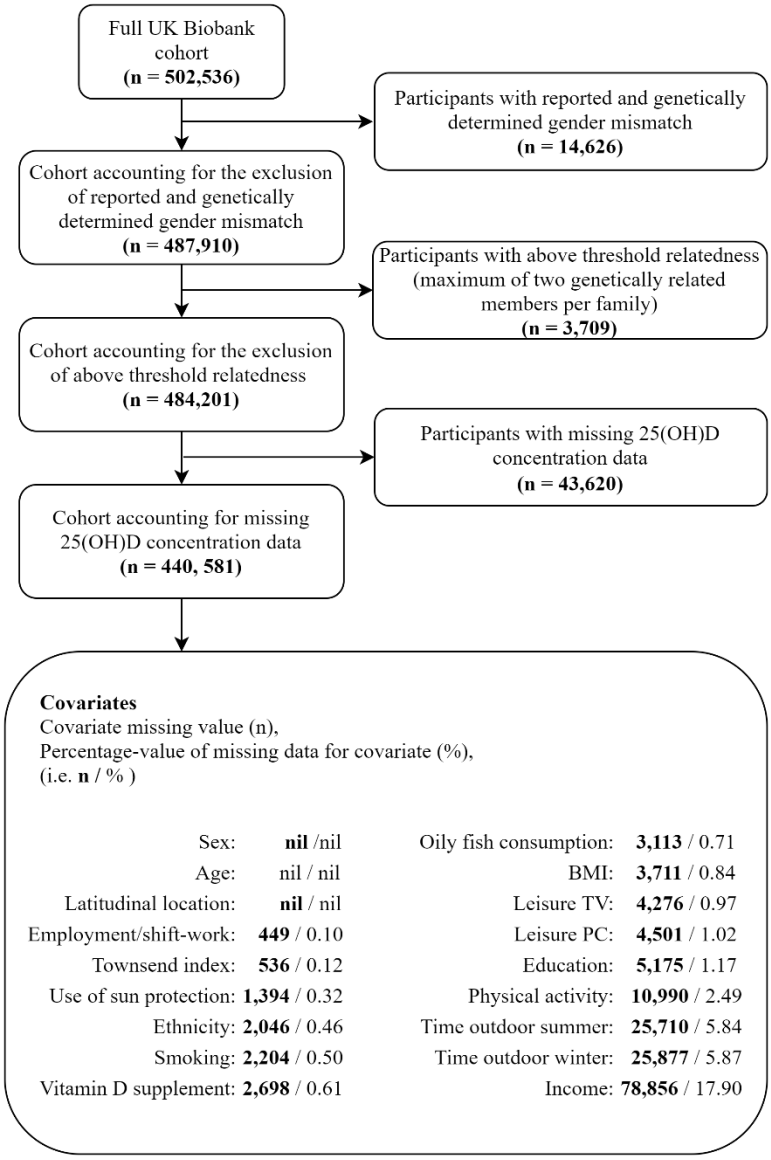
Since the National Academy of Medicine (NAM) (formerly Institute of Medicine) released their 2010 update on vitamin D dietary reference intakes, a decade of varied, and sometimes controversial commentary has ensued regarding optimal 25(OH)D serum concentrations for an increasing range of health issues [7,8,9,10]. Although debate and exploration of these matters are essential, such dialogue should not overshadow the continuing reality of severe deficiency, as evidenced by the persistence of rickets in the twenty-first century, with 2011 rickets-attributable hospitalization rates in the UK being the highest in 50 years [11].

There are long-held concerns regarding vitamin D deficiency in the UK, both at the population level and within particular at-risk groups, such as South Asians and other immigrant populations

[12,13,14,15]. In contrast to many other high-latitude Western countries, there is no mandatory vitamin D food fortification program in the UK that is designed to mitigate deficiency at a population level [16], aside from the small additions of vitamin D to margarine. Not surprisingly, recent reviews indicate that the UK has lower overall vitamin D status when compared to much of Western Europe [17], in addition to having the lowest level of adherence to infant vitamin D supplementation guidelines [18]. In this study, we examine ethnic and social distributions in the prevalence of vitamin D deficiency (25(OH)D < 25nml/L) using data from over 440,000 UK participants.

2. Methods

The UK Biobank is a large-scale prospective cohort containing an expansive collection of phenotypic and genotypic data. The study was established to provide high-quality data to the scientific community for the betterment of the public's health [19]. The UK Biobank recruited 502,316 men and women aged 40-69 years (10% response rate) [20], from within the United Kingdom (England, Scotland and Wales), between March 2006 and July 2010 [19]. Participants had to reside within 10 miles of a UK Biobank assessment centre, and were requested to take part in questionnaire surveys and physical assessments, and to provide biological samples [21]. The present analysis was limited to UK Biobank participants with data on serum 25(OH)D concentration (n = 440,581) **(Figure 1)**.



Study 1 - Figure 1: Study population and participant inclusion criteria.

2.1. Measurement of vitamin D status

Serum 25(OH)D concentration was measured using the LIAISON XL 25(OH)D assay (DiaSorin, Stillwater, USA), from blood samples (hours of fasting time: mean 3.8, SD 2.4) collected at baseline. The LIAISON XL is a fully automated chemiluminescence immunoassay device with magnetic microparticle separation technology for the direct measurement of 25(OH)D from

human serum [22]. Functional sensitivity, the concentration at which coefficient of variation (CV) exceeds 20%, was ≤ 4.0 ng/ml, as determined over multiple runs in accordance with the Clinical Laboratory and Standards Institute (CLSI) EP17-A protocols [22]. Assay precision (CLSI EP5-A2), conducted using 6 samples and 2 controls gave CV 2.3% (intra-run) and CV 7.8% (total) [22]. Due to unexpected issues with sample dilution [23], we conducted sensitivity analyses restricting to the sample which were not affected by related problems (430,413 participants, 9,960 excluded).

2.2. Covariates

Covariates were chosen for their relevance to outdoor exposure (latitude, use of sun protection, outdoor time - winter/summer, employment/shiftwork, and leisure time - TV/PC use) [2], vitamin D intake (vitamin D supplementation and oily fish consumption) [24, 25], or their independent potential to affect 25(OH)D concentrations (smoking, body mass index [BMI] and physical activity) [26,27]. All covariates were self-reported at baseline, with the exception of latitude and socioeconomic status, which were derived from assessment center location and residential data, respectively.

Socioeconomic status was assessed using the Townsend deprivation index [28]. Each participant was assigned a score corresponding to socioeconomic data attributed to their postcode; this was derived from the preceding UK census of population and housing [29]. (*Sentence edited for thesis*) Ethnicity was self-reported at baseline and the data collapsed into five primary ethnic groups, with each containing subcategorization choices, 1) White European (White, British, Irish, and any other White background), 2) Chinese, 3) (*corrected for thesis*) Asian or Asian British (South Asian options: Indian, Pakistani, Bangladeshi; and any other Asian background), 4) (*corrected for*

thesis) Black or Black British (Caribbean, African, and any other Black background), and 5) mixed ("Other" ethnic group, White and Black Caribbean, White and Black African, White and Asian, and any other mixed background). *(Added for thesis)* The 'any other background' options for White, Asian, Black, and mixed constituted 3.45%, 19.6%, 1.48%, and 35.9%, respectively. Season was determined based on month of blood sampling and determined as summer (July-Sept), autumn (Oct-Dec), winter (Jan-March) and spring (April-June).

Body mass index was assessed at baseline from height and weight measures (calculated as body weight [kilograms] divided by height squared [meters]) [30]. For our study, we divided the cohort into ethnicity-specific BMI quartiles, with data presented for lowest 25%, middle 50% and highest 25% (ethnic-specific BMI ranges: White European 12.1-24.09, 24.1-29.8, 29.8-74.7 kg/m^2, Chinese 16.3-21.6, 21.7-26.1, 26.2-45.7 kg/m^2, Asian 14.9-24.19, 24.2-29.4, 29.4-60.0 kg/m^2, Black African 16.2-25.8, 25.8-32.2, 32.2-68.1 kg/m^2, and mixed 14.6-24.1, 24.1-30.3, 30.3-56.8 kg/m^2).

2.3. Statistical methods

Serum 25(OH)D concentration was categized as <25.0, 25.0-49.9, 50.0-74.9, and ≥75.0 nmol/L. For analyses using 25(OH)D as a continuous indicator, we used natural logarithmic transformation to adjust for skewed distribution. Linear and logistic regression models were fitted to examine the associations of log 25(OH)D and risk for vitamin D deficiency with individual, dietary and lifestyle indicators. All models were weighed by 1 – kinship coefficient [31] to account for relatedness, and adjusted for age, sex, smoking, physical activity, month of measurement, assessment centre, Townsend index, sample aliquots for 25(OH)D measurement and fasting. For each dietary or lifestyle indicator, we further examined if the association pattern

with log 25(OH)D varied by seasons by introducing dietary/lifestyle-indicator-by-season interaction terms to the model. Evidence of interaction was examined using the likelihood ratio test comparing the log-likelihood of nested models with and without the interaction term(s). The same interaction analysis was performed for latitudinal locations and ethnic groups, followed by stratified analyses where relevant. Sensitivity analyses were conducted, but as this had no implications on the results, data are presented for the full sample. STATA, version 14.1 (StataCorp LP, College Station, Texas, USA) was used for all analyses.

2.4. Ethics

Ethics approval for the UK Biobank was granted by the National Information Governance Board for Health and Social Care and North West Multicenter Research Ethics Committee (11/NW/0382) [32]. The present analysis operates under UK Biobank application 20175.

3. Results

Of the 440,581 participants included in this analysis, 415,903 (94.4%) were White European, 1,383 (0.3%) Chinese, 7,880 (1.8%) Asian, 6,896 (1.6%) Black African, and 6,473 (1.5%) mixed ethnicity **(Table 1)**. The overall mean 25(OH)D concentration was 43.8 nmol/L, with 7.5% of participants in summer/autumn and 18.8% in winter/spring having 25(OH)D < 25 nmol/L. There were large differences in the average 25(OH)D concentration by ethnic group, ranging from 44.9 nmol/L in White Europeans to 24.3 nmol/L in Asian ancestry participants. Asian participants tended to consume less oily fish and use vitamin D supplements less frequently compared to other ethnicities, while intakes of both oily fish and vitamin D supplementation were highest in Black African participants (**Table 1**).

Study 1 - Table 1: Demographic characteristics of UK Biobank participants.

		25(OH)D	Vit D supplementation[a]		Oily Fish[b]		Use of sun protection[c]		Outdoor time winter[d]		Outdoor time summer[e]	
		Geometric mean	No	Yes	Never	>1 per wk	Avoid sun	Mostly	>1hr	≥3hrs	>1hr	≥3hrs
N =		440,581	355,990	81,893	47,748	79,138	2,646	245,498	82,698	96,553	18,424	267,912
	N(%)	Mean (95% CI)	%	%	%	%	%	%	%	%	%	%
All	440,581 (100)	43.8 (43.8, 43.9)	80.80	18.59	10.91	18.09	0.60	55.90	19.94	23.28	4.44	64.58
Sex												
Men	204,927 (46.5)	43.7 (43.6, 43.8)	81.56	18.44	11.22	17.66	0.53	43.96	15.65	30.19	3.63	67.94
Women	235,654 (53.5)	44.0 (43.9, 44.0)	75.58	24.42	10.65	18.46	0.67	66.26	23.76	17.15	5.17	61.58
P[g]		0.01		4.33×10^{-261}		3.81×10^{-31}		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$
Age												
<60	250,909 (57.0)	42.0 (41.9, 42.1)	78.62	21.38	13.11	14.98	0.57	57.13	23.95	19.25	5.56	56.98
≥60	189,672 (43.1)	46.4 (46.3, 46.5)	78.21	21.79	8.01	22.20	0.65	54.27	14.60	28.66	2.96	74.64
P[g]		$< 1.0 \times 10^{-300}$		0.40		$< 1.0 \times 10^{-300}$		1.21×10^{-41}		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$
Ethnicity												
White European	415,903 (94.4)	44.9 (44.9, 45.0)	78.82	21.18	10.75	17.86	0.53	57.78	19.64	23.09	4.19	64.95
Chinese	1,383 (0.3)	32.1 (31.3, 32.8)	72.50	27.50	4.34	18.69	2.41	27.21	43.04	13.52	19.51	36.15
Asian	7,880 (1.8)	24.3 (24.0, 24.5)	73.37	26.63	27.01	14.58	2.08	16.42	28.27	22.96	11.10	50.14
Black African	6,896 (1.6)	30.3 (29.9, 30.6)	66.82	33.18	5.17	30.59	2.02	14.83	21.47	35.45	6.89	67.95
mixed ethnicity	6,473 (1.5)	32.9 (32.5, 33.3)	72.03	27.97	9.61	22.39	1.47	33.01	23.69	26.08	7.25	60.42
Missing	2,046 (0.5)	36.0 (35.1, 36.8)	79.18	20.82	10.60	24.59	1.79	41.33	20.82	24.90	5.64	62.93
P[g]		$< 1.0 \times 10^{-300}$		7.18×10^{-88}		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$		1.17×10^{-120}		1.52×10^{-201}
BMI[f]												
Lowest 25%	109,184 (24.8)	46.8 (46.7, 46.9)	76.10	23.90	10.56	18.52	0.45	60.48	21.27	19.15	4.62	60.46
Middle 50%	218,453 (49.6)	45.1 (45.0, 45.2)	78.65	21.35	10.22	18.13	0.47	55.64	18.65	24.32	3.84	66.06
Highest 25%	109,233 (24.8)	39.1 (39.0, 39.3)	80.28	19.72	12.57	17.47	0.97	52.22	21.06	25.35	5.41	65.82
Missing	3,711 (0.8)	35.0 (34.4, 35.6)	78.66	21.34	13.60	22.11	2.21	42.14	25.11	24.10	7.24	61.33
P[g]		$< 1.0 \times 10^{-300}$		6.85×10^{-32}		7.25×10^{-59}		9.93×10^{-159}		1.17×10^{-15}		2.93×10^{-52}
Latitudinal location												
≤51°	155,928 (35.4)	44.8 (44.7, 44.9)	76.76	23.24	9.57	18.32	0.64	54.11	19.94	21.59	4.70	61.64
52°-53°	201,429 (45.7)	44.2 (44.1, 44.3)	79.05	20.95	11.17	18.51	0.59	56.36	20.34	23.87	4.40	65.75
54°- ≥55°	83,224 (18.9)	41.3 (41.2, 41.5)	80.05	19.95	12.82	16.63	0.56	58.12	18.98	25.04	4.04	67.33
P[g]		6.67×10^{-181}		6.11×10^{-47}		3.02×10^{-107}		1.14×10^{-109}		9.97×10^{-41}		1.91×10^{-156}

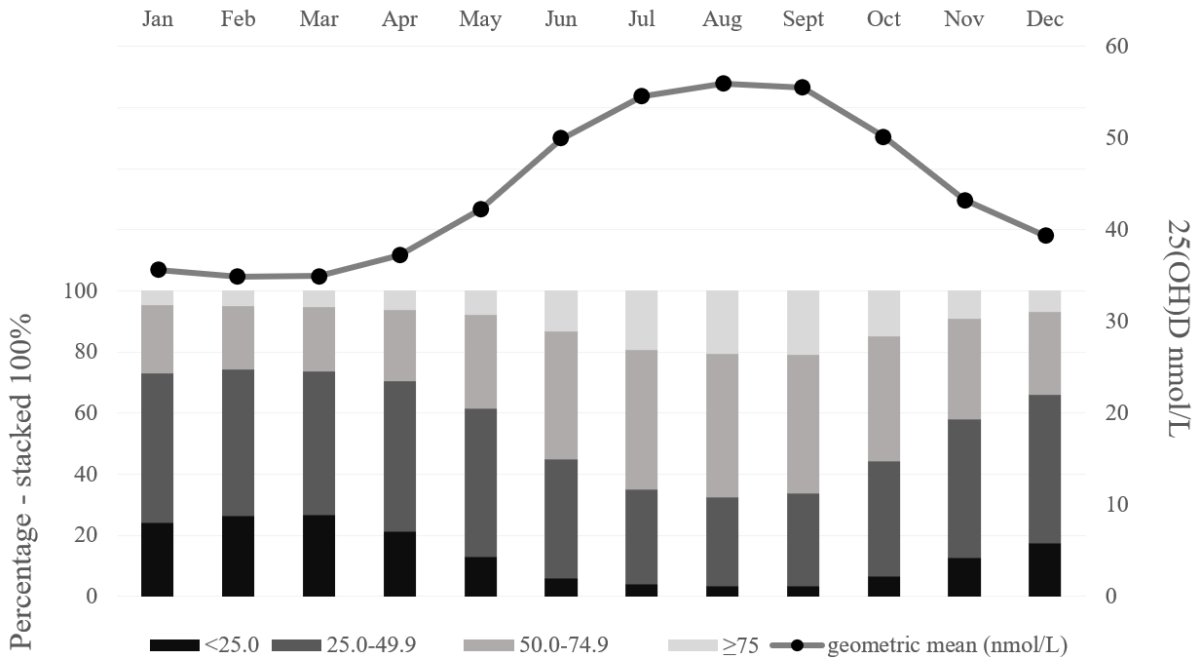
Smoking												
Non-smokers	240,002 (54.5)	44.0 (43.9, 44.1)	78.11	21.89	10.76	17.71	0.60	57.82	21.38	21.03	4.84	61.40
Ex-smokers	152,475 (34.6)	45.4 (45.3, 45.5)	78.27	21.73	9.62	19.35	0.51	55.78	18.18	25.03	3.80	67.84
Current smokers	45,900 (10.4)	38.7 (38.5, 38.9)	80.75	19.25	15.94	15.79	0.91	46.41	18.35	28.95	4.52	70.04
Missing	2,204 (0.5)	39.2 (38.3, 40.1)	79.54	20.46	14.43	20.14	0.87	50.84	17.90	34.01	4.41	73.53
P[g]		$< 1.0 \times 10^{-300}$		5.87×10^{-20}		3.18×10^{-136}		1.03×10^{-305}		3.93×10^{-169}		8.31×10^{-299}
Physical activity												
Low	132,941 (30.2)	40.4 (40.3, 40.5)	82.72	17.28	12.27	14.76	0.91	54.00	32.84	14.79	8.07	54.17
Moderate	212,201 (48.2)	44.9 (44.8, 45.0)	80.81	19.19	9.78	18.72	0.43	57.84	16.43	20.76	3.06	64.52
High	84,449 (19.2)	48.3 (48.1, 48.4)	79.82	20.18	10.68	22.18	0.34	55.35	7.17	42.25	1.30	80.97
Missing	10,990 (2.5)	36.2 (35.8, 36.5)	85.09	14.91	18.94	14.35	2.32	44.77	47.15	18.91	17.11	53.06
P[g]		$< 1.0 \times 10^{-300}$		1.48×10^{-86}		$< 1.0 \times 10^{-300}$		2.21×10^{-167}		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$
Townsend index												
Q1 ↓ deprivation	110,010 (25.0)	47.0 (46.9, 47.1)	78.43	21.57	8.38	17.76	0.33	59.83	21.11	20.36	3.86	64.40
Q2	110,006 (25.0)	46.2 (46.1, 46.4)	78.67	21.33	9.50	18.01	0.38	58.65	20.16	22.04	3.83	65.76
Q3	109,998 (25.0)	43.8 (43.7, 44.0)	78.32	21.68	11.19	17.74	0.54	56.02	19.53	23.47	4.32	64.27
Q4 ↑ deprivation	110,031 (25.0)	38.8 (38.7, 38.9)	78.37	21.63	14.2	18.86	1.16	49.05	18.92	27.41	5.79	63.89
Missing	536 (0.1)	42.7 (41.0, 44.5)	76.76	23.24	14.53	14.34	0.94	52.82	20.99	22.77	6.75	56.35
P[g]		$< 1.0 \times 10^{-300}$		0.23		2.05 x 10-67		$< 1.0 \times 10^{-300}$		$< 1.0 \times 10^{-300}$		5.56×10^{-11}

[a] Missing n= 2,698 (0.61%); [b]once per week or less n= 310,582 (70.49%), missing n= 3,113 (0.71%); [c]never/rarely/sometimes n= 191,043 (43.36%), missing n= 1,394 (0.32%); [d]1-2 hours n= 235,453 (53.44%), missing n= 25,877 (5.87%); [e]1-2 hours n= 128,535 (29.17%), missing n= 25,710 (5.84%).

[f]Constructed using ethnic specific BMI quartiles.

[g]All models were weighed by 1-kinship coefficient, and adjusted for sex, age, latitude, month and year of measurement, aliquot and fasting upon serum acquisition.

Serum 25(OH)D concentrations followed a seasonal pattern, with a trough in February-March and peak in August-September **(Figure 2).**



Study 1 - Figure 2: Distribution of categorical 25(OH)D concentration (stacked bar chart) and geometric mean of 25(OH)D concentration (line chart) by months.

Season modified the associations between all covariates and 25(OH)D concentrations **(Table 2,** all $P_{interaction} \leq 0.01$). There was a clear seasonal contrast in the prevalence of vitamin D deficiency among White Europeans, with smaller differences in other ethnic groups. From the Asian participants, more than one-half were affected throughout the year. A consistently high rate was also seen in Black African participants of whom a minimum of 30.78% were affected throughout the year. For the most affluent group, the contrast in the prevalence of vitamin D deficiency between winter and summer was larger, while the most socioeconomically deprived group showed smaller differences.

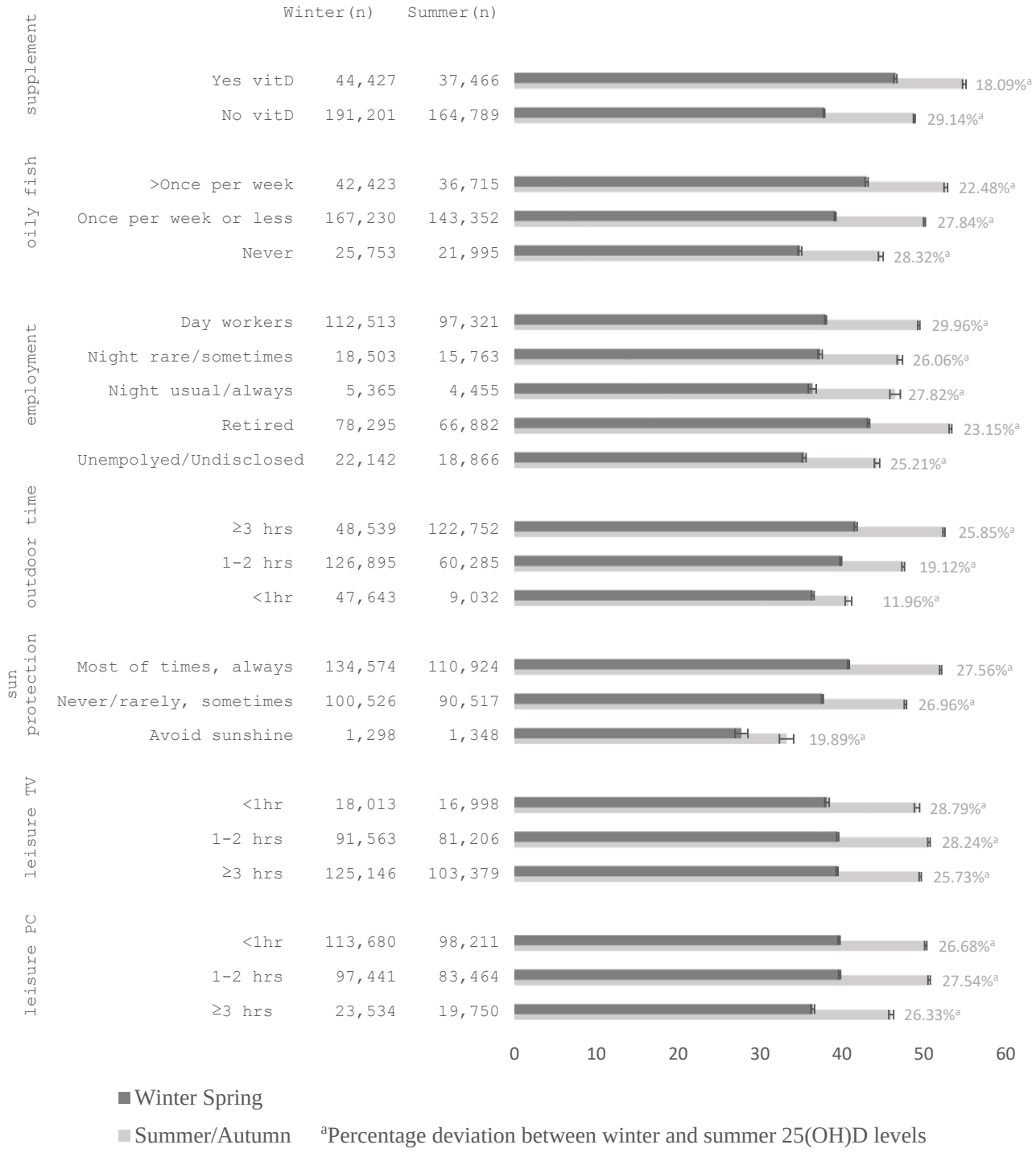
Study 1 - Table 2: 25(OH)D concentration in UK Biobank participants, by season and demographic characteristics.

		Summer and Autumn 25(OH)D Concentrations					Winter and Spring 25(OH)D Concentrations				
	Subjects N(%)	<25.0 %	25.0-49.9 %	50.0 -74.9 %	≥75.0 %	Geometric mean (95% CI)	<25.0 %	25.0-49.9 %	50.0 -74.9 %	≥75.0 %	Geometric mean (95% CI)
All	440,581 (100)	7.50	36.59	40.32	15.59	49.8 (49.7, 49.9)	18.83	46.45	27.36	7.36	39.3 (39.2, 39.4)
Sex											
Male	204,927 (46.5)	7.43	35.86	40.52	16.19	50.2 (50.1, 50.4)	19.20	47.18	26.88	6.74	38.8 (38.7, 39.0)
Female	235,654 (53.5)	7.57	37.23	40.14	15.06	49.5 (49.4, 49.6)	18.51	45.81	27.77	7.90	39.7 (39.6, 39.8)
P[b]					7.24×10^{-28}	6.28×10^{-26}				1.31×10^{-29}	3.29×10^{-33}
Age											
<60	250,909 (57.0)	8.99	38.10	38.12	14.79	48.3 (48.2, 48.5)	21.89	47.72	24.13	6.26	37.3 (37.2, 37.4)
≥60	189,672 (43.1)	5.53	34.59	43.24	16.64	51.9 (51.8, 52.1)	14.80	44.77	31.62	8.81	42.1 (42.0, 42.3)
P[b]					1.09×10^{-242}	$< 1.0 \times 10^{-300}$				$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$
Ethnicity											
White European	415,903 (94.4)	5.90	35.98	41.79	16.33	51.3 (51.2, 51.4)	17.45	46.62	28.26	7.68	40.1 (40.0, 40.2)
Chinese	1,383 (0.3)	20.68	59.13	18.58	1.62	34.6 (33.5, 35.7)	33.12	53.14	12.57	1.18	30.1 (29.2, 31.1)
Asian	7,880 (1.8)	50.80	39.93	8.19	1.08	25.2 (24.8, 25.6)	57.18	34.21	7.48	1.13	23.4 (23.0, 23.7)
Black African	6,896 (1.6)	30.78	52.41	14.14	2.67	31.8 (31.3, 32.3)	38.47	47.62	12.10	1.80	29.1 (28.7, 29.6)
mixed ethnicity	6,473 (1.5)	22.48	49.38	23.47	4.67	36.4 (35.8, 37.0)	36.53	46.25	14.81	2.42	30.2 (29.8, 30.7)
Missing	2,046 (0.5)	20.29	41.56	27.63	10.53	39.5 (38.2, 40.9)	30.07	45.86	19.14	4.94	33.3 (32.3, 34.4)
P[b]					$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$				$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$
BMI[a]											
Lowest 25%	109,184 (24.8)	6.36	31.23	41.67	20.73	53.1 (52.9, 53.3)	16.83	42.63	30.34	10.20	41.7 (41.5, 41.8)
Middle 50%	218,453 (49.6)	6.29	35.07	42.43	16.21	51.2 (51.1, 51.3)	16.69	46.55	29.05	7.71	40.5 (40.4, 40.6)
Highest 25%	109,233 (24.8)	10.73	45.22	35.02	9.03	44.5 (44.3, 44.6)	24.51	50.01	21.47	4.02	35.2 (35.1, 35.4)
Missing	3,711 (0.8)	22.46	42.60	25.91	9.03	37.9 (36.9, 38.9)	31.88	43.69	19.37	5.05	33.0 (32.2, 33.7)
P[b]					$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$				$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$
Latitudinal location											
≤51°	155,928 (35.4)	7.01	35.71	41.60	15.68	50.4 (50.3, 50.6)	17.44	47.22	28.15	7.20	39.8 (39.7, 40.0)
52°-53°	201,429 (45.7)	7.19	36.13	40.71	15.97	50.2 (50.1, 50.4)	17.50	46.01	28.60	7.88	40.2 (40.1, 40.3)
54°- ≥55°	83,224 (18.9)	9.14	39.29	37.00	14.58	47.9 (47.7, 48.1)	24.95	46.22	22.58	6.25	36.0 (35.9, 36.2)
P[b]					2.01×10^{-31}	1.48×10^{-35}				1.25×10^{-131}	3.6×10^{-162}
Townsend index											
Q1 ↓ deprivation	110,010 (25.0)	4.37	33.05	44.69	17.89	53.5 (53.3, 53.7)	14.85	45.96	30.82	8.37	41.8 (41.6, 41.9)
Q2	110,006 (25.0)	4.88	34.30	43.35	17.46	52.7 (52.5, 52.8)	15.31	46.71	29.95	8.02	41.3 (41.1, 41.4)
Q3	109,998 (25.0)	7.37	37.39	39.94	15.30	49.7 (49.5, 49.9)	18.60	47.10	27.02	7.28	39.2 (39.0, 39.4)
Q4 ↑ deprivation	110,031 (25.0)	13.82	42.01	32.79	11.39	43.7 (43.5, 43.9)	26.21	46.02	21.92	5.85	35.4 (35.2, 35.5)
Missing	536 (0.1)	10.26	38.46	34.07	17.22	47.7 (45.2, 50.4)	20.15	47.53	26.24	6.08	38.0 (35.8, 40.3)
P[b]					$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$				$< 1.0 \times 10^{-300}$	$< 1.0 \times 10^{-300}$

[a]Constructed using ethnic specific BMI quartile ranges.
[b]All models were weighed by 1-kinship coefficient, and adjusted for sex, age, latitude, month and year of measurement, aliquot and fasting upon serum acquisition.

Greater time-spent-outdoors was associated with higher 25(OH)D concentrations in summer months, with a similar but weaker trend seen in winter **(Figure 3,** all $P_{interaction} \leq 0.05$). The seasonal differences in 25(OH)D concentrations were smaller in participants who used vitamin D supplements and who ate oily fish regularly compared to others ($P_{interaction} \leq 7.04 \times 10^{-35}$ for both comparisons). There was a strong north-south gradient in the prevalence of vitamin D deficiency across the seasons **(Figure 4,** $P_{interaction}$ 7.10×10^{-32}). Nearly one-third of participants living ≥ 55 degrees latitude (Edinburgh and Glasgow assessment centers) had 25(OH)D concentrations < 25 nmol/L in winter, while 7.38% of participants also remained deficient during the summer. Lower income was associated with higher prevalence of vitamin D deficiency in all ethnic groups except for Black Africans, while educational status was associated with deficiency in Asians but not in others **(Figure 5,** all $P_{interaction} \leq 8.36 \times 10^{-7}$)

The associations between lifestyle factors and the odds of vitamin D deficiency varied by ethnic group ($P_{interaction} < 3 \times 10^{-5}$ for all comparisons, **Table 3**, **Supplementary Table 1,** appendix A). Similar patterns were observed at a 25(OH)D <50 nmol/L cutoff ($P_{interaction} < 8.34 \times 10^{-3}$ for all comparisons, **Supplementary Table 2**, appendix A), while no ethnicity-covariate interactions were seen under 75 nmol/L (data not shown). Time spent outdoors, leisure time TV and PC use were all strongly associated with the odds of deficiency in White Europeans while weaker or no associations were seen for other ethnicities. Oily fish consumption and vitamin D supplementation were associated with the odds of deficiency across all ethnicities, with the estimates for regular oily fish consumption ranging from OR = 0.29 (95% CI, 0.23-0.38) in Black Africans to OR = 0.45 (95% CI, 0.43-0.47) in White Europeans, and for supplementation, from OR = 0.36 (95% CI, 0.35-0.37) in White European to OR = 0.45 (95% CI, 0.39-0.51) in participants of Black African ancestry.



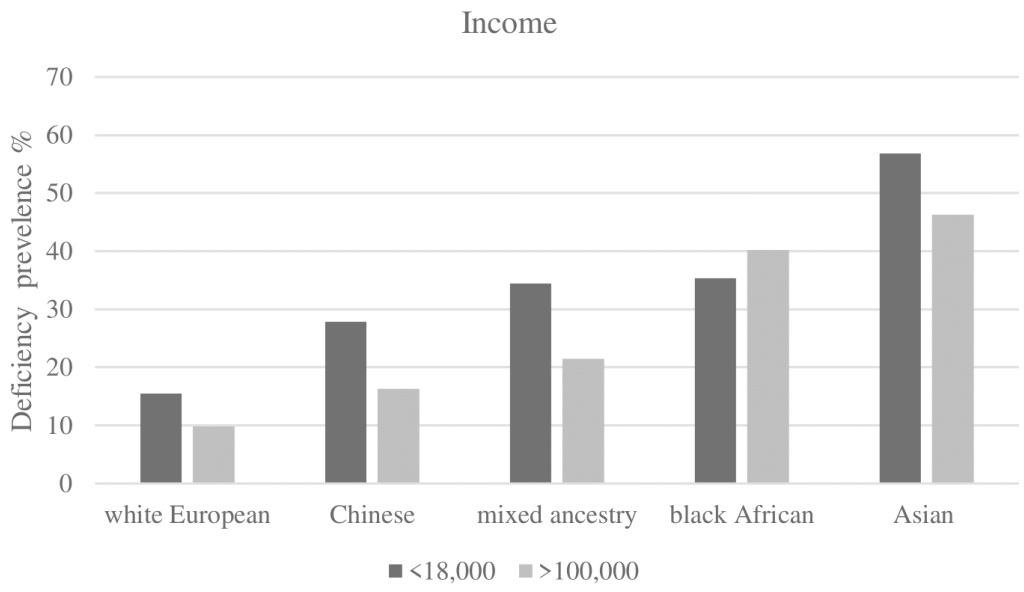
Study 1 - Figure 3: Distribution of 25(OH)D concentration by seasons and socioeconomic, dietary, or lifestyle indicators.



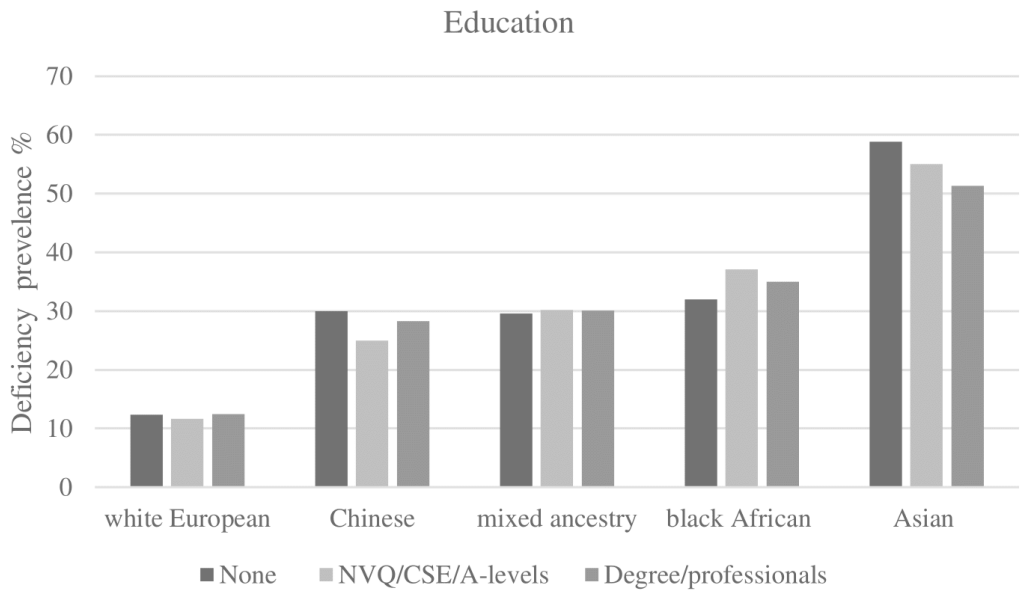
Latitudinal zone 25(OH)D serum concentration range percentages (≤51°, 52-53°, 54°, ≥55° respectively)					
<25 nmol/L	Winter	23.96, 25.01, 27.60, 33.63	<50 nmol/L	Winter	72.51, 72.79, 74.46, 79.72
	Spring	12.35, 11.92, 14.28, 24.15		Spring	58.51, 56.61, 58.82, 72.26
	Summer	03.83, 03.51, 02.81, 07.38		Summer	33.06, 33.79, 32.18, 48.35
	Autumn	10.06, 11.68, 10.88, 17.98		Autumn	51.98, 54.93, 54.55, 66.79

Study 1 - Figure 4: Distribution of categorical 25(OH)D concentration (A) <25 nmol/L, (B) <50 nmol/L, across seasons and UK latitudinal zoning.

A



B



Study 1 - Figure 5: Distribution of 25(OH)D deficiency (<25nmol/L) by ethnic groupings and income (A) and by ethnic groupings and education (B). Educational level terms: National Vocational Qualification (NVQ), Certificate of Secondary Education (CSE), and A-level

A-levels = common name for the General Certificate of Education Advanced Level certificate.

Study 1 - Table 3: Association of 25(OH)D deficiency (<25nmol/L) with selected risk factors by ethnic groups

	Subjects N(%)	White Euro[a] OR: 95% CI	Chinese[b] OR: 95% CI	Asian[c] OR: 95% CI	Black African[d] OR: 95% CI	Mixed[e] OR: 95% CI
BMI[f,o]						
Lowest 25%	109,184 (24.78)	1.02: 1.00, 1.05	0.53: 0.37, 0.75	0.70: 0.62, 0.79	1.13: 0.99, 1.28	0.85: 0.73, 0.98
Middle 50%	218,453 (49.58)	Reference	Reference	Reference	Reference	Reference
Highest 25%	109,233 (24.79)	1.67: 1.63, 1.71	1.18: 0.86, 1.60	1.26: 1.11, 1.42	1.17: 1.03, 1.34	1.35: 1.17, 1.55
P		$< 1.0 \times 10^{-300}$	2×10^{-4}	6.30×10^{-16}	0.04	1.37×10^{-7}
Latitudinal location						
≤51°	155,928 (35.39)	Reference	Reference	Reference	Reference	Reference
52°-53°	201,429 (45.72)	1.15: 1.12, 1.18	1.08: 0.82, 1.44	1.39: 1.25, 1.55	1.33: 1.18, 1.49	1.12: 0.99, 1.27
54°- ≥55°	83,224 (18.89)	1.65: 1.60, 1.70	1.19: 0.78, 1.80	1.33: 1.07, 1.66	1.64: 1.13, 2.37	1.30: 1.04, 1.63
P		1.33×10^{-286}	0.68	2.53×10^{-9}	8.26×10^{-7}	0.03
Vitamin D Supplements[g]						
No	355,990 (80.80)	Reference	Reference	Reference	Reference	Reference
Yes	81,893 (18.59)	0.36: 0.35, 0.37	0.30: 0.21, 0.44	0.40: 0.36, 0.46	0.45: 0.39, 0.51	0.52: 0.44, 0.60
P		$< 1.0 \times 10^{-300}$	9.53×10^{-10}	3.52×10^{-48}	1.84×10^{-32}	4.54×10^{-17}
Oily Fish[h]						
Never	47,748 (10.84)	Reference	Reference	Reference	Reference	Reference
1 per wk or less	310,582 (70.49)	0.67: 0.65, 0.69	0.44: 0.24, 0.79	0.71: 0.63, 0.80	0.59: 0.46, 0.75	0.81: 0.66, 0.99
>1 per wk	79,138 (17.96)	0.45: 0.43, 0.47	0.28: 0.14, 0.54	0.41: 0.35, 0.49	0.29: 0.23, 0.38	0.50: 0.39, 0.63
P		$< 1.0 \times 10^{-300}$	7×10^{-4}	3.97×10^{-25}	1.08×10^{-32}	2.47×10^{-11}
Time Outdoor Summer[i]						
<1hr	18,424 (4.18)	Reference	Reference	Reference	Reference	Reference
1-2 hrs	128,535 (33.36)	0.65: 0.62, 0.68	0.95: 0.65, 1.38	0.83: 0.7, 1.00	0.97: 0.76, 1.25	0.60: 0.47, 0.78
≥3hrs	267,912 (94.16)	0.40: 0.38, 0.42	0.58: 0.39, 0.87	0.72: 0.6, 0.86	0.89: 0.70, 1.12	0.62: 0.48, 0.79
P		$< 1.0 \times 10^{-300}$	4×10^{-3}	4×10^{-4}	0.31	2×10^{-4}
Time Outdoor Winter[j]						
<1hr	82,698 (18.77)	Reference	Reference	Reference	Reference	Reference
1-2 hrs	235,453 (53.44)	0.80: 0.78, 0.82	0.80: 0.59, 1.08	0.86: 0.76, 0.97	1.07: 0.92, 1.25	0.87: 0.74, 1.01
≥3hrs	96,553 (22.91)	0.65: 0.62, 0.67	0.62: 0.38, 1.00	0.81: 0.69, 0.94	1.03: 0.88, 1.21	0.90: 0.76, 1.08
P		2.28×10^{-132}	0.10	0.01	0.64	0.20
Leisure TV[k]						
<1hr	35,011 (7.95)	1.20: 1.16, 1.25	0.99: 0.65, 1.51	0.90: 0.75, 1.06	0.86: 0.70, 1.06	0.93: 0.76, 1.13
1-2 hrs	172,769 (39.21)	1.02: 1.00, 1.04	0.86: 0.64, 1.14	1.08: 0.97, 1.20	0.93: 0.82, 1.04	1.05: 0.93, 1.20
≥3hrs	228,525 (51.87)	Reference	Reference	Reference	Reference	Reference
P		3.74×10^{-19}	0.52	0.07	0.23	0.42
Leisure Computer[l]						
<1hr	211,891 (48.09)	0.80: 0.77, 0.83	0.83: 0.57, 1.20	1.11: 0.96, 1.30	0.92: 0.79, 1.08	0.84: 0.71, 1.00
1-2 hrs	180,905 (41.06)	0.79: 0.77, 0.82	0.70: 0.49, 0.99	0.97: 0.83, 1.13	0.95: 0.82, 1.10	0.94: 0.80, 1.11
≥3hrs	43,284 (9.82)	Reference	Reference	Reference	Reference	Reference
P		4.24×10^{-40}	0.12	0.04	0.59	0.09
Use of Sun Protection[m]						
Avoid sunshine	2,646 (0.60)	3.63: 3.23, 4.08	1.91: 0.82, 4.47	1.18: 0.81, 1.70	1.02: 0.68, 1.51	1.62: 1.00, 2.63
Never - sometimes	191,043 (43.36)	Reference	Reference	Reference	Reference	Reference
Mostly/always	245,498 (55.72)	0.76: 0.75, 0.78	0.63: 0.46, 0.86	0.88: 0.77, 1.00	0.87: 0.74, 1.01	0.84: 0.73, 0.95
P		2.77×10^{-259}	3×10^{-3}	0.10	0.19	2×10^{-3}
Smoking[n]						
Non-smokers	240,002 (54.47)	0.54: 0.52, 0.56	0.61: 0.38, 0.98	0.99: 0.83, 1.17	0.64: 0.54, 0.75	0.73: 0.62, 0.86
Ex-smokers	152,475 (34.61)	0.53: 0.51, 0.54	0.60: 0.34, 1.05	0.79: 0.64, 0.96	0.79: 0.65, 0.96	0.64: 0.53, 0.76
Current smokers	45,900 (10.42)	Reference	Reference	Reference	Reference	Reference
P		$< 1.0 \times 10^{-300}$	0.11	6×10^{-3}	6.36×10^{-8}	6.62×10^{-6}

Ethnicity N for: [a]415,903 (94.40%); [b]1,383 (0.31%); [c]7,880 (1.79%); [d]6,896 (1.57%); [e]6,473 (1.47%); Missing 2,046 (0.46%).
Missing N for: [f]3,711 (0.84%); [g]2,698 (0.61%); [h]3,113 (0.71%); [i]25,710 (5.84%): [j]25,877 (5.87%); [k]4,276 (0.97%); [l]4,501 (1.02%); [m]1,394 (0.32%); [n]2,204 (0.50%). [o]Constructed using ethnic specific BMI quartiles. All models were weighed by 1–kinship coefficient, and adjusted for age, sex, latitudinal location, 25(OH)D acquisition month, 25(OH)D aliquot, fasting upon serum acquisition, Townsend deprivation index, BMI, physical activity, smoking.

4. Discussion

The National Academy of Medicine and The Endocrine Society respectively recommend 25(OH)D concentrations should be at least 50 nmol/L [7] and 75 nmol/L [33] for maximal bone health. Our analysis showed that the average vitamin D status in the UK Biobank is below the more conservative of these recommendations, and we uncovered striking rates of vitamin D deficiency. We confirmed the strong patterns by season and latitude [34] and also showed high rates of deficiency within certain socioeconomic subgroups. In line with previous research [13, 14, 35], we report very high risks of deficiency in non-White populations. Indeed, it was alarming to find >50% of Asian participants and >30% of Black African participants to have year-round vitamin D deficiency, which is a finding warranting urgent consideration and action.

Asian participants in our study were found to engage in vitamin D promoting behaviors the least, which is congruent with previous research and commentary [36,37]. However, with UK cross-ethnic qualitative analyses demonstrating relatively greater knowledge of vitamin D in Asian populations [36], our analysis may further indicate a possible disconnect between vitamin D awareness and action in this population. Our study also suggests environmental factors (i.e. time spent outdoors) may have less impact in reducing vitamin D deficiency among Asians. These findings emphasize the need to explore the psychocultural determinants of engaging in vitamin D promoting behaviors in Asian minority groups living in the UK.

The greater prevalence of vitamin D deficiency in higher latitude locations was anticipated due to the incrementally lower UVB exposure associated with increasing latitudinal degrees [38]; however, the extent of vitamin D deficiency (one-third) in the northernmost locations of the UK over winter was alarming. Alongside latitude, there may be other explanations for the

strong north-south pattern in the prevalence of vitamin D deficiency in the UK. Northern England and Scotland have historically demonstrated higher socioeconomic deprivation [39,40], and in our analyses we observed social patterns in the prevalence of vitamin D deficiency. Our study also shows that participants from higher latitudinal locations engage in vitamin D supplementation and dietary-based behaviors to a lesser extent than people living in the southern parts of the UK, as might be expected with affordability and accessibility challenges [41]. These findings may offer some explanation for the consistently higher rates of mortality observed among UK northern populations [42]. Low 25(OH)D concentrations have been associated with increased mortality [43,44], which may in part contribute to regional health disparities between the southern and norther parts of the UK [45].

Our research highlights the need for effective public health interventions to reduce the prevalence of vitamin D deficiency in the UK, at both the population level and within at-risk groups, specifically Asian populations. At the date of this publication, the topic of vitamin D fortification in the UK is gaining unprecedented traction, with calls for the fortification of wheat-flour [46]; in fact, some fortified bread products are now available to the UK public via limited high-end supermarkets. Dietary analysis of ethnic minorities in the UK, however, has highlighted that although some South Asian groups were found to consume wheat-based staples, in way of roti and chapattis, Indians from the Gujarat region and Bangladeshis tended to primarily use millet and rice, respectively [47]. The question therefore remains as to whether fortification of any single item alone would be enough to redress abject rates of severe vitamin D deficiency across ethnicities. As others have recently indicated, fortification measures in the UK will likely need to consider specific at-risk groups and vulnerable life

stages, and proceed in conjunction with robust and accountable supplementation programs [48].

4.1. Strengths and limitations

Our study benefits from a large sample size with ethnic group diversity and the collection of comprehensive sociodemographic and lifestyle information. This level of granularity has allowed us to explore the determinants for a relatively severe form of vitamin D deficiency, defined as 25(OH)D <25 nmol/L. A primary limitation of our study is the lack of representativeness between the UK Biobank and the general UK population, with variations across a range of physical, sociodemographic, lifestyle and health-related characteristics underpinning an established 'healthy volunteer' bias [49,50]. This type of bias may explain why older participants appeared to be more active and spend more time outdoors than their general population counterparts [51] or indeed, the working-aged individuals within the UK Biobank cohort. However, possible preferential self-selection of healthy volunteers to the cohort suggests that the abject deficiency rates reported in our analysis may be even greater in magnitude within the general population. Furthermore, previous work suggests that despite the lack of representativeness with respect to disease prevalence and incidence rates, the UK Biobank can yield valid assessment of exposure-outcome relationships [49].

Additionally, although our analyses looked at ethnic differences in health behaviors and their association with vitamin D deficiency, we lacked explicit insight regarding culturally specific eating habits, which if available, could have bolstered inferences made from dietary analyses. Furthermore, the use of a single test of serum 25(OH)D concentration may raise some questions about the reliability of the vitamin D measurements, with repeat measures otherwise

being able to afford greater insight into individual changes over time. The UK Biobank's sample dilution issues [23] were not seen to pose concern for our study, with sensitivity analysis demonstrating no notable differences in the primary findings.

5. Conclusion

Vitamin D deficiency remains a major public health issue throughout the UK, particularly in lower socioeconomic areas and Asian and Black African populations. These findings justify the provision of effective and targeted public health measures aimed at reducing vitamin D deficiency at a population level, and in vulnerable subgroups.

Funding

The study was supported by the National Health and Medical Research Council (NHMRC), grant number 11123603. Joshua P Sutherland studentship is funded by Australian Research Training Program Scholarship.

Availability of data and material

This research has been conducted using the UK Biobank resource under application number 20175. All data is available to approved users upon application to the UK Biobank.

Author Contributions

EH conceived the study and designed the research question. JS wrote the first draft and analyzed the data with guidance from EH and AZ. All authors revised the paper, interpreted results, and approved the final manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgements

We thank all UK Biobank participants involved and Dr Anwar Mulugeta (University of South Australia) for support in data management.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://www.sciencedirect.com/science/article/pii/S0261561420306397?via%3Dihub

References:

1. Bikle DD. Extraskeletal actions of vitamin D. Annals of the New York Academy of Sciences. 2016;1376(1):29-52. https://doi.org/10.1111/nyas.13219.

2. Nair R, Maseeh A. Vitamin D: The "sunshine" vitamin. Journal of pharmacology pharmacotherapeutics. 2012;3(2):118. DOI:10.4103/0976-500X.95506.

3. Wheeler BJ, Snoddy AME, Munns C, Simm P, Siafarikas A, Jefferies C. A Brief History of Nutritional Rickets. Frontiers in Endocrinology. 2019;10. https://doi.org/10.3389/fendo.2019.00795.

4. Mula-Abed W-A. 25- Hydroxyvitamin D: Explosion in Clinical Interest and Laboratory Requests. Oman Medical Journal. 2009. https://doi.org/10.5001/omj.2009.49.

5. Cashman KD. Vitamin D in childhood and adolescence. Postgraduate Medical Journal. 2007;83(978):230-5. https://dx.doi.org/10.1136%2Fpgmj.2006.052787.

6. Pilz S, Zittermann A, Trummer C, Theiler-Schwetz V, Lerchbaum E, Keppel MH, et al. Vitamin D testing and treatment: a narrative review of current evidence. Endocrine connections. 2019;8(2):R27-R43. https://dx.doi.org/10.1530%2FEC-18-0432.

7. Heaney RP, Holick MF. Why the IOM recommendations for vitamin D are deficient. Journal of Bone and Mineral Research. 2011;26(3):455-7. https://doi.org/10.1002/jbmr.328**.**

8. Minisola S, Ferrone F, Danese V, Cecchetti V, Pepe J, Cipriani C, et al. Controversies Surrounding Vitamin D: Focus on Supplementation and Cancer. International Journal of Environmental Research and Public Health. 2019;16(2):189. https://dx.doi.org/10.3390%2Fijerph16020189.

9. Tanner SB, Harwell SA. More than healthy bones: a review of vitamin D in muscle health. Therapeutic Advances in Musculoskeletal Disease. 2015;7(4):152-9. https://dx.doi.org/10.1177%2F1759720X15588521.

10. Pfotenhauer KM, Shubrook JH. Vitamin D Deficiency, Its Role in Health and Disease, and Current Supplementation Recommendations. The Journal of the American Osteopathic Association. 2017;117(5):301. https://doi.org/10.7556/jaoa.2017.055.

11. Goldacre M, Hall N, Yeates DGR. Hospitalisation for children with rickets in England: a historical perspective. The Lancet. 2014;383(9917):597-8. http://doi.org/10.1016/s0140-6736(14)60211-7.

12. Hyppönen E, Power C. Hypovitaminosis D in British adults at age 45 y: nationwide cohort study of dietary and lifestyle predictors. The American Journal of Clinical Nutrition. 2007;85(3):860-8. https://doi.org/10.1093/ajcn/85.3.860.

13. Lowe NM, Bhojani I. Special considerations for vitamin D in the south Asian population in the UK. Therapeutic Advances in Musculoskeletal Disease. 2017. https://doi.org/10.1177%2F1759720X17704430.

14. Preece M, McIntosh W, Tomlinson S, Ford J, Dunnigan M, O'Riordan J. Vitamin-D deficiency among Asian immigrants to Britain. The Lancet. 1973;301(7809):907-10. https://doi.org/10.1016/S0140-6736(73)91361-5.

15. Bivins R. Ideology and disease identity: the politics of rickets, 1929-1982. Medical Humanities. 2014;40(1):3-10. https://dx.doi.org/10.1136%2Fmedhum-2013-010400.

16. Itkonen S, Erkkola M, Lamberg-Allardt C. Vitamin D Fortification of Fluid Milk Products and Their Contribution to Vitamin D Intake and Vitamin D Status in Observational Studies—A Review. Nutrients. 2018;10(8):1054. https://dx.doi.org/10.3390%2Fnu10081054.

17. Lips P, Cashman KD, Lamberg-Allardt C, Bischoff-Ferrari HA, Obermayer-Pietsch B, Bianchi ML, et al. Current vitamin D status in European and Middle East countries and strategies to prevent vitamin D deficiency: a position statement of the European Calcified Tissue Society. European Journal of Endocrinology. 2019:P23-P54. https://doi.org/10.1530/EJE-18-0736.

18. Uday S, Kongjonaj A, Aguiar M, Tulchinsky T, Högler W. Variations in infant and childhood vitamin D supplementation programmes across Europe and factors influencing adherence. Endocrine connections. 2017;6(8):667-75. https://dx.doi.org/10.1530%2FEC-17-0193.

19. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine. 2015;12(3). https://doi.org/10.1371/journal.pmed.1001779.

20. Watts G. UK Biobank gets 10% response rate as it starts recruiting volunteers. BMJ. 2007;334(7595):659. https://dx.doi.org/10.1136%2Fbmj.39167.407616.DB.

21. UK Biobank. UK Biobank: Protocol for a large-scale prospective epidemilogical resource - Protocol No: UKBB-PROT-09-06. 2007. https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.

22. DiaSorin. LIAISON - 25 OH Vitamin D TOTAL Assay 2019. https://www.diasorin.com/sites/default/files/allegati_prodotti/ese_brochure_25_oh_vit._d_total_low.pdf.

23. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications). 2019. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf.

24. Macdonald HM, Wood AD, Tang JC, Fraser WD. Comparison of vitamin D2 and vitamin D3 supplementation in increasing serum 25-hydroxyvitamin D status: a systematic review and meta-analysis. The American journal of clinical nutrition. 2012;96(5):1152-3. https://doi.org/10.3945/ajcn.112.046110.

25. L ehmann U, Gjessing HR, Hirche F, Mueller-Belecke A, Gudbrandsen OA, Ueland PM, et al. Efficacy of fish intake on vitamin D status: a meta-analysis of randomized controlled trials. The American journla of clinical nutrition. 2015;102(4):837-47. https://doi.org/10.3945/ajcn.114.105395.

26. N wosu BU, Kum-Nji P. Tobacco smoke exposure is an independent predictor of vitamin D deficiency in US children. PLOS ONE. 2018;13(10). https://dx.doi.org/10.1371%2Fjournal.pone.0205342.

27. Bertrand KA, Giovannucci E, Liu Y, Malspeis S, Eliassen AH, Wu K, et al. Determinants of plasma 25-hydroxyvitamin D and development of prediction models in three US cohorts. British Journal of Nutrition. 2012;108(10):1889-96. https://dx.doi.org/10.1017%2FS0007114511007409.

28. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North: Routledge; 1988. https://doi.org/10.1016/0168-8510(88)90006-1.

29. Vickers D, Rees P. Creating the UK National Statistics 2001 output area classification. Journal of the Royal Statistical Society. 2007;170(2):379-403. https://doi.org/10.1111/j.1467-985X.2007.00466.x.

30. Nuttall FQ. Body Mass Index. Nutrition Today. 2015;50(3):117-28. https://dx.doi.org/10.1097%2FNT.0000000000000092.

31. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867-73. https://doi.org/10.1093/bioinformatics/btq559.

32. UK Biobank. The Ethics and Governance Council 2019. https://www.ukbiobank.ac.uk/the-ethics-and-governance-council/.

33. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, Treatment, and Prevention of Vitamin D Deficiency: an Endocrine Society Clinical Practice Guideline. The Journal of Clinical Endocrinology & Metabolism. 2011;96(7):1911-30. https://doi.org/10.1210/jc.2011-0385.

34. Kroll MH, Bi C, Garber CC, Kaufman HW, Liu D, Caston-Balderrama A, et al. Temporal Relationship between Vitamin D Status and Parathyroid Hormone in the United States. PLOS ONE. 2015;10(3):e0118108. https://doi.org/10.1371/journal.pone.0118108.

35. Shaw NJ. Vitamin D deficiency in UK Asian families: activating a new concern. Archives of Disease in Childhood. 2002;86(3):147-9. http://dx.doi.org/10.1136/adc.86.3.147.

36. Webb A, Aseem S, Kift R, Rhodes L, Farrar M. Target the message: A qualitative study exploring knowledge and cultural attitudes to sunlight and vitamin D in Greater Manchester, UK. British Journal of Dermatology. 2016;175(6):1401-3. https://doi.org/10.1111/bjd.14800.

37. Kift R, Berry JL, Vail A, Durkin MT, Rhodes LE, Webb AR. Lifestyle factors including less cutaneous sun exposure contribute to starkly lower vitamin D levels in U.K. South Asians compared with the white population. British Journal of Dermatology. 2013;169(6):1272-8. https://doi.org/10.1111/bjd.12518.

38. Engelsen O, Brustad M, Aksnes L, Lund E. Daily Duration of Vitamin D Synthesis in Human Skin with Relation to Latitude, Total Ozone, Altitude, Ground Cover, Aerosols and Cloud Thickness. Photochemistry and Photobiology. 2005;81(6):1287. https://doi.org/10.1562/2004-11-19-rn-375.

39. Addison M, Kaner E, Johnstone P, Hillier-Brown F, Moffatt S, Russell S, et al. Equal North: how can we reduce health inequalities in the North of England? A prioritization exercise with researchers, policymakers and practitioners. Journal of Public Health. 2019;41(4):652-64. https://dx.doi.org/10.1093%2Fpubmed%2Ffdy170.

40. Yousaf S, Bonsall A. UK Townsend Deprivation Scores from 2011 Census. Economic and Social Research Council (ESRC), UK Data Service; 2017. http://s3-eu-west-1.amazonaws.com/statistics.digitalresources.jisc.ac.uk/dkan/files/Townsend_Deprivation_Scores/UK%20Townsend%20Deprivation%20Scores%20from%202011%20census%20data.pdf.

41. Corscadden L, Levesque JF, Lewis V, Strumpf E, Breton M, Russell G. Factors associated with multiple barriers to access to primary care: an international analysis. International Journal for Equity in Health. 2018;17(1). https://doi.org/10.1186/s12939-018-0740-1.

42. Schofield L, Walsh D, Munoz-Arroyo R, McCartney G, Buchanan D, Lawder R, et al. Dying younger in Scotland: Trends in mortality and deprivation relative to England and Wales, 1981–2011. Health and Place. 2016;40:106-15. https://doi.org/10.1016/j.healthplace.2016.05.007.

43. Melamed ML, Michos ED, Post W, Astor B. 25-Hydroxyvitamin D Levels and the Risk of Mortality in the General Population. Archives of Internal Medicine. 2008;168(15):1629-37. https://doi.org/10.1001/archinte.168.15.1629.

44. Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. Bmj. 2014;349:g6330. https://doi.org/10.1136/bmj.g6330.

45. Bambra C, Barr B, Milne E. North and South: addressing the English health divide. Journal of Public Health. 2014;36(2):183-6. https://doi.org/10.1093/pubmed/fdu029.

46. Aguiar M, Andronis L, Pallan M, Högler W, Frew E. The economic case for prevention of population vitamin D deficiency: a modelling study using data from England and Wales. European Journal of Clinical Nutrition. 2019. https://doi.org/10.1038/s41430-019-0486-x.

47. Leung G, Stanner S. Diets of minority ethnic groups in the UK: influence on chronic disease risk and implications for prevention. Nutrition Bulletin. 2011;36(2):161-98. https://doi.org/10.1111/j.1467-3010.2011.01889.x.

48. Uday S, Fratzl-Zelman N, Roschger P, Klaushofer K, Chikermane A, Saraff V, et al. Cardiac, bone and growth plate manifestations in hypocalcemic infants: revealing the hidden body of the vitamin D deficiency iceberg. BMC Pediatrics. 2018;18(1). https://doi.org/10.1186/s12887-018-1159-y.

49. UK Biobank. Access matter: representativeness of the UK Biobank resource; 2017. http://www.ukbiobank.ac.uk/.

50. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American Journal of Epidemiology. 2017;186(9):1026-34. https://doi.org/10.1093/aje/kwx246.

51. Stamatakis E, Ekelund U, Wareham NJ. Temporal trends in physical activity in England: The Health Survey for England 1991 to 2004. Preventive Medicine. 2007;45(6):416-23. https://doi.org/10.1016/j.ypmed.2006.12.014.

Supplementary Table 1. Association of 25(OH)D deficiency (<25nmol/L) with selected risk factors by ethnic groups (with ethnic specific prevalence and case/number ratios).

		White Euro[a]		Chinese[b]		Asian[c]		Black African[d]		Mixed[e]	
	Subjects N(%)	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI
BMI[f,o]											
Lowest 25%	109,184 (24.78)	10.41 (10672 / 102496)	1.02: 1.00, 1.05	18.62 (62 / 333)	0.53: 0.37, 0.75	45.03 (842 / 1870)	0.70: 0.62, 0.79	37.08 (600 / 1618)	1.13: 0.99, 1.28	25.18 (394 / 1565)	0.85: 0.73, 0.98
Middle 50%	218,453 (49.58)	10.26 (20743 / 202108)	Reference	28.85 (189 / 655)	Reference	54.1 (1972 / 3645)	Reference	33.20 (1080 / 3253)	Reference	28.76 (881 / 3063)	Reference
Highest 25%	109,233 (24.79)	16.52 (16333 / 98841)	1.67: 1.63, 1.71	30.97 (105 / 339)	1.18: 0.86, 1.60	59.27 (1020 / 1721)	1.26: 1.11, 1.42	35.64 (572 / 1605)	1.17: 1.03, 1.34	35.8 (537 / 1500)	1.35: 1.17, 1.55
P			$< 1.0 \times 10^{-300}$		2×10^{-4}		6.30×10^{-16}		0.04		1.37×10^{-7}
Latitudinal location											
≤51°	155,928 (35.39)	9.67 (13354 / 138110)	Reference	24.82 (171 / 689)	Reference	48.2 (2062 / 4278)	Reference	31.87 (1271 / 3988)	Reference	28.08 (1004 / 3576)	Reference
52°-53°	201,429 (45.72)	11.43 (21259 / 186033)	1.15: 1.12, 1.18	27.24 (134 / 492)	1.08: 0.82, 1.44	60.01 (1535 / 2558)	1.39: 1.25, 1.55	39.09 (921 / 2356)	1.33: 1.18, 1.49	31.1 (655 / 2106)	1.12: 0.99, 1.27
54°- ≥55°	83,224 (18.89)	16.56 (13135 / 79302)	1.65: 1.60, 1.70	34.93 (51 / 146)	1.19: 0.78, 1.80	59.25 (237 / 400)	1.33: 1.07, 1.66	45.45 (60 / 132)	1.64: 1.13, 2.37	34.3 (153 / 446)	1.30: 1.04, 1.63
P			1.33×10^{-286}		0.68		2.53×10^{-9}		8.26×10^{-7}		0.03
Vitamin D Supplements[g]											
No	355,990 (80.80)	13.25 (43405 / 327552)	Reference	31.07 (308 / 991)	Reference	58 (3191 / 5501)	Reference	39.48 (1823 / 4618)	Reference	32.48 (1513 / 4658)	Reference
Yes	81,893 (18.59)	5.55 (4134 / 74412)	0.36: 0.35, 0.37	13.46 (42 / 312)	0.30: 0.21, 0.44	35.25 (562 / 1594)	0.40: 0.36, 0.46	22.1 (388 / 1756)	0.45: 0.39, 0.51	19.84 (278 / 1401)	0.52: 0.44, 0.60
P			$< 1.0 \times 10^{-300}$		9.53×10^{-10}		3.52×10^{-48}		1.84×10^{-32}		4.54×10^{-17}
Oily Fish[h]											
Never	47,748 (10.84)	18.51 (7850 / 42403)	Reference	45.28 (24 / 53)	Reference	58.35 (1080 / 1851)	Reference	52.52 (167 / 318)	Reference	35.63 (202 / 567)	Reference
1 per wk or less	310,582 (70.49)	11.85 (34037 / 287260)	0.67: 0.65, 0.69	27.54 (279 / 1013)	0.44: 0.24, 0.79	53.41 (2229 / 4173)	0.71: 0.63, 0.80	38.92 (1582 / 4065)	0.59: 0.46, 0.75	31 (1273 / 4107)	0.81: 0.66, 0.99
>1 per wk	79,138 (17.96)	7.73 (5574 / 72096)	0.45: 0.43, 0.47	18.67 (45 / 241)	0.28: 0.14, 0.54	40.12 (416 / 1037)	0.41: 0.35, 0.49	22.77 (441 / 1937)	0.29: 0.23, 0.38	21.48 (290 / 1350)	0.50: 0.39, 0.63
P			$< 1.0 \times 10^{-300}$		7×10^{-4}		3.97×10^{-25}		1.08×10^{-32}		2.47×10^{-11}
Time Outdoor Summer[i]											
<1hr	18,424 (4.18)	20.21 (3046 / 15072)	Reference	31.00 (71 / 229)	Reference	57.54 (393 / 683)	Reference	36.68 (139 / 379)	Reference	36.95 (143 / 387)	Reference
1-2 hrs	128,535 (33.36)	14.12 (16701 / 118278)	0.65: 0.62, 0.68	28.28 (153 / 541)	0.95: 0.65, 1.38	53.53 (1401 / 2617)	0.83: 0.7, 1.00	35.72 (513 / 1436)	0.97: 0.76, 1.25	27.38 (497 / 1815)	0.60: 0.47, 0.78
≥3hrs	267,912 (94.16)	9.82 (24503 / 249448)	0.40: 0.38, 0.42	21.00 (92 / 438)	0.58: 0.39, 0.87	50.64 (1706 / 3369)	0.72: 0.6, 0.86	33.95 (1311 / 3862)	0.89: 0.70, 1.12	29.63 (1011 / 3412)	0.62: 0.48, 0.79
P			$< 1.0 \times 10^{-300}$		4×10^{-3}		0.0004		0.31		2×10^{-4}
Time Outdoor Winter[j]											
<1hr	82,698 (18.77)	15.87 (11582 / 72990)	Reference	30.41 (156 / 513)	Reference	56.35 (1011 / 1794)	Reference	35.05 (415 / 1184)	Reference	31.96 (411 / 1286)	Reference
1-2 hrs	235,453 (53.44)	11.33 (25046 / 221153)	0.80: 0.78, 0.82	24.29 (128 / 527)	0.80: 0.59, 1.08	50.98 (1664 / 3264)	0.86: 0.76, 0.97	34.95 (858 / 2455)	1.07: 0.92, 1.25	27.93 (794 / 2843)	0.87: 0.74, 1.01
≥3hrs	96,553 (22.91)	8.99 (7968 / 88597)	0.65: 0.62, 0.67	19.51 (32 / 164)	0.62: 0.38, 1.00	50.69 (770 / 1519)	0.81: 0.69, 0.94	33.67 (672 / 1996)	1.03: 0.88, 1.21	29.84 (436 / 1461)	0.90: 0.76, 1.08
P			2.28×10^{-132}		0.10		0.01		0.64		0.20
Leisure TV[k]											
<1hr	35,011 (7.95)	12.22 (3880 / 31751)	1.20: 1.16, 1.25	29.83 (54 / 181)	0.99: 0.65, 1.51	49.86 (365 / 732)	0.90: 0.75, 1.06	32.33 (172 / 532)	0.86: 0.70, 1.06	25.87 (193 / 746)	0.93: 0.76, 1.13
1-2 hrs	172,769 (39.21)	11.07 (17667 / 159568)	1.02: 1.00, 1.04	25.33 (152 / 600)	0.86: 0.64, 1.14	54.09 (1823 / 3370)	1.08: 0.97, 1.20	33.6 (759 / 2259)	0.93: 0.82, 1.04	29.3 (733 / 2502)	1.05: 0.93, 1.20
≥3hrs	228,525 (51.87)	12.28 (25756 / 209662)	Reference	27.62 (145 / 525)	Reference	52.31 (1573 / 3007)	Reference	35.91 (1255 / 3495)	Reference	30.66 (845 / 2756)	Reference
P			3.74×10^{-19}		0.52		0.07		0.23		0.42
Leisure Computer[l]											
<1hr	211,891 (48.09)	11.76 (22970 / 195394)	0.80: 0.77, 0.83	27.13 (124 / 457)	0.83: 0.57, 1.20	53.37 (1773 / 3322)	1.11: 0.96, 1.30	32.17 (838 / 2605)	0.92: 0.79, 1.08	27.57 (663 / 2405)	0.84: 0.71, 1.00
1-2 hrs	180,905 (41.06)	11.20 (18787 / 167792)	0.79: 0.77, 0.82	24.41 (145 / 594)	0.70: 0.49, 0.99	51.54 (1338 / 2596)	0.97: 0.83, 1.13	36.6 (882 / 2410)	0.95: 0.82, 1.10	30.01 (761 / 2536)	0.94: 0.80, 1.11
≥3hrs	43,284 (9.82)	14.91 (5630 / 37766)	Reference	33.2 (82 / 247)	Reference	52.62 (513 / 975)	Reference	38.18 (470 / 1231)	Reference	32.82 (343 / 1045)	Reference
P			4.24×10^{-40}		0.12		0.04		0.59		0.09
Use of Sun Protection[m]											
Avoid sunshine	2,646 (0.60)	34.04 (672 / 1974)	3.63: 3.23, 4.08	38.71 (12 / 31)	1.91: 0.82, 4.47	55.4 (77 / 139)	1.18: 0.81, 1.70	33.06 (40 / 121)	1.02: 0.68, 1.51	45.24 (38 / 84)	1.62: 1.00, 2.63
Never - sometimes	191,043 (43.36)	12.95 (21650 / 167220)	Reference	28.48 (262 / 920)	Reference	53.58 (3076 / 5741)	Reference	35.23 (1862 / 5285)	Reference	31.03 (1223 / 3941)	Reference
Mostly/always	245,498 (55.72)	10.84 (25371 / 234070)	0.76: 0.75, 0.78	20.92 (77 / 368)	0.63: 0.46, 0.86	48.27 (572 / 1185)	0.88: 0.77, 1.00	32.63 (309 / 947)	0.87: 0.74, 1.01	25.54 (519 / 2032)	0.84: 0.73, 0.95
P			2.77×10^{-259}		3×10^{-3}		0.10		0.19		2×10^{-3}
Smoking[n]											
Non-smokers	240,002 (54.47)	11.06 (24192 / 218689)	0.54: 0.52, 0.56	25.98 (272 / 1047)	0.61: 0.38, 0.98	52.98 (2920 / 5512)	0.99: 0.83, 1.17	32.1 (1462 / 4554)	0.64: 0.54, 0.75	29.36 (1004 / 3420)	0.73: 0.62, 0.86
Ex-smokers	152,475 (34.61)	10.78 (15491 / 143744)	0.53: 0.51, 0.54	26.23 (48 / 183)	0.60: 0.34, 1.05	48.98 (502 / 1025)	0.79: 0.64, 0.96	37.4 (426 / 1139)	0.79: 0.65, 0.96	25.78 (457 / 1773)	0.64: 0.53, 0.76
Current smokers	45,900 (10.42)	19.66 (8065 / 41012)	Reference	37.11 (36 / 97)	Reference	58.94 (412 / 699)	Reference	46.49 (364 / 783)	Reference	37.54 (351 / 935)	Reference
P			$< 1.0 \times 10^{-300}$		0.11		6×10^{-3}		6.36×10^{-8}		6.62×10^{-6}

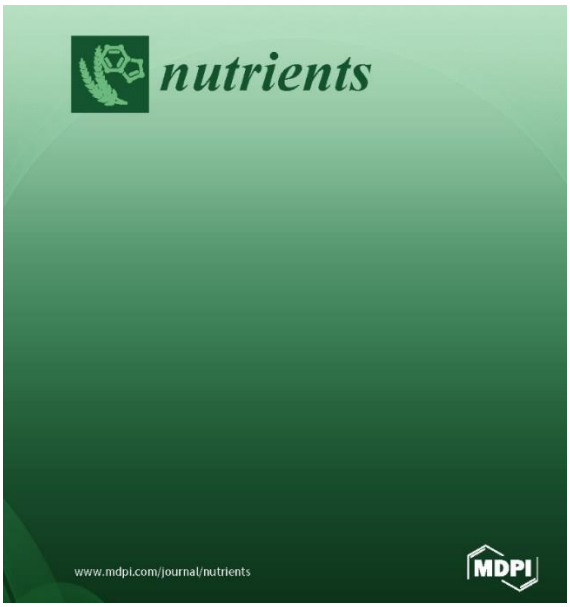
Ethnicity N for: [a]415,903 (94.40%); [b]1,383 (0.31%); [c]7,880 (1.79%); [d]6,896 (1.57%); [e]6,473 (1.47%); Missing 2,046 (0.46%).
Missing N for: [f]3,711 (0.84%); [g]2,698 (0.61%); [h]3,113 (0.71%); [i]25,710 (5.84%): [j]25,877 (5.87%); [k]4,276 (0.97%); [l]4,501 (1.02%); [m]1,394 (0.32%); [n]2,204 (0.50%).
[o]Constructed using ethnic specific BMI quartiles
All models were weighed by 1–kinship coefficient, and adjusted for age, sex, latitudinal location, 25(OH)D acquisition month, 25(OH)D aliquot, fasting upon serum acquisition, Townsend deprivation index, BMI, physical activity, smoking.

Supplementary Table 2. Association of 25(OH)D deficiency (<50nmol/L) with selected risk factors by ethnic groups (with ethnic specific prevalence and case/number ratios).

		White Euro[a]		Chinese[b]		Asian[c]		Black African[d]		Mixed[e]	
	Subjects N(%)	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI	Prevalence % (cases / n)	OR: 95% CI
BMI[f,o]											
Lowest 25%	109,184 (24.78)	10.42 (10610 / 101810)	0.83: 0.82, 0.85	19.10 (64 / 335)	0.64: 0.45, 0.91	44.90 (827 / 1842)	0.57: 0.47, 0.68	37.19 (604 / 1624)	0.82: 0.70, 0.97	25.11 (389 / 1549)	0.72: 0.62, 0.84
Middle 50%	218,453 (49.58)	10.26 (20784 / 202617)	Reference	28.72 (189 / 658)	Reference	53.92 (1954 / 3624)	Reference	33.14 (1076 / 3247)	Reference	28.78 (885 / 3075)	Reference
Highest 25%	109,233 (24.79)	16.52 (16354 / 99018)	1.73: 1.70, 1.76	30.84 (103 / 334)	1.26: 0.84, 1.89	59.49 (1053 / 1770)	1.71: 1.34, 2.17	35.64 (572 / 1605)	1.43: 1.19, 1.72	35.77 (538 / 1504)	1.41: 1.19, 1.67
P			< 1.0 x 10^{-300}		0.01		1.47 x 10^{-18}		5.55 x 10^{-7}		5.55 x 10^{-7}
Latitudinal location											
≤51°	155,928 (35.39)	9.67 (13354 / 138110)	Reference	24.82 (171 / 689)	Reference	48.20 (2062 / 4278)	Reference	31.87 (1271 / 3988)	Reference	28.08 (1004 / 3576)	Reference
52°-53°	201,429 (45.72)	11.43 (21259 / 186033)	1.06: 1.05, 1.08	27.24 (134 / 492)	1.11: 0.80, 1.56	60.01 (1535 / 2558)	1.36: 1.12, 1.65	39.09 (921 / 2356)	1.35: 1.16, 1.57	31.1 (655 / 2106)	1.07: 0.93, 1.23
54°- ≥55°	83,224 (18.89)	16.56 (13135 / 79302)	1.31: 1.29, 1.34	34.93 (51 / 146)	1.32: 0.75, 2.35	59.25 (237 / 400)	1.26: 0.84, 1.89	45.45 (60 / 132)	1.33: 0.77, 2.29	34.3 (153 / 446)	1.3: 1.00, 1.69
P			2.15 x 10-174		0.57		0.01		5.52 x 10-4		0.13
Vitamin D Supplements[g]											
No	355,990 (80.80)	13.25 (43405 / 327552)	Reference	31.08 (308 / 991)	Reference	58.01 (3191 / 5501)	Reference	39.48 (1823 / 4618)	Reference	32.48 (1513 / 4658)	Reference
Yes	81,893 (18.59)	5.56 (4134 / 74412)	0.50: 0.49, 0.5	13.46 (42 / 312)	0.32: 0.23, 0.45	35.26 (562 / 1594)	0.51: 0.43, 0.60	22.10 (388 / 1756)	0.47: 0.40, 0.54	19.84 (278 / 1401)	0.61: 0.53, 0.70
P			< 1.0 x 10-300		2.14 x 10-11		3.37 x 10-14		1.58 x 10-24		1.56 x 10-11
Oily Fish[h]											
Never	47,748 (10.84)	18.51 (7850 / 42403)	Reference	45.28 (24 / 53)	Reference	58.35 (1080 / 1851)	Reference	52.52 (167 / 318)	Reference	35.63 (202 / 567)	Reference
1 per wk or less	310,582 (70.49)	11.85 (34037 / 287260)	0.82: 0.80, 0.84	27.54 (279 / 1013)	0.50: 0.19, 1.35	53.41 (2229 / 4173)	1.00: 0.82, 1.21	38.92 (1582 / 4065)	0.70: 0.47, 1.04	31.00 (1273 / 4107)	0.91: 0.72, 1.15
>1 per wk	79,138 (17.96)	7.73 (5574 / 72096)	0.68: 0.66, 0.70	18.67 (45 / 241)	0.35: 0.12, 0.96	40.12 (416 / 1037)	0.70: 0.54, 0.91	22.77 (441 / 1937)	0.40: 0.26, 0.59	21.48 (290 / 1350)	0.64: 0.49, 0.82
P			1.1 x 10-183		0.04		0.01		4.09 x 10-15		3.98 x 10-6
Time Outdoor Summer[i]											
<1hr	18,424 (4.18)	20.21 (3046 / 15072)	Reference	31.00 (71 / 229)	Reference	57.54 (393 / 683)	Reference	36.68 (139 / 379)	Reference	36.95 (143 / 387)	Reference
1-2 hrs	128,535 (33.36)	14.12 (16701 / 118278)	0.67: 0.65, 0.70	28.28 (153 / 541)	0.89: 0.55, 1.44	53.53 (1401 / 2617)	0.92: 0.67, 1.26	35.72 (513 / 1436)	1.05: 0.76, 1.45	27.38 (497 / 1815)	0.77: 0.57, 1.04
≥3hrs	267,912 (94.16)	9.82 (24503 / 249448)	0.43: 0.41, 0.44	21.00 (92 / 438)	0.60: 0.37, 0.99	50.64 (1706 / 3369)	0.82: 0.60, 1.12	33.95 (1311 / 3862)	1.09: 0.8, 1.47	29.63 (1011 / 3412)	0.71: 0.53, 0.94
P			< 1.0 x 10-300		0.05		0.31		0.83		0.05
Time Outdoor Winter[j]											
<1hr	82,698 (18.77)	15.87 (11582 / 72990)	Reference	30.41 (156 / 513)	Reference	56.35 (1011 / 1794)	Reference	35.05 (415 / 1184)	Reference	31.96 (411 / 1286)	Reference
1-2 hrs	235,453 (53.44)	11.33 (25046 / 221153)	0.83: 0.81, 0.85	24.29 (128 / 527)	0.76: 0.53, 1.10	50.98 (1664 / 3264)	0.76: 0.61, 0.94	34.95 (858 / 2455)	1.02: 0.83, 1.25	27.93 (794 / 2843)	0.88: 0.74, 1.04
≥3hrs	96,553 (22.91)	8.99 (7968 / 88597)	0.68: 0.66, 0.70	19.51 (32 / 164)	0.65: 0.39, 1.06	50.69 (770 / 1519)	0.72: 0.55, 0.93	33.67 (672 / 1996)	1.11: 0.9, 1.38	29.84 (436 / 1461)	0.92: 0.75, 1.12
P			1.18 x 10-234		0.17		0.02		0.52		0.33
Leisure TV[k]											
<1hr	35,011 (7.95)	12.22 (3880 / 31751)	1.27: 1.24, 1.30	29.83 (54 / 181)	1.26: 0.73, 2.17	49.86 (365 / 732)	0.96: 0.73, 1.26	32.33 (172 / 532)	0.86: 0.67, 1.10	25.87 (193 / 746)	0.86: 0.71, 1.05
1-2 hrs	172,769 (39.21)	11.07 (17667 / 159568)	1.09: 1.07, 1.11	25.33 (152 / 600)	0.81: 0.58, 1.15	54.09 (1823 / 3370)	1.28: 1.07, 1.53	33.60 (759 / 2259)	1.01: 0.87, 1.18	29.30 (733 / 2502)	1.06: 0.92, 1.22
≥3hrs	228,525 (51.87)	12.28 (25756 / 209662)	Reference	27.62 (145 / 525)	Reference	52.31 (1573 / 3007)	Reference	35.91 (1255 / 3495)	Reference	30.66 (845 / 2756)	Reference
P			5.18 x 10-79		0.19		0.01		0.43		0.11
Leisure Computer[l]											
<1hr	211,891 (48.09)	11.76 (22970 / 195394)	0.81: 0.79, 0.83	27.13 (124 / 457)	0.85: 0.54, 1.34	53.37 (1773 / 3322)	1.19: 0.92, 1.53	32.17 (838 / 2605)	1.12: 0.92, 1.36	27.57 (663 / 2405)	0.82: 0.68, 0.99
1-2 hrs	180,905 (41.06)	11.20 (18787 / 167792)	0.82: 0.80, 0.85	24.41 (145 / 594)	0.84: 0.54, 1.31	51.54 (1338 / 2596)	1.25: 0.96, 1.62	36.60 (882 / 2410)	1.26: 1.03, 1.54	30.01 (761 / 2536)	0.87: 0.72, 1.05
≥3hrs	43,284 (9.82)	14.91 (5630 / 37766)	Reference	33.20 (82 / 247)	Reference	52.62 (513 / 975)	Reference	38.18 (470 / 1231)	Reference	32.82 (343 / 1045)	Reference
P			1.55 x 10-60		0.73		0.25		0.07		0.13
Use of Sun Protection[m]											
Avoid sunshine	2,646 (0.60)	34.04 (672 / 1974)	3.21: 2.84, 3.62	38.71 (12 / 31)	3.03: 0.72, 12.88	55.40 (77 / 139)	0.89: 0.50, 1.61	33.06 (40 / 121)	0.75: 0.47, 1.19	45.24 (38 / 84)	1.02: 0.59, 1.79
Never - sometimes	191,043 (43.36)	12.95 (21650 / 167220)	Reference	28.48 (262 / 920)	Reference	53.58 (3076 / 5741)	Reference	35.23 (1862 / 5285)	Reference	31.03 (1223 / 3941)	Reference
Mostly/always	245,498 (55.72)	10.84 (25371 / 234070)	0.81: 0.80, 0.83	20.92 (77 / 368)	0.67: 0.48, 0.94	48.27 (572 / 1185)	0.80: 0.64, 0.99	32.63 (309 / 947)	0.83: 0.68, 1.00	25.54 (519 / 2032)	0.75: 0.66, 0.87
P			8.78 x 10-271		0.02		0.11		0.08		3.07 x 10-4
Smoking[n]											
Non-smokers	240,002 (54.47)	11.06 (24192 / 218689)	0.71: 0.69, 0.73	25.98 (272 / 1047)	0.58: 0.28, 1.20	52.98 (2920 / 5512)	0.87: 0.62, 1.22	32.10 (1462 / 4554)	0.72: 0.56, 0.92	29.36 (1004 / 3420)	0.87: 0.72, 1.06
Ex-smokers	152,475 (34.61)	10.78 (15491 / 143744)	0.66: 0.65, 0.68	26.23 (48 / 183)	0.44: 0.20, 0.95	48.98 (502 / 1025)	0.75: 0.51, 1.11	37.40 (426 / 1139)	0.97: 0.73, 1.30	25.78 (457 / 1773)	0.73: 0.59, 0.89
Current smokers	45,900 (10.42)	19.66 (8065 / 41012)	Reference	37.11 (36 / 97)	Reference	58.94 (412 / 699)	Reference	46.49 (364 / 783)	Reference	37.54 (351 / 935)	Reference
P			6.64 x 10-235		0.10		0.32		9.29 x 10-4		4.09 x 10-3

Ethnicity N for: [a]415,903 (94.40%); [b]1,383 (0.31%); [c]7,880 (1.79%); [d]6,896 (1.57%); [e]6,473 (1.47%); Missing 2,046 (0.46%).
Missing N for: [f]3,711 (0.84%); [g]2,698 (0.61%); [h]3,113 (0.71%); [i]25,710 (5.84%): [j]25,877 (5.87%); [k]4,276 (0.97%); [l]4,501 (1.02%); [m]1,394 (0.32%); [n]2,204 (0.50%).
[o]Constructed using ethnic specific BMI quartiles
All models were weighed by 1–kinship coefficient, and adjusted for age, sex, latitudinal location, 25(OH)D acquisition month, 25(OH)D aliquot, fasting upon serum acquisition, Townsend deprivation index, BMI, physical activity, smoking.

Muscle Traits, Sarcopenia, and Sarcopenic Obesity: A Vitamin D Mendelian Randomization Study



Nutrients (MDPI) has published this manuscript.

Impact factor: 5.9 (2023)

Rank in field: 10th out of 34 Quartile 1 journals in Nutrition & Dietetics.

Manuscript

Supplementary Material

MDPI Open Access Information and Policy: All articles published by MDPI are made immediately available worldwide under an open access license. This means everyone is free to re-use the published material if proper accreditation/citation of the original publication is given.

Study 2 Manuscript

List of study 2 manuscript figures:

Study 2 - Figure 1: Phenotypic (A,C,E,G) and genetic (B,D,F,H) associations of 25(OH)D with grip strength (A,B), probable sarcopenia (C,D), non-obese probable sarcopenia (E,F), and arm skeletal muscle mass (G,H) in the UK Biobank, with genetic associations (B,D,F,H) projected on the measured 25(OH)D scale. 91

Study 2 - Figure 2: Mendelian randomization sensitivity analyses of genetically predicted 25(OH)D and grip strength, skeletal muscle mass, and probable sarcopenia in the UK biobank. 94

List of study 2 manuscript tables:

Study 2 - Table 1: Demographic characteristics of UK Biobank participants. 90

Muscle Traits, Sarcopenia, and Sarcopenic Obesity: A Vitamin D Mendelian Randomization Study

Joshua P. Sutherland,[1] Ang Zhou,[1,2,3,] and Elina Hyppönen, [1,2,*]

1. Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide 5000, Australia; joshua.sutherland@mymail.unisa.edu.au (J.P.S.); ang.zhou@unisa.edu.au (A.Z.).

2. South Australian Health and Medical Research Institute, Adelaide 5000, Australia

3. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

*Correspondence: elina.hypponen@unisa.edu.au (E.H.); Tel: +61-8-830-22518

Abstract:

(1) **Background:** Observational studies associate vitamin D deficiency with muscle disorders, while some clinical trial data support a minor association between the vitamin and skeletal muscle performance in healthy subjects. Vitamin D receptor knockout mice studies confirm the relationship between vitamin D and skeletal muscle; however, causal inference in humans is challenging due to the ethical implications of including vitamin D-deficient participants in randomized trials. This study uses genetic methods to safely explore causal underpinnings for the relationship between 25(OH)D concentrations and skeletal muscle-related traits, including grip strength and combined arm skeletal muscle mass, and extends this analysis to suspected pathophysiology in the form of probable sarcopenia and sarcopenic obesity.

(2) **Methods:** We conducted Mendelian randomization (MR) analyses in up to 307,281 participants from the UK Biobank of whom 25,414 had probable sarcopenia and 16,520 had sarcopenic obesity. In total, 35 variants were used to instrument 25(OH)D and MR analyses conducted using multiple approaches.

(3) **Results:** Genetic analyses provided support for a relationship between genetically predicted higher 25(OH)D and skeletal muscle traits, with linear MR analyses for grip strength showing 0.11 kg (95% CI, 0.04-0.19) greater contractile force per 10 unit higher 25(OH)D, while there was a modest association with skeletal muscle mass (0.01 kg (95% CI, 0.003-0.02) greater muscle mass). For probable sarcopenia risk, there was suggestive evidence for lower odds by higher 25(OH)D (OR = 0.96, 95% CI, 0.92-1.00); however, this did not reflect an association with sarcopenic obesity (OR = 0.97, 95% CI, 0.93-1.02), but was seen in probable sarcopenia cases who were not obese (OR = 0.92, 95% CI, 0.86-0.98). Results were similar across multiple MR approaches.

(4) **Conclusions:** Our study supports a causal relationship between 25(OH)D and skeletal muscle health. While evidence for benefit did not extend to lower risk of sarcopenic obesity, effective vitamin D-deficiency prevention strategies may help reduce age-related muscle weakness.

Keywords: vitamin D; skeletal muscle; grip strength; sarcopenia; sarcopenic obesity; vitamin D deficiency; nutritional status; public health.

1. Introduction

Vitamin D is a crucial component in the regulation of calcium homeostasis and bone metabolism; however, evolving research demonstrates wide ranging roles for the vitamin in other tissues, including skeletal muscle tissue [1]. Interest in vitamin D and skeletal muscle tissue has arisen from observations of myopathy and hypotonia concurring with rickets and osteomalacia, which are diseases remediated by reversing severe vitamin D deficiency [2] (25-hydroxy vitamin D (25(OH)D) < 25 nmol/L [3]). Several decades of observational studies have supported an association between vitamin D deficiency and impaired skeletal muscle health, including sarcopenia and sarcopenic obesity [4–7], which are common concerns in the elderly, reflecting age-related loss of skeletal muscle integrity. Experiments using the myocyte-specific vitamin D receptor (VDR) knockout mouse mice [8,9] provide further support for the role of vitamin D-related signaling pathways in skeletal muscle development and in determining lean mass and muscle function, highlighting biological relevance and the need for research in this area.

Clinical supplementation trials, aimed at optimizing 25(OH)D, have supported minor strength improvements in healthy athletes [10]; however, vitamin D supplementation trials for sarcopenia have presented conflicting results [6], with a subset of evidence suggesting co-administration with supplemental protein may have particular benefit for improving muscle strength [7]. Crucially, vitamin D supplementation trials often lack inclusion of severely deficient participants due to ethical concerns or recruitment challenges [11], making it difficult to use non-representative and low-powered clinical trial data to facilitate causal inference in this area. There is additionally a notable lack of clinical data examining the effects of supplemental vitamin D in those with sarcopenic obesity [7].

A further line of evidence can be obtained from studies using Mendelian randomization (MR), which uses genetic variants that approximate the exposure as an instrumental variable or "proxy indicator" and can provide evidence of causal effects in situations where it is unethical or infeasible to conduct RCTs [12]. Observational research with vitamin D and skeletal muscle tissue has confounding challenges, such as the association between outdoor time (i.e., vitamin D stimulating UVB exposure) and increased physical activity [13] (i.e., with physical activity having independent effect on myocyte expression), while paradoxical observations can arise due to skeletal muscle tissue itself acting as a 25(OH)D storage site [14]. The MR can largely overcome these challenges. To date, there are several MR studies that have looked into the association between vitamin D and muscle traits (grip strength [15,16]—a tool often used to provide rapid indication of an individual's general muscle strength [17]—and skeletal muscle mass [18]), while evidence for association with skeletal muscle pathology (i.e., sarcopenia) is inconclusive [19] with only one multi-exposure MR study having used a comparatively low powered 25(OH)D instrument. These prior muscle trait studies provide promising support for an association between higher 25(OH)D and general muscle health; however, the role of 25(OH)D in sarcopenia warrants a more conclusive investigation, with an adjoining focus on sarcopenic obesity. In this study we use information from up to 307,281 UK Biobank participants to examine evidence for a causal role of 25(OH)D in skeletal muscle health. Importantly, in addition to grip strength and muscle mass, we include exploration of pathophysiology, as reflected by associations with probable sarcopenia [20,21] and sarcopenic obesity.

2. Methods

The UK Biobank is a large-scale, prospective cohort comprising a baseline assessment of 502,316 individuals aged 37–73 years, undertaken from March 2006 and July 2010 in England, Scotland, and Wales [22]. Volunteer participants partook in questionnaire surveys, underwent physical assessments, and provided biological samples for biomarker and genetic assays. Primary analyses in our study were limited to unrelated participants of European ancestry, who had measures for serum 25(OH)D concentrations and available data for the muscle trait outcomes (n = 307,281); these factors and information pertaining to missing data are outlined in Figure S1 [23].

2.1. Grip Strength, Muscle Mass, and Probable Sarcopenia

(*Sentence edited for thesis*) Grip strength assessment and bioimpedance analysis of lean skeletal muscle mass and whole-body fat mass were performed at the initial baseline assessment centre visit and used a Jamar J00105 hydraulic hand dynamometer and Tanita BC418MA body composition analyser (in conjunction with inputted height data, acquired from a Seca 202 stadiometer), respectively. For our grip strength analyses, we used the left-and-right-hand average contractile force, measured in kilograms (kg). We use a previously developed definition for probable sarcopenia [20,21], and defined cases as those participants who had grip strength < 27 kg of force (males) and < 16 kg of force (females). Whole-body fat mass was used to calculate body fat percentage (whole-body fat mass divided by body weight and multiplied by 100) and the conventional cut-off points of ≥ 25% for males and ≥ 35% for females were used to determine obesity [24]. Analyses on probable sarcopenic obesity and non-obesity probable sarcopenia used participants without sarcopenia as the control group. Analysis of lean skeletal muscle mass used combined left-and-right-arm data.

2.2. Measured and Genetically Instrumented 25(OH)D

The LIAISON XL 25(OH)D assay (DiaSorin, Stillwater , OK, USA) was used to measure baseline serum 25(OH)D concentrations, as described in the Supplementary Materials [23]. Covariates used in adjustments were acquired at baseline from self-reported, touchscreen questionaries (age, sex, physical activity, smoking, alcohol, birth location), residential address data (Townsend deprivation index) [25], physical assessments (height and waist circumference), with these and additional covariates (assessment centre, month of 25(OH)D acquisition, genetic-related information, and 25(OH)D measurement nuisance factors) outlined in full in the Supplementary Materials (pp. 3–4) [23]. We constructed a weighted genetic score by assembling 35 common autosomal SNPs that were discovered from a genome-wide association analysis on measured 25(OH)D concentration [26,27] and replicated in the SUNLIGHT consortium [28]. The construction of the vitamin D genetic score, and the selection of the variants used, are described in the Supplementary Materials (p. 5) [23]. Ethics approval for the UK Biobank was granted by the National Information Governance Board for Health and Social Care and North West Multicentre Research Ethics Committee (11/NW/0382). The present analysis operates under UK Biobank application 20175.

2.3. Statistical Methods

The main phenotypic analyses were conducted using linear and logistic regression for continuous (grip strength and skeletal muscle mass) and binary outcomes (probable sarcopenia, and probable sarcopenia with or without obesity), respectively. We fitted two models, a simple and a full model; in the simple model, we adjusted for study structure factors—age, sex, height, and assessment centre—as well as nuisance factors, which could affect the measured 25(OH)D concentrations (blood sampling month, fasting time before sample acquisition, and aliquot); in

the full model, we further adjusted for lifestyle factors, including waist circumference, physical activity (low, moderate, and high), alcohol (daily or almost daily, once or twice a week, three of four times a week, one to three times a month, special occasions only, and never), smoking (non-smokers, former smokers, and current smokers), education (none, national vocational qualification/certificate of secondary education/A-levels, and degree/professional), and Townsend deprivation index (quartile division of index: least to most deprivation). Height and waist circumference were chosen for phenotypic adjustments – to account for body shape and obesity – rather than using body mass index (BMI), due to BMI's calculation explicitly including measure of skeletal muscle mass via body weight. To examine the shape of the association between 25(OH)D and the muscle trait outcomes, we used a likelihood ratio test and compared model fit between the best-fitting fractional polynomial model and the linear model [29].

We performed linear MR analyses to examine genetic evidence for causality in the association of measured 25(OH)D concentrations and muscle traits, including grip strength, arm skeletal muscle mass, probable sarcopenia, and probable sarcopenia with or without obesity (Supplementary Materials, pp. 6–7 and Figures S2 and S3) [23], with the genetic score-based one-sample MR analyses presented as primary findings. We also examined the presence of non-linearity of association using two available MR methods [30,31] (Supplementary Materials, pp. 8–9 and Figure S3) [23]. Linear and non-linear analyses are detailed in the Supplementary Materials [23]. As our analysis included three (correlated) outcomes, to ensure robust evidence of association, we applied Bonferroni correction to control type I error rate, taking a p-value below $0.05/3 = 0.02$ as evidence for an association.

Valid causal inference in MR analysis is reliant on three key assumptions [32] (Figure S2) [23]; in the context of our study, they can be outlined as the following: (1) the vitamin D

genetic score associates with measured 25(OH)D concentrations; (2) the vitamin D genetic score has no direct effect on muscle traits; (3) the vitamin D genetic score does not associate with confounders of measured 25(OH)D and muscle traits. In relation to the first MR assumption, the association of the vitamin D genetic score was examined with 25(OH)D concentrations in the UK Biobank (Figure S4) [23]. Horizontal pleiotropy is a phenomenon wherein the genetic instrument associates with outcomes via pathways other than the exposure of interest; if present, it will violate the second and/or third MR assumption and will resultantly bias the analyses. We implemented the following strategies to evaluate any potential horizontal pleiotropy: (i) We assessed the association of the vitamin D genetic score with a series of potential confounders in the UK Biobank, which included the Townsend deprivation index, body mass index, alcohol intake, smoking, physical activity, and education (Table S1 and Figure S5) [23]. (ii) We assigned variants to four functional blocks (renal, blood, lipids/metabolic, and unclassified) and conducted a leave-block-out analysis, designed to assess if our findings were driven by any specific blocks of potentially pleiotropic variants (in particular pleiotropy by lipid traits) [33] (Supplementary Materials, p. 8, Tables S2 and S3) [23]. (iii) In the MR analyses, we employed five sensitivity analyses, including IVW, MR-Egger, Weighted Median MR, Weighted Mode MR, and MR-PRESSO (Supplementary Materials, pp. 6–7) [23] with relatively independent assumptions on pleiotropy.

R version 3.6.1 was used for non-linear MR (‘nlmr’ and ‘SUMnlmr’ packages) and linear MR sensitivity analyses (‘TwoSampleMR’ and ‘MRPRESSO’ packages), and STATA version 14.1 (Stata-Corp LP, College Station, TX, USA) was used for all other analyses.

3. Results

Table 1 shows baseline participant characteristics and distribution of 25(OH)D, 25(OH)D < 25 nmol/L, grip strength, probable sarcopenia cases, sarcopenic obesity cases, and arm skeletal muscle mass. The average 25(OH)D concentration was 49.82 nmol/L, with 11.70% of the participants (n = 36,009) having 10–24.9 nmol/L. Concentrations for average 25(OH)D were higher in participants living in the southern-most latitudinal zone, those with less socioeconomic deprivation, non-smokers, lowest BMIs, and those with higher levels of physical activity. Grip strength was higher in those with the greatest level of physical activity, educational status, and the lowest social deprivation. Arm skeletal muscle mass was higher in males and in those < 60 years of age. At baseline, 8.28% (n = 25,414) had probable sarcopenia, with an overrepresentation of females and those over 60 years of age (10% and 12%, respectively), as well as of those with the highest level of BMI, social deprivation, and sedentary behaviour. Sarcopenic obesity was notably more common in women compared to men, those aged > 60 years vs. < 60, and those with less compared to more physical activity. Table S4 [23] additionally presents data for height and waist circumference; similar trends were seen between BMI and waist circumference, while height differences had notable impact on skeletal muscle mass (greater mass in the tallest participants) and probable sarcopenia (greater prevalence in the shorter participants). Figure S6 [23] shows positive probable sarcopenia cases to have 0.59 kg less of arm skeletal muscle mass, relative to negative cases.

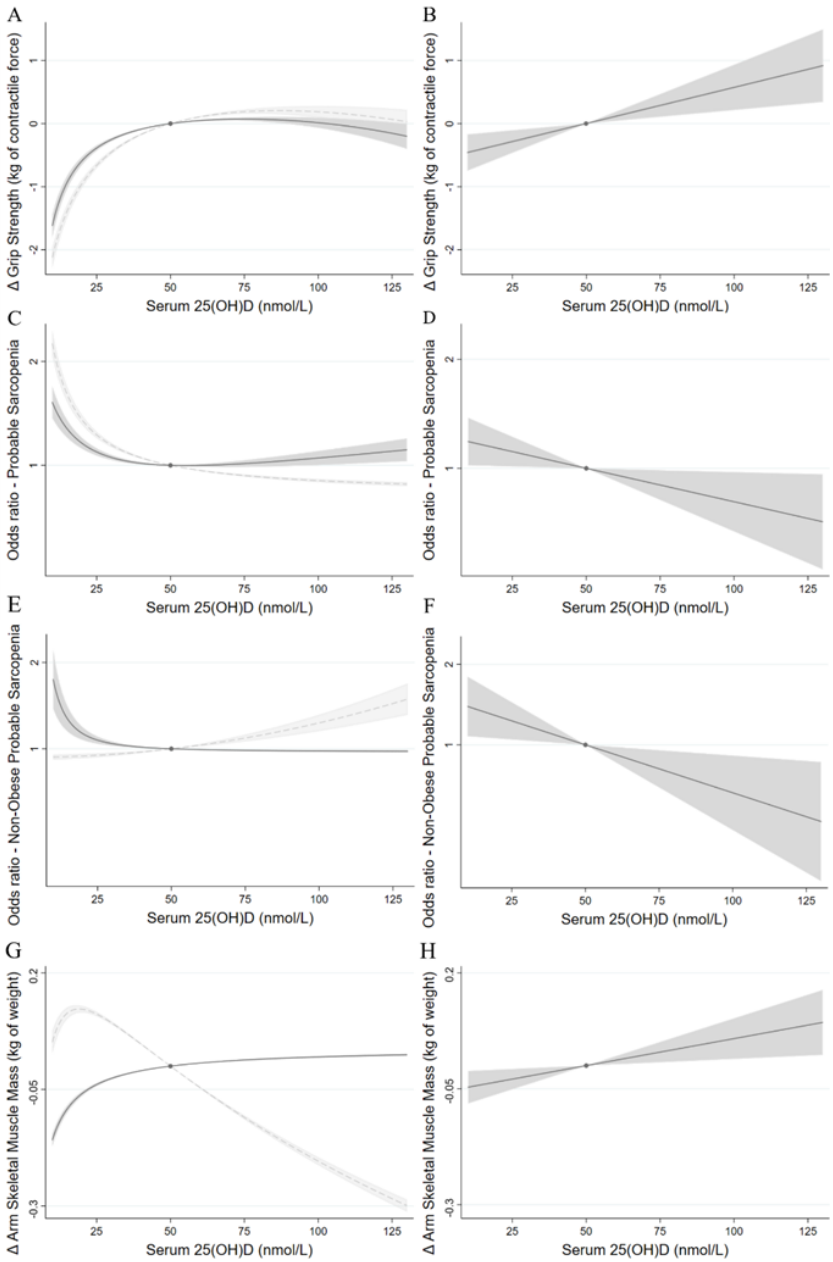
Study 2 - Table 1: Demographic characteristics of UK Biobank participants.

		25(OH)D	25(OH)D <25 nmol/L	Grip Strength [a]	Probable Sarcopenia	Sarcopenic Obesity	Arm Skeletal Muscle Mass [b]
	N (%)	*N* = 307,281 Mean (SD)	*N* = 36,009 %	*N* = 306,967 [c] Mean (SD)	*N* = 25,414 %	*N* = 16,520 %	*N* = 302,112 [d] Mean (SD)
All	307,281	49.82 (20.96)	11.70	31.04 (11.03)	8.28	5.54	5.51 (1.58)
Sex							
Men	144,538 (47.4)	49.86 (21.03)	11.60	39.61 (8.74)	6.39	3.98	6.92 (1.09)
Women	162,743 (53.0)	49.78 (20.89)	11.80	23.43 (6.22)	9.96	6.94	4.26 (0.60)
Age							
<60	169,594 (55.2)	48.38 (21.12)	13.46	32.48 (11.18)	5.46	3.40	5.56 (1.66)
≥60	137,687 (44.8)	51.57 (20.62)	9.54	29.28 (10.58)	11.75	8.23	5.44 (1.48)
BMI							
Low 25%, 12.1–24.0	76,633 (24.9)	53.04 (22.09)	10.19	28.62 (9.62)	7.61	1.21	4.53 (1.13)
Mid 50%, 24.1–29.8	153,312 (50.0)	50.95 (20.68)	10.05	32.19 (11.17)	7.55	5.46	5.57 (1.44)
High 25% 29.8–74.7	76,667 (25.0)	44.42 (19.23)	16.40	31.23 (11.63)	10.18	9.77	6.36 (1.71)
Missing	669 (0.2)	40.69 (21.11)	26.01	25.51 (12.86)	33.73	4.33	5.34 (1.48)
Location [e]							
South, ≤51° Lat	102,226 (33.3)	51.43 (20.49)	9.29	30.77 (10.84)	8.35	5.20	5.53 (1.58)
Mid, 52°–53° Lat	144,470 (47.0)	49.93 (20.88)	11.35	31.10 (11.11)	8.42	5.80	5.51 (1.58)
North, 54°–≥55° Lat	60,585 (19.7)	46.82 (21.59)	16.62	31.36 (11.16)	7.83	2.49	5.45 (1.58)
Smoking							
Non-smokers	167,537 (54.5)	50.04 (20.63)	10.92	30.48 (11.00)	8.07	5.22	5.36 (1.55)
Ex-smokers	108,015 (35.2)	50.77 (21.04)	10.66	31.61 (10.99)	8.37	6.05	5.69 (1.60)
Current smokers	30,673 (10.0)	45.22 (21.80)	19.63	32.23 (11.17)	8.91	5.31	5.66 (1.60)
Missing	1056 (0.3)	50.17 (21.77)	12.41	28.95 (11.14)	14.58	11.22	5.57 (1.60)
Alcohol							
Daily	65,476 (21.3)	51.22 (21.51)	11.03	33.09 (10.82)	6.51	3.93	5.71 (1.56)
1 to 4 times wk	155,474 (50.6)	50.87 (20.80)	10.19	31.85 (11.01)	7.11	4.67	5.58 (1.60)
1 to 3 times mo	34,061 (11.1)	48.08 (20.26)	12.91	29.56 (10.73)	8.46	5.85	5.35 (1.58)
Special occasion	32,125 (10.5)	46.18 (20.40)	15.74	27.05 (10.43)	12.85	9.37	5.13 (1.48)
Never	19,934 (6.5)	45.91 (20.95)	17.06	27.07 (10.66)	15.47	11.11	5.19 (1.45)
Missing	211 (0.07)	44.97 (21.05)	16.59	28.54 (11.32)	17.54	13.86	5.39 (1.57)
Physical activity							
Low	91,911 (29.9)	46.30 (20.20)	15.18	29.78 (10.96)	10.39	7.49	5.47 (1.59)
Moderate	149,064 (48.5)	50.60 (20.80)	10.50	31.31 (10.88)	7.17	4.65	5.47 (1.57)
High	59,518 (19.4)	54.01 (21.41)	8.11	32.65 (11.09)	6.37	3.78	5.62 (1.60)
Missing	6788 (2.2)	43.33 (21.32)	22.48	28.32 (12.19)	20.78	14.80	5.77 (1.70)
Education							
None	52,119 (17.0)	50.40 (21.43)	11.89	28.56 (10.98)	14.47	10.74	5.43 (1.55)
NVQ/CSE/A-Lev.	109,007 (35.5)	50.55 (21.19)	11.22	31.02 (11.18)	7.90	5.33	5.51 (1.61)
Deg./professional	143,586 (46.7)	49.04 (20.57)	12.00	31.99 (10.79)	6.23	3.79	5.54 (1.57)
Missing	2569 (0.84)	50.33 (20.90)	11.60	29.48 (11.22)	13.36	9.04	5.52 (1.59)
Townsend index							
Q1 Deprivation low	76,746 (25.0)	51.91 (20.71)	9.16	31.71 (11.06)	6.60	4.21	5.50 (1.57)
Q2	76,745 (25.0)	51.50 (20.70)	9.35	31.31 (11.10)	7.37	4.82	5.49 (1.58)
Q3	76,719 (25.0)	49.91 (20.79)	11.20	30.92 (10.98)	8.25	5.55	5.50 (1.59)
Q4 Deprivation high	76,711 (25.0)	45.94 (21.08)	17.11	30.24 (10.93)	10.90	7.62	5.54 (1.59)
Missing	360 (0.1)	50.02 (20.53)	11.39	31.37 (11.16)	7.22	4.84	5.61 (1.66)

Kg, kilograms; BMI, body mass index; NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; Q, quartiles; SD, standard deviation. a Average grip strength, min 0 kg to max 85 kg of contractile force, b Bioimpedance analysis-derived skeletal muscle mass of the combined arm, min 2 kg to 35.9 kg in weight, c Missing n= 314 (0.10%), d Missing n = 5169 (1.68%). e Location categorization derived from assessment centre aggregation by longitudinal zoning, as described in the Supplementary Materials (p. 4). p-values for all models <0.01; all models were adjusted for sex, age, assessment centre, and nuisance factors that could affect serum 25(OH)D measurements, including the month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

3.1. Grip Strength

In our phenotypic analyses, we observed a non-linear inverse relationship between measured 25(OH)D and grip strength (Figure 1A), with both the simple and fully adjusted models following similarly in shape.



Study 2 - Figure 1: Phenotypic (A,C,E,G) and genetic (B,D,F,H) associations of 25(OH)D with grip strength (A,B), probable sarcopenia (C,D), non-obese probable sarcopenia (E,F), and arm skeletal muscle mass (G,H) in the UK Biobank, with genetic associations (B,D,F,H) projected on the measured 25(OH)D scale.

Figure 1 legend:

Shaded areas reflect 95% confidence intervals, and the dot represents the reference point of 50 nmol/L. (Genetic analysis: B,D,F,H): the x axis – noted at serum 25(OH)D – refers to the measured 25(OH)D scale upon which the genetic analysis has been mapped. (Phenotypic analysis: A,C,E,G): simple models (light grey dashed line) were adjusted for sex, age, height, assessment centre, and nuisance factors that could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement, with full models (dark grey solid line) additionally adjusted for educational status, the Townsend depravation index, waist circumference, physical activity, alcohol, and smoking. (Genetic analysis: B,D,F,H): adjusted for age, sex, assessment centre, birth locations, SNP array, top 40 genetic principal components, and nuisance factors that could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. The notable change in the phenotypic association with skeletal muscle mass by confounder adjustment was largely due to the inclusion of waist circumference, which was strongly correlated with both 25(OH)D and muscle mass ($r = -0.1609$, $r = 0.689$, respectively; $p < 0.1 \times 10–4$ for both).

Grip strength decreased precipitously below 25 nmol/L 25(OH)D, with the adjusted association appearing to plateau from 50 nmol/L onwards. For grip strength, compared to the reference (25(OH)D 50 nmol/L), contractile force was −0.39 kg (95% CI, −0.43-−0.35) lower at 25 nmol/l and −1.62 kg (95% CI, −1.80-−1.44) lower at 10 nmol. In genetic analysis, each 10 nmol/L higher 25(OH)D was associated with 0.11 kg (95% CI, 0.04-0.19) greater contractile force (Figure 1B). Statistical evidence for non-linearity was observed using the residual method but not with the double-ranked method (corrected $P = 5.28 \times 10-9$ and $P = 0.27$, respectively, Figure S7A,B) [23], with a visually similar association regardless of the approach.

3.2. Probable Sarcopenia

We observed a non-linear relationship between measured 25(OH)D and sarcopenia, with the simple and fully adjusted phenotypic models both following closely in shape (Figure 1C). The highest odds ratios (OR) for probable sarcopenia were observed < 25 nmol/L of measured 25(OH)D. Compared to 50 nmol/L of measured 25(OH)D, the fully adjusted OR of probable sarcopenia was 9% higher for those at 25 nmol/L (OR = 1.09, 95% CI, 1.07-1.12) and 52% higher at 10 nmol/L (OR = 1.52, 95% CI, 1.37-1.69). In the genetic analysis (Figure 1D), there was some evidence for lower odds of probable sarcopenia by higher 25(OH)D (OR = 0.96, 95%

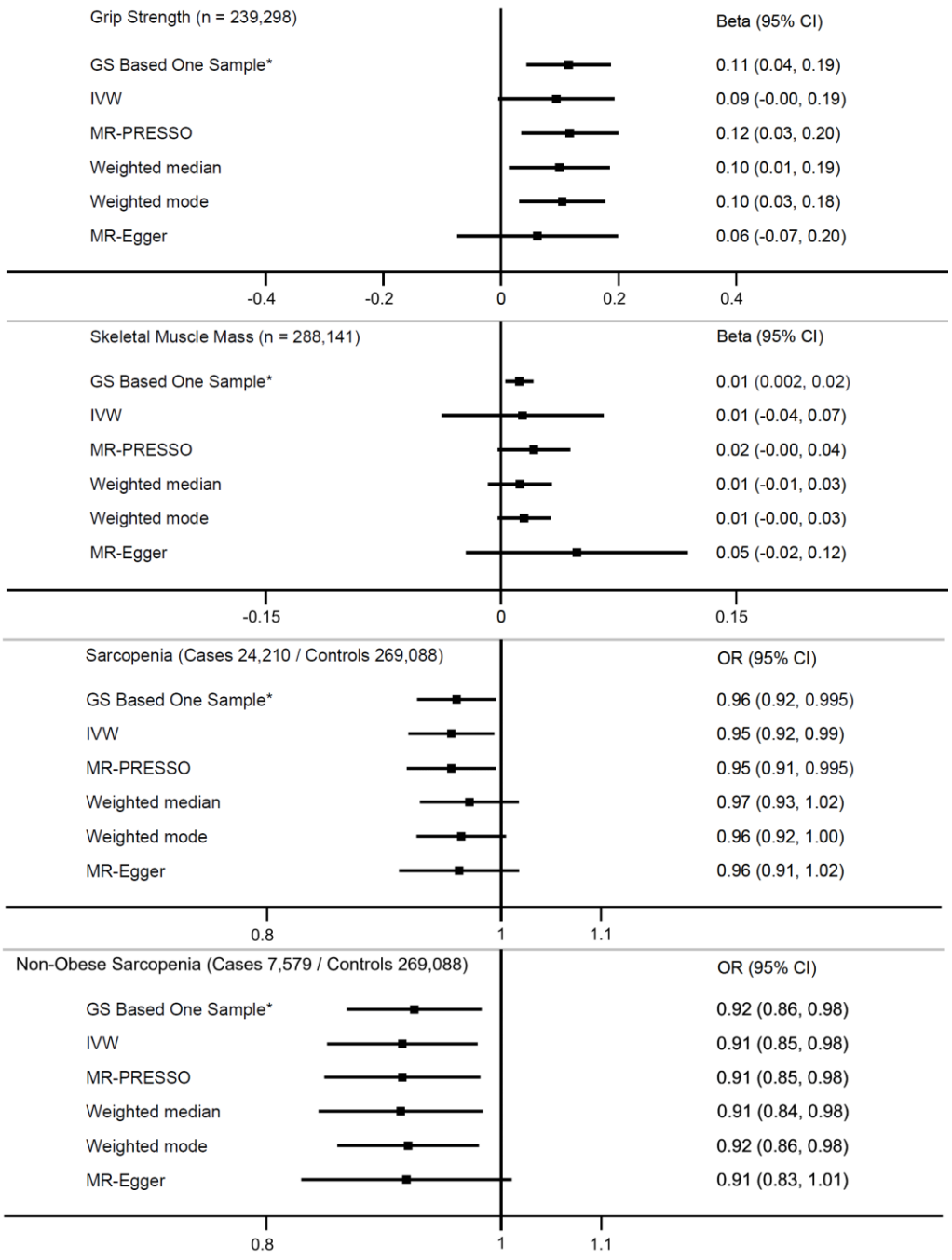
CI, 0.92-1.00). There was no evidence for non-linearity in the association between 25(OH)D and sarcopenia in the genetic analyses (corrected P = 0.06 with the residual method and P = 0.7 for the ranked method). Phenotypic analyses suggested that the association between 25(OH)D and sarcopenia is modified by obesity (P for interaction 1.368 × 10−12). In the genetic analysis, there was no evidence for lower odds by higher 25(OH)D for sarcopenic obesity (OR = 0.97, 95% CI, 0.93-1.02), while analyses on probable sarcopenia not concurring with obesity (i.e., body fat < 25% in males and < 35% in females) retained the evidence seen in the overall probable sarcopenia assessment, with a stronger odds ratio (OR = 0.92, 95% CI, 0.86-0.98) (Figure 1H).

3.3. Arm Skeletal Muscle Mass

As seen in the phenotypic analyses (Figure 1G), adjustment for covariates had a strong effect on the observed association between 25(OH)D and arm skeletal muscle mass, with muscle mass decreasing below 25 nmol/L 25(OH)D and appearing to plateau from 50 nmol/L onwards in the adjusted analysis. In genetic analysis (Figure 1H), evidence for a linear association was present (P-non-linearity > 0.2 with both approaches), with each 10 nmol/L higher 25(OH)D associated with a 0.01 kg of weight (95% CI, 0.003-0.02) greater muscle mass.

3.4. Sensitivity Analyses

Linear analyses using different MR approaches supported the primary MR relationships (Figure 2). Similar results were obtained from sensitivity analyses produced using the 122 version of the vitamin D genetic score and when using the leave-block-out approach (Table S3) [23].



Study 2 - Figure 2: Mendelian randomization sensitivity analyses of genetically predicted 25(OH)D and grip strength, skeletal muscle mass, and probable sarcopenia in the UK biobank.

* In the GS-based one-sample approach, the weight of SNPs are based on independent SNP-25(OH)D betta coefficients from the Sunlight Consortium [28]. Genetic analysis adjusted for age, sex, assessment centre, birth locations, SNP array, top 40 genetic principal components, and nuisance factors that could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. Outlier SNPs for MR PRESSO include the following: Grip Strength, rs1047891, rs212100, rs6782190; Skeletal Muscle Mass, rs1047891, rs35408430, rs6782190, rs7528419, rs76798800.

4. Discussion

Age-related decline of muscle strength and reduction in muscle composition are increasing concerns in ageing populations. In addition to an expected decrease of up to 30% of skeletal muscle mass by age 80 years [34], a subset of the general population is affected by above average sarcopenic conditions that can exacerbate a range of concurrent chronic health conditions, as well as increase the likelihood of suffering acute injuries [35]. The economic burden of muscle weakness has been a focus of research for several decades [36], with an example being the southern England-based Hertfordshire Cohort Study, where excess costs associated with muscle weakness—ascertained using sex-specific cut-offs for weak grip strength—were estimated to contribute GBP 2.5 billion to healthcare expenditure in the UK annually [37]. The increasing recognition of sarcopenia and muscle weak-ness as major contributors of disease burden has led researchers to call for a greater focus on the role and costs such challenges will have on healthcare systems moving forward in-to the future [38–41].

The considerable size of the UK Biobank has allowed our study to safely and ethically present evidence for a causal relationship between 25(OH)D concentration and suspected skeletal muscle pathophysiology in the form of probable sarcopenia and sarcopenic obesity, helping to inform this relatively underexplored area. Importantly, our study reaffirms the muscle trait MR associations (physiological – grip strength [15,16], and anatomical – skeletal muscle mass [18]), the mounting evidence of which collectively support the VDR knockout studies [8,9], which demonstrated skeletal muscle developmental issues, decreased lean mass, and reduced grip strength in affected mice.

The lack of supportive evidence in our study for the reducing of sarcopenic obesity risk by higher 25(OH)D was surprising, given the robust correlation seen in phenotypic research

between vitamin D deficiency and both obesity – and non-obesity-related sarcopenia [7]. Our results—which show the evidence for probable sarcopenia remaining strongest in those without obesity—potentially suggest that vitamin D's contribution to sarcopenia prevention/remediation is less able to overcome the conditions associated with obesity. Sarcopenic obesity has been associated with varying increases of inflammatory factors—specifically IL-6, C-reactive protein, IL-1 receptor antagonist, and soluble IL-6 receptor [42]—and while vitamin D has been shown to reduce pathological levels of inflammation (particularly interleukin 6 [43]), it could arise that resulting shifts in metabolic demand (resulting from obesity-related inflammation [44]) might overwhelm the influence vitamin D has on maintaining musculoskeletal homeostatic processes.

In the same way that emerging clinical trials (with non-obese sarcopenia sufferers) are demonstrating some consistent successes with the combining of vitamin D and branch chain amino acids [7], future clinical trials focusing on sarcopenic obesity could possibly consider co-administration of vitamin D with other anti-inflammatory nutraceutical or pharmaceutical agents. While obese individuals are known to already have an excess of branch chain amino acids (relative to healthy weight individuals [45]), further clinical trials for sarcopenic obesity could elucidate if co-administration of supplemental protein is appropriate in this group, and whether co-administration with additional anti-inflammatory factors support potential benefits from vitamin D in the presence of obesity and increased inflammation.

In addition to sarcopenia, there are other vitamin D-related aetiologies which may explain grip strength weakness, including adrenal dysregulation [46], metabolic syn-drome [47,48], and anaemia [49,50]. For instance, the relation between 25(OH)D and serum ferritin [51] (and subsequently anaemia and grip strength [49,50]), as well as vitamin D and its role in endocrinology and metabolism more generally [52], may be areas via which the vitamin has

indirect effects on skeletal muscle function and pathophysiology. Similarly, sarcopenia also predicts mortality, which itself is causally associated with 25(OH)D deficiency [53–55]. Further research focusing on broader mechanisms linking vitamin D and skeletal muscle, and their interactions with wider pathophysiology, is warranted.

Prior in vitro research and the VDR knockout studies provide evidence for active vitamin D having an effect on both skeletal muscle myocyte mass and function, with the latter likely being driven by myocyte mitochondrial regulation [56] and non-transcriptional membrane-associated signaling mechanisms that govern the contractile force of the muscle unit [57]. This dualistic effect of active vitamin D on skeletal muscle myocytes likely accounts for why we see a robust grip strength association in our study despite the modest anatomical variation in skeletal muscle mass by 25(OH)D concentrations. Although the evidence for a causal link between 25(OH)D and skeletal muscle mass observed in this study is notable, our study would suggest that vitamin D-mediated skeletal muscle mass stimulation is likely to be only one factor related to skeletal muscle development and maintenance, and improvements to mass in the healthy are minor and should not be interpreted as a reason to pursue excess of 25(OH)D sufficiency. More research is needed to establish whether ensuring sufficient vitamin D intakes may help to prevent age-related decline in muscle strength, and the extent to which related effects may depend upon the presence of excess adiposity.

Strengths and Limitations

A primary strength of our study is the genetic approach, which has facilitated safe exploration of causal relationships pertaining to differences in 25(OH)D concentrations, including the exploration of associations in participants with deficiency, in whom it would otherwise be unethical to conduct properly controlled RCTs. Our study is possible because of 25(OH)D serum

measurements and genotyping in 307,281 individuals by the UK Biobank, which provided sufficient power for analyses on both physiological and an-atomical skeletal muscle outcomes. While we do not know if the probable sarcopenia measure used in this study is deterministic of a clinical diagnosis in the assessed population, the use of grip strength cut-offs is supported by the European Working Group on Sarcopenia in Older People as a clinical screening tool [20], and further research [37] has affirmed the public health relevance of these cut-offs. While future studies using more complex clustering of sarcopenia measures would be advantageous, our results remain highly pertinent given the emerging public health relevance being attributed to weak grip strength measures alone [58]. Although the grip strength cut-offs could reflect issues not directly related to skeletal muscle dysfunction, our study supports consideration of vitamin D deficiency as a determinant of grip strength, even in instances outside of clinically diagnosed sarcopenia.

We restricted the analyses to participants of white European descent, and while this reduces potential bias from population stratification, transferability to other ethnic groups may be limited. With a 5% response rate, the UK Biobank is not representative of the general UK population [59]. However, prior studies using the UK Biobank have replicated expected exposure–disease associations [59], with the MR approach shown to be less affected by selection bias than other types of observational studies, suggesting limited influence on our findings [60]. The MR method provides support for a causal association; however, estimates can be biased by horizontal pleiotropy [12]. Importantly, the instrument used for our analyses was constrained to variants with replicated evidence for an association with 25(OH)D [61]. Furthermore, our analyses include the use of several pleiotropy robust approaches, and there was no evidence for an association between the genetic score and potential confounders, either in the overall analyses

or within stratums of residual serum 25(OH)D concentrations. As is the case with all MR analyses, genetic instruments approximate average effects over the life course. Therefore, it is possible that the true associations may be more complex than presented here, although as muscle degradation reflects a long-term process, there is an advantage to the MR approach in this respect.

5. Conclusions

Our study supports a causal relationship between 25(OH)D and skeletal muscle health. While evidence for the benefits did not extend to a lower risk of sarcopenic obesity, effective vitamin D-deficiency prevention strategies may help reduce age-related muscle weakness.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/2072-6643/15/12/2703. **Figure S1**: Study population and participant inclusion criteria; **Figure S2**: Mendelian randomization assumption diagram; **Figure S3**: Schematic representation of Mendelian randomization analyses performed in the current study; **Figure S4**: Distribution of the vitamin D genetic score (mapped on an X-axis by increasing number of alleles) in the UK Biobank (A), and the association of the vitamin D genetic score (in six groups – lowest to highest number of alleles) with measured 25(OH)D concentrations in the UK Biobank (B); **Figure S5**:Association of the vitamin D genetic score with potential confounders across 100 strata of residuals of measured 25(OH)D; **Figure S6**: Mean kilogram difference in arm skeletal muscle mass by probable sarcopenia status in UK Biobank participants; **Figure S7**: Residual method (A) and doubly rank method (B) non-linear Mendelian randomization analysis of genetically predicted 25(OH)D on grip strength, projected on the measured 25(OH)D scale; **Figure S8**: Selection of variants for the genetic instrument for measured 25(OH)D concentrations; **Figure S9**: Distribution of the GRS-measured 25(OH)D association across 100 strata of residual measured 25(OH)D; **Table S1**: Association of the

vitamin D genetic score with potential confounders in the UK Biobank; **Table S2**: Functional blocks used in the leave-block-out analyses; **Table S3**: 122 SNP genetic score and leave block out analysis; **Table S4**: Demographic characteristics of UK Biobank participants – including height and waist circumference; **Table S5**. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations. References [25,26,28,30–32,62–73] are cited in supplementary file.

Author Contributions: J.P.S. wrote the first draft and analysed the data. EH and AZ advised on analyses, revised the paper, and supervised the study. All authors interpreted results and have read and agreed to the published version of the manuscript.

Funding: The study was supported by the National Health and Medical Research Council (NHMRC), grant number 11123603. Joshua P. Sutherland studentship is funded by Australian Research Training Program Scholarship. Neither funding source had any role in this study's design, conduct, or re-porting.

Institutional Review Board Statement:

Informed Consent Statement: Informed consent was obtained by the UK Biobank from all sub-jects whose data was used in this study.

Data Availability Statement: Protocol: not available. Data: This research was conducted using the UK Biobank resource under application number 20175. An analytic code will be made available upon request, and all data and code will be available to approved users upon application to the UK Biobank.

Acknowledgments: We thank all UK Biobank participants involved, and Anwar Mulugeta (University of South Australia) for support in data management.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ceglia, L. Vitamin D and its role in skeletal muscle. *Curr. Opin. Clin. Nutr. Metab. Care* **2009**, *12*, 628–633.

2. Uday, S.; Högler, W. Nutritional Rickets and Osteomalacia in the Twenty-first Century: Revised Concepts, Public Health, and Prevention Strategies. *Curr. Osteoporos. Rep.* **2017**, *15*, 293–302.

3. Pilz, S.; Zittermann, A.; Trummer, C.; Theiler-Schwetz, V.; Lerchbaum, E.; Keppel, M.H.; Grübler, M.R.; März, W.; Pandis, M. Vitamin D testing and treatment: A narrative review of current evidence. *Endocr. Connect.* **2019**, *8*, R27–R43.

4. Iannuzzi-Sucich, M.; Prestwood, K.M.; Kenny, A.M. Prevalence of Sarcopenia and Predictors of Skeletal Muscle Mass in Healthy, Older Men and Women. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **2002**, *57*, M772–M777.

5. Torres, C.F.; Forbes, G.B.; Decancq, G.H. Muscle weakness in infants with rickets: Distribution, course, and recovery. *Pediatr. Neurol.* **1986**, *2*, 95–98.

6. Remelli, F.; Vitali, A.; Zurlo, A.; Volpato, S. Vitamin D Deficiency and Sarcopenia in Older Persons. *Nutrients* **2019**, *11*, 2861.

7. Di Filippo, L.; De Lorenzo, R.; Giustina, A.; Rovere-Querini, P.; Conte, C. Vitamin D in Osteosarcopenic Obesity. *Nutrients* **2022**, *14*, 1816.

8. Endo, I.; Inoue, D.; Mitsui, T.; Umaki, Y.; Akaike, M.; Yoshizawa, T.; Kato, S.; Matsumoto, T. Deletion of Vitamin D Receptor Gene in Mice Results in Abnormal Skeletal Muscle Development with Deregulated Expression of Myoregulatory Transcription Factors. *Endocrinology* **2003**, *144*, 5138–5144.

9. Girgis, C.M.; Cha, K.M.; So, B.; Tsang, M.; Chen, J.; Houweling, P.J.; Schindeler, A.; Stokes, R.; Swarbrick, M.M.; Evesson, F.J.; Cooper, S.T. Mice with myocyte deletion of vitamin D receptor have sarcopenia and impaired muscle function. *J. Cachexia Sarcopenia Muscle* **2019**, *10*, 1228–1240.

10. Chiang, C.-M.; Ismaeel, A.; Griffis, R.B.; Weems, S. Effects of Vitamin D Supplementation on Muscle Strength in Athletes: A Systematic Review. *J. Strength Cond. Res.* **2017**, *31*, 566–574.

11. Bolland, M.J.; Grey, A.; Avenell, A. Assessment of research waste part 2: Wrong study populations- an exemplar of baseline vitamin D status of participants in trials of vitamin D supplementation. *BMC Med. Res. Methodol.* **2018**, *18*, 101.

12. Davies, N.M.; Holmes, M.V.; Davey Smith, G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **2018**, *362*, k601.

13. Schaefer, L.; Plotnikoff, R.C.; Majumdar, S.R.; Mollard, R.; Woo, M.; Sadman, R.; Rinaldi, R.L.; Boulé; N; Torrance, B.; Ball, G.D.; et al. Outdoor Time Is Associated with Physical Activity, Sedentary Time, and Cardiorespiratory Fitness in Youth. *J. Pediatr.* **2014**, *165*, 516–521.

14. Mason, R.S.; Rybchyn, M.S.; Abboud, M.; Brennan-Speranza, T.C.; Fraser, D.R. The Role of Skeletal Muscle in Maintaining Vitamin D Status in Winter. *Curr. Dev. Nutr.* **2019**, *3*, nzz087.

15. Pinto Pereira, S.M.; Garfield, V.; Norris, T.; Burgess, S.; Williams, D.M.; Dodds, R.; Sayer, A.A.; Robinson, S.M.; Cooper, R. Linear and Nonlinear Associations Between Vitamin D and Grip Strength: A Mendelian Randomization Study in UK Biobank. *J. Gerontol. Ser. A* 2022, *in press*.

16. Mazidi, M.; Davies, I.G.; Penson, P.; Rikkonen, T.; Isanejad, M. Lifetime serum concentration of 25-hydroxyvitamin D 25 (OH) is associated with hand grip strengths: Insight from a Mendelian randomisation. *Age Ageing* **2022**, *51*, afac079.

17. Wind, A.E.; Takken, T.; Helders, P.J.M.; Engelbert, R.H.H. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *Eur. J. Pediatr.* **2010**, *169*, 281–287.

18. Kirwan, R.; Isanejad, M.; Davies, I.G.; Mazidi, M. Genetically Determined Serum 25-Hydroxyvitamin D Is Associated with Total, Trunk, and Arm Fat-Free Mass: A Mendelian Randomization Study. *J. Nutr. Health Aging* **2022**, *26*, 46–51.

19. Sha, T.; Li, W.; He, H.; Wu, J.; Wang, Y.; Li, H. Causal relationship of genetically predicted serum micronutrients levels with sarcopenia: A Mendelian randomization study. *Front. Nutr.* **2022**, *9*, 913155.

20. Cruz-Jentoft, A.J.; Bahat, G.; Bauer, J.; Boirie, Y.; Bruyère, O.; Cederholm, T.; Cooper, C.; Landi, F.; Rolland, Y.; Sayer, A.A.; et al. Sarcopenia: Revised European consensus on definition and diagnosis. *Age Ageing* **2019**, *48*, 16–31.

21. Dodds, R.M.; Granic, A.; Robinson, S.M.; Sayer, A.A. Sarcopenia, long-term conditions, and multimorbidity: Findings from UK Biobank participants. *J. Cachexia Sarcopenia Muscle* **2020**, *11*, 62–68.

22. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **2015**, *12*, e1001779.

23. Sutherland, J.; Zhou, A.; Hypponen, E. Muscle traits, sarcopenia, and sarcopenic obesity: A vitamin D Mendelian randomization study–Supplementary Appendix. Can be uploaded to external server if required.

24. Deurenberg, P.; Yap, M.; Van Staveren, W. Body mass index and percent body fat: A meta analysis among different ethnic groups. *Int. J. Obes.* **1998**, *22*, 1164–1171.

25. Townsend, P.; Phillimore, P.; Beattie, A. *Health and Deprivation: Inequality and the North*; Routledge: London, UK, 1988.

26. Revez, J.A.; Lin, T.; Qiao, Z.; Xue, A.; Holtz, Y.; Zhu, Z.; Zeng, J.; Wang, H.; Sidorenko, J.; Kemper, K.E.; et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. *Nat. Commun.* **2020**, *11*, 1647.

27. Zhou, A.; Selvanayagam, J.B.; Hyppönen, E. Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. *Eur. Heart J.* **2022**, *43*, 1731–1739.

28. Jiang, X.; O'Reilly, P.F.; Aschard, H.; Hsu, Y.-H.; Richards, J.B.; Dupuis, J.; Ingelsson, E.; Karasik, D.; Pilz, S.; Berry, D.; et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* **2018**, *9*, 260.

29. Royston, P.; Ambler, G.; Sauerbrei, W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int. J. Epidemiol.* **1999**, *28*, 964–974.

30. Staley, J.R.; Burgess, S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet. Epidemiol.* **2017**, *41*, 341–352.

31. Burgess, S.; Mason, A.M.; Grant, A.J.; Slob, E.A.; Gkatzionis, A.; Zuber, V.; Patel, A.; Tian, H.; Liu, C.; Haynes, W.G.; et al. Using genetic association data to guide drug discovery and development: Review of methods and applications. *Am. J. Hum. Genet.* **2023**, *110*, 195–214.

32. Burgess, S.; Davey Smith, G.; Davies, N.M.; Dudbridge, F.; Gill, D.; Glymour, M.M.; Hartwig, F.P.; Holmes, M.V.; Minelli, C.; Relton, C.L.; et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* **2020**, *4*, 186.

33. Burgess, S.; Gill, D. Genetic evidence for vitamin D and cardiovascular disease: Choice of variants is critical. *Eur. Heart J.* **2022**, *43*, 1740–1742.

34. McCormick, R.; Vasilaki, A. Age-related changes in skeletal muscle: Changes to life-style as a therapy. *Biogerontology* **2018**, *19*, 519–536.

35. Yeung, S.S.Y.; Reijnierse, E.M.; Pham, V.K.; Trappenburg, M.C.; Lim, W.K.; Meskers, C.G.M.; Maier, A.B. Sarcopenia and its association with falls and fractures in older adults: A systematic review and meta-analysis. *J. Cachexia Sarcopenia Muscle* **2019**, *10*, 485–500.

36. Bruyère, O.; Beaudart, C.; Ethgen, O.; Reginster, J.-Y.; Locquet, M. The health economics burden of sarcopenia: A systematic review. *Maturitas* **2019**, *119*, 61–69.

37. Pinedo-Villanueva, R.; Westbury, L.D.; Syddall, H.E.; Sanchez-Santos, M.T.; Dennison, E.M.; Robinson, S.M.; Cooper, C. Health Care Costs Associated with Muscle Weakness: A UK Population-Based Estimate. *Calcif. Tissue Int.* **2019**, *104*, 137–144.

38. Steffl, M.; Sima, J.; Shiells, K.; Holmerova, I. The increase in health care costs associated with muscle weakness in older people without long-term illnesses in the Czech Republic: Results from the Survey of Health, Ageing and Retirement in Europe (SHARE). *Clin. Interv. Aging* **2017**, *12*, 2003–2007.

39. Trantham, L.; Sikirica, M.V.; Candrilli, S.D.; Benson, V.S.; Mohan, D.; Neil, D.; Joshi, A.V. Healthcare costs and utilization associated with muscle weakness diagnosis codes in patients with chronic obstructive pulmonary disease: A United States claims analysis. *J. Med. Econ.* **2019**, *22*, 319–327.

40. Beaudart, C.; Rizzoli, R.; Bruyère, O.; Reginster, J.-Y.; Biver, E. Sarcopenia: Burden and challenges for public health. *Arch. Public Health* **2014**, *72*, 45.

41. Sousa, A.S.; Guerra, R.S.; Fonseca, I.; Pichel, F.; Ferreira, S.; Amaral, T.F. Financial impact of sarcopenia on hospitalization costs. *Eur. J. Clin. Nutr.* **2016**, *70*, 1046–1051.

42. Schrager, M.A.; Metter, E.J.; Simonsick, E.; Ble, A.; Bandinelli, S.; Lauretani, F.; Ferrucci, L. Sarcopenic obesity and inflammation in the InCHIANTI study. *J. Appl. Physiol.* **2007**, *102*, 919–925.

43. Silberstein, M. COVID-19 and IL-6: Why vitamin D (probably) helps but tocilizumab might not. *Eur. J. Pharmacol.* **2021**, *899*, 174031.

44. Kominsky, D.J.; Campbell, E.L.; Colgan, S.P. Metabolic shifts in immunity and inflammation. *J. Immunol.* **2010**, *184*, 4062–4068.

45. Siddik, M.A.B.; Shin, A.C. Recent Progress on Branched-Chain Amino Acids in Obesity, Diabetes, and Beyond. *Endocrinol. Metab.* **2019**, *34*, 234.

46. Katsuhara, S.; Yokomoto-Umakoshi, M.; Umakoshi, H.; Matsuda, Y.; Iwahashi, N.; Kaneko, H.; Ogata, M.; Fukumoto, T.; Terada, E.; Sakamoto, R.; et al. Impact of Cortisol on Reduction in Muscle Strength and Mass: A Mendelian Randomization Study. *J. Clin. Endocrinol. Amp; Metab.* **2022**, *107*, e1477–e1287.

47. Ko, D.H.; Kim, Y.K. The Prevalence of Metabolic Syndrome According to Grip Strength in Teenagers. *Children* **2021**, *8*, 108.

48. Li, D.; Guo, G.; Xia, L.; Yang, X.; Zhang, B.; Liu, F.; Ma, J.; Hu, Z.; Li, Y.; Li, W.; Jiang, J. Relative handgrip strength is inversely associated with metabolic profile and metabolic disease in the general population in China. *Front. Physiol.* **2018**, *9*, 59.

49. Sutandyo, N.; Rinaldi, I.; Sari, N.K.; Winston, K. Prevalence of Anemia and Factors Associated with Handgrip Strength in Indonesian Elderly Population. *Cureus* **2022**, *14*, e25290.

50. Gi, Y.-M.; Jung, B.; Kim, K.-W.; Cho, J.-H.; Ha, I.-H. Low handgrip strength is closely associated with anemia among adults: A cross-sectional study using Korea National Health and Nutrition Examination Survey (KNHANES). *PLoS ONE* **2020**, *15*, e0218058.

51. Malczewska-Lenczowska, J.; Sitkowski, D.; Surała, O.; Orysiak, J.; Szczepańska, B.; Witek, K. The Association between Iron and Vitamin D Status in Female Elite Athletes. *Nutrients* **2018**, *10*, 167.

52. Sakuma, K.; Yamaguchi, A. Sarcopenia and Age-Related Endocrine Function. *Int. J. Endocrinol.* **2012**, *2012*, 127362.

53. Zhang, X.; Wang, C.; Dou, Q.; Zhang, W.; Yang, Y.; Xie, X. Sarcopenia as a predictor of all-cause mortality among older nursing home residents: A systematic review and meta-analysis. *BMJ Open* **2018**, *8*, e021252.

54. Sutherland, J.P.; Zhou, A.; Hyppönen, E. Vitamin D Deficiency Increases Mortality Risk in the UK Biobank: A Nonlinear Mendelian Randomization Study. *Ann. Intern. Med.* **2022**, *175*, 1552–1559.

55. Sofianopoulou, E.; Kaptoge, S.K.; Afzal, S.; Jiang, T.; Gill, D.; Gundersen, T.E.; Bolton, T.R.; Allara, E.; Arnold, M.G.; Mason, A.M.; et al. Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: Observational and Mendelian randomisation analyses. *Lancet Diabetes Endocrinol.* **2021**, *9*, 837–846.

56. Dzik, K.P.; Kaczor, J.J. Mechanisms of vitamin D on skeletal muscle function: Oxidative stress, energy metabolism and anabolic state. *Eur. J. Appl. Physiol.* **2019**, *119*, 825–839.

57. Abrams, G.D.; Feldman, D.; Safran, M.R. Effects of vitamin D on skeletal muscle and athletic performance. *JAAOS—J. Am. Acad. Orthop. Surg.* **2018**, *26*, 278–285.

58. Bohannon, R.W. Grip Strength: An Indispensable Biomarker for Older Adults. *Clin. Interv. Aging.* **2019**, *14*, 1681–1691.

59. Batty, G.D.; Gale, C.R.; Kivimäki, M.; Deary, I.J.; Bell, S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: Prospective cohort study and individual participant meta-analysis. *BMJ* **2020**, *368*, m131.

60. Gkatzionis, A.; Burgess, S. Contextualizing selection bias in Mendelian randomization: How bad is it likely to be? *Int. J. Epidemiol.* **2019**, *48*, 691–701.

61. Hyppönen, E.; Vimaleswaran, K.S.; Zhou, A. Genetic Determinants of 25-Hydroxyvitamin D Concentrations and Their Relevance to Public Health. *Nutrients* **2022**, *14*, 4408.

62. DiaSorin. LIAISON—25 OH Vitamin D TOTAL Assay 2019. Available online: https://www.diasorin.com/sites/default/files/allegati_prodotti/25_oh_vit._d_total_m0870004213_e.pdf (accessed on 8 June 2020).

63. Vickers, D.; Rees, P. Creating the UK National Statistics 2001 output area classification. *J. R. Stat. Soc.* **2007**, *170*, 379-403.

64. Nuttall, F.Q. Body Mass Index. *Nutr. Today* **2015**, *50*, 117–128.

65. Burgess, S.; Thompson, S.G. Use of allele scores as instrumental variables for Mendelian randomization. *Int. J. Epidemiol.* **2013**, *42*, 1134–1144.

66. Palmer, T.M.; Sterne, J.A.; Harbord, R.M.; Lawlor, D.A.; Sheehan, N.A.; Meng, S.; Granell, R.; Smith, G.D.; Didelez, V. Instrumental Variable Estimation of Causal Risk Ratios and

Causal Odds Ratios in Mendelian Randomization Analyses. *Am. J. Epidemiol.* **2011**, *173*, 1392–1403.

67. Burgess, S.; Davies, N.M. Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet. Epidemiol.* **2016**, *40*, 597–608.

68. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data. *Genet. Epidemiol.* **2013**, *37*, 658–665.

69. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525.

70. Verbanck, M.; Chen, C.Y.; Neale, B.; Do, R. Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases. *Nat. Genet.* **2018**, *50*, 693–698.

71. Hartwig, F.P.; Davey Smith, G.; Bowden, J. Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption. *Int. J. Epidemiol.* **2017**, *46*, 1985–1998.

72. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314.

73. Kamat, M.A.; Blackshaw, J.A.; Young, R.; Surendran, P.; Burgess, S.; Danesh, J.; Butterworth, A.S.; Staley, J.R. PhenoScanner V2: An expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **2019**, *35*, 4851–4853.

Study 2 Supplementary

Muscle traits, sarcopenia, and sarcopenic obesity: a vitamin D Mendelian randomization study – Supplementary Appendix.

Table of Contents:

Measurement of vitamin D status	**03**
Covariates	**03**
Genetic instruments using 35 variants	**05**
Genetic instrument using 122 vitamin D GWAS variants	**05**
Linear Mendelian randomization	**06**
Leave-block-out analysis	**08**
Non-linear Mendelian randomization	**08**
Figure S1. Study population and participant inclusion criteria.	**10**
Figure S2. Mendelian randomization assumption diagram.	**11**
Figure S3. Schematic representation of Mendelian randomization analyses performed in the current study.	**12**
Figure S4. Distribution of the vitamin D genetic score (mapped on an X-axis by increasing number of alleles) in the UK Biobank (A), and the association of the vitamin D genetic score (in six groups – lowest to highest number of alleles) with measured 25(OH)D concentrations in the UK Biobank (B).	**13**
Figure S5. Association of the vitamin D genetic score with potential confounders across 100 strata of residuals of measured 25(OH)D.	**14**
Figure S6. Mean kilogram difference in arm skeletal muscle mass by probable sarcopenia status in UK Biobank participants.	**15**
Figure S7. Residual method (A) and doubly rank method (B) non-linear Mendelian randomization analysis of genetically predicted 25(OH)D on grip strength, projected on the measured 25(OH)D scale.	**16**
Figure S8. Selection of variants for the genetic instrument for measured 25(OH)D concentrations.	**17**
Figure S9. Distribution of the GRS-measured 25(OH)D association across 100 strata of residual measured 25(OH)D.	**18**
Table S1. Association of the vitamin D genetic score with potential confounders in the UK Biobank.	**19**
Table S2. Functional blocks used in the leave-block-out analyses.	**20**
Table S3. 122 SNP genetic score and leave block out analysis.	**21**
Table S4. Demographic characteristics of UK Biobank participants – including height and waist circumference.	**22**
Table S5. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.	**23**
References	**24**

Authors:

Joshua P Sutherland[1], Ang Zhou[1,2], Elina Hyppönen[1,2]

[1]Australian Centre for Precision Health, Clinical and Health Sciences, University of South Australia, Adelaide, Australia.

[2]South Australian Health and Medical Research Institute, Adelaide, Australia

Corresponding author: Prof Elina Hyppönen; Australian Centre for Precision Health, Clinical and Health Sciences, University of South Australia, Adelaide, Australia, GPO Box 2471, Adelaide SA 5001;

+61 8 830 22518; Elina.Hypponen@unisa.edu.au

Supplementary Methods:

Measurement of vitamin D status

The LIAISON XL 25(OH)D assay (DiaSorin, Stillwater, USA) is a magnetic microparticle separation technology, fully automated, chemiluminescence immunoassay device.[1] Baseline serum 25(OH)D concentrations were measured from blood samples with a mean fasting time of 3.8 hours (SD 2.4). Assay precision (CLSI EP5-A2), produced a coefficient of variation (CV) of 2.3% (intra-run) and a CV 7.8% (total), conducted using 6 samples and 2 controls.[1] The LIAISON XL has a functional sensitivity of ≤=4.0 ng/ml, in accordance with Clinical Laboratory and Standards Institute (CLSI) EP17-A protocols.[1] 25(OH)D reflect total concentrations, including both 25(OH)D_2 and 25(OH)D_3.

Covariates

As described in the primary manuscript, covariates used in adjustments were acquired at baseline – at the assessment centre most closely located to a respective participant's residential location – from self-reported, touchscreen questionaries (age, sex, physical activity, smoking, alcohol, birth location), residential address data (Townsend deprivation index)[2] physical assessments (height and waist circumference). The assessment centre at which baseline assessment occurred was itself used as a covariate, constructed as per the following: Leeds (n= 28,105), Bristol (n=27,983), Newcastle (n=23,755), Nottingham (n=21,864), Liverpool (n=20,438), Sheffield (n=19,352), Reading (n=19,297), Hounslow (n=13,377), Bury (n=18,690), Croydon (n=13,556), Birmingham (n=14,252), Middlesborough (n=13,806), Stoke (n=12,785), Glasgow (n=11,697), Cardiff (n=11,959), Edinburgh (n=11,327), Oxford (n=9,348), Manchester (n=8,524), Barts (n=5,257), Swansea (n=1,449), and Wrexham (n=460). Height data was acquired from a Seca 202 stadiometer. For Townsend deprivation index, each participant was assigned a score corresponding to socioeconomic data attributed to their postcode; this was derived from the preceding UK census of population and housing.[2,3] 25(OH)D and the remaining genetic related covariates were derived from baseline biological sample acquisition data (month in which blood sample

was taken, fasting time before blood sample was taken, and top 40 genetic principal components) and subsequent laboratory data [sample aliquots for 25(OH)D measurement and Single Nucleotide Polymorphism (SNP) array data]. Finally, two variables were used exclusively for presenting demographic characteristic purposes (**main text Table 1 and Table S4**), these being longitudinal location and BMI. Due to the relevance of UVB exposure in vitamin D analyses, assessment centre information was presented as longitudinal location categories: <=51° (Croydon, Barts, Hounslow, Reading, Oxford, Bristol, Cardiff, Swansea), 52°-53° (Birmingham, Stoke, Nottingham, Wrexham, Liverpool, Bury, Manchester, Sheffield, Leeds), and 54°-≥55° (Middlesborough, Newcastle, Glasgow, Edinburgh). Body mass index (BMI) was derived from height and weight measures collected at baseline (calculated as body weight [kilograms] divided by height squared [meters]).[4] Height and waist circumference were chosen for phenotypic adjustments – to account for body shape and obesity – rather than using body mass index (BMI), due to BMI's calculation explicitly including measure of skeletal muscle mass via body weight; BMI was nonetheless used for descriptive purposes, due its common use and familiarity. All phenotypic analyses were restricted to the sample with full information on covariates included in the full model (skeletal muscle mass: observations n = 291,847; grip strength: observations n = 296,013; probable sarcopenia controls n = 272,555 / cases n = 23,458); and sarcopenic non-obesity controls n = 272,555 / cases n = 7,580. For the genetic analyses, we included sex, age, vitamin D month, assessment location, birth location, principal components, fasting, vitamin D aliquot, and SNP array (skeletal muscle mass: observations n = 288,697; grip strength observations n = 293,298; probable sarcopenia controls n = 269,088 / cases n = 24,210; and sarcopenic non-obesity controls n = 269,088 / cases n = 7,579). (**Figure S1**).

Genetic instruments using 35 variants

We constructed a weighted genetic score wherein common autosomal SNPs (i.e., minor allele frequency >5%) were collated from a genome-wide association analysis (GWAS) for measured 25(OH)D serum concentration in UK Biobank; 143 independent loci were identified.[5] Subsequently, we replicated our SNP analyses in the previous SUNLIGHT consortium GWAS,[6] and this left 35 SNPs that showed a consistent direction and a p value of <0.05 (**Figure S8**). The process of restricting our instrument to replicated SNPs has enabled us to ensure the observed GWAS signals are robust. Additionally, we were able to use the SUNLIGHT consortium meta-analyses derived effect estimates [for the SNP association with measured 25(OH)D] as the individual SNP weights; this allowed us to avoid bias due to sample overlap.[7] For our vitamin D genetic score, the independently derived weighted average of the number of measured 25(OH)D-increasing alleles for an individual were computed, and then multiplied by the number of available variants.

Genetic instrument using 122 vitamin D GWAS variants

We constructed an alternative genetic instrument for sensitivity analyses, in which a broader set of variants (122 autosomal single nucleotide polymorphisms, SNPs) that associated with measured 25(OH)D concentrations were included (Figure S8). As these 122 SNPs were discovered in the UK Biobank,[5] we used the 10-fold cross-validation approach[7] to construct the 122 SNP version of the vitamin D genetic score; this was done to mitigate any potential bias caused by using the internal weights. This process involved the random division of samples into 10 equal sub-samples, with weights for each sub-sample taken from the other 9 sub-samples. Information for the 122 SNPs can be found in **Table S5**.

Linear Mendelian randomization:

In the primary Genetic Score One Sample method, linear MR estimate is computed using the ratio of coefficients method[8] where the vitamin D genetic score – measured 25(OH)D and the vitamin D genetic score – outcome association estimates are computed and then used to calculate the ratio estimator. The vitamin D genetic score-25(OH)D and vitamin D genetic score-outcome associations included adjustment for: age, sex, assessment centre, birth locations, SNP array, top 40 genetic principal components, and nuisance factors which could affect measured 25(OH)D, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

In order to address potential horizontal pleiotropy, we also use five SNP-based two-sample methods, including inverse-variance weighted method (IVW), Mendelian randomization Egger method, weighted median Mendelian randomization method, weighted mode method, and Mendelian randomization pleiotropy residual sum and outlier (MR Presso) test. These five methods are complementary to one another, in that each one is robust to different patterns of horizontal pleiotropy; to this end, a good agreement across the methods is suggestive of robust evidence of causal association. In contrast to the above outlined genetic score-based analysis, the SNP-based two-sample methods use SNP-level information and takes SNP-exposure and SNP-outcome association estimates as inputs (**Figure S3**). In our two-sample MR analyses, SNP-25(OH)D and SNP–outcome association estimates were all taken from UK Biobank. For grip strength and skeletal muscle mass, we applied the split-sample strategy to avoid bias due to overlapping samples.[9] More specifically, the full sample is randomly split into two subsamples of equal size (Samples A and B); the MR estimate is calculated for each sample (MR_A and MR_B) and then combined to obtain the overall estimate ($MR_{combined}$). MR_A is computed using SNP–outcome association estimates from Sample A and SNP–exposure association estimates from Sample B, whereas SNP-outcome association estimates from Sample B and SNP–exposure association estimates from Sample A are used to compute MR_B. MR_A and MR_B are then combined in a fixed-effects meta-analysis to compute $MR_{combined}$. For the sarcopenia analyses, we implemented a different strategy to avoid

potential bias due to overlapping samples,[9] where we computed SNP-Probable sarcopenia association estimates in the full sample but took SNP-25(OH)D association estimates from controls only. All five two-sample MR methods were conducted using the 35 vitamin D variants that were used to construct the vitamin D genetic score (Table S5). In IVW analysis, it is assumed that there is no pleiotropy of variants, or the net pleiotropy is zero. The estimator is obtained by the weighted regression of the effect estimates of variant-outcome association on the effect estimates of variant-exposure association,[10] with the weight being the inverse variance of variant's ratio estimate. Like the IVW method, Mendelian randomization Egger also expects that the regression model additionally included an intercept term, which represents the average pleiotropic effect.[11] Conversely, the Mendelian randomization Egger method can account for directional pleiotropy, however, additional untestable assumptions are introduced, which state that pleiotropic effect must be independent of variant-exposure association. In Mendelian randomization Presso,[12] the IVW method is performed iteratively, wherein each variant is, in turn, omitted from the analysis. The residual sum of squares (RSS) is calculated at each iteration and used as a heterogeneity measure for the ratio estimate. The corresponding variant is considered as an outlier, if the observed RSS is substantially large in the null distribution of the expected RSSs. The causal estimate in MR Presso is obtained by the IVW method with all outlying variants removed. In the weighted mode method, variants (which are weighted by the inverse variance of their ratio estimates) are clustered into groups by the similarity of their ratio estimates.[13] The clustering is performed using a normal kernel density function, with its spread depending on a bandwidth parameter. The causal effect is estimated based on the cluster with the largest weight. An unbiased causal effect will be returned in the weighted mode method if the variants within the cluster with the largest weight are valid instruments. In the weighted median-based approach,[14] variants are ranked by their ratio estimates, with the causal estimate being the median of the distribution. The weighted median method will return a consistent estimate if at least 50% of the weight comes from valid instruments – median is taken from a distribution of the ratio estimates in which variants with more precise ratio estimates receive more weight.

Leave-block-out analysis

In our leave-block-out analysis, variants were grouped into functional blocks based on the traits that they are associated with from the PhenoScanner search. For sensitivity purposes, we then removed one block of variants at a time and repeated the non-linear Mendelian randomization analysis with the vitamin D genetic score that was constructed using the non-removed, remaining variants. This process was iterated through all functional blocks.

PhenoScanner V2 is a publicly available database that has been curated from the results of large-scale genetic association studies.[15] On 09/09/2021, we searched for evidence of variant-disease/trait associations in the PhenoScanner V2 database for all 35 variants. At the GWAS threshold (P = <5E-08), 20 variants were associated with traits besides serum 25(OH)D (**Table S2**). We grouped variants into 4 functional blocks, based on their associated traits, including blocks for blood, lipids/metabolic, renal traits – and for variants whose associated traits did not fall into one of these 3 blocks, they were grouped together as the 'unclassified' block (Table S2-3).

Non-linear Mendelian randomization:

We examined non-linearity of genetic association using two available methods, including the residual method[16] and doubly-ranked method.[17] Both approaches involve stratifying the sample into 100 strata, but employ different stratification strategies. For the residual method, [16] strata are formed using 100 quantiles of residual measured 25(OH)D, which is defined as measured 25(OH)D minus the variation induced by the genetic score (correlation between residual measured 25(OH)D and measured 25(OH)D, r= 0.986).

For the doubly ranked method, [17] strata are formed by first ranking participants into pre-strata based on their level of genetic score and then by ranking participants within each pre-stratum according to their level of measured 25(OH)D. In both non-linear MR methods, stratum-specific MR estimate is computed using the ratio of coefficients method (as described in the linear MR analysis, see section linear MR), and then meta-regressed on the mean stratum measured 25(OH)D concentration. The meta-regression was performed by fitting a range of fractional polynomial exposure-outcome models of degree 1 and degree 2, with the best-fitting model being selected based on the likelihood ratio test. Conventionally, if the best-fitting fractional polynomial model of degree 1 has a better fit than the linear model[16] (i.e. $P_{Non\text{-}linear} < 0.05$), then a non-linear effect is evident – however, as per our use of the Bonferroni correction (outlined in main text) we used a stricter threshold for our analyses. In residual method non-linear MR analysis it is assumed that the association of genetic instrument with the exposure is constant over the entire distribution of exposure [here, measured 25(OH)D], and to test this assumption we computed the estimate for the association of the vitamin D genetic score and measured 25(OH)D in each of 100 strata and then examined the heterogeneity between strata using the trend test and Cochran Q test.[16] As we observed an apparent violation in the constant genetic effect assumption arising from 1st and 100th strata (**Figure S9**), when performing the non-linear analysis using the residual method, we excluded these two outlying strata to minimize the impact of any potential violation in the constant genetic effect assumption.

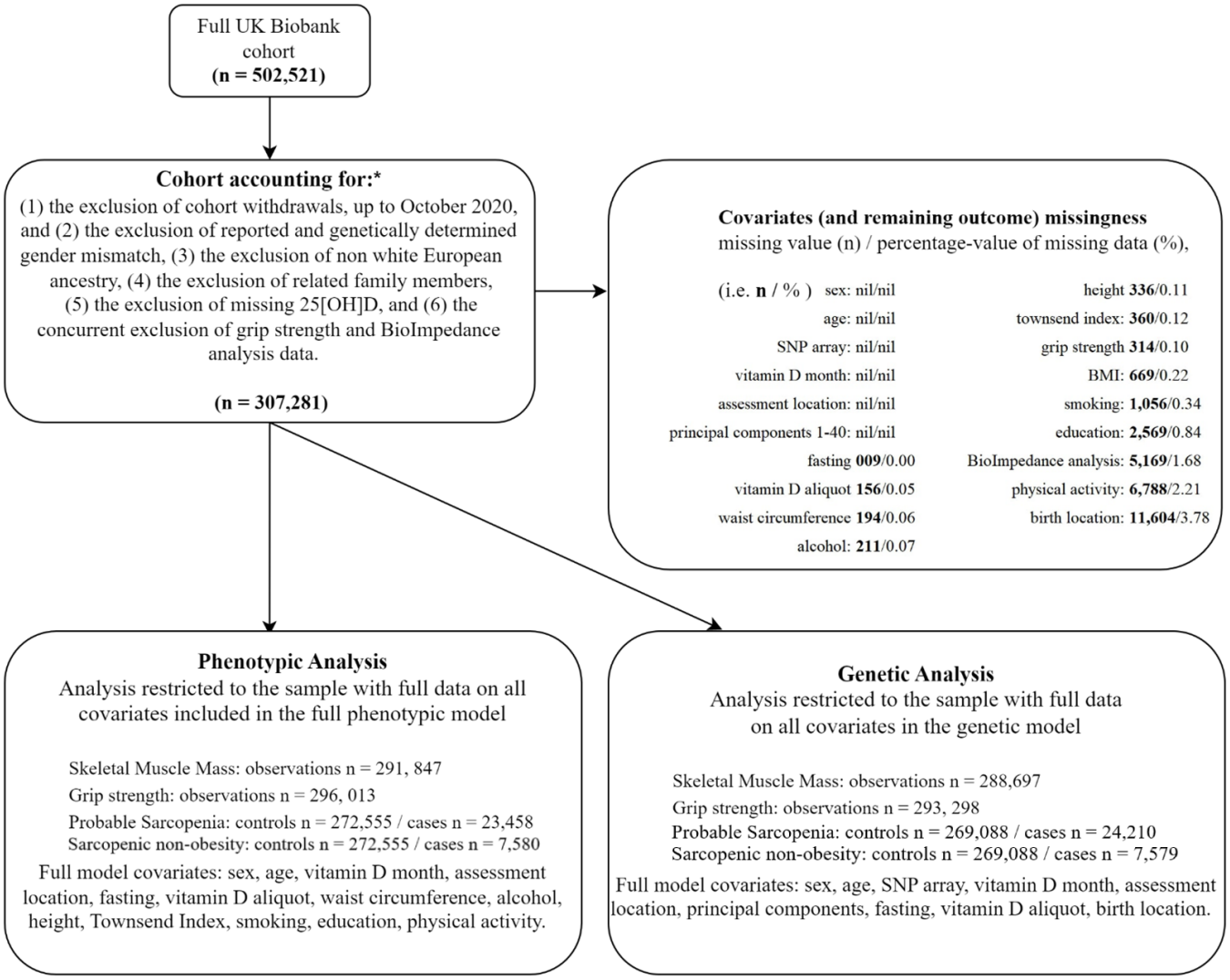


Figure S1. Study population and participant inclusion criteria *Exclusion was done in a sequential order.

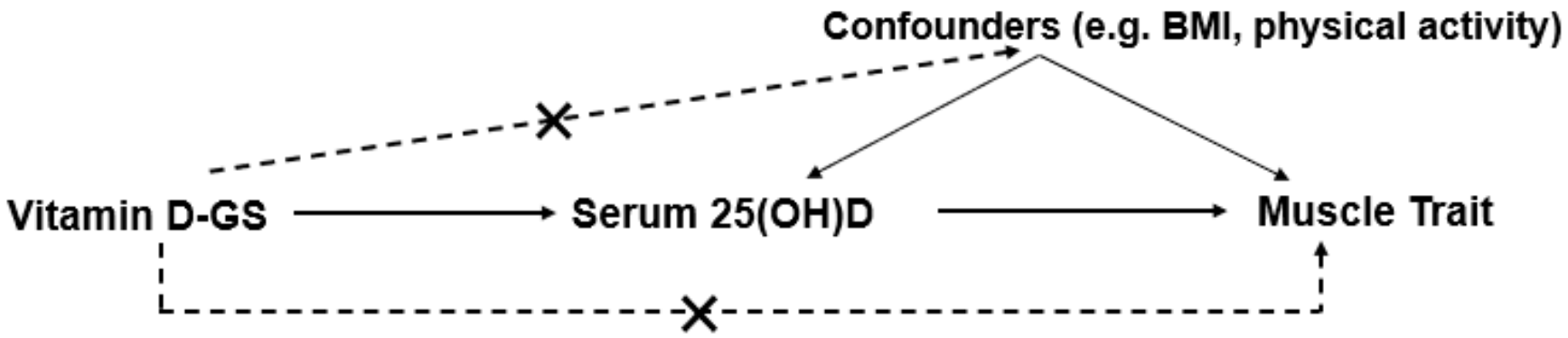


Figure S2. Mendelian randomization assumption diagram.

Valid causal inference from the Mendelian randomization analysis relies on 3 key assumptions:[18] 1) vitamin D genetic score associates with measured 25(OH)D concentrations; 2) vitamin D genetic score has no direct effect on outcome; 3) vitamin D genetic score does not associate with confounders of measured 25(OH)D and outcome. Vitamin D-GS: vitamin D genetic score.

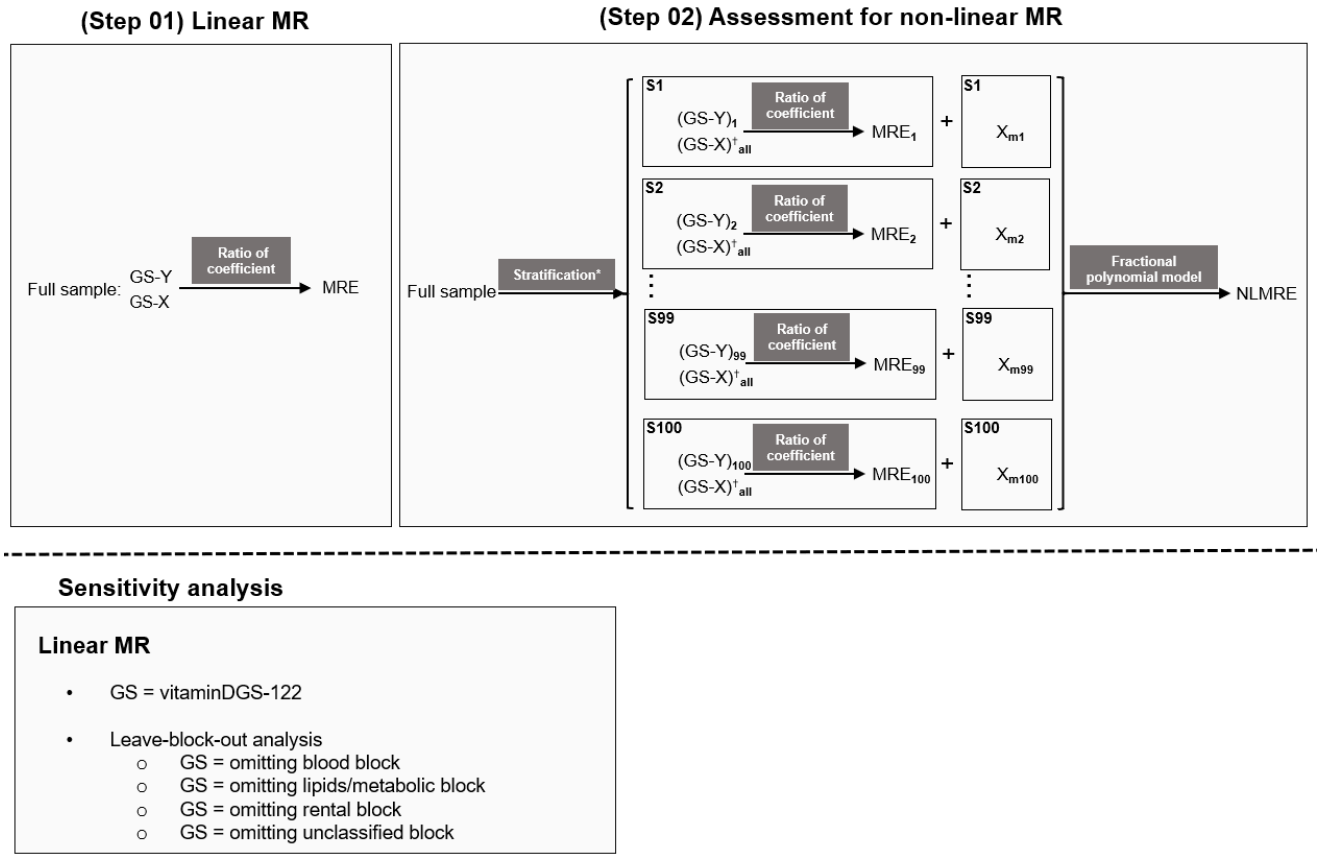
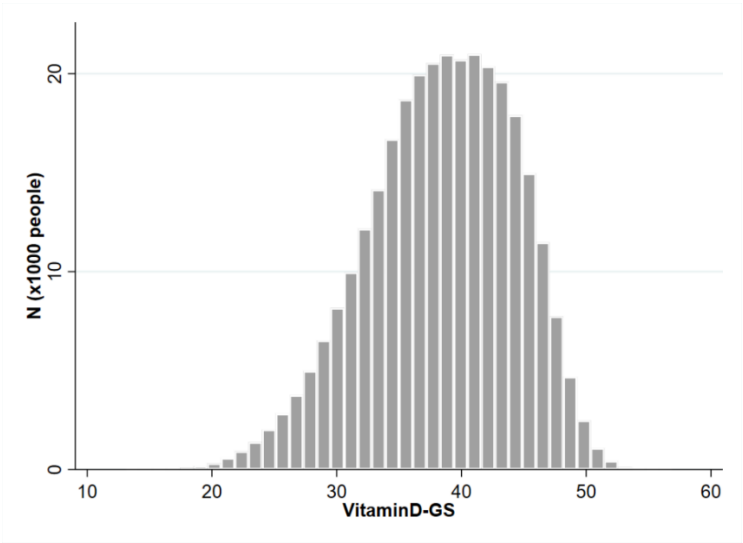


Figure S3. Schematic representation of Mendelian randomization analyses performed in the current study.

MR: Mendelian randomization; GS: genetic score; SNP: single nucleotide polymorphism; S_n: the nth stratum; MRE: MR estimates; NLMRE: non-linear MR estimates; GS-X: beta of GS-exposure association; GS-Y: beta of GS-outcome association; SNP-X: beta of SNP-exposure association; SNP-Y: beta of SNP-outcome association; *The residual method and doubly ranked method use different stratification methods. For the residual method, strata were formed using quartiles of residual measured 25(OH)D. For the doubly ranked method, strata were formed by first ranking participants into pre-strata based on their level of genetic score and then by ranking participants within each pre-stratum according to their level of measured 25(OH)D. Stratification using residual measured 25(OH)D; In non-linear MR: $(GS\text{-}X)^{\dagger}_{all}$: for the residual method GS-X is taken from the full sample, while for the doubly ranked method, GS-X is taken from the corresponding stratum; $(GS\text{-}Y)_n$: GS-Y from the nth stratum; X_{mn}: median of exposure in the nth stratum; In secondary analysis, stratified MR: S1: < 25 nmol/L; S2: 25-49.9 nmol/L; S3: 50-74.9 nmol/L; S4: >=75 nmol/L. In sensitivity analysis, stratified MR: $S1^{\$}$: < 25 nmol/L; $S2^{\$}$: >=25 nmol/L; S3: 50-74.9 nmol/L; S4: >=75 nmol/L $(SNP\text{-}Y)_n$: SNP-Y from the nth stratum; $(SNP\text{-}X)_{control}$: for sarcopenia analyses, SNP-X from non sarcopenia case individuals in the full sample; †Two-sample methods include Inverse Variance Weighted MR, MR-Egger, Weighted median MR, Weighted mode MR, and MR PRESSO.

A



B

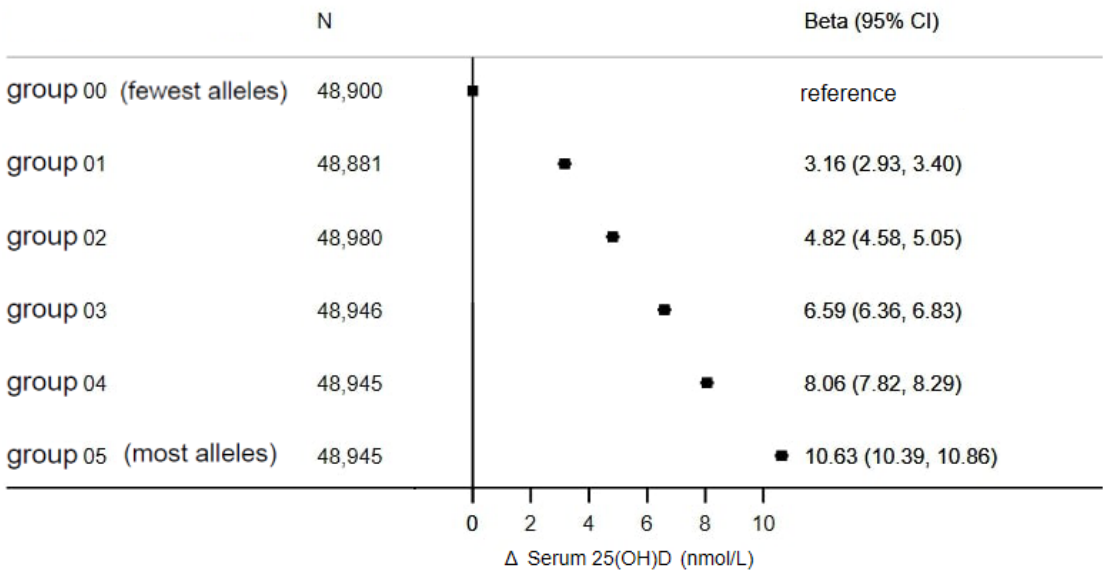


Figure S4. Distribution of the vitamin D genetic score (mapped on an X-axis by increasing number of alleles) in the UK Biobank (A), and the association of the vitamin D genetic score (in six groups – lowest to highest number of alleles) with measured 25(OH)D concentrations in the UK Biobank (B).

The association was tested using the linear regression, with the model adjusted for age, sex, genotyping array, birth location, assessment center location, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. Error bars are 95% confidence intervals.
Vitamin D-GS: Vitamin D Genetic Score

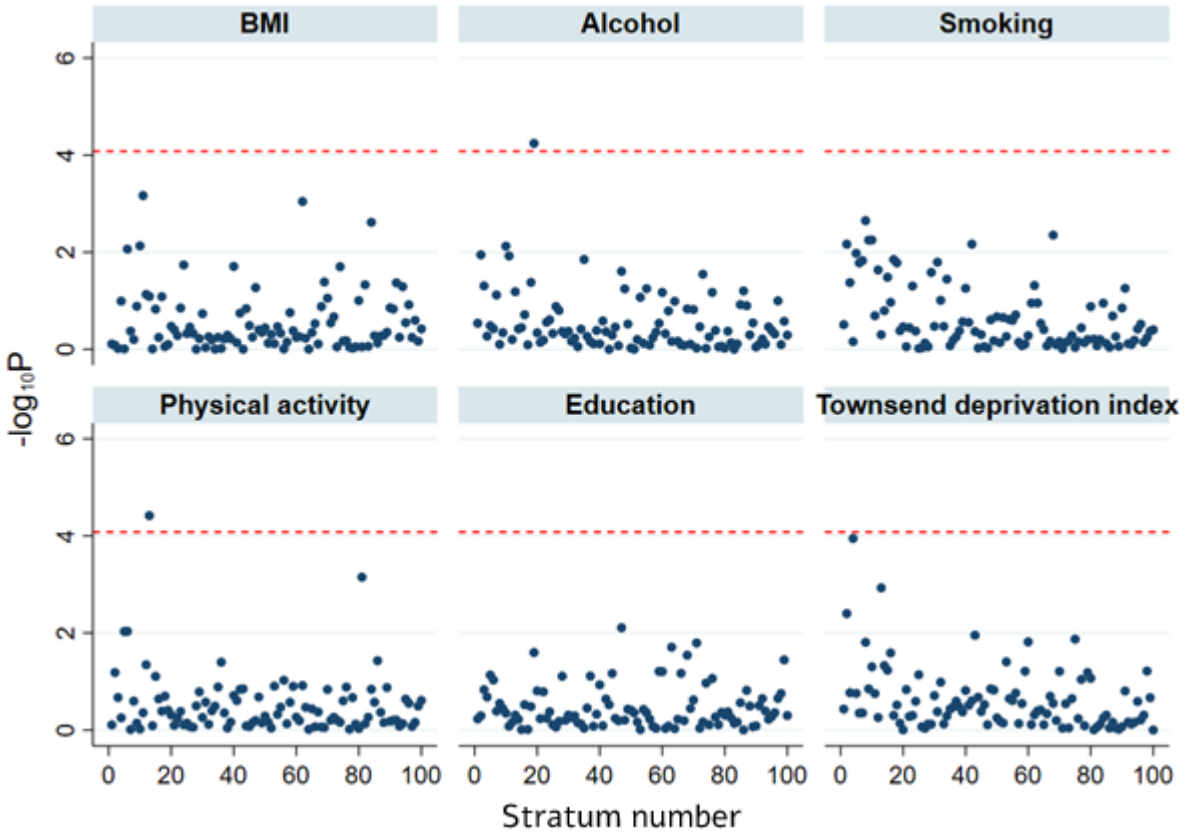


Figure S5. Association of the vitamin D genetic score with potential confounders across 100 strata of residual measured 25(OH)D.

Potential confounders include BMI, alcohol intake, smoking status, physical activity level, education attainment and Townsend deprivation index. $P_{threshold}$ = 0.05/(6 confounders x 100 strata) = 8.3E-5. Dashed line represents -$\log_{10}$P at $P_{threshold}$.



Figure S6. Mean kilogram difference in arm skeletal muscle mass by probable sarcopenia status in UK Biobank participants.
SMM: Skeletal Muscle Mass

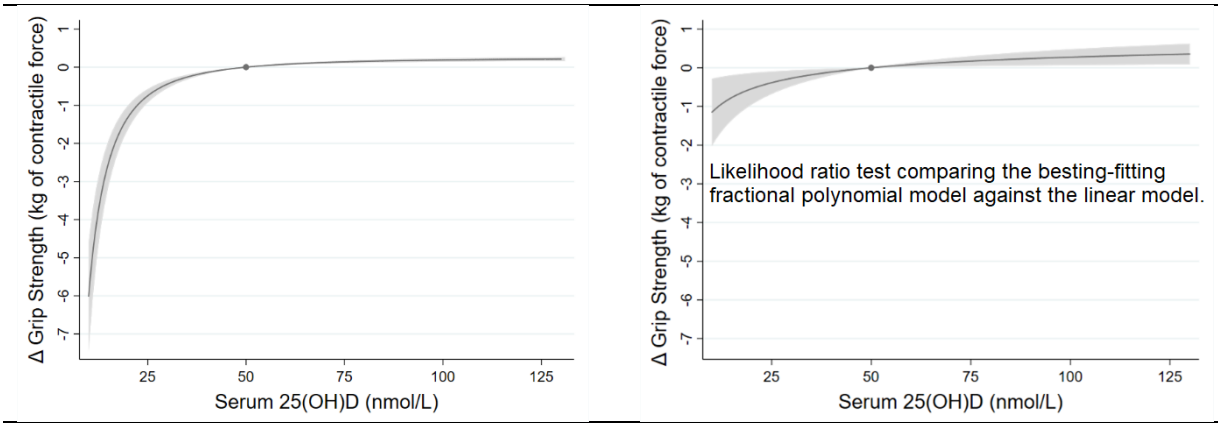


Figure S7. Residual method (A) and doubly rank method (B) non-linear Mendelian randomization analysis of genetically predicted 25(OH)D on grip strength, projected on the measured 25(OH)D scale.

In both methods the sample is stratified into 100 strata. For the residual method we removed the 1st and 100th strata because an apparent violation in the constant genetic effect assumption, as per Figure S9. Refer to Figure 1, main text, for full accompanying information on genetic analysis model adjustments and figure elements.

*The doubly ranked method, it may have a limited statistical power to detect an effect at low 25(OH)D concentration. We estimated statistical power (with respect to grip strength) across 25(OH)D distribution by calculating stratum-specific power of detecting a linear genetic effect across 4 strata of 25(OH)D. Statistical power (PMID: 24608958) of detecting a 0.022 SD change in grip strength by SD increase in 25(OH)D from the lowest to the highest stratum was 18.1%, 29.3%, 39.9% and 40.2%, respectively. The estimate for the association for which the power was calculated was taken from the linear MR analysis and re-scaled to reflect SD change in grip strength per SD increase in 25(OH)D (SD for 25(OH)D = 20.96 nmol/L, SD for grip strength = 11.03 kilograms of contractile force). The differences in statistical power across strata are driven by differences in instrument strength (R2), which are 0.031, 0.057, 0.083, and 0.083 from the lowest to the highest stratum, respectively.

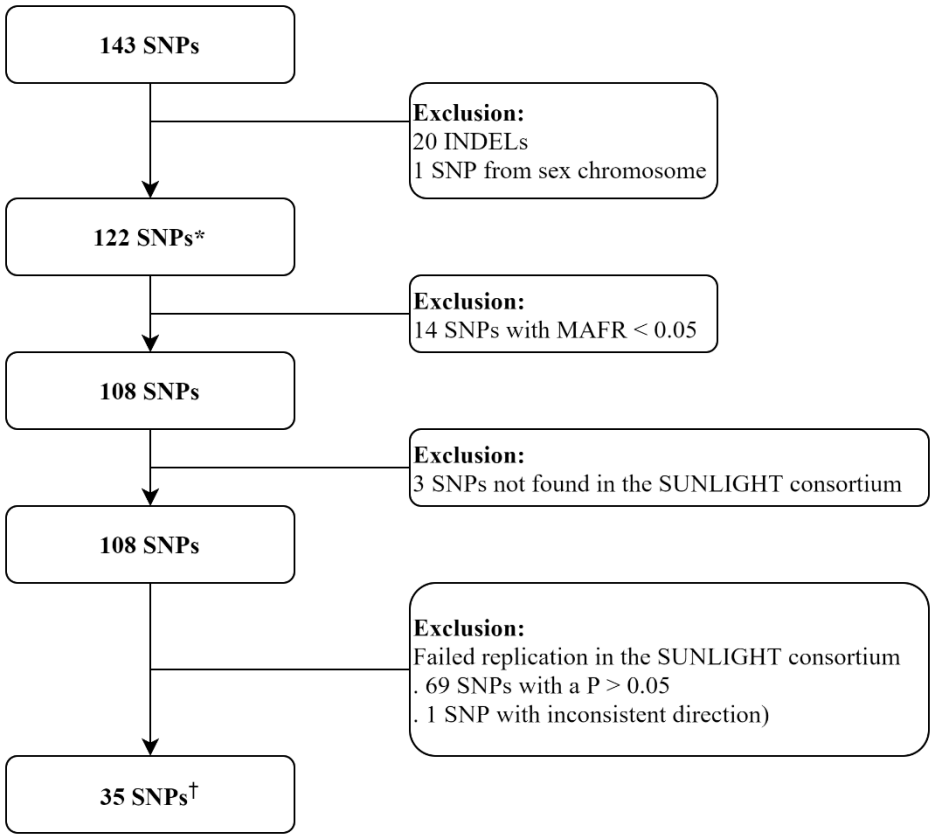


Figure S8. Selection of variants for the genetic instrument for measured 25(OH)D concentrations.

SNP: single nucleotide polymorphism. *SNPs used for the vitamin D genetic score in the primary analysis; †SNPs used for the 122 SNP version of the vitamin D genetic score in the sensitivity analysis. INDEL: insertion and deletion; MAF: minor allele frequency.

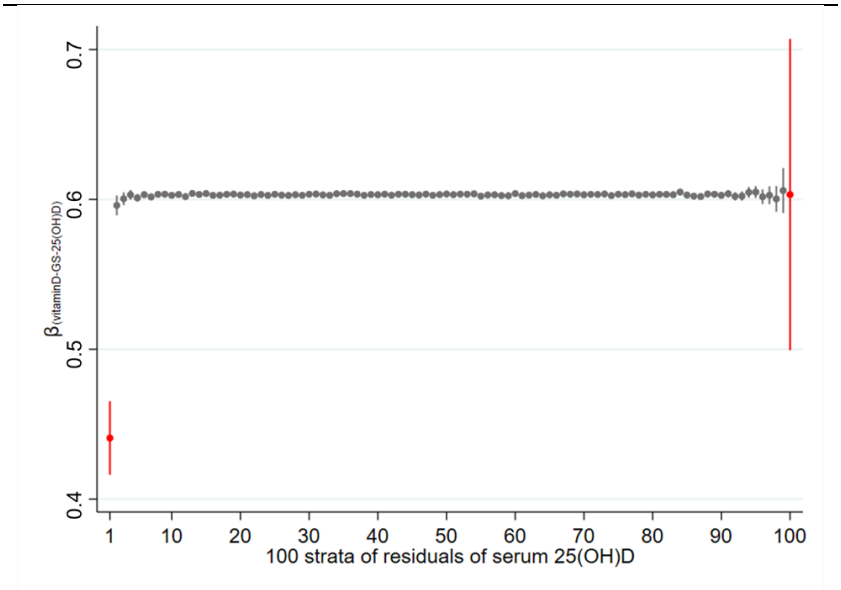


Figure S9. Distribution of the GRS-measured 25(OH)D association across 100 strata of residual measured 25(OH)D. Vitamin D-GS: vitamin D genetic score

Table S1. Association of the vitamin D genetic score with potential confounders in the UK Biobank.

		Vitamin D-GS
	N (%)	Mean (S.D.)
BMI		
Lowest 25% - 12.1-24.1	76,640 (24.9)	38.41 (5.87)
Middle 50% - 24.1-29.8	153,321 (50.0)	38.42 (5.88)
Highest 25% - 29.8-74.7	76,675 (25.0)	38.45 (5.86)
P[1]		0.88
Smoking		
Non-smokers	167,703 (54.7)	38.42 (5.86)
Ex-smokers	108,118 (35.3)	38.44 (5.89)
Current smokers	30,719 (10.0)	38.39 (5.89)
P[1]		0.18
Alcohol intake		
Daily	65,542 (21.3)	38.34 (5.88)
1 to 4 times week	155,608 (50.6)	38.45 (5.87)
1 to 3 times month	34,098 (11.1)	38.39 (5.87)
Special occasion	32,179 (10.5)	38.36 (5.90)
Never	19,963 (6.5)	38.51 (5.85)
P[1]		0.02
Physical activity		
Low	92,012 (30.6)	38.44 (5.87)
Moderate	149,205 (49.6)	38.42 (5.88)
High	59,561 (19.8)	38.42 (5.87)
P[1]		0.67
Education		
None	52,193 (17.1)	38.44 (5.86)
NVQ/CSE/A-levels	109,099 (35.8)	38.44 (5.88)
Degree/professional	143,735 (47.1)	38.42 (5.88)
P[1]		0.55
Townsend deprivation index quartiles		
Q1 lowest	76,793 (25.0)	38.41 (5.86)
Q2	76,821 (25.0)	38.46 (5.89)
Q3	76,815 (25.0)	38.42 (5.87)
Q4 highest	76,811 (25.0)	38.43 (5.88)
P[1]		0.38

NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; SD, standard deviation; Q, quartiles. Threshold for statistical significance = 0.05/6(confounders) = 0.008. P[1] Values have been adjusted for age, sex, genotyping array, birth location, and assessment center. Vitamin D-GS: vitamin D genetic score

Table S2. Functional blocks used in the leave-block-out analyses.

Functional block[a]	SNP	Gene	Traits
Blood traits	rs78151190	CARMIL1 (LRRC16A)	Various, incl. hemoglobin, platelet count, reticulocyte count
	rs1352846	GC	White cell count, granulocyte count
	rs1047891	CPS1	Various, incl. amino acid levels, platelet count, hemoglobin, metabolite levels
	rs72997623	DGAT2	Hemoglobin
	rs7528419	CELSR2	Blood protein levels, progranulin levels
	rs6672758	GALNT2	Platelet count, red cell distribution
Lipids/Metabolic	rs76798800	DCST2	Various, incl. fat-free mass, birth weight, height
	rs78151190	CARMIL1 (LRRC16A)	HbA1c, pulse, blood pressure
	rs77924615	PDILT	Blood pressure
	rs261291	ALDH1A2	Cholesterol, lipid metabolism
	rs727857	LINC01122	Fat mass, BMI, impedance, weight
	rs1047891	CPS1	Various, incl. weight, impedance, fat- free mass cholesterol
	rs12794714	CYP2R1	Hip circumference
	rs72997623	DGAT2	HDL cholesterol
	rs7528419	CELSR2	Angina pectoris, coronary artery disease, cholesterol, statin use
	rs75741381	VGF	Impedance
	rs6782190	CADM2	Fat free mass, BMI, impedance
Renal	rs77924615	PDILT	Glomerular filtration rate
	rs1047891	CPS1	Creatinine, chronic kidney disease
	rs17216707	CYP24A1	Creatinine, glomerular filtration rate
Unclassified	rs6782190	CADM2	Nervous feelings, alcohol intake, smoking, risk taking, number of children fathered
	rs61891388	RP11-867G23.13	Qualifications
	rs12803256	AP002387.1	Population differentiation
	rs1047891	CPS1	Headache
	rs212100	SULT2A1	Cholelithiasis, dehydroepiandrosterone sulphate
	rs7522116	FOXO6	Intelligence, qualifications
	rs78151190	CARMIL1 (LRRC16A)	Disorders of mineral metabolism
	rs261291	ALDH1A2	Age-related macular degeneration
	rs12056768	LINC00536	Hair or balding pattern: pattern 4
	rs1933064	FLG-AS1	Atopic dermatitis

[a]Functional blocks identified using trait associations identified through PhenoScanner V2[16].

Table S3. 122 SNP genetic score and leave block out analysis.

Vitamin D genetic score / leave-out block	SNPs excluded from the 35 SNP genetic score	Grip strength, Beta: 95% CI	P Value	Skeletal Muscle Mass, Beta: 95% CI	P Value
122 SNP Genetic score	**N/A**	0.01: 0.05, 0.16	0.0004	0.004: -0.003, 0.01	0.24
35 SNP Genetic score	**N/A**	0.11: 0.04, 0.19	0.002	0.01: 0.003, 0.02	0.009
Blood traits block	rs1047891, rs1352846, rs6672758, rs72997623, rs7528419,rs78151190	0.17: 0.06, 0.27	0.002	0.01: 0.002, 0.03	0.0001
Lipids / Metabolic block	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs7574138, rs6782190	0.12: 0.04, 0.20	0.002	0.01: 0.005, 0.02	0.004
Renal block	rs1047891, rs17216707, rs77924615	0.12: 0.05, 0.19	0.001	0.01: 0.003, 0.02	0.007
Unclassified block	rs6782190, rs61891388, rs12803256, rs1047891, rs212100, rs7522116, rs78151190, rs261291, rs12056768, rs1933064	0.10: 0.02, 0.18	0.01	0.01: 0.003, 0.02	0.01
Renal or Metabolic block	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs75741381, rs6782190, rs77924615	0.12: 0.05, 0.20	0.002	0.01: 0.002, 0.02	0.02

Adjustment includes age, sex, assessment center, birth location, SNP array, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

Table S4. Demographic characteristics of UK Biobank participants – including height and waist circumference.

		25(OH)D	25(OH)D <25 nmol/L	Grip strength[a]	Probable sarcopenia	Sarcopenic obesity	Arm skeletal muscle mass[b]
		N= 307,281	*N*=36,009	*N*=306,967[c]	*N*=25,414	*N*=16,520	*N*=302,112[d]
	N(%)	Mean (SD)	%	Mean (SD)	%	%	Mean (SD)
All	307,281	49.82 (20.96)	11.70	31.04 (11.03)	8.28	5.54	5.51 (1.58)
Sex							
Men	144,538 (47.4)	49.86 (21.03)	11.60	39.61 (8.74)	6.39	3.98	6.92 (1.09)
Women	162,743 (53.0)	49.78 (20.89)	11.80	23.43 (6.22)	9.96	6.94	4.26 (0.60)
Age							
<60	169,594 (55.2)	48.38 (21.12)	13.46	32.48 (11.18)	5.46	3.40	5.56 (1.66)
≥60	137,687 (44.8)	51.57 (20.62)	9.54	29.28 (10.58)	11.75	8.23	5.44 (1.48)
Height							
Low, 75-163.9 cm	94,144 (30.67)	49.60 (21.00)	12.11	22.54 (6.41)	13.31	9.27	4.14 (0.65)
Mid, 164-172.9 cm	102,958 (33.54)	50.00 (21.00)	11.62	29.46 (9.02)	7.66	5.25	5.17 (1.15)
High, 173-209 cm	109,843 (35.79)	49.89 (21.00)	11.38	39.84 (9.31)	4.44	2.70	7.00 (1.21)
Missing	336 (0.11)	41.01 (21.00)	26.49	23.52 (13.70)	45.51	9.90	5.34 (1.48)
Waist Circumference							
Low, 20-83.9 cm	99,707 (32.45)	53.02 (21.67)	9.52	25.96 (8.27)	7.90	2.80	4.24 (0.83)
Mid, 84-95.9 cm	103,030 (33.53)	50.64 (20.75)	10.50	32.22 (11.22)	7.83	5.51	5.51 (1.26)
High, 96-197 cm	104,350 (33.96)	45.96 (19.82)	14.95	34.76 (11.33)	9.02	8.09	6.73 (1.45)
Missing	194 (0.06)	39.97 (21.33)	29.90	23.28 (13.05)	38.34	4.03	5.90 (1.54)
BMI							
Low 25%, 12.1-24.0	76,633 (24.9)	53.04 (22.09)	10.19	28.62 (9.62)	7.61	1.21	4.53 (1.13)
Mid 50%, 24.1-29.8	153,312 (50.0)	50.95 (20.68)	10.05	32.19 (11.17)	7.55	5.46	5.57 (1.44)
High 25% 29.8-74.7	76,667 (25.0)	44.42 (19.23)	16.40	31.23 (11.63)	10.18	9.77	6.36 (1.71)
Missing	669 (0.2)	40.69 (21.11)	26.01	25.51 (12.86)	33.73	4.33	5.34 (1.48)
Location[e]							
South, ≤51° Lat	102,226 (33.3)	51.43 (20.49)	9.29	30.77 (10.84)	8.35	5.20	5.53 (1.58)
Mid, 52°-53° Lat	144,470 (47.0)	49.93 (20.88)	11.35	31.10 (11.11)	8.42	5.80	5.51 (1.58)
North, 54°- ≥55° Lat.	60,585 (19.7)	46.82 (21.59)	16.62	31.36 (11.16)	7.83	2.49	5.45 (1.58)
Smoking							
Non-smokers	167,537 (54.5)	50.04 (20.63)	10.92	30.48 (11.00)	8.07	5.22	5.36 (1.55)
Ex-smokers	108,015 (35.2)	50.77 (21.04)	10.66	31.61 (10.99)	8.37	6.05	5.69 (1.60)
Current smokers	30,673 (10.0)	45.22 (21.80)	19.63	32.23 (11.17)	8.91	5.31	5.66 (1.60)
Missing	1,056 (0.3)	50.17 (21.77)	12.41	28.95 (11.14)	14.58	11.22	5.57 (1.60)
Alcohol							
Daily	65,476 (21.3)	51.22 (21.51)	11.03	33.09 (10.82)	6.51	3.93	5.71 (1.56)
1 to 4 times wk	155,474 (50.6)	50.87 (20.80)	10.19	31.85 (11.01)	7.11	4.67	5.58 (1.60)
1 to 3 times mo	34,061 (11.1)	48.08 (20.26)	12.91	29.56 (10.73)	8.46	5.85	5.35 (1.58)
Special occasion	32,125 (10.5)	46.18 (20.40)	15.74	27.05 (10.43)	12.85	9.37	5.13 (1.48)
Never	19,934 (6.5)	45.91 (20.95)	17.06	27.07 (10.66)	15.47	11.11	5.19 (1.45)
Missing	211 (0.07)	44.97 (21.05)	16.59	28.54 (11.32)	17.54	13.86	5.39 (1.57)
Physical activity							
Low	91,911 (29.9)	46.30 (20.20)	15.18	29.78 (10.96)	10.39	7.49	5.47 (1.59)
Moderate	149,064 (48.5)	50.60 (20.80)	10.50	31.31 (10.88)	7.17	4.65	5.47 (1.57)
High	59,518 (19.4)	54.01 (21.41)	8.11	32.65 (11.09)	6.37	3.78	5.62 (1.60)
Missing	6,788 (2.2)	43.33 (21.32)	22.48	28.32 (12.19)	20.78	14.80	5.77 (1.70)
Education							
None	52,119 (17.0)	50.40 (21.43)	11.89	28.56 (10.98)	14.47	10.74	5.43 (1.55)
NVQ/CSE/A-Lev.	109,007 (35.5)	50.55 (21.19)	11.22	31.02 (11.18)	7.90	5.33	5.51 (1.61)
Deg./professional	143,586 (46.7)	49.04 (20.57)	12.00	31.99 (10.79)	6.23	3.79	5.54 (1.57)
Missing	2,569 (0.84)	50.33 (20.90)	11.60	29.48 (11.22)	13.36	9.04	5.52 (1.59)
Townsend index							
Q1 Deprivation, least	76,746 (25.0)	51.91 (20.71)	9.16	31.71 (11.06)	6.60	4.21	5.50 (1.57)
Q2	76,745 (25.0)	51.50 (20.70)	9.35	31.31 (11.10)	7.37	4.82	5.49 (1.58)
Q3	76,719 (25.0)	49.91 (20.79)	11.20	30.92 (10.98)	8.25	5.55	5.50 (1.59)
Q4 Deprivation, most	76,711 (25.0)	45.94 (21.08)	17.11	30.24 (10.93)	10.90	7.62	5.54 (1.59)
Missing	360 (0.1)	50.02 (20.53)	11.39	31.37 (11.16)	7.22	4.84	5.61 (1.66)

Kg, kilograms; BMI, body mass index; NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; Q, quartiles; SD, standard deviation. [a] Average grip strength, min 0kg to max 85kg of contractile force, [b] Bioimpedance analysis derived skeletal muscle mass of the combined arm, min 2kg to 35.9 kg in weight, [c] Missing n= 314 (0.10%), [d] Missing n= 5,169 (1.68%). [e]Location categorization derived from assessment centre aggregation by longitudinal zoning, as described in supplementary methods (page 4). P values for all models <0.01; all models were adjusted for sex, age, assessment centre, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

Table S5. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.

See Supplementary Table 01 of the primary thesis:

Supplementary Table 1: Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations. 321

Supplementary References:

1. DiaSorin. LIAISON - 25 OH Vitamin D TOTAL Assay 2019. (https://www.diasorin.com/sites/default/files/allegati_prodotti/25_oh_vit._d_total_m0870004213_e.pdf)

2. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North: Routledge; 1988.

3. Vickers D, Rees P. Creating the UK National Statistics 2001 output area classification. Journal of the Royal Statistical Society. 2007;170(2):379-403.

4. Nuttall FQ. Body Mass Index. Nutrition Today. 2015;50(3):117-28.

5. Revez JA, Lin T, Qiao Z, Xue A, Holtz Y, Zhu Z, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nat Commun. 2020;11(1):1647.

6. Jiang X, O'Reilly PF, Aschard HA-Ohoo, Hsu YH, Richards JBA-Ohoo, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun. 2018;09(1):2041-1723.

7. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. International journal of epidemiology. 2013;42(4):1134-44.

8. Palmer, T. M., J. A. Sterne, R. M. Harbord, D. A. Lawlor, N. A. Sheehan, S. Meng, R. Granell, G. D. Smith, and V. Didelez. 2011. "Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses." Am J Epidemiol 173 (12): 1392–1403.

9. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genetic Epidemiology. 2016;40(7):597-608.

10. Burgess, S., A. Butterworth, and S. G. Thompson. 2013. "Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data." Genet Epidemiol 37 (7): 658–65.

11. Bowden, J., G. Davey Smith, and S. Burgess. 2015. "Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression." Int J Epidemiol 44 (2): 512–25.

12. Verbanck, Marie, Chia-Yen Chen, Benjamin Neale, and Ron Do. 2018. "Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases." Nature Genetics 50 (5): 693–98.

13. Hartwig, F. P., G. Davey Smith, and J. Bowden. 2017. "Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption." Int J Epidemiol 46 (6): 1985–98.

14. Bowden, J., G. Davey Smith, P. C. Haycock, and S. Burgess. 2016. "Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator." Genet Epidemiol 40 (4): 304–14.

15. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS, Staley JR. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics. 2019 Nov 1;35(22):4851-4853

16. Staley, James R., and Stephen Burgess. 2017. "Semiparametric Methods for Estimation of a Nonlinear Exposure-Outcome Relationship Using Instrumental Variables with Application to Mendelian Randomization." Genetic Epidemiology 41 (4): 341–52.

17. Burgess S, Mason AM, Grant AJ, Slob EA, Gkatzionis A, Zuber V, Patel A, Tian H, Liu C, Haynes WG, Hovingh GK. Using genetic association data to guide drug discovery and development: Review of methods and applications. The American Journal of Human Genetics. 2023 Feb 2;110(2):195-214.

18. Burgess, S., Davey Smith, G., Davies, N.M., Dudbridge, F., Gill, D., Glymour, M.M., Hartwig, F.P., Holmes, M.V., Minelli, C., Relton, C.L., Theodoratou, E., 2020. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research 4, 186

Chapter Seven: Study Three

Vitamin D Deficiency Increases Mortality Risk in the UK Biobank

A Nonlinear Mendelian Randomization Study



Annals of Internal Medicine has published this manuscript.

Impact factor: 39.2 (2022)

Rank in field: 8th out of 40 Quartile 1 journals in Internal Medicine.

Manuscript

Supplementary Material

Addendum

Annals of Internal Medicine: (Excerpted from *Information for Authors*, Published Date 03/27/2024) Authors are granted the rights after publication in Annals to reuse the published article without requesting permission from ACP. These authors' rights are to… include the article, or portions thereof, in their thesis… In all these cases for reuse, authors will give proper credit to the original publication in Annals as follows: Reproduced with permission from Author(s).

Title: Vitamin D Deficiency Increases Mortality Risk in the UK Biobank, A Nonlinear Mendelian Randomization Study. **Annals of Internal Medicine**: Year; 2022 vol: 175 pp-pp 1552-1559. URL https://doi.org/10.7326/M21-3324 ©American College of Physicians.

Study 3 Manuscript

List of study 3 manuscript figures:

Study 3 - Figure 1: Phenotypic association of 25(OH)D with all-cause (A), cancer (B), cardiovascular disease (CVD), and (D) respiratory mortality in the UK Biobank. 147

Study 3 - Figure 2: Non-linear (A, B, and C) and linear (D) Mendelian randomization analyses for the association of genetically predicted 25(OH)D with all-cause (A), cancer (B), cardiovascular disease (CVD), and (D) respiratory mortality in the UK Biobank. 149

List of study 3 manuscript tables:

Study 3 - Table 1: Demographic Characteristics of UK Biobank Participants 146

Vitamin D Deficiency Increases Mortality Risk in the UK Biobank

A Nonlinear Mendelian Randomization Study

Joshua P. Sutherland, BHSc Nut Med (Hons)

Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, South Australia, Australia

Ang Zhou, PhD

Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, and South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia

Elina Hyppönen, PhD

Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, and South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia

Background: Low vitamin D status is associated with increased mortality, but randomized trials on severely deficient participants are lacking.

Objective: To assess genetic evidence for the causal role of low vitamin D status in mortality.

Design: Nonlinear Mendelian randomization analyses.

Setting: UK Biobank, a large-scale, prospective cohort from England, Scotland, and Wales with participants recruited between March 2006 and July 2010.

Participants: 307 601 unrelated UK Biobank participants of White European ancestry (aged 37 to 73 years at recruitment) with available measurements of 25-hydroxyvitamin D (25-(OH)D) and genetic data.

Measurements: Genetically predicted 25-(OH)D was estimated using 35 confirmed variants of 25-(OH)D. All-cause and cause-specific mortality (cardiovascular disease [CVD], cancer, and respiratory) were recorded up to June 2020.

Results: There were 18 700 deaths during the 14 years of follow-up. The association of genetically predicted 25-(OH)D with all-cause mortality was L-shaped (P for nonlinearity < 0.001), and risk for death decreased steeply with increasing concentrations until 50 nmol/L. Evidence for an association was also seen in analyses of mortality from cancer, CVD, and respiratory diseases ($P \leq 0.033$ for all outcomes). Odds of all-cause mortality in the genetic analysis were estimated to increase by 25% (odds ratio = 1.25, 95% CI, 1.16-1.35) for participants with a measured 25-(OH)D concentration of 25 nmol/L compared with 50 nmol/L.

Limitations: Analyses were restricted to a White European population. A genetic approach is best suited to providing proof of principle on causality, whereas the strength of the association is approximate.

Conclusion: Our study supports a causal relationship between vitamin D deficiency and mortality. Additional research needs to identify strategies that meet the National Academy of Medicine's guideline of greater than 50 nmol/L and that reduce the premature risk for death associated with low vitamin D levels.

Primary Funding Source: National Health and Medical Research Council.

Introduction

The effects of vitamin D supplementation on mortality remain largely unexplored in the context of deficiency. This is because randomized controlled trials (RCTs) often do not recruit people with vitamin D deficiency or, because of ethical reasons, are prevented from doing so (1, 2). This leaves studies to investigate the effects of intakes that may be in surplus to the actual nutritional requirement.

Observational analyses and RCTs have considered the effects of vitamin D supplementation on a range of pathologies (3), with ongoing debate about the true health implications arising from deficiency. Wide-ranging health effects might be expected because vitamin D is a nutrient and prohormone that is mostly derived from exposure to type B ultraviolet radiation from the sun, and vitamin D receptors are found in most major organs and human tissues (4). More than 70 RCTs have looked at the effects of vitamin D supplementation on mortality (5-8). Some modest associations with all-cause mortality have been shown (6, 7); however, meta-analyses of RCTs have typically supported survival benefits only for specific groups, such as elderly persons and those with cancer (5, 6). A key problem with RCTs of vitamin D supplementation is the failure to include participants with vitamin D deficiency, and further limitations relate to heterogeneous dosing methods, inadequate participant diversity, and short follow-up periods (9).

Mendelian randomization is a genetic approach that uses genetic variants that approximate the exposure as an instrument or “proxy indicator”; it can be used to provide causal evidence for exposures where RCTs are either unethical or infeasible (10). Mendelian randomization allows us to overcome key challenges, such as reverse causation and confounding, of other observational approaches. The Mendelian randomization approach has been recently expanded to allow for nonlinear associations (11), making it possible to explore threshold effects. This is

particularly helpful in the context of nutritional exposures because it provides a novel, noninvasive approach to seek evidence on benefits that might only be seen in the context of rectifying an existing deficiency.

Until recently, Mendelian randomization studies investigating the effects of 25-hydroxyvitamin D (25-(OH)D), a marker of nutritional vitamin D status, have used linear Mendelian randomization analyses, with mixed findings for effects on mortality (12-15). Although some studies have inferred benefits by higher genetically predicted concentrations of 25-(OH)D, other large studies on mortality from all causes (12), cancer (12, 13), and cardiovascular disease (CVD) (12, 14) have not provided evidence for a causal relationship. However, recent Mendelian randomization findings (16) provide evidence for a threshold effect between 25-(OH)D and mortality. In this study, we used an expanded genetic instrument, in conjunction with information from up to 307 601 UK Biobank participants, to examine evidence for a nonlinear causal role of 25-(OH)D in all-cause and cause-specific mortality (cancer and CVD), while providing novel insight into respiratory mortality and applicability across ethnic groups. Where appropriate, we used a nonlinear approach.

Methods

The UK Biobank is a large-scale, prospective cohort that includes 502 316 participants from England, Scotland, and Wales who were aged 37 to 73 years at recruitment between March 2006 and July 2010 (17). Participants were invited to take survey questionnaires, have physical assessments, and provide biological samples. Our primary analysis was limited to unrelated participants of White European ancestry with measurements of serum 25-(OH)D concentration (n = 307 601) (Supplement Figure 1, available at Annals.org). Mortality data (n = 18 700) were obtained from NHS Digital and the NHS Central Register (18). The latest recorded death in our

study occurred in June 2020. Primary causes of death were defined using the International Classification of Diseases, 10th Revision, codes for cancer (C00 to D48), CVD (I00 to I89), and respiratory diseases (J09 to J18, J20 to J22, and J40 to J47) (19). The LIAISON XL 25-(OH)D assay (DiaSorin) was used to determine measured baseline concentrations of serum 25-(OH)D, as described in the Supplement (available at Annals.org). All covariates were derived at baseline from self-reported, touchscreen questionaries, aside from location and socioeconomic status, which were ascertained from assessment center and residential data, respectively. The Townsend deprivation index was used to establish socioeconomic status (20). Body mass index was assessed at baseline from height and weight measurements (calculated as body weight in kilograms divided by height in meters squared). We constructed a weighted vitamin D genetic score by collating 35 common autosomal single-nucleotide polymorphisms that were discovered in the UK Biobank from a genome-wide association analysis on measured 25-(OH)D concentrations (21) and replicated in the SUNLIGHT (Study of Underlying Genetic Determinants of Vitamin D and Highly Related Traits) consortium (22). Selection of variants and construction of the vitamin D genetic score are described in the Supplement. Ethics approval for the UK Biobank was granted by the National Information Governance Board for Health and Social Care and the North West Multi-centre Research Ethics Committee (11/NW/0382). The present analysis operates under UK Biobank application 20175.

Statistical Analysis

The phenotypic analyses were done using logistic regression, with all models adjusted for age, sex, assessment center, education, Townsend index, body mass index, physical activity, alcohol, smoking, and nuisance factors that could affect the measured 25-(OH)D concentrations (blood sampling month, fasting time before sample acquisition, and aliquot). We used a likelihood ratio

test to compare model fit between the best-fitting fractional polynomial model and the linear model, taking *P* less than 0.05 to indicate a nonlinear association (23).

We used genetically predicted concentrations of 25-(OH)D and did linear and nonlinear Mendelian randomization analyses on mortality, with the latter capable of assessing evidence for curvature (pages 4 to 6 of the Supplement and Supplement Figures 2 and 3, available at Annals.org). We present nonlinear Mendelian randomization analyses as the primary findings if evidence for curvature was observed; otherwise, results from linear Mendelian randomization analyses are considered to be the primary findings. As a secondary analysis, we did stratified Mendelian randomization analysis (page 6 of the Supplement and Supplement Figure 3) where we segmented the cohort into 4 strata of residual measured 25-(OH)D concentrations—less than 25 nmol/L, 25 to 49.9 nmol/L, 50 to 74.9 nmol/L, and 75 nmol/L or higher—and examined evidence of a linear association within each stratum. Residual measured concentrations of 25-(OH)D for each participant were determined as the measured 25-(OH)D minus the difference in concentrations attributed to genetic variants (correlation between residual measured and measured 25-(OH)D; $r = 0.986$). Further sensitivity analyses were done restricting the population to non-White ethnic groups (Supplement Table 1, available at Annals.org). Linear, nonlinear, and stratified Mendelian randomization analyses are detailed in the Supplement Methods.

Valid causal inference from the Mendelian randomization analysis relies on 3 key assumptions (24) (Supplement Figure 2), which in the context of our study can be described as follows: vitamin D genetic score associates with measured 25-(OH)D concentrations, vitamin D genetic score has no direct effect on mortality, and vitamin D genetic score does not associate with confounders of measured 25-(OH)D and mortality. To assess the first Mendelian randomization assumption, we examined the association of the vitamin D genetic score with

measured 25-(OH)D concentrations in the UK Biobank. To assess the second and third assumptions, we examined horizontal pleiotropy. Horizontal pleiotropy describes the situation when the genetic instrument associates with outcomes via pathways other than exposure of interest; it will violate the second or third Mendelian randomization assumption and consequently biases the Mendelian randomization analysis. To gauge potential horizontal pleiotropy, we examined the association of the vitamin D genetic score with several potential confounders in the UK Biobank, including body mass index, smoking, alcohol intake, physical activity, education, and Townsend deprivation index. We also did an analysis in which we left out specific blocks of potentially pleiotropic variants (in particular, pleiotropy by lipid traits) to examine if our primary finding is driven by any particular block (25). Functional blocks were based on the traits that variants were associated with in the PhenoScanner search (26) (page 9 of the Supplement and Supplement Table 2, available at Annals.org). We identified 4 functional blocks, including blocks for "blood," "lipids/metabolic," and "renal" traits; variants whose associated traits did not fall into 1 of these 3 blocks were grouped together as the "unclassified" block (Supplement Table 3, available at Annals.org). Further, in the stratified Mendelian randomization analysis where inclusion of pleiotropy-robust methods based on single-nucleotide polymorphisms is possible, we used 5 methods with largely independent assumptions on pleiotropy to compute stratum-specific Mendelian randomization estimates. These 5 Mendelian randomization methods are inverse variance–weighted Mendelian randomization, Mendelian randomization–Egger, weighted median Mendelian randomization, weighted mode Mendelian randomization, and Mendelian randomization pleiotropy residual sum and outlier; the Supplement (pages 7 to 9) details these methods. In strata where associations were evident, we also computed E-values to evaluate the sensitivity of the observed associations to unmeasured

confounding (Supplement). The nonlinear Mendelian randomization analysis additionally assumes that the association between vitamin D genetic score and measured 25-(OH)D concentration is constant over the entire distribution of measured 25-(OH)D (page 5 of the Supplement). To assess this assumption, we examined the heterogeneity of the associations between vitamin D genetic score and measured 25-(OH)D concentration across 100 strata of residual measured 25-(OH)D (11) (page 5 of the Supplement and Supplement Figure 4, available at Annals.org). Supplement Table 1 outlines the ethnic sensitivity analysis. We used R, version 3.6.1 (R Foundation), for nonlinear Mendelian randomization (NLMR package) and linear Mendelian randomization (TwoSampleMR and MRPRESSO packages) sensitivity analyses and Stata, version 14.1 (StataCorp), for all other analyses.

Role of the Funding Source

Neither the National Health and Medical Research Council nor the Australian Research Training Program Scholarship had any role in this study's design, conduct, or reporting.

Results

The Table shows baseline participant characteristics and the distribution of measured 25-(OH)D concentrations and vitamin D deficiency (measured 25-(OH)D concentration <25 nmol/L). The average measured concentration of 25-(OH)D was 45.2 nmol/L, and 11.71% of participants (*n* = 36 009) had concentrations between 10.0 and 24.9 nmol/L. Higher average measured concentrations were seen in participants living in southern areas; nonsmokers; and those with higher physical activity, less socioeconomic deprivation, and lower body mass index. During follow-up, 6.08% (*n* = 18 700) of participants died, with overrepresentation among those

with no declared educational status and those who smoked at the time of the baseline survey (mortality of 11% and 12%, respectively) (Supplement Table 4, available at Annals.org).

Study 3 - Table 1: Demographic Characteristics of UK Biobank Participants

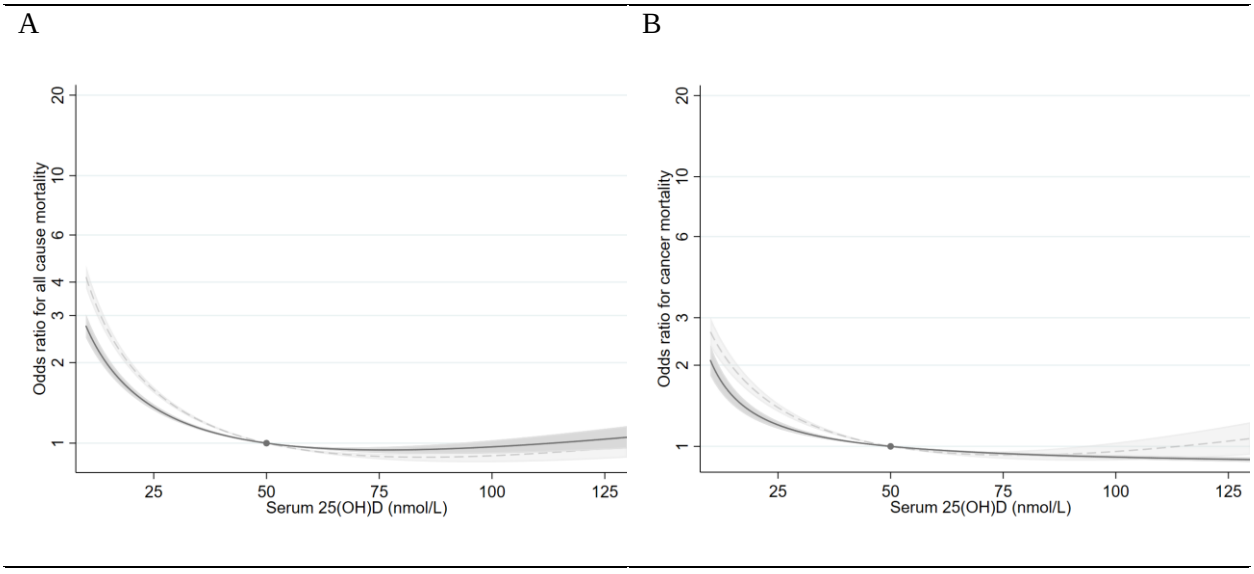
Characteristic	Participants, *n (%)*	Geometric Mean 25-(OH)D Concentration (95% CI) (*n* = 307 601), *nmol/L*	25-(OH)D Concentration <25 nmol/L (*n* = 36 009), *%*
All	307 601 (100)	45.2 (45.1–45.3)	11.71
Sex			
Men	144 680 (47.0)	45.2 (45.1–45.4)	11.60
Women	162 921 (53.0)	45.2 (45.1–45.3)	11.80
Age			
<60 y	169 756 (55.2)	43.7 (43.6–43.8)	13.46
≥60 y	137 845 (44.8)	47.2 (47.1–47.3)	9.54
BMI			
Lowest 25% (12.1–24.1 kg/m^2)	76 640 (24.9)	48.1 (48.0–48.3)	10.19
Middle 50% (24.1–29.8 kg/m^2)	153 321 (49.8)	46.5 (46.4–46.6)	10.05
Highest 25% (29.8–74.7 kg/m^2)	76 675 (24.9)	40.2 (40.1–40.3)	16.40
Missing	965 (0.3)	37.6 (36.4–38.9)	22.90
Location			
South (≤51° latitude)	102 335 (33.3)	47.1 (47.0–47.2)	9.29
Middle (52°–53° latitude)	144 654 (47.0)	45.4 (45.3–45.5)	11.36
North (≥54° latitude)	60 612 (19.7)	41.8 (41.6–41.9)	16.62
Smoking			
Nonsmokers	167 703 (54.5)	45.6 (45.5–45.7)	10.92
Former smokers	108 118 (35.2)	46.2 (46.1–46.3)	10.67
Current smokers	30 719 (10.0)	40.0 (39.7–40.2)	19.63
Missing	1061 (0.3)	45.1 (43.8–46.4)	12.54
Alcohol			
Daily	65 542 (21.3)	46.5 (46.6–46.3)	11.03
1–4 times/wk	155 608 (50.6)	46.4 (46.5–46.3)	10.20
1–3 times/mo	34 098 (11.1)	43.7 (43.9–43.5)	12.91
Special occasion	32 179 (10.5)	41.6 (41.3–41.8)	15.76
Never	19 963 (6.5)	41.0 (40.8–41.3)	17.06
Missing	211 (0.07)	40.1 (42.9–37.5)	16.59
Physical activity			
Low	92 012 (29.9)	41.8 (41.7–41.9)	15.18
Moderate	149 205 (48.5)	46.1 (46.0–46.2)	10.50
High	59 561 (19.4)	49.5 (49.3–49.7)	8.11
Missing	6823 (2.2)	38.2 (37.7–38.6)	22.51
Education			
None	52 193 (17.0)	45.6 (45.4–45.8)	11.90
NVQ/CSE/A-levels	109 099 (35.5)	45.9 (45.8–46.0)	11.23
Degree/professional	143 735 (46.7)	44.6 (44.5–44.7)	12.00
Missing	2574 (0.84)	45.7 (44.9–46.5)	11.58
Townsend index			
Quartile 1 (lowest deprivation)	76 793 (25.0)	47.6 (47.4–47.7)	9.16
Quartile 2	76 821 (25.0)	47.1 (47.0–47.3)	9.35
Quartile 3	76 815 (25.0)	45.4 (45.3–45.5)	11.20
Quartile 4 (highest deprivation)	76 811 (25.0)	41.1 (40.9–41.2)	17.11
Missing	361 (0.1)	45.6 (43.5–47.7)	11.63

25-(OH)D = 25-hydroxyvitamin D; A-levels = advanced levels; BMI = body mass index; CSE = Certificate of Secondary Education; NVQ = National Vocational Qualification.

Phenotypic Analyses

Nonlinear inverse relationships with measured 25-(OH)D were similar among all-cause mortality and the various cause-specific mortality rates (Figure 1). Crude and fully adjusted models followed similar patterns, with some attenuation by adjustment. For all 4 outcomes, odds ratios (ORs) were highest for concentrations lower than 25 nmol/L and adjusted associations seemed to plateau between 50 and 75 nmol/L, with little to no further reduction in mortality with measured 25-(OH)D values of 75 to 125 nmol/L. The fully adjusted odds of all-cause mortality were 36% higher for participants at 25 nmol/L compared with 50 nmol/L (OR = 1.36, 95% CI, 1.33-1.40).

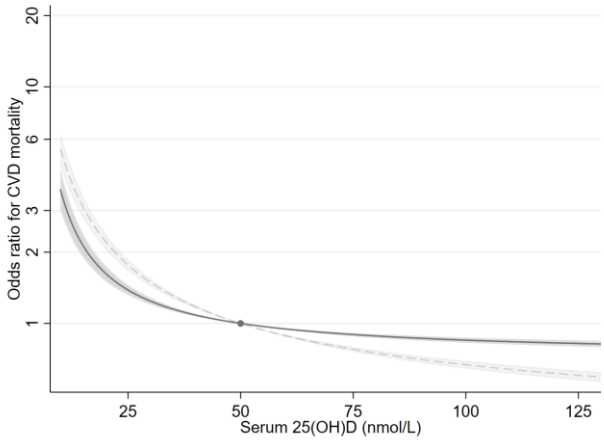
Study 3 - Figure 1: Phenotypic association of 25(OH)D with all-cause (A), cancer (B), cardiovascular disease (CVD), and (D) respiratory mortality in the UK Biobank.



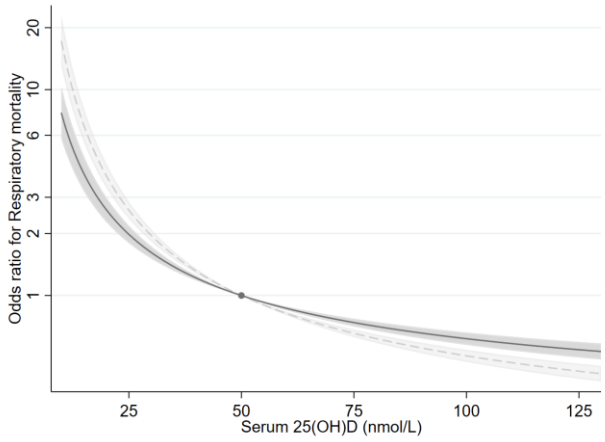
C and D continued on next page

Continuation of Figure 1

C



D

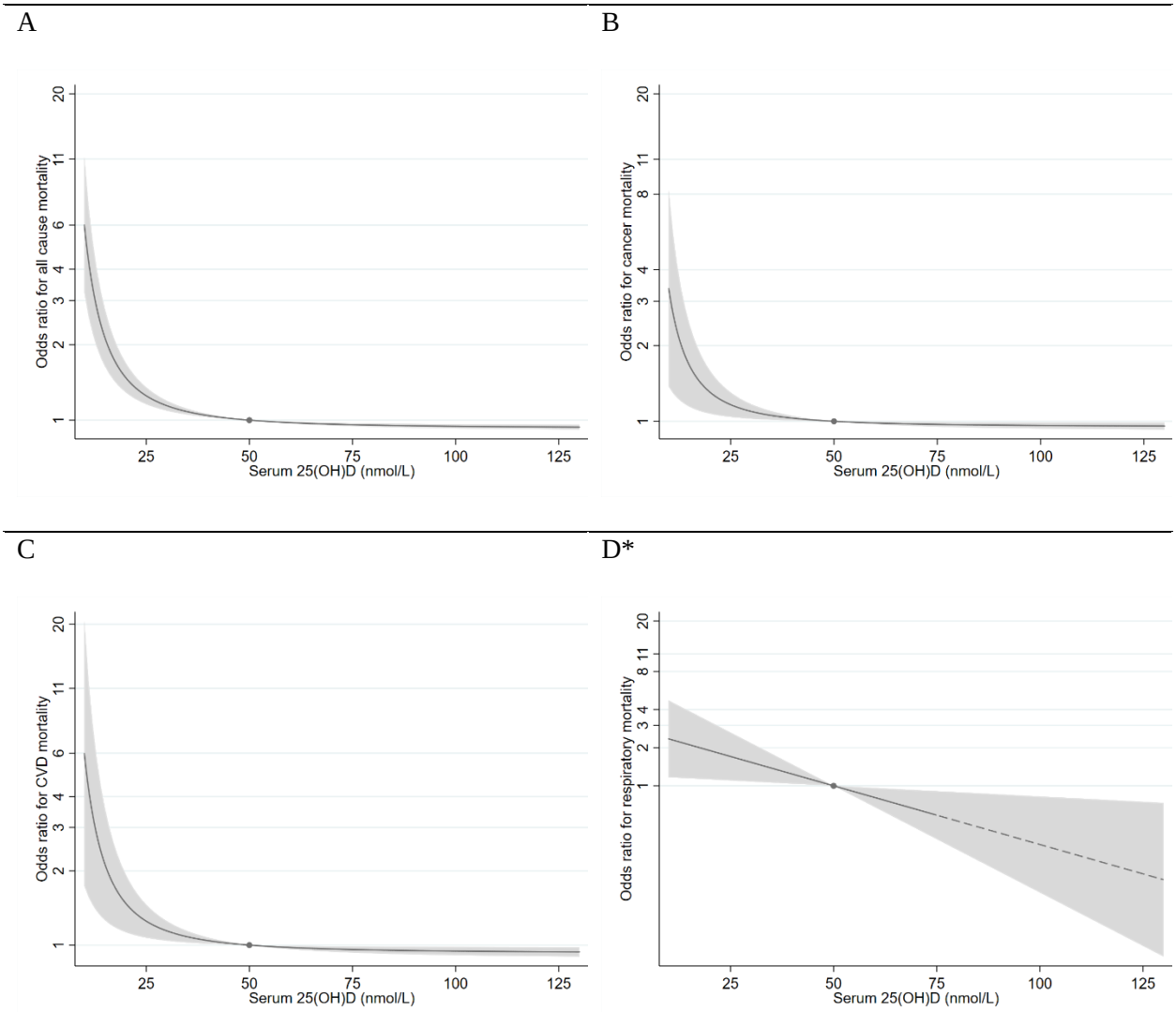


Shaded areas reflect 95% CIs. The dot represents the reference point of a measured 25-(OH)D serum concentration of 50 nmol/L. Simple models (light gray) were adjusted for sex, age, assessment center, and nuisance factors that could affect 25-(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. Full models (dark gray) were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking. 25-(OH)D = 25-hydroxyvitamin D; CVD = cardiovascular disease; OR = odds ratio.

Mendelian Randomization

As shown in Figure 2, genetically predicted 25-(OH)D had an L-shaped association with all-cause (*P* for nonlinearity < 0.001), cancer, and CVD mortality (*P* for nonlinearity ≤ 0.033) when displayed across the measured 25-(OH)D concentrations shown on the *x*-axis. Genetically predicted 25-(OH)D had the strongest association with these outcomes in participants with measured 25-(OH)D concentrations below 25 nmol/L, and the associations plateaued by 50 nmol/L. For respiratory diseases, there was no significant curvature, but we observed evidence for a linear association (OR = 0.81, 95% CI, 0.68-0.96) per 10-nmol/L increase in genetically predicted 25-(OH)D concentration) (Figure 2, *D*).

Study 3 - Figure 2: Non-linear (A, B, and C) and linear (D) Mendelian randomization analyses for the association of genetically predicted 25(OH)D with all-cause (A), cancer (B), cardiovascular disease (CVD), and (D) respiratory mortality in the UK Biobank.



The x-axis refers to the measured 25-(OH)D scale on which the genetically predicted analysis has been mapped; the dot represents the reference point of 50 nmol/L. Shaded areas reflect 95% CIs. Adjusted for age, sex, assessment center, birth location, single-nucleotide polymorphism array, top 40 genetic principals, and nuisance factors that could affect serum 25-(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. 25-(OH)D = 25-hydroxyvitamin D; CVD = cardiovascular disease; OR = odds ratio. * Dashed line denotes the beginning of data that cannot be interpreted because of insuffient statistical power from there onward.

However, in the stratified analyses, we observed no evidence for an association between genetically predicted 25-(OH)D and respiratory disease mortality in the stratum including participants with the highest measured 25-(OH)D concentrations (> 75 nmol/L: OR = 1.04, 95% CI, 0.52-2.09; 69 cases) (Supplement Table 5, available at Annals.org). Compared with a

measured 25-(OH)D concentration of 50 nmol/L, we estimated that the genetically predicted odds of all-cause mortality would increase by 6-fold (OR = 6.00, 95% CI, 3.22-11.17) for participants at 10 nmol/L and by 25% (OR = 1.25, 95% CI, 1.16-1.35) for those at 25 nmol/L (Figure 2). Compared with a measured 25-(OH)D concentration of 50 nmol/L, participants at 10 nmol/L had genetically predicted ORs of 5.98, 95% CI, 1.73-20.59) for CVD mortality, OR = 3.37 (95% CI, 1.37-8.28) for cancer mortality, and OR = 12.44 (95% CI, 4.32-35.85) for respiratory mortality. For the comparison of measured 25-(OH)D concentrations of 25 nmol/L versus 50 nmol/L, these outcomes had ORs of 1.25 (95% CI, 1.07-1.46), 1.16 (95% CI, 1.04-1.30), and 1.96 (95% CI, 1.88-4.67), respectively.

Sensitivity Analyses

Sensitivity analyses excluding the first and 100th strata of residual measured 25-(OH)D (Supplement Figure 5, available at Annals.org) and using the version of the vitamin D genetic score with 122 single-nucleotide polymorphisms (Supplement Figure 6, available at Annals.org) provided similar results. Findings were also similar in analyses in which blocks were left out (Supplement Tables 2 and 3) and in stratified Mendelian randomization analyses using multiple Mendelian randomization approaches, with the Egger intercept test detecting no pleiotropy (Supplement Table 6, available at Annals.org). Finally, analyses using information on the subsample of participants who were of non-White ethnic origin (*n* = 20 837) also provided consistent results supportive of a causal effect of genetically predicted 25-(OH)D on all-cause mortality in those with low measured 25-(OH)D concentrations (Supplement Table 1).

Discussion

The causal effects of low vitamin D status in a general population are challenging to establish, with RCTs either failing to recruit people with severe deficiency or, because of ethical reasons, being prevented from doing so (1, 2). We examined the association of genetically predicted 25-(OH)D with mortality, observing evidence for a causal relationship across all included mortality outcomes. Of note, the genetic evidence supporting a relationship between higher 25-(OH)D concentrations and mortality was largely restricted to persons with measured concentrations below 50 nmol/L, which reflects the cutoff endorsed by the U.S. National Academy of Medicine (27). The strongest effects were seen for persons with measured 25-(OH)D concentrations in the severe deficiency range (<25 nmol/L). Although the past decade has seen benefits in some settings through increases in food fortification and updates on policy guidelines (28, 29), recent estimates for the prevalence of severe deficiency range from 5% to 50%, with rates varying by geographic location and population characteristics (28, 30-32). Therefore, our study affirms the potential for a notable effect on premature death and the continued need for efforts to abolish vitamin D deficiency.

Prior data are limited on the efficacy of vitamin D supplementation in preventing death in persons with vitamin D deficiency. In fact, a recent meta-analysis that included 9 trials of vitamin D supplementation in critically ill persons (33)—which did not provide evidence for benefit—contained only 1 RCT that was restricted to vitamin D–deficient participants (34) and another study with a dedicated subgroup with severe deficiency (35). These 2 studies, independent of the meta-analysis, were among those to show some positive effect of vitamin D supplementation on mortality. All studies also used large one-off (or "bolus") doses, although evidence suggests that this type of dosing may interfere with the synthesis and breakdown of

enzymes regulating vitamin D activity (36). Our data suggest that although remediation of vitamin D deficiency is essential, supplementation is unlikely to have notable benefits for preventing death when given in surplus to the nutritional requirement.

The human genome contains at least 13 000 known receptor-binding sites for vitamin D (37), and most tissues and cell types are responsive to 1,25-$(OH)_2D_3$ (the active vitamin D hormone), suggesting universal action (4). In line with this, we observed effects on all major causes of death investigated. The association of vitamin D deficiency with CVD, respiratory illness, and cancer is biologically plausible. Of note for CVD, 1,25-$(OH)_2D_3$ is a demonstrated homeostatic regulator of the renin–angiotensin–aldosterone system, which is related to blood pressure, whereas impaired vitamin D metabolism has an adverse effect on macrophage differentiation, potentially increasing vascular cholesterol retention (38). For the respiratory tract, epithelial cells have been shown to convert 25-(OH)D from serum into 1,25-$(OH)_2D_3$ and subsequently produce antimicrobial peptides – one possible mechanism through which vitamin D reduces risk for respiratory tract infection (39). Effects of 1,25-$(OH)_2D_3$ on cancer may be mediated through established influences on cellular proliferation, differentiation, and apoptosis (40, 41). However, in our analyses done in the context of vitamin D deficiency, the genetically predicted effects of increasing 25-(OH)D concentrations on mortality were unspecific. Although many factors may influence mortality, our results may also suggest that general frailty – such as that acquired at life-threatening stages of chronic diseases – is related to vitamin D, with higher 25-(OH)D concentrations promoting the ability to retain at least the minimal physiologic reserves required to sustain life.

A primary strength of our study is its genetic approach, which has allowed us to safely explore the effects of raising 25-(OH)D in persons with very low concentrations, in contrast with

RCTs, where participants would be subjected to potential harm if left deficient. Our study has been made possible by the availability of measured 25-(OH)D concentrations and genotyping in 307 601 participants, with enough cases to allow analyses on all-cause, CVD, cancer, and respiratory mortality. This is the first study to our knowledge to use the nonlinear Mendelian randomization approach, and we were powered to include more specific causes of death than prior stratified Mendelian randomization work undertaken in this same population (17). We also included sensitivity analyses on non-White ethnic groups, which provides proof of principle for external validity outside White ethnic populations. Although the Mendelian randomization approach can provide support for a causal association, estimates can be biased by horizontal pleiotropy, which occurs when variants influence outcomes through pathways other than via the exposure (10). We restricted our instrument to variants with replicated evidence for an association with measured 25-(OH)D concentrations and replicated our analyses using several pleiotropy-robust approaches and other sensitivity analyses, confirming the stability of our findings. In addition, we found no evidence for an association between the vitamin D genetic score and potential confounders across the cohort or within the strata of residual measured concentrations of 25-(OH)D. A reasonably large E-value for the association between vitamin D genetic score and mortality seen in the stratum with residual measured 25-(OH)D concentrations less than 25 nmol/L (page 10 of the Supplement) provides further assurance that the observed adverse effects for vitamin D deficiency are unlikely to be explained by residual confounding. With all Mendelian randomization analyses, genetic instruments approximate average effects over the life course; therefore, the true shape and strength of association may be more complex than presented here. The UK Biobank is not representative of the general population of the United Kingdom (5% response rate) (42). However, earlier publications from the UK Biobank

have replicated expected associations between exposure and disease (43), and studies using the Mendelian randomization approach have been shown to be less affected by selection bias, suggesting limited influence on our findings (43).

In conclusion, our study supports a causal relationship between vitamin D deficiency and mortality. Additional research needs to identify strategies that meet the National Academy of Medicine's guideline concentration of greater than 50 nmol/L and that reduce the premature risk for death associated with low vitamin D levels.

Acknowledgment: The authors thank all UK Biobank participants involved, as well as Dr. Anwar Mulugeta (University of South Australia) for support in data management.

Financial Support: By grant 11123603 from the National Health and Medical Research Council. Mr. Sutherland's studentship is funded by an Australian Research Training Program Scholarship.

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M21-3324.

Reproducible Research Statement: *Study protocol:* Not available. *Data set and statistical code:* This research has been conducted using the UK Biobank resource under application 20175. All data and code will be available to approved users on application to the UK Biobank.

Corresponding Author: Elina Hyppönen, PhD, Australian Centre for Precision Health, University of South Australia, c/o South Australian Health and Medical Research Institute, GPO Box 2471, Adelaide, SA 5001, Australia; e-mail, Elina.Hypponen@unisa.edu.au.

Author Contributions: Conception and design: E. Hyppönen.

Analysis and interpretation of the data: J.P. Sutherland, A. Zhou, E. Hyppönen.

Drafting of the article: J.P. Sutherland, E. Hyppönen, A. Zhou.

Critical revision for important intellectual content: E. Hyppönen, A. Zhou.

Final approval of the article: E. Hyppönen, J.P. Sutherland, A. Zhou.

Provision of study materials or patients: E. Hyppönen.

Statistical expertise: A. Zhou, E. Hyppönen

Obtaining of funding: E. Hyppönen.

Administrative, technical, or logistic support: E. Hyppönen, J.P. Sutherland.

Collection and assembly of data: J.P. Sutherland, A. Zhou.

References

1. **Cooper C, Harvey NC, Bishop NJ, et al; MAVIDOS Study Group.** Maternal gestational vitamin D supplementation and offspring bone health (MAVIDOS): a multicentre, double-blind, randomised placebo-controlled trial. Lancet Diabetes Endocrinol. 2016;4:393-402. [PMID: 26944421] doi:10.1016/S2213-8587(16)00044-9

2. **Bolland MJ, Grey A, Avenell A.** Assessment of research waste part 2: wrong study populations- an exemplar of baseline vitamin D status of participants in trials of vitamin D supplementation. BMC Med Res Methodol. 2018;18:101. [PMID: 30285729] doi:10.1186/s12874-018-0555-1

3. **Wang H, Chen W, Li D, et al.** Vitamin D and chronic diseases. Aging Dis. 2017;8:346-353. [PMID: 28580189] doi:10.14336/AD.2016.1021

4. **Wang Y, Zhu J, DeLuca HF.** Where is the vitamin D receptor? Arch Biochem Biophys. 2012;523:123-33. [PMID: 22503810] doi:10.1016/j.abb.2012.04.001

5. **Zhang Y, Fang F, Tang J, et al.** Association between vitamin D supplementation and mortality: systematic review and meta-analysis. BMJ. 2019;366:l4673. [PMID: 31405892] doi:10.1136/bmj.l4673

6. **Bjelakovic G, Gluud LL, Nikolova D, et al.** Vitamin D supplementation for prevention of cancer in adults. Cochrane Database Syst Rev. 2014:CD007469. [PMID: 24953955] doi:10.1002/14651858.CD007469.pub2

7. **Bolland MJ, Grey A, Gamble GD, et al.** The effect of vitamin D supplementation on skeletal, vascular, or cancer outcomes: a trial sequential meta-analysis. Lancet Diabetes Endocrinol. 2014;2:307-320. [PMID: 24703049] doi:10.1016/S2213-8587(13)70212-2

8. **Manson JE, Cook NR, Lee IM, et al; VITAL Research Group.** Vitamin D supplements and prevention of cancer and cardiovascular disease. N Engl J Med. 2019;380:33-44. [PMID: 30415629] doi:10.1056/NEJMoa1809944

9. **Camargo CA Jr, Martineau AR.** Vitamin D to prevent COVID-19: recommendations for the design of clinical trials. FEBS J. 2020;287:3689-3692. [PMID: 33448695] doi:10.1111/febs.15534

10. **Davies NM, Holmes MV, Davey Smith G.** Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601. [PMID: 30002074] doi:10.1136/bmj.k601

11. **Staley JR, Burgess S.** Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. Genet Epidemiol. 2017;41:341-352. [PMID: 28317167] doi:10.1002/gepi.22041

12. **Afzal S, Brøndum-Jacobsen P, Bojesen SE, et al.** Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. BMJ. 2014;349:g6330. [PMID: 25406188] doi:10.1136/bmj.g6330

13. **Ong JS, Gharahkhani P, An J, et al.** Vitamin D and overall cancer risk and cancer mortality: a Mendelian randomization study. Hum Mol Genet. 2018;27:4315-4322. [PMID: 30508204] doi:10.1093/hmg/ddy307

14. **Huang T, Afzal S, Yu C, et al; China Kadoorie Biobank Collaborative Group.** Vitamin D and cause-specific vascular disease and mortality: a Mendelian randomisation study involving 99,012 Chinese and 106,911 European adults. BMC Med. 2019;17:160. [PMID: 31466528] doi:10.1186/s12916-019-1401-y

15. **Meng X, Li X, Timofeeva MN, et al.** Phenome-wide Mendelian-randomization study of genetically determined vitamin D on multiple health outcomes using the UK Biobank study. Int J Epidemiol. 2019;48:1425-1434. [PMID: 31518429] doi:10.1093/ije/dyz182

16. **Emerging Risk Factors Collaboration/EPIC-CVD/Vitamin D Studies Collaboration.** Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses. Lancet Diabetes Endocrinol. 2021;9:837-846. [PMID: 34717822] doi:10.1016/S2213-8587(21)00263-1

17. **Sudlow C, Gallacher J, Allen N, et al.** UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779. [PMID: 25826379] doi:10.1371/journal.pmed.1001779

18. **UK Biobank.** Mortality data: linkage to death registries, version 2.0. June 2020. Accessed at https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/DeathLinkage.pdf on 17th March 2021.

19. **Steindel SJ.** International Classification of Diseases, 10th Edition, Clinical Modification and procedure coding system: descriptive overview of the next generation HIPAA code sets. J Am Med Inform Assoc. 2010;17:274-82. [PMID: 20442144] doi:10.1136/jamia.2009.001230

20. **Townsend P, Phillimore P, Beattie A.** Health and Deprivation: Inequality and the North. Routledge; 1988.

21. **Revez JA, Lin T, Qiao Z, et al.** Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nat Commun. 2020;11:1647. [PMID: 32242144] doi:10.1038/s41467-020-15421-7

22. **Jiang X, O'Reilly PF, Aschard H, et al.** Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun. 2018;9:260. [PMID: 29343764] doi:10.1038/s41467-017-02662-2

23. **Royston P, Ambler G, Sauerbrei W.** The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol. 1999;28:964-74. [PMID: 10597998]

24. **Burgess S, Davey Smith G, Davies NM, et al.** Guidelines for performing Mendelian randomization investigations. Wellcome Open Res. 2019;4:186. [PMID: 32760811] doi:10.12688/wellcomeopenres.15555.2

25. **Burgess S, Gill D.** Genetic evidence for vitamin D and cardiovascular disease: choice of variants is critical [Editorial]. Eur Heart J. 2022;43:1740-1742. [PMID: 34972215] doi:10.1093/eurheartj/ehab870

26. **Kamat MA, Blackshaw JA, Young R, et al.** PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics. 2019;35:4851-4853. [PMID: 31233103] doi:10.1093/bioinformatics/btz469

27. **Ross AC, Taylor CL, Yaktine AL, et al, eds; Institute of Medicine.** Dietary Reference Intakes for Calcium and Vitamin D. National Academies Pr; 2011.

28. **Munasinghe LL, Yuan Y, Willows ND, et al.** Vitamin D deficiency and sufficiency among Canadian children residing at high latitude following the revision of the RDA of vitamin D intake in 2010. Br J Nutr. 2017;117:457-465. [PMID: 28245892] doi:10.1017/S0007114517000320

29. **Raulio S, Erlund I, Männistö S, et al.** Successful nutrition policy: improvement of vitamin D intake and status in Finnish adults over the last decade. Eur J Public Health. 2017;27:268-273. [PMID: 28339536] doi:10.1093/eurpub/ckw154

30. **Chacham S, Rajput S, Gurnurkar S, et al.** Prevalence of vitamin D deficiency among infants in northern India: a hospital based prospective study. Cureus. 2020;12:e11353. [PMID: 33304688] doi:10.7759/cureus.11353

31. **Amrein K, Scherkl M, Hoffmann M, et al.** Vitamin D deficiency 2.0: an update on the current status worldwide. Eur J Clin Nutr. 2020;74:1498-1513. [PMID: 31959942] doi:10.1038/s41430-020-0558-y

32. **Sutherland JP, Zhou A, Leach MJ, et al.** Differences and determinants of vitamin D deficiency among UK Biobank participants: a cross-ethnic and socioeconomic study. Clin Nutr. 2021;40:3436-3447. [PMID: 33309415] doi:10.1016/j.clnu.2020.11.019

33. **Lan SH, Lai CC, Chang SP, et al.** Vitamin D supplementation and the outcomes of critically ill adult patients: a systematic review and meta-analysis of randomized controlled trials. Sci Rep. 2020;10:14261. [PMID: 32868842] doi:10.1038/s41598-020-71271-9

34. **Hasanloei MAV, Rahimlou M, Eivazloo A, et al.** Effect of oral versus intramuscular vitamin D replacement on oxidative stress and outcomes in traumatic mechanical ventilated patients admitted to intensive care unit. Nutr Clin Pract. 2020;35:548-558. [PMID: 31486158] doi:10.1002/ncp.10404

35. **Amrein K, Schnedl C, Holl A, et al.** Effect of high-dose vitamin D3 on hospital length of stay in critically ill patients with vitamin D deficiency: the VITdAL-ICU randomized clinical trial. JAMA. 2014;312:1520-30. [PMID: 25268295] doi:10.1001/jama.2014.13204

36. **Vieth R.** How to optimize vitamin D supplementation to prevent cancer, based on cellular adaptation and hydroxylase enzymology. Anticancer Res. 2009;29:3675-84. [PMID: 19667164]

37. **Carlberg C.** Genome-wide (over)view on the actions of vitamin D. Front Physiol. 2014;5:167. [PMID: 24808867] doi:10.3389/fphys.2014.00167

38. **Al Mheid I, Quyyumi AA.** Vitamin D and cardiovascular disease: controversy unresolved. J Am Coll Cardiol. 2017;70:89-100. [PMID: 28662812] doi:10.1016/j.jacc.2017.05.031

39. **Martineau AR, Jolliffe DA, Hooper RL, et al.** Vitamin D supplementation to prevent acute respiratory tract infections: systematic review and meta-analysis of individual participant data. BMJ. 2017;356:i6583. [PMID: 28202713] doi:10.1136/bmj.i6583

40. **Carlberg C, Muñoz A.** An update on vitamin D signaling and cancer. Semin Cancer Biol. 2022;79:217-230. [PMID: 32485310] doi:10.1016/j.semcancer.2020.05.018

41. **Chakraborti CK.** Vitamin D as a promising anticancer agent. Indian J Pharmacol. 2011;43:113-20. [PMID: 21572642] doi:10.4103/0253-7613.77335

42. **Batty GD, Gale CR, Kivimäki M, et al.** Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. BMJ. 2020;368:m131. [PMID: 32051121] doi:10.1136/bmj.m131

43. **Gkatzionis A, Burgess S.** Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? Int J Epidemiol. 2019;48:691-701. [PMID: 30325422] doi:10.1093/ije/dyy202

Study 4 Supplementary

Supplementary data for: Vitamin D deficiency increases mortality risk in the UK Biobank. A nonlinear Mendelian randomization study

Table of Contents:

Measurement of vitamin D status	**03**
Genetic instruments using 35 variants	**03**
Genetic instrument using 122 vitamin D GWAS variants	**04**
Linear, Non-linear, and Stratified Mendelian randomization (including SNP-based two-sample methods)	**04**
• Linear Mendelian randomization	04
• Non-linear Mendelian randomization	05
• Stratified Mendelian randomization	05
• SNP-based two-sample method	06
Leave-block-out analysis	**09**
E Value	**10**
Supplement Figure 1. Study population and participant inclusion criteria.	**12**
Supplement Figure 2. Mendelian randomization assumption diagram.	**13**
Supplement Figure 3. Schematic representation of Mendelian randomization analyses performed in the current study.	**14**
Supplement Figure 4. Distribution of βs for the Vitamin D Genetic Score-measured 25(OH)D association across 100 strata of residuals of measured 25(OH)D.	**15**
Supplement Figure 5. Non-linear Mendelian randomization analysis for the association of genetically predicted 25(OH)D with (A) all-cause mortality risk, (B) cancer mortality, and (C) cardiovascular disease (CVD) mortality in the UK Biobank, projected on the measured 25(OH)D scale; after excluding local causal estimates from the first and 100th* stratum in the primary analysis.	**16**
Supplement Figure 6. Non-linear Mendelian randomization analysis for the association of genetically predicted 25(OH)D with (A) all-cause mortality risk, (B) cancer mortality, and (C) cardiovascular disease (CVD) mortality in the UK Biobank, projected on the measured 25(OH)D scale; using the 122 SNP version of the Vitamin D Genetic Score.	**17**
Supplement Figure 7. Association of the Vitamin D Genetic Score with potential confounders across 100 strata of residuals of measured 25(OH)D.	**18**
Supplement Figure 8. Distribution of the Vitamin D Genetic Score and the association of sextiles of the Vitamin D Genetic Score with measured 25(OH)D concentration in the UK Biobank, illustrating the validity of the genetic instrument.	**19**
Supplement Figure 9. Selection of variants for the genetic instrument for measured 25(OH)D.	**20**
Supplement Table 1. Ethnic sensitivity – Stratified Mendelian randomization of genetically predicted 25(OH)D and all-cause mortality in the UK Biobank.	**21**
Supplement Table 2. Functional blocks used in the leave-block-out analyses.	**22**
Supplement Table 3. Leave-block-out analysis for non-linear Mendelian randomization analysis of genetically predicted 25(OH)D on risk for all-cause mortality.	**23**
Supplement Table 4. Demographic characteristics of UK Biobank participants.	**24**
Supplement Table 5. Linear Mendelian Randomization analyses of higher genetically predicted 25(OH)D and mortality in the UK biobank, by overall and stratified residual measured 25(OH)D categories.	**25**
Supplement Table 6. Stratified Mendelian Randomization sensitivity analyses of predicted 25(OH)D and all-cause mortality in the UK biobank.	**26**
Supplement Table 7. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.	**27**
Supplement Table 8. Association of the Vitamin D Genetic Score with potential confounders in the UK Biobank.	**28**
References	**29**

Authors:

Joshua P Sutherland[1], Ang Zhou[1,2], Elina Hyppönen[1,2]

1Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, Australia.

2South Australian Health and Medical Research Institute, Adelaide, Australia

Corresponding author: Prof Elina Hyppönen; Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, Australia, GPO Box 2471, Adelaide SA 5001; +61 8 830 22518; Elina.Hypponen@unisa.edu.au

Supplement Methods:

Measurement of vitamin D status

The LIAISON XL 25(OH)D assay (DiaSorin, Stillwater, USA) – a fully automated chemiluminescence immunoassay device with magnetic microparticle separation technology[1] – was used to analyses baseline serum 25(OH)D concentrations, measured from blood samples with a mean fasting time of 3.8 hours (SD 2.4). The LIAISON XL has a functional sensitivity of ≤ 4.0 ng/ml, in accordance with Clinical Laboratory and Standards Institute (CLSI) EP17-A protocols.[1] Assay precision (CLSI EP5-A2), produced a coefficient of variation (CV) of 2.3% (intra-run) and a CV 7.8% (total), conducted using 6 samples and 2 controls.[1]

Genetic instruments using 35 variants

We constructed a weighted genetic score by collating common autosomal SNPs (i.e. minor allele frequency >5%) from a genome-wide association analysis (GWAS) for measured 25(OH)D serum concentration in UK Biobank, wherein 143 independent loci were identified.[2] We subsequently replicated our SNP analyses in the previous SUNLIGHT consortium GWAS,[3] leaving 35 SNPs demonstrating consistent direction and a p value of <0.05 (**Supplement Figure 9**). By restricting our instrument to replicated SNPs we were able to ensure the robustness of the observed GWAS signals. We were also able to use the effect estimates for the SNP association with measured 25(OH)D from the SUNLIGHT consortium meta-analyses as the individual SNP weights, allowing us to avoid bias due to sample overlap.[4] Our Vitamin D Genetic Score was constructed by first computing the independently derived weighted average of the number of measured 25(OH)D-increasing alleles for an individual, and then multiplying it by the number of available variants.

Genetic instrument using 122 vitamin D GWAS variants

For sensitivity analyses, we constructed an alternative genetic instrument, which included a broader set of variants (122 autosomal single nucleotide polymorphisms, SNPs) associated with measured 25(OH)D concentrations. (**Supplement Figure 9**). Information for the 122 SNPs can be found in **Supplement Table 7**. As these 122 SNPs were discovered in the UK Biobank,[2] we used the 10-fold cross-validation approach[4] to construct the 122 SNP version of the Vitamin D Genetic Score, in order to mitigate any potential bias caused by using the internal weights. This involved randomly dividing the sample into 10 equal sub-samples, with weights for each sub-sample taken from the other 9 sub-samples.

Linear, Non-linear, and Stratified Mendelian randomization (including SNP-based two-sample methods)

Linear Mendelian randomization:

Linear Mendelian randomization estimate is computed using the ratio of coefficients method,[5] where the Vitamin D Genetic Score – measured 25(OH)D (γ_1, equation 1) and the Vitamin D Genetic Score – mortality association estimates (Γ_1, equation 2) are computed and then used to calculate the ratio estimator, $\hat{\beta} = \widehat{\Gamma_1} \big/ \widehat{\gamma_1}$.

(1)

Measured 25(OH)D$_i$ $= \gamma_0 + \gamma_1$ Vitamin D Genetic Score$_i + \gamma_{2j}$ covariate$_j + \epsilon_i,\ \epsilon_i \sim N(0, \delta^2)$

Logit(mortality$_i$) $= \Gamma_0 + \Gamma_1$ Vitamin D Genetic Score$_i$ + Γ_{2j} covariate$_j$

(2)

where,

- the subscript i denotes an individual;
- the subscript j denotes a covariate;
- the list of covariates in the model includes age, sex, assessment centre, birth locations, SNP array, top 40 genetic principal components, and nuisance factors which could affect measured 25(OH)D, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

Non-linear Mendelian randomization:

In the non-linear Mendelian randomization analysis, the sample is first stratified into 100 quantile strata of residual measured 25(OH)D (50 strata is used for CVD due to relatively smaller CVD mortality case numbers; reducing strata from 100 to 50 did not affect findings for all-cause and cancer mortality). Stratum-specific Mendelian randomization estimate is computed using the ratio of coefficients method (as described in the linear Mendelian randomization analysis, see section linear Mendelian randomization), and then meta-regressed on the mean stratum measured 25(OH)D concentration. The meta-regression was performed by fitting a range of fractional polynomial exposure-outcome models of degree 1 (equation 3) and degree 2 (equation 4), with the best-fitting model being selected based on the likelihood ratio test. A non-linear effect is evident if the best-fitting fractional polynomial model of degree 1 has a better fit than the linear model[6] (i.e. $P_{Non\text{-}linear} < 0.05$). It is worth noting that the stratification of the sample is based on residual 25(OH)D, which is defined as participants' measured 25(OH)D concentration minus the centered genetic contribution by the Vitamin D Genetic Score rather than the raw measured 25(OH)D concentration. This is to avoid collider bias.[7,8] However, there remains a strong correlation between residual measured 25(OH)D and measured 25(OH)D (r= 0.986). Non-linear Mendelian randomization analyses assume that the association of genetic instrument with the exposure is constant over the entire distribution of exposure (here, measured 25(OH)D), and to test this assumption we computed the Vitamin D Genetic Score and measured 25(OH)D estimates in each of 100 strata and then examined the heterogeneity between strata using the trend test and Cochran Q test[6] (**Supplement Figure 4**).

Fractional polynomials of degree 1 are defined as:

$$f(x) = \beta_0 + \beta_1 x^p \quad (3)$$

where, p ∈ {-2, -1, -0.5, 0, 0.5, 1, 2, 3}, and power of 0 refers to the natural-log function.

Fractional polynomials of degree 2 are defined as:

$$f(x) = \begin{cases} \beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2} \ if \ p_1 \neq p_2 \\ \beta_0 + \beta_1 x^p + \beta_2 x^p \log(x) \ if \ p_1 = p_2 = p \end{cases} \quad (4)$$

where, p_1 and p_2 ∈ {-2, -1, -0.5, 0, 0.5, 1, 2, 3}, and power of 0 refers to the natural-log function.

Stratified Mendelian randomization:

In the stratified Mendelian randomization analysis, the sample is stratified into 4 strata of residual measured 25(OH)D, including < 25 nmol/L, 25-49.9 nmol/L, 50-74.9 nmol/L and ≥ 75 nmol/L. Stratum-specific Mendelian randomization estimate is computed using the Vitamin D Genetic Score and the ratio of coefficients method as described in the linear Mendelian randomization analysis (see section linear Mendelian randomization).

SNP-based two-sample methods:

To address potential horizontal pleiotropy, we also compute stratum-specific Mendelian randomization estimates using five SNP-based two-sample methods, including inverse-variance weighted method (IVW), Mendelian randomization Egger method, weighted median Mendelian randomization method, weighted mode method, and Mendelian randomization pleiotropy residual sum and outlier (Mendelian randomization Presso) test.

These five methods are complementary to each other in the sense that they are robust to different patterns of horizontal pleiotropy, and therefore a good agreement across them suggests strong evidence of causal association. In contrast to the genetic score-based analysis, the SNP-based two-sample methods use SNP-level information and take SNP-exposure and SNP-outcome association estimates as inputs (**Supplement Figure 3**). To compute stratum-specific Mendelian randomization estimate, we take the SNP-mortality estimate from the corresponding stratum, and SNP-measured 25(OH)D association estimate from the non-deceased individuals of the whole cohort. Exclusion of deceased individuals for the SNP-measured 25(OH)D association estimate is to avoid potential bias due to overlapping samples.[9] All five two-sample Mendelian randomization methods were conducted using the 35 vitamin D variants that were used to construct the Vitamin D Genetic Score (Supplement Table 7.). Further, since evidence of association using the genetic score based approach (Supplement Figure 3) was only observed for the stratum with residual measured 25(OH)D < 25nmol/L, we restricted SNP-based two-sample analyses to two stratums of residual measured 25(OH)D, below or above 25 nmol/L (Supplement Figure 3).

Five SNP-based two-sample methods are described as follows:

- **IVW:** IVW estimator (β, equation 5) is obtained by the weighted regression of the effect estimates of variant-outcome association ($\hat{\Gamma}_j$, equation 5) on the effect estimates of variant-exposure association ($\hat{\gamma}_j$, equation 5),[10] with the weight being the inverse variance of variant's ratio estimate. IVW assumes that there is no pleiotropy of variants, or the net pleiotropy is zero.

$$\hat{\Gamma}_j = \beta\,\hat{\gamma}_j + \epsilon_j,\ \epsilon_j \sim N\left(0, \delta_j^2\right) \quad (5)$$

where, the subscript j denotes a variant.

- **Mendelian randomization Egger:** Mendelian randomization Egger is similar to the IVW method, except that the regression model additionally included an intercept term, β_0 (equation 6), which represents the average pleiotropic effect.[11] The Mendelian randomization Egger method can account for directional pleiotropy (when $\beta_0 \neq 0$), however, it introduces an additional untestable assumption, which states that pleiotropic effect must be independent of variant-exposure association.

$$\hat{\Gamma}_j = \beta_0 + \beta\,\hat{\gamma}_j + \epsilon_j,\ \epsilon_j \sim N\left(0, \delta_j^2\right) \quad (6)$$

where, the subscript j denotes a variant.

- **Mendelian randomization Presso:** In Mendelian randomization Presso,[12] the IVW method is performed iteratively with each variant omitted from the analysis in turn. At each iteration, the residual sum of squares (RSS) is calculated and used as a heterogeneity measure for the ratio estimate. If the observed RSS is substantially large in the null distribution of the expected RSSs, then the corresponding variant is considered as an outlier. The causal estimate is obtained by the IVW method with all outlying variants removed.

- **Weighted mode Mendelian Randomization method:** In the weighted mode method, variants (which are weighted by the inverse variance of their ratio estimates) are clustered into groups by the similarity of their ratio estimates.[13] The clustering is performed using a normal kernel density function, with its spread depending on a bandwidth parameter. The causal effect is estimated based on the cluster with the largest weight. The weighted mode method will return an unbiased causal effect if the variants within the cluster with the largest weight are valid instruments.

- **Weighted median Mendelian randomization method:** In the weighted median-based approach,[14] variants are ranked by their ratio estimates, with the causal estimate being the median of the distribution. For the weighted version of the median method, median is taken from a distribution of the ratio estimates in which variants with more precise ratio estimates receive more weight. The weighted median method will return a consistent estimate if at least 50% of the weight comes from valid instruments.

Leave-block-out analysis

For the leave-block-out analysis, we first grouped variants into functional blocks based on the traits that they are associated with from the PhenoScanner search, and then we removed one block of variants at a time and repeated the non-linear Mendelian randomization analysis with the Vitamin D Genetic Score constructed using the remaining variants. This process was iterated through all functional blocks.

PhenoScanner V2 is a curated database of publicly available results from large-scale genetic association studies.[16] We searched for evidence of variant-disease/trait associations in the PhenoScanner V2 database for all 35 variants (accessed on 09/09/2021). At the GWAS threshold (P <5E-08), 20 variants were associated with traits besides serum 25(OH)D (**Supplement Table 2**). Based on their associated traits, we grouped variants into 4 functional blocks, including blocks for blood, lipids/metabolic, and renal traits, with variants whose associated traits did not fall into one of these 3 blocks grouped together as the 'unclassified' block (**Supplement Table 3**).

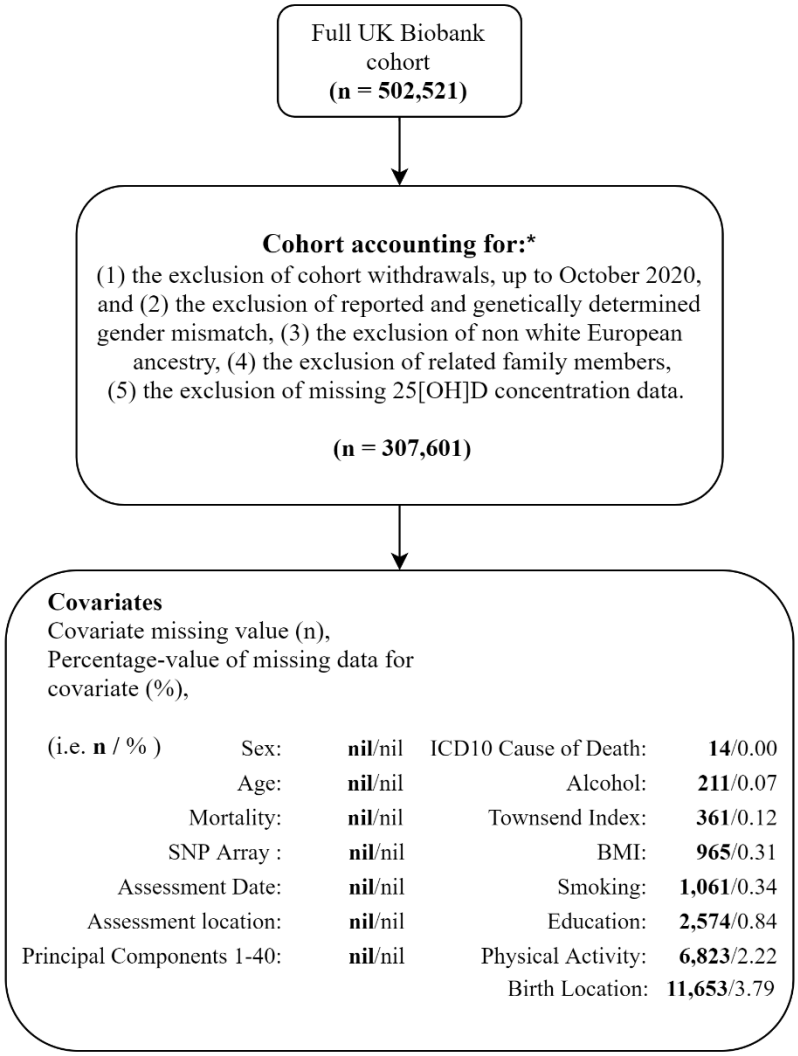
E-value

We computed E-value to evaluate sensitivity of our primary finding to unmeasured confounding. E-value is defined as the minimum strength of the association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to completely explain way an observed exposure-outcome association.[17] It is calculated as following: $RR + \sqrt{RR(RR-1)}$, where RR is the observed risk ratio in the study. When RR is less than 1, it needs to be inverted first before applying the formula.[17]

For our analysis, as E-value is computed on linear effect estimates,[17] it cannot be directly applied to the non-linear effect observed in our primary analysis where the Vitamin D Genetic Score and mortality association estimate varies by measured 25(OH)D concentration. To accommodate this, we have based our E-value analysis on the linear estimate of the Vitamin D Genetic Score - mortality association in the stratum of residual measured 25(OH)D below 25nmol/L, where the odd ratio for the risk of all-cause mortality is 0.92 per 1SD (~ 5.8 alleles) increase in the Vitamin D Genetic Score. The corresponding E-value is 1.40.

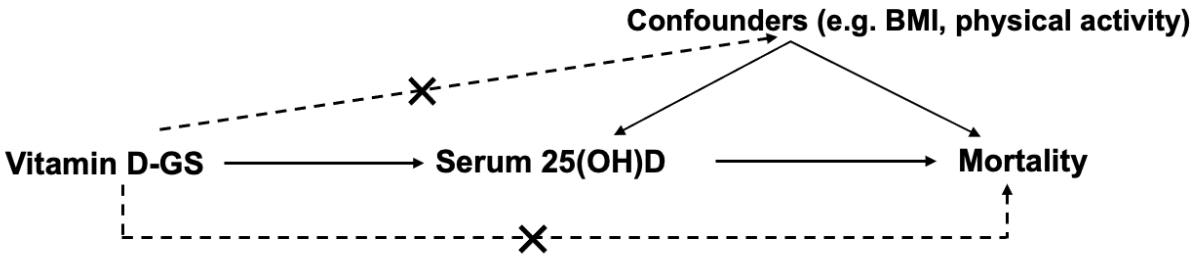
Interpreting E-value in terms of unmeasured confounding can be challenging, as it is context-specific, depending on outcome and exposure under consideration and what measured confounders that have been adjusted for.[18] In genetic analysis, population stratification is considered a major source of confounding,[19] and hence can be utilized to gain some sense as to what constitutes a small or large E-value for our Mendelian randomization analysis. We re-computed E-value using crude Vitamin D Genetic Score - mortality association estimate without accounting for indicators of population stratification, including birth locations and top 40 genetic principal components.

A near identical E-value (=1.39) is obtained, suggesting that the E value of 1.40 is reasonably large and that we can reasonably conclude that residual confounding is unlikely to explain away a Vitamin D Genetic Score - mortality association in the stratum of residual measured 25(OH)D below 25nmol/L.



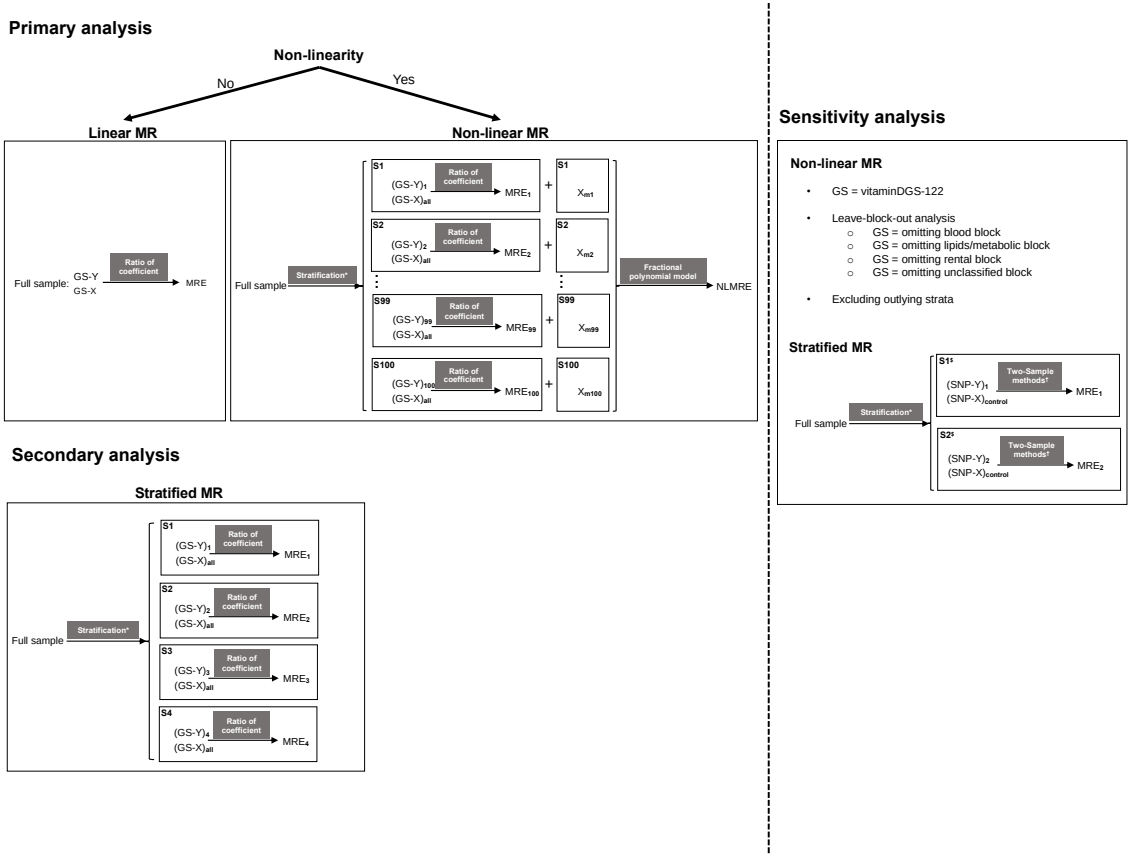
Supplement Figure 1. Study population and participant inclusion criteria

*Exclusion was done in a sequential order. * Ethnic sensitivity analysis required the reversal of restriction based on white European ancestry, resulting in the inclusion of an additional n = 20,837 / 5.65% individuals, comprising the non-white European category for analysis.



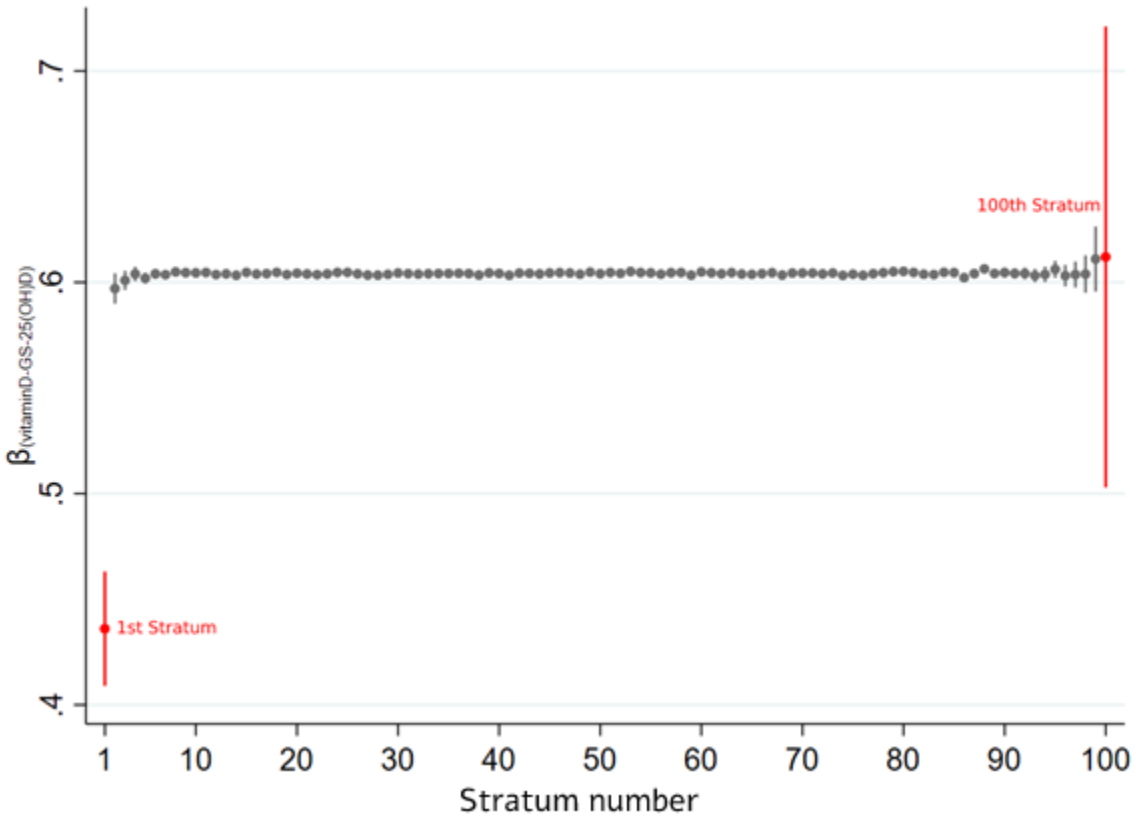
Supplement Figure 2. Mendelian randomization assumption diagram.

Valid causal inference from the Mendelian randomization analysis relies on 3 key assumptions:[20]1) Vitamin D Genetic Score associates with measured 25(OH)D concentrations; 2) Vitamin D Genetic Score has no direct effect on mortality; 3) Vitamin D Genetic Score does not associate with confounders of measured 25(OH)D and mortality. Vitamin D-GS: Vitamin D Genetic Score

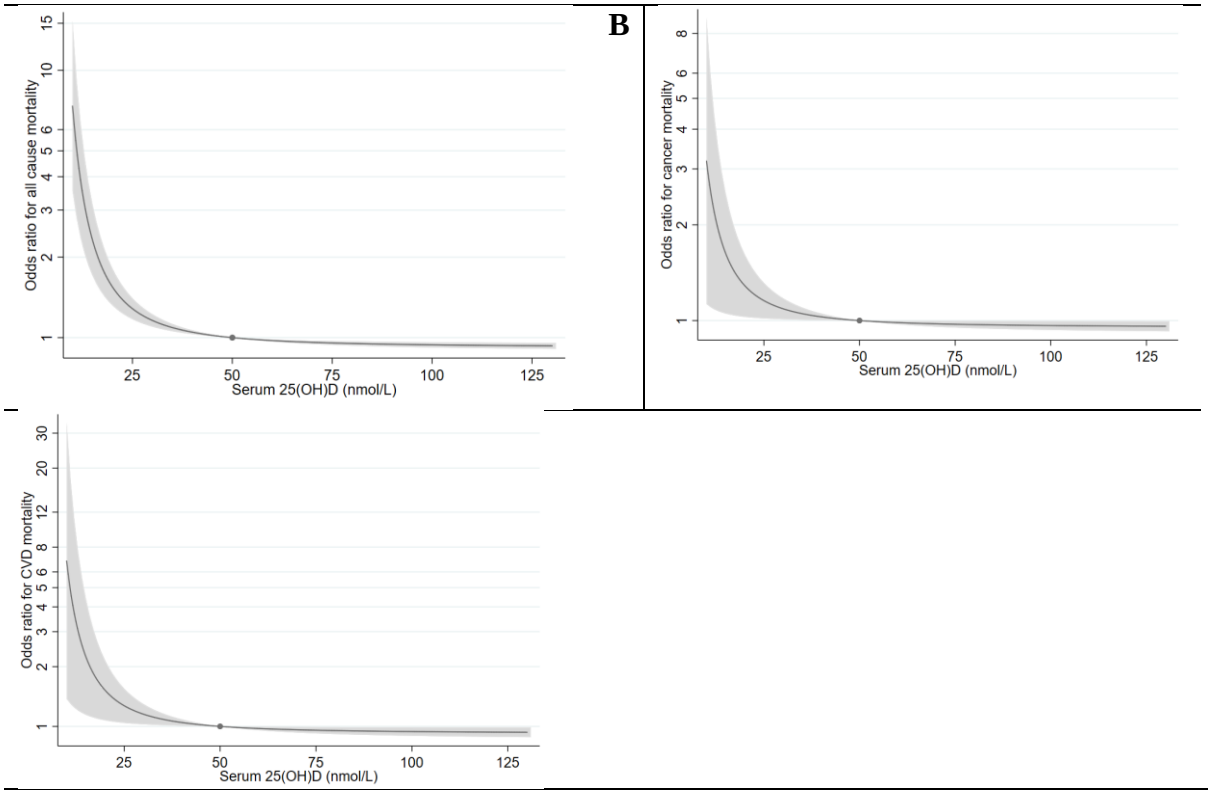


Supplement Figure 3. Schematic representation of Mendelian randomization analyses performed in the current study.

MR: Mendelian randomization; GS: genetic score; SNP: single nucleotide polymorphism; Sn: the nth stratum; MRE: MR estimates; NLMRE: non-linear MR estimates; GS-X: beta of GS-exposure association; GS-Y: beta of GS-outcome association; SNP-X: beta of SNP-exposure association; SNP-Y: beta of SNP-outcome association; *Stratification using residual measured 25(OH)D; In primary analysis, non-linear MR: (GS-X)all: GS-X from the full sample; (GS-Y)n: GS-Y from the nth stratum; Xmn: median of exposure in the nth stratum; In secondary analysis, stratified MR: S1: < 25 nmol/L; S2: 25-49.9 nmol/L; S3: 50-74.9 nmol/L; S4: >=75 nmol/L. In sensitivity analysis, stratified MR: S1$: < 25 nmol/L; S2$: >=25 nmol/L; (SNP-Y)n: SNP-Y from the nth stratum; (SNP-X)control: SNP-X from non-deceased individuals in the full sample; †Two-sample methods include Inverse Variance Weighted MR, MR-Egger, Weighted median MR, Weighted mode MR, and MR PRESSO.

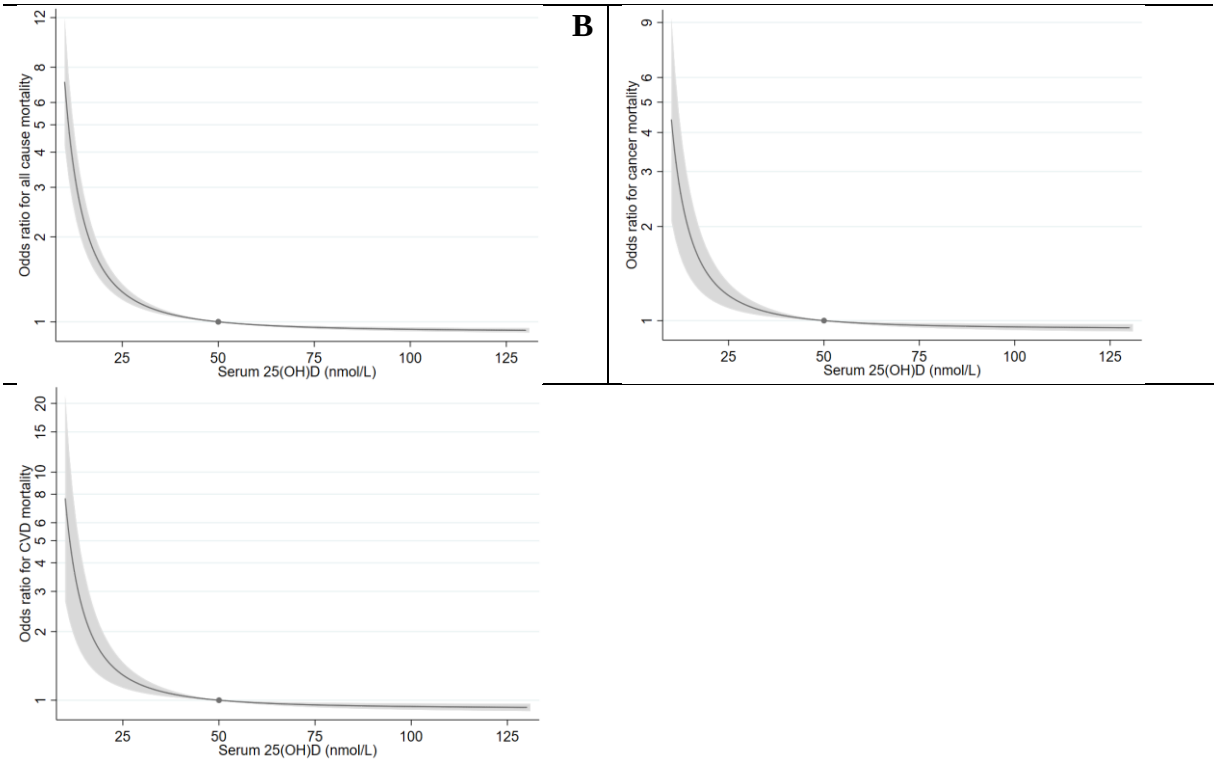


Supplement Figure 4. Distribution of βs for the association of Vitamin D Genetic Score with measured 25(OH)D across 100 strata of residual measured 25(OH)D. Vitamin D-GS: Vitamin D Genetic Score



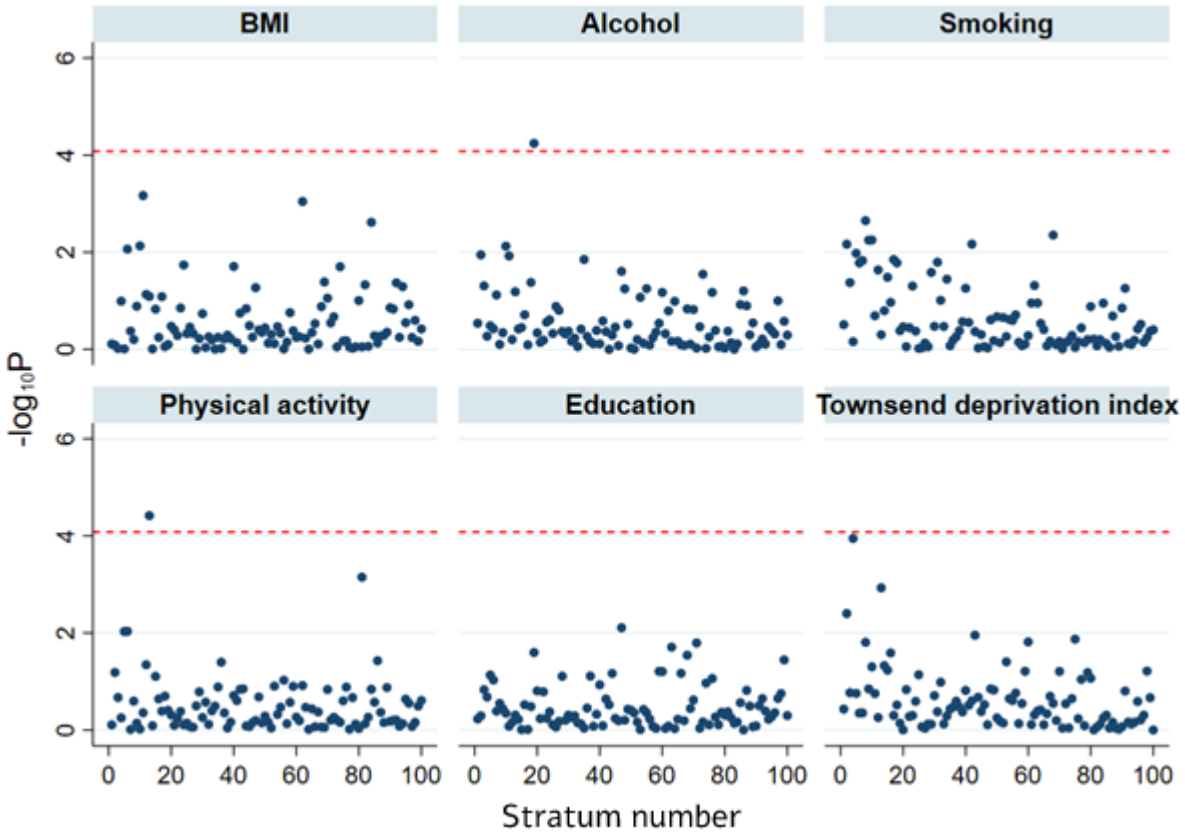
Supplement Figure 5. Non-linear Mendelian randomization analysis for the association of genetically predicted 25(OH)D with (A) all-cause mortality risk, (B) cancer mortality, and (C) cardiovascular disease (CVD) mortality in the UK Biobank, projected on the measured 25(OH)D scale; after excluding local causal estimates from the first and 100^{th*} stratum in the primary analysis.

The x axis – noted at serum 25(OH)D – refers to the measured 25(OH)D scale upon which the genetically predicted analysis has been mapped; the dot represents the reference point of 50 nmol/L. The shaded areas represent the 95% confidence intervals. The adjustment includes age, sex, assessment center, birth location, SNP array, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.
CVD: cardiovascular disease.*1st and 50th stratum for CVD.



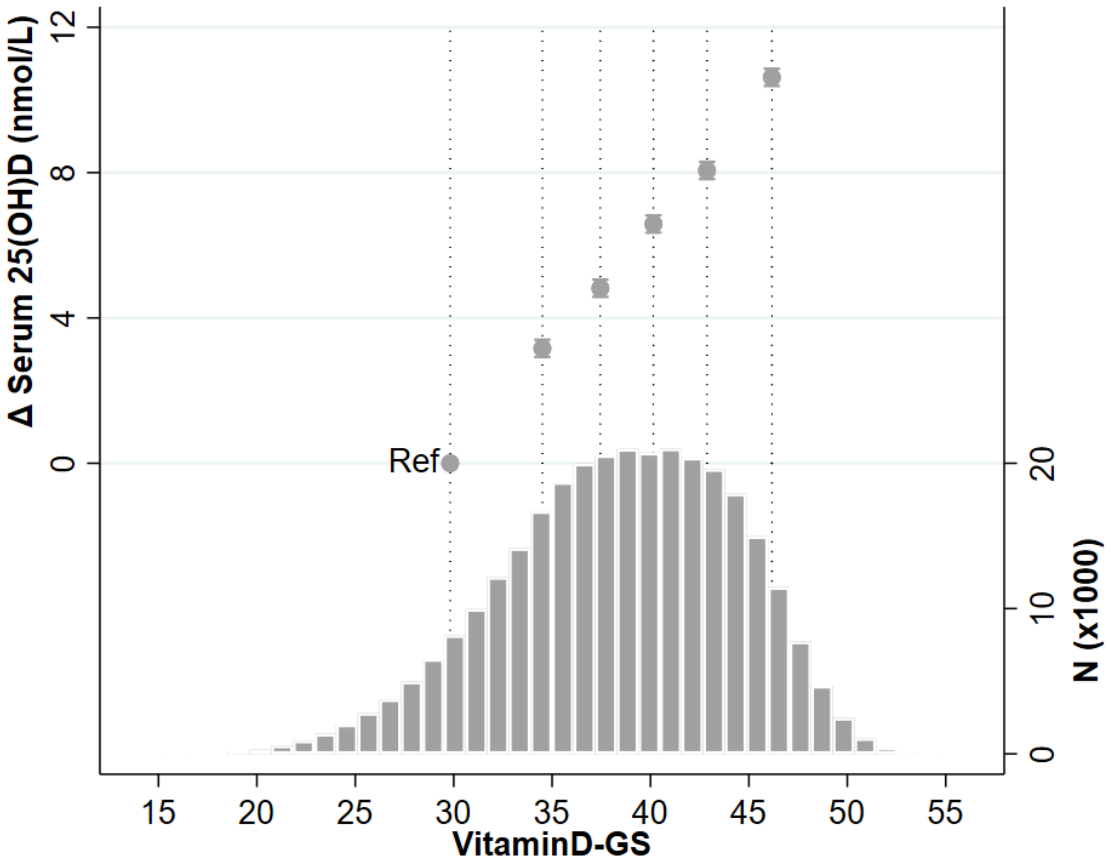
Supplement Figure 6. Non-linear Mendelian randomization analysis for the association of genetically predicted 25(OH)D with (A) all-cause mortality risk, (B) cancer mortality, and (C) cardiovascular disease (CVD) mortality in the UK Biobank, projected on the measured 25(OH)D scale; using the 122 SNP version of the Vitamin D Genetic Score.

The x axis – noted at serum 25(OH)D – refers to the measured 25(OH)D scale upon which the genetically predicted analysis has been mapped; the dot represents the reference point of 50 nmol/L. The shaded areas represent the 95% confidence intervals. The shaded areas represent the 95% confidence intervals. The adjustment includes age, sex, assessment center, birth location, SNP array, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. CVD: cardiovascular disease.



Supplement Figure 7. Association of the Vitamin D Genetic Score with potential confounders across 100 strata of residual measured 25(OH)D.

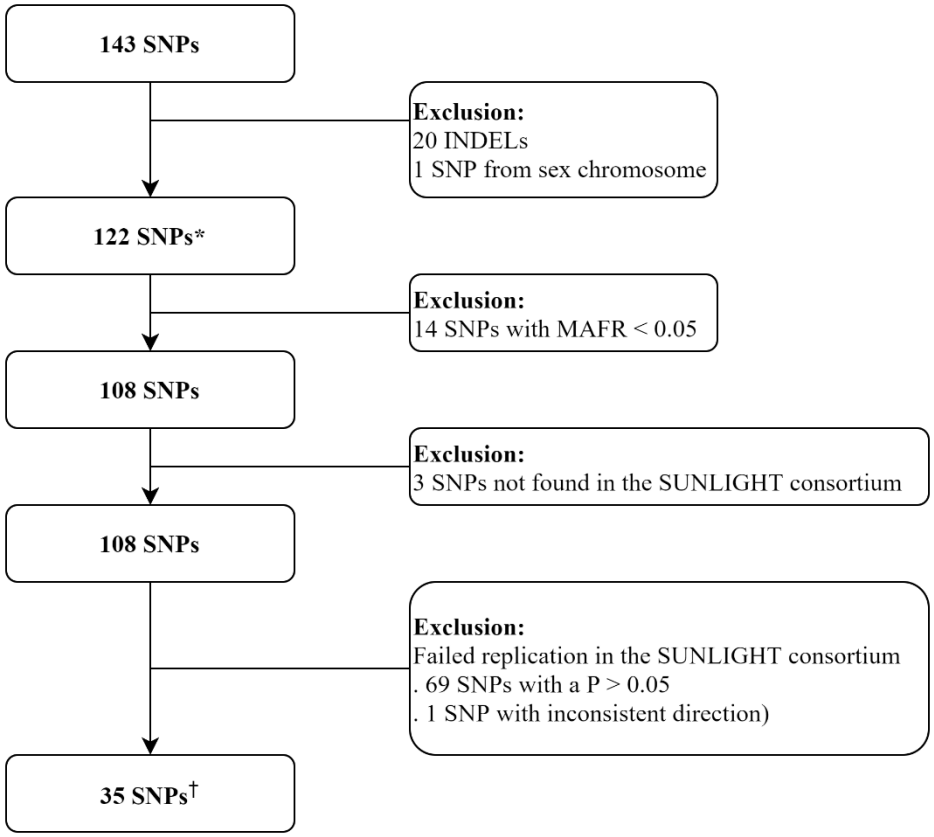
Potential confounders include BMI, alcohol intake, smoking status, physical activity level, education attainment and Townsend deprivation index. $P_{threshold}$ = 0.05/(6 confounders x 100 strata) = 8.3E-5. Dashed line represents -$\log_{10}$P at $P_{threshold}$.



Supplement Figure 8. Distribution of the Vitamin D Genetic Score and the association of sextiles of the Vitamin D Genetic Score with measured 25(OH)D concentration in the UK Biobank, illustrating the validity of the genetic instrument.

The association was tested using the linear regression, with the model adjusted for age, sex, genotyping array, birth location, assessment center location, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. Error bars are 95% confidence intervals.

Vitamin D-GS: Vitamin D Genetic Score



Supplement Figure 9. Selection of variants for the genetic instrument for measured 25(OH)D concentrations.

SNP: single nucleotide polymorphism. *SNPs used for the Vitamin D Genetic Score in the primary analysis; †SNPs used for the 122 SNP version of the Vitamin D Genetic Score in the sensitivity analysis. INDEL: insertion and deletion; MAF: minor allele frequency.

Supplement Table 1. Ethnic sensitivity analysis – Stratified Mendelian randomization of genetically predicted 25(OH)D and all-cause mortality in the UK Biobank.

	All-Cause mortality in non-white[b] ethnicities.	
25(OH)D	OR: 95% CI[a]	P
Overall		
non-stratified	0.66: 0.38, 1.12	
n/cases	20824/834	0.123
Residual categories		
<25 nmol/L	0.35: 0.15, 0.81	
n/cases	8089/355	0.014
25-49.9 nmol/L	0.67: 0.30, 1.56	
n/cases	9556/334	0.36
50-74.9 nmol/L	1.50: 0.36, 6.30	
n/cases	2563/114	0.58
≥75 nmol/L	0.04: 0.0007, 2.24	
n/cases	299/31	0.117

[a]OR per 10nmol/L increase in genetically predicted 25(OH)D. Adjusted for age, sex, ethnicity, assessment centre, SNP array, top 15 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. [b]Non-white ethnicities constituted: (1) Chinese, (2) Asian [Asian British, Indian, Pakistani, Bangladeshi, and any other Asian background], (3) Black African [Black British, Caribbean, African, and any other Black background], and (4) mixed ["Other" ethnic group, White and Black Caribbean, White and Black African, White and Asian, and any other mixed background].

Supplement Table 2. Functional blocks used in the leave-block-out analyses.

Functional block[a]	SNP	Gene	Traits
Blood traits	rs78151190	CARMIL1 (LRRC16A)	Various, incl. hemoglobin, platelet count, reticulocyte count
	rs1352846	GC	White cell count, granulocyte count
	rs1047891	CPS1	Various, incl. amino acid levels, platelet count, hemoglobin, metabolite levels
	rs72997623	DGAT2	Hemoglobin
	rs7528419	CELSR2	Blood protein levels, progranulin levels
	rs6672758	GALNT2	Platelet count, red cell distribution
Lipids/Metabolic	rs76798800	DCST2	Various, incl. fat-free mass, birth weight, height
	rs78151190	CARMIL1 (LRRC16A)	HbA1c, pulse, blood pressure
	rs77924615	PDILT	Blood pressure
	rs261291	ALDH1A2	Cholesterol, lipid metabolism
	rs727857	LINC01122	Fat mass, BMI, impedance, weight
	rs1047891	CPS1	Various, incl. weight, impedance, fat- free mass cholesterol
	rs12794714	CYP2R1	Hip circumference
	rs72997623	DGAT2	HDL cholesterol
	rs7528419	CELSR2	Angina pectoris, coronary artery disease, cholesterol, statin use
	rs75741381	VGF	Impedance
	rs6782190	CADM2	Fat free mass, BMI, impedance
Renal	rs77924615	PDILT	Glomerular filtration rate
	rs1047891	CPS1	Creatinine, chronic kidney disease
	rs17216707	CYP24A1	Creatinine, glomerular filtration rate
Unclassified	rs6782190	CADM2	Nervous feelings, alcohol intake, smoking, risk taking, number of children fathered
	rs61891388	RP11-867G23.13	Qualifications
	rs12803256	AP002387.1	Population differentiation
	rs1047891	CPS1	Headache
	rs212100	SULT2A1	Cholelithiasis, dehydroepiandrosterone sulphate
	rs7522116	FOXO6	Intelligence, qualifications
	rs78151190	CARMIL1 (LRRC16A)	Disorders of mineral metabolism
	rs261291	ALDH1A2	Age-related macular degeneration
	rs12056768	LINC00536	Hair or balding pattern: pattern 4
	rs1933064	FLG-AS1	Atopic dermatitis

[a]Functional blocks identified using trait associations identified through PhenoScanner V2[16].

Supplement Table 3. Leave-block-out analysis for non-linear Mendelian randomization analysis of genetically predicted 25(OH)D on risk for all-cause mortality.

Functional blocks excluded from the Vitamin D Genetic Score	SNPs in the functional block	$P_{Non\text{-}linear}$
Blood traits	rs1047891, rs1352846, rs6672758, rs72997623, rs7528419 and rs78151190	4.33E-05
Lipids/Metabolic	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs75741381 and rs6782190	5.72E-04
Renal	rs1047891, rs17216707 and rs77924615	6.25E-07
Renal or Metabolic	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs75741381, rs6782190 and rs77924615	1.88E-04
Unclassified	rs6782190, rs61891388, rs12803256, rs1047891, rs212100, rs7522116, rs78151190, rs261291, rs12056768 and rs1933064	1.08E-05

[a]Likelihood ratio test comparing the besting-fitting fractional polynomial model of degree 1 against the linear model

Supplement Table 4. Demographic characteristics of UK Biobank participants.

		Mortality	Cause Specific Mortality[a]		
		Deceased	Cancer	CVD	Respiratory
		N=18,700	N=10,230	N=3,895	N=1,016
	N(%)	%	%	%	%
All	307,601 (100)	6.08	3.43	1.33	0.35
Sex					
Men	144,680 (47.0)	8.16	3.98	1.95	0.45
Women	162,921 (53.0)	4.57	2.75	0.66	0.22
Age					
<60	169,756 (55.2)	3.12	1.77	0.56	0.14
≥60	137,845 (44.8)	10.13	5.24	2.14	0.57
BMI					
Low 25%[b]	76,640 (24.9)	5.23	2.86	0.75	0.43
Mid 50%[c]	153,321 (49.8)	5.86	3.25	1.15	0.26
High 25%[d]	76,675 (24.9)	7.93	3.93	1.97	0.36
Missing	965 (0.3)	19.27	4.97	4.25	1.66
Location					
South[e]	102,335 (33.3)	5.54	3.00	1.06	0.27
Mid[f]	144,654 (47.0)	6.46	3.39	1.33	0.37
North[g]	60,612 (19.7)	7.01	3.74	1.47	0.33
Smoking					
Non-smokers	167,703 (54.5)	4.35	2.40	0.82	0.13
Ex-smokers	108,118 (35.2)	7.59	4.01	1.57	0.40
Current smokers	30,719 (10.0)	11.80	5.89	2.57	1.13
Missing	1,061 (0.3)	11.78	5.09	2.64	1.60
Alcohol					
Daily	65,542 (21.3)	6.92	3.67	1.40	0.36
1 to 4 times wk	155,608 (50.6)	5.34	2.98	1.07	0.23
1 to 3 times mo	34,098 (11.1)	5.37	2.84	1.08	0.31
Special occasion	32,179 (10.5)	7.80	3.98	1.57	0.49
Never	19,963 (6.5)	10.23	4.68	2.14	0.82
Missing	211 (0.07)	12.32	5.21	1.90	1.42
Physical activity					
Low	92,012 (29.9)	7.15	3.65	1.50	0.44
Moderate	149,205 (48.5)	5.55	3.10	1.09	0.24
High	59,561 (19.4)	5.72	3.14	1.13	0.24
Missing	6,823 (2.2)	14.36	5.61	3.22	1.63
Education					
None	52,193 (17.0)	11.01	5.31	2.46	0.87
NVQ/CSE/A-Levels	109,099 (35.5)	5.65	3.03	1.14	0.27
Degree/professional	143,735 (46.7)	4.91	2.80	0.91	0.17
Missing	2,574 (0.84)	10.80	5.24	2.21	0.70
Townsend index					
Q1 Deprivation ↓	76,793 (25.0)	5.27	3.05	0.97	0.17
Q2	76,821 (25.0)	5.56	3.19	0.99	0.21
Q3	76,815 (25.0)	6.00	3.21	1.25	0.30
Q4 Deprivation ↑	76,811 (25.0)	8.21	3.86	1.85	0.63
Missing	361 (0.1)	4.71	1.66	1.11	0.00

NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; SD, standard deviation; Q, quartiles. [a] Missing n= 14 (0.005%). [b]BMI 12.1-24.1, [c]BMI 24.1-29.8, [d]BMI, 29.8-74.7. [e]≤51° Lat., [f]52°-53° Lat., [g]54°- ≥55° Lat.

Supplement Table 5. Linear Mendelian randomization analyses of higher genetically predicted 25(OH)D and mortality in the UK biobank, overall and by stratified residual measured 25(OH)D categories.

	All Cause	Cancer	CVD	Respiratory[b]
25(OH)D	OR: 95% CI[a]. P	OR: 95% CI[a]. P	OR: 95% CI[a]. P	OR: 95% CI[a]. P
Overall				
non-stratified	1.01: 0.97, 1.06. 0.56	1.01: 1.06, 0.95. 0.73	1.02: 0.93, 1.12. 0.63	0.81: 0.68, 0.96. 0.017
n/cases	293860/18289	285310/9739	279276/3705	277837/985
Residual categories				
<25 nmol/L	0.79: 0.70, 0.89. 9×10^{-5}	0.82: 0.69, 0.96. 0.016	0.69: 0.54, 0.85. 1×10^{-3}	0.68: 0.47, 1.004. 0.052
n/cases	33434/3024	31833/1423	31077/674	30880/246
25-49.9 nmol/L	0.98: 0.91, 1.04. 0.48	0.97: 0.89, 1.06. 0.51	1.05: 0.92, 1.22. 0.44	0.74: 0.57, 0.97. 0.029
n/cases	121098/7566	117504/3972	115121/1589	114574/420
50-74.9 nmol/L	1.01: 0.94, 1.09. 0.74	1.03: 0.93, 1.13. 0.61	1.05: 0.89, 1.24. 0.57	0.66: 0.47, 0.93. 0.016
n/cases	105475/5843	102898/3266	100758/1126	100235/250
≥75 nmol/L	1.08: 0.94, 1.24. 0.28	1.05: 0.87, 1.25. 0.62	0.93: 0.67, 1.28. 0.65	1.04: 0.52, 2.09. 0.91
n/cases	33853/1856	33075/1078	32313/316	28460/69

[a]OR per 10nmol/L increase in genetically predicted 25(OH)D. Adjusted for age, sex, assessment centre, birth locations, SNP array, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

Supplement Table 6. Stratified Mendelian Randomization sensitivity analyses of genetically predicted 25(OH)D and all-cause mortality in the UK biobank.

	<25 nmol/L	**≥25 nmol/L**
n/cases	33434/3024	33735/3077
	OR: 95% CI[a]. P	OR: 95% CI[a]. P
Inverse Variance Weighted	0.79: 0.70, 0.90. 1.69×10^{-4}	1.00: 0.96, 1.04. 0.97
MR Egger[b]	0.80: 0.68, 0.95. 0.017	1.00: 0.94, 1.06. 0.95
Weighted median	0.84: 0.73, 0.97. 0.017	0.99: 0.94, 1.05. 0.79
Weighted mode	0.79: 0.69, 0.89. 4.28×10^{-4}	1.00: 0.95, 1.04. 0.84
MR Presso	0.79: 0.70, 0.90. 6.37×10^{-4}	1.00: 0.96, 1.04. 0.97

[a]OR per 10nmol/L increase in genetically predicted 25(OH)D. [b]Egger Intercept P value: < 25 nmol/L 0.81; ≥ 25 nmol/L 0.95. Adjusted for: age, sex, assessment centre, vitamin D month, aliquot, fasting, birth location, SNP array and top 40 genetic principal components. Stratification by residual measured 25(OH)D concentrations.

Supplement Table 7. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.

See Supplementary Table 01 of the primary thesis:

Supplementary Table 1: Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations. 321

Supplement Table 8. Association of the Vitamin D Genetic Score with potential confounders in the UK Biobank.

		Vitamin D-GS
	N (%)	Mean (S.D.)
BMI		
Lowest 25% - 12.1-24.1	76,640 (24.9)	38.41 (5.87)
Middle 50% - 24.1-29.8	153,321 (50.0)	38.42 (5.88)
Highest 25% - 29.8-74.7	76,675 (25.0)	38.45 (5.86)
P[1]		0.88
Smoking		
Non-smokers	167,703 (54.7)	38.42 (5.86)
Ex-smokers	108,118 (35.3)	38.44 (5.89)
Current smokers	30,719 (10.0)	38.39 (5.89)
P[1]		0.18
Alcohol intake		
Daily	65,542 (21.3)	38.34 (5.88)
1 to 4 times week	155,608 (50.6)	38.45 (5.87)
1 to 3 times month	34,098 (11.1)	38.39 (5.87)
Special occasion	32,179 (10.5)	38.36 (5.90)
Never	19,963 (6.5)	38.51 (5.85)
P[1]		0.02
Physical activity		
Low	92,012 (30.6)	38.44 (5.87)
Moderate	149,205 (49.6)	38.42 (5.88)
High	59,561 (19.8)	38.42 (5.87)
P[1]		0.67
Education		
None	52,193 (17.1)	38.44 (5.86)
NVQ/CSE/A-levels	109,099 (35.8)	38.44 (5.88)
Degree/professional	143,735 (47.1)	38.42 (5.88)
P[1]		0.55
Townsend deprivation index quartiles		
Q1 lowest	76,793 (25.0)	38.41 (5.86)
Q2	76,821 (25.0)	38.46 (5.89)
Q3	76,815 (25.0)	38.42 (5.87)
Q4 highest	76,811 (25.0)	38.43 (5.88)
P[1]		0.38

NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; SD, standard deviation; Q, quartiles. Threshold for statistical significance = 0.05/6(confounders) = 0.008. P[1] Values have been adjusted for age, sex, genotyping array, birth location, and assessment center. Vitamin D-GS: Vitamin D Genetic Score

Supplementary References:

1. DiaSorin. LIAISON - 25 OH Vitamin D TOTAL Assay 2019. (https://www.diasorin.com/sites/default/files/allegati_prodotti/25_oh_vit._d_total_m0870004213_e.pdf)

2. Revez JA, Lin T, Qiao Z, Xue A, Holtz Y, Zhu Z, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nat Commun. 2020;11(1):1647.

3. Jiang X, O'Reilly PF, Aschard HA-Ohoo, Hsu YH, Richards JBA-Ohoo, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun. 2018;09(1):2041-1723.

4. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. International journal of epidemiology. 2013;42(4):1134-44.

5. Palmer, T. M., J. A. Sterne, R. M. Harbord, D. A. Lawlor, N. A. Sheehan, S. Meng, R. Granell, G. D. Smith, and V. Didelez. 2011. “Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses.” Am J Epidemiol 173 (12): 1392–1403.

6. Staley, James R., and Stephen Burgess. 2017. “Semiparametric Methods for Estimation of a Nonlinear Exposure-Outcome Relationship Using Instrumental Variables with Application to Mendelian Randomization.” Genetic Epidemiology 41 (4): 341–52.

7. Cole SR, Platt RW, Schisterman EF et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol 2010;39:417–20.

8. Dudbridge F, Allen RJ, Sheehan NA et al. Adjustment for index event bias in genome-wide association studies of subsequent events. Nat Commun 2019;10:15611561.

9. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genetic Epidemiology. 2016;40(7):597-608.

10. Burgess, S., A. Butterworth, and S. G. Thompson. 2013. “Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data.” Genet Epidemiol 37 (7): 658–65.

11. Bowden, J., G. Davey Smith, and S. Burgess. 2015. “Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression.” Int J Epidemiol 44 (2): 512–25.

12. Verbanck, Marie, Chia-Yen Chen, Benjamin Neale, and Ron Do. 2018. “Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases.” Nature Genetics 50 (5): 693–98.

13. Hartwig, F. P., G. Davey Smith, and J. Bowden. 2017. “Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption.” Int J Epidemiol 46 (6): 1985–98.

14. Bowden, J., G. Davey Smith, P. C. Haycock, and S. Burgess. 2016. “Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.” Genet Epidemiol 40 (4): 304–14.

15. Barendregt JJ, Veerman JL. Categorical versus continuous risk factors and the calculation of potential impact fractions. Journal of Epidemiology & Community Health. 2010;64(3):209-12.

16. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS, Staley JR. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics. 2019 Nov 1;35(22):4851-4853

17 Swanson SA, VanderWeele TJ. E-Values for Mendelian Randomization. Epidemiology. 2020 May;31(3):e23-e24. doi: 10.1097/EDE.0000000000001164. PMID: 31996542.

18 Vanderweele, T.J., Mathur, M.B., 2020. Commentary: Developing best-practice guidelines for the reporting of E-values. International Journal of Epidemiology 49, 1495–1497.

19 Hellwege, J.N., Keaton, J.M., Giri, A., Gao, X., Velez Edwards, D.R., Edwards, T.L., 2017. Population Stratification in Genetic Association Studies. Current Protocols in Human Genetics 95, 1.22.1–1.22.23.

20. Burgess, S., Davey Smith, G., Davies, N.M., Dudbridge, F., Gill, D., Glymour, M.M., Hartwig, F.P., Holmes, M.V., Minelli, C., Relton, C.L., Theodoratou, E., 2020. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research 4, 186.

Study 3 Addendum

Addendum to published study:

Vitamin D Deficiency Increases Mortality Risk in the UK Biobank – A Nonlinear Mendelian Randomization Study

https://doi.org/10.7326/M21-3324

Joshua P Sutherland (BHSc Nut Med, Hons)[1], Ang Zhou (PhD)[1,2,3], Elina Hyppönen (PhD)[1,2]

1. Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, Australia.
2. South Australian Health and Medical Research Institute, Adelaide, Australia
3. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

Joshua P Sutherland: ORCID 0000-0003-4102-3164

Ang Zhou: ORCID 0000-0001-7088-0777

Elina Hyppönen: ORCID 0000-0003-3670-9399

Our study entitled Vitamin D Deficiency Increases Mortality Risk in the UK Biobank – A Nonlinear Mendelian Randomization Study [https://doi.org/10.7326/M21-3324] used both linear and non-linear Mendelian randomization [MR] methodology to explore the relationship between genetically predicted higher 25(OH)D and mortality in the UK Biobank (Sutherland, Zhou & Hyppönen 2022). The work presented a linear association for respiratory related mortality, and non-linear associations for all-cause, cardiovascular disease and cancer mortality. At the time of this study's undertaking, the work employed state-of-the-art methodology, which included the use of the residual method for nonlinear Mendelian randomization analysis (Staley & Burgess 2017), all of which was applied in good faith. In this addendum we address the critique of the residual method, and present a reanalysis of the data using doubly-ranked non-linear MR analyses (Tian et al. 2023), which is a new method developed after the publication of our paper.

In non-linear MR analysis, the initial step involves stratifying the cohort into subgroups of varying levels of exposure. Subsequently, stratum-specific linear MR estimates are calculated and then 'stitched' together to construct the overall exposure-outcome association curve. The initial stratification step is critical - directly stratifying on the exposure can lead to collider bias, as exposure is a common effect of genetic instruments and confounders in the MR setting. Two stratification methods have been proposed to circumvent the related bias: a residual method (Staley & Burgess 2017) and a doubly-ranked method (Tian et al. 2023). The residual method was developed earlier and was considered the state-of-the-art stratification method at the time of undertaking our research (Sutherland, Zhou & Hyppönen 2022). In the residual method, strata are created by segmenting individuals into quantiles of residual exposure. Residual exposure is the residual taken from the regression of exposure on the genetic instrument, representing the non-genetic proportion of the exposure. In contrast, the doubly-ranked method involves initially grouping the participants into pre-strata based on their genetic score, followed by further ranking the data by exposure levels within each pre-stratum. This exposure ranking is then used to stratify the

cohort. For the residual method, as it relies on the linear regression model to obtain residual exposure, this imposes a strong assumption on the effect of the genetic instrument on the exposure, and requires that the effect of the genetic instrument be linear and constant for all individuals in the population (Burgess 2023). Violation of this constant genetic effect assumption can lead to biased results, as demonstrated through extensive simulation analyses (Burgess 2023; Tian et al. 2023). Unlike the residual method, the doubly-ranked stratification method refrains from making any parametric assumptions regarding to the association of genetic instrument with exposure. Rather, it posits that an individual's counterfactual exposure ranking remains consistent across all levels of the genetic instrument, which is a much weaker assumption than the constant genetic effect assumption (Tian et al. 2023). In simulation analyses, the doubly-ranked method has been shown to be less sensitive to the violation of the constant genetic assumption than the residual method (Burgess 2023; Tian et al. 2023).

In our original study using the residual method (Sutherland, Zhou & Hyppönen 2022), to ensure the validity of the constant genetic effect assumption, we examined heterogeneity in the observed association of the genetic instrument with 25(OH)D levels across strata of 'residual' 25(OH)D, and we further confirmed in sensitivity analyses that any observed heterogeneity did not affect our findings. While we followed the best practice [at the time], in our analyses, the appropriateness of these assumption checks has now been called in question. Recent simulation studies (Tian et al. 2023) have demonstrated that even when the genetic effect truly varies across different subgroups of the population, the observed genetic association with the exposure can misleadingly appear uniform across strata identified using the residual method. This implies that the assumption of a constant genetic effect may have been breached in our non-linear MR analyses but remained unnoticed, potentially introducing bias into our findings. Indeed, there is now evidence that the genetic influence on 25(OH)D levels exhibits variability within the population (Burgess 2023). Further reservations regarding the residual method stem from observations in negative control outcome analyses – where the residual method identified a spurious non-linear relationships between genetically predicted 25(OH)D levels and both age and sex (Hamilton et al. 2023).

While we did not observe evidence for non-linear association with sex in analyses using the instrument and data as in our original paper, we did confirm a spurious non-linear association with age, aggravating the concerns for potential bias. We have now re-run the non-linear MR analyses using the doubly-ranked stratification method, including checks on the constant genetic effect assumption. We have defined the population, the instrument, and the outcomes as in our original analyses and adjusted for covariates as previously described. In **Figure 01**, we confirm that there is heterogeneity in genetic association with 25(OH)D across the strata, which suggests that the constant genetic effect assumption is violated also when using the instrument as defined by us. This justifies the re-analysis using the doubly-ranked method. We have also examined the properties of the strata by investigating the distribution of the 25(OH)D GRS and 25(OH)D across strata. Ideally, the strata should reflect the level of exposure while being independent of genetic score, which would avoid collider bias upon stratification. In **Figure 02-03** we show that the strata are positively correlated with the average 25(OH)D level, and in **Figure 04** that they are also independent of the 25(OH)D GRS. This provides reassurance that collider bias is unlikely to have operated in this setting.

In the reanalysis using the doubly-ranked method presented in Table 01, we do not see evidence for a non-linear association between 25(OH)D and all cause, CVD or cancer mortality, and we confirmed that this finding is not sensitive to the number of strata. This is consistent with the association between the 25(OH)D GRS and all-cause mortality fluctuating around zero across strata, as we show in **Figure 05**. We have also performed the power analysis. Assuming there exists a 20% [i.e. OR = 1.8] or smaller true protective effect for mortality per 1 SD increase in 25(OH)D level within the deficiency range, our analyses would be underpowered to detect such an effect in strata where mean 25(OH)D levels fall within the deficiency range [Figure 06-07]. For an effect size of 30% protection or smaller, these analyses are also underpowered in the deficiency range within the 100 strata breakdown, while the 10 strata breakdown would be relatively well powered for larger effect sizes. Our original analyses on all-cause

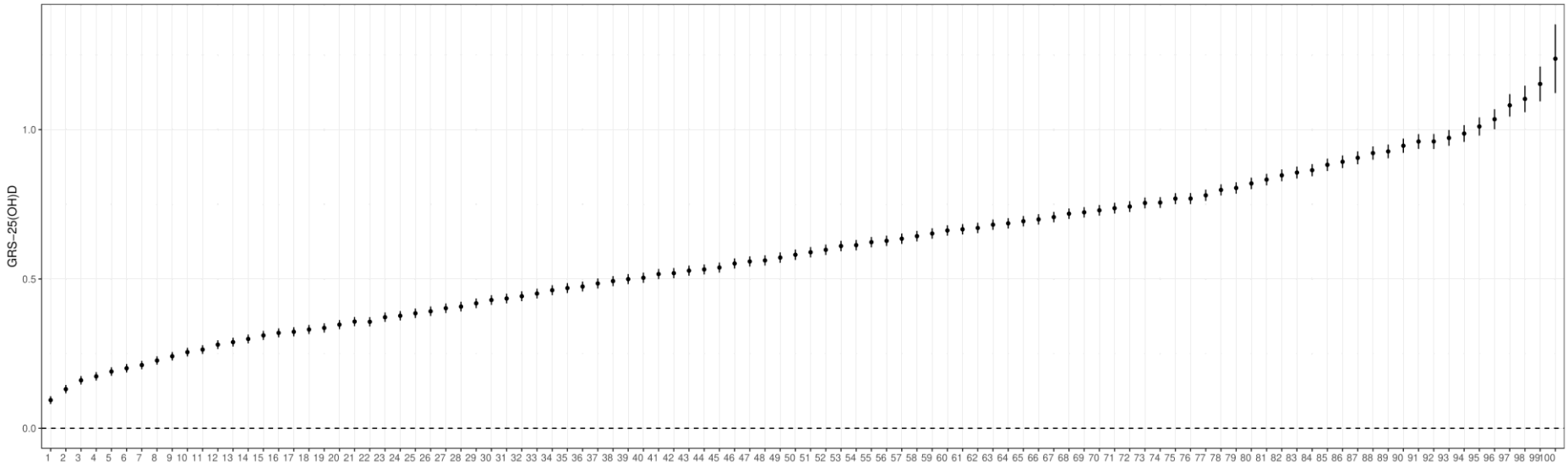
mortality, using the residual method, suggested that the non-linearity was stemming from a 21% mortality reduction, occurring in the <25 nmol/L category. Therefore, while the analyses using the doubly-ranked method did not provide evidence for an effect of 25(OH)D on mortality, the data is not sufficient to discount an effect that would be restricted to the correction of severe deficiency.

In summary, as reported in our original paper (Sutherland, Zhou & Hyppönen 2022) in the observational analyses we observe strong non-linear associations between 25(OH)D and all cause and cause specific mortality, with the risk of death notably higher for individuals with the lowest concentrations. Furthermore, our linear MR analyses provide support for a protective association between higher 25(OH)D and lower respiratory disease mortality. Re-analysis using the doubly-ranked method shows that for all cause, CVD or cancer mortality, genetic evidence for an association with 25(OH)D is not observed in either linear or non-linear analyses. While a beneficial effect arising from correction of severe vitamin D deficiency is biologically plausible and compatible with evidence from clinical trials (Argano et al. 2023; Bjelakovic et al. 2014; Cao et al. 2023; Zhang, R et al. 2022; Zhang, Y et al. 2019), with the exception of mortality from respiratory diseases, MR studies conducted to date have failed to provide robust evidence to support for underlying causality.

Addendum Table 1. Doubly-ranked Method Nonlinear MR test for nonlinearity across the varying number of strata.

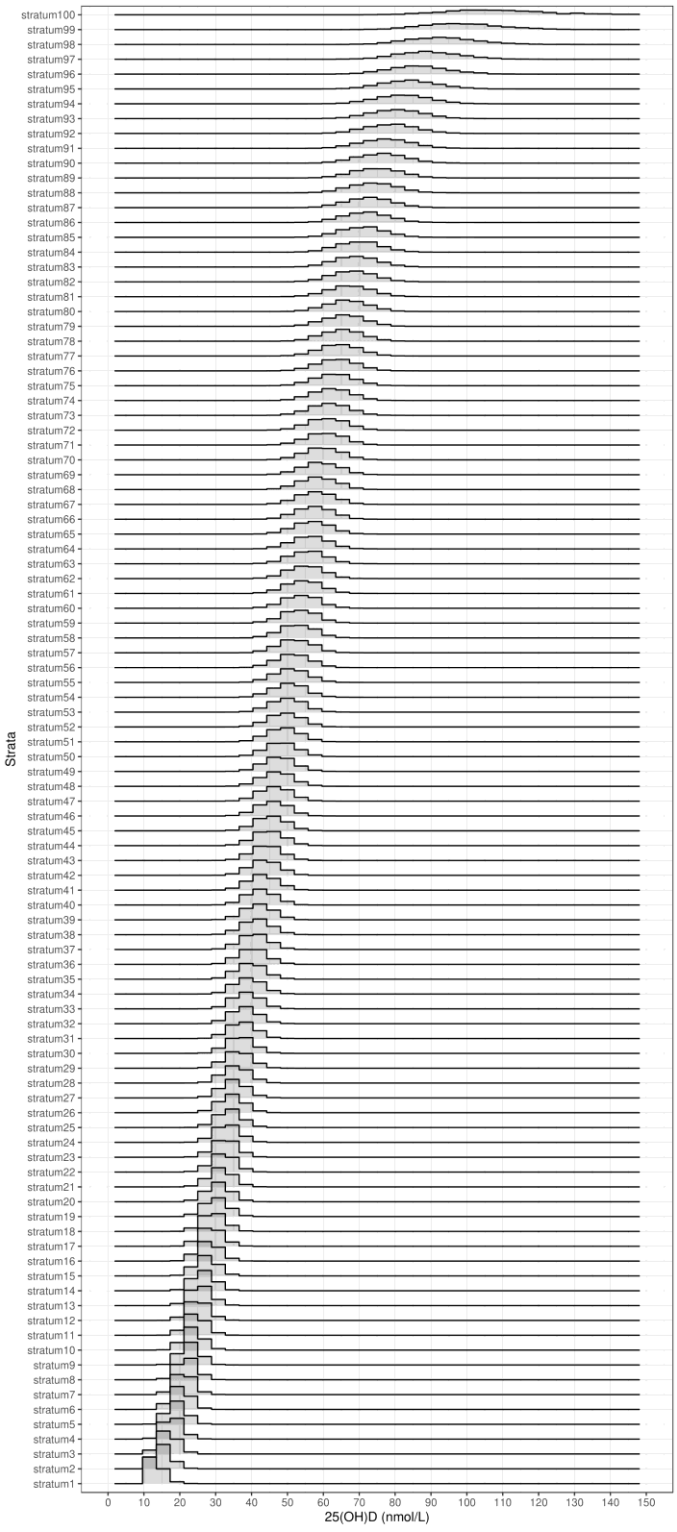
Outcome	Number of strata	Nonlinearity p-value
All-cause mortality	100	0.25
	5	0.41
	10	0.34
	25	0.35
	50	0.34
CVD mortality	100	0.75
	5	0.39
	10	0.33
	25	0.65
	50	0.69
Cancer mortality	100	0.73
	5	0.49
	10	0.40
	25	0.40
	50	0.27

The analyses were adjusted for age, sex, assessment centre, birth location, types of genotyping array, and the first 40 PCs.

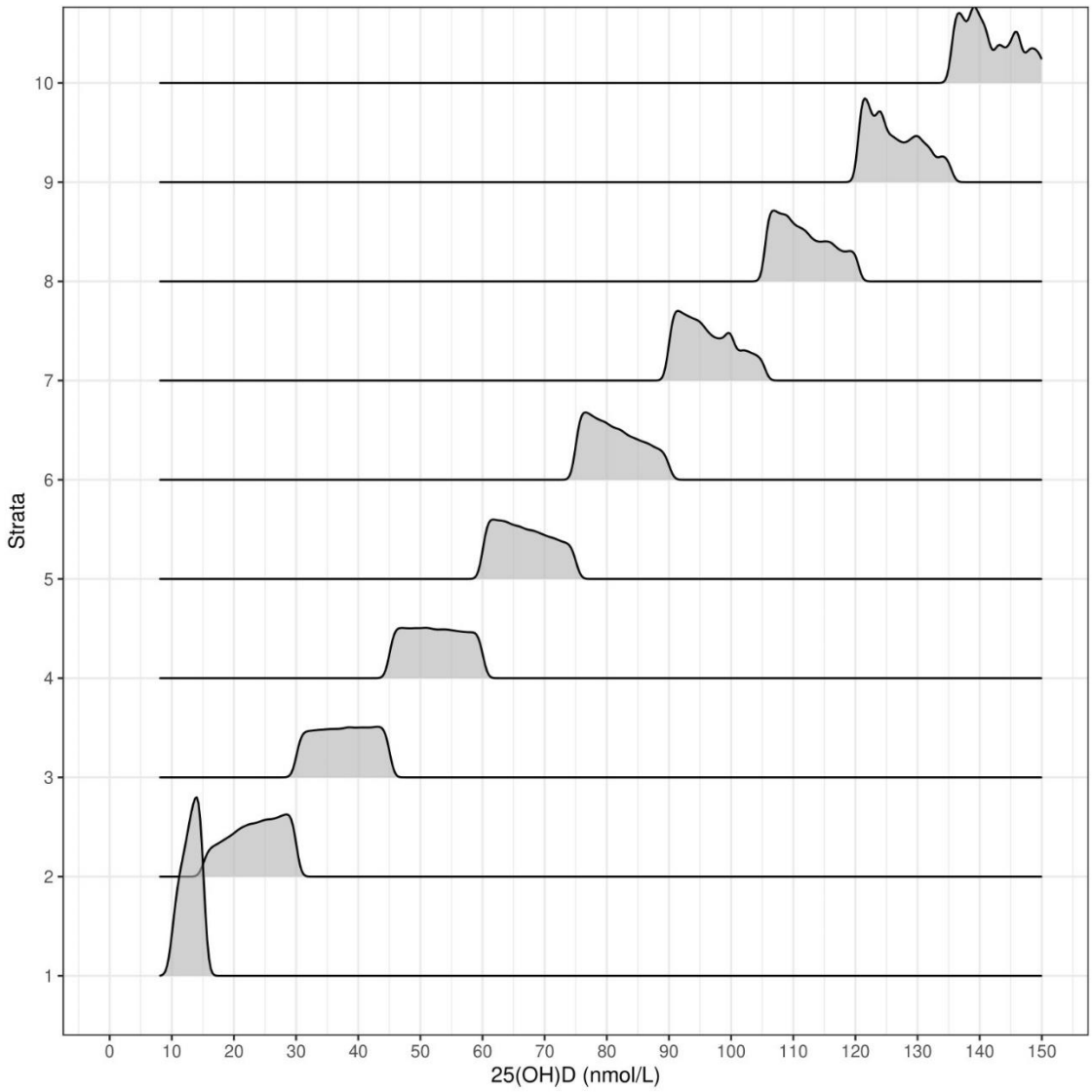


Addendum Figure 1. The estimated genetic association with 25(OH)D at each stratum by the doubly-ranked method.

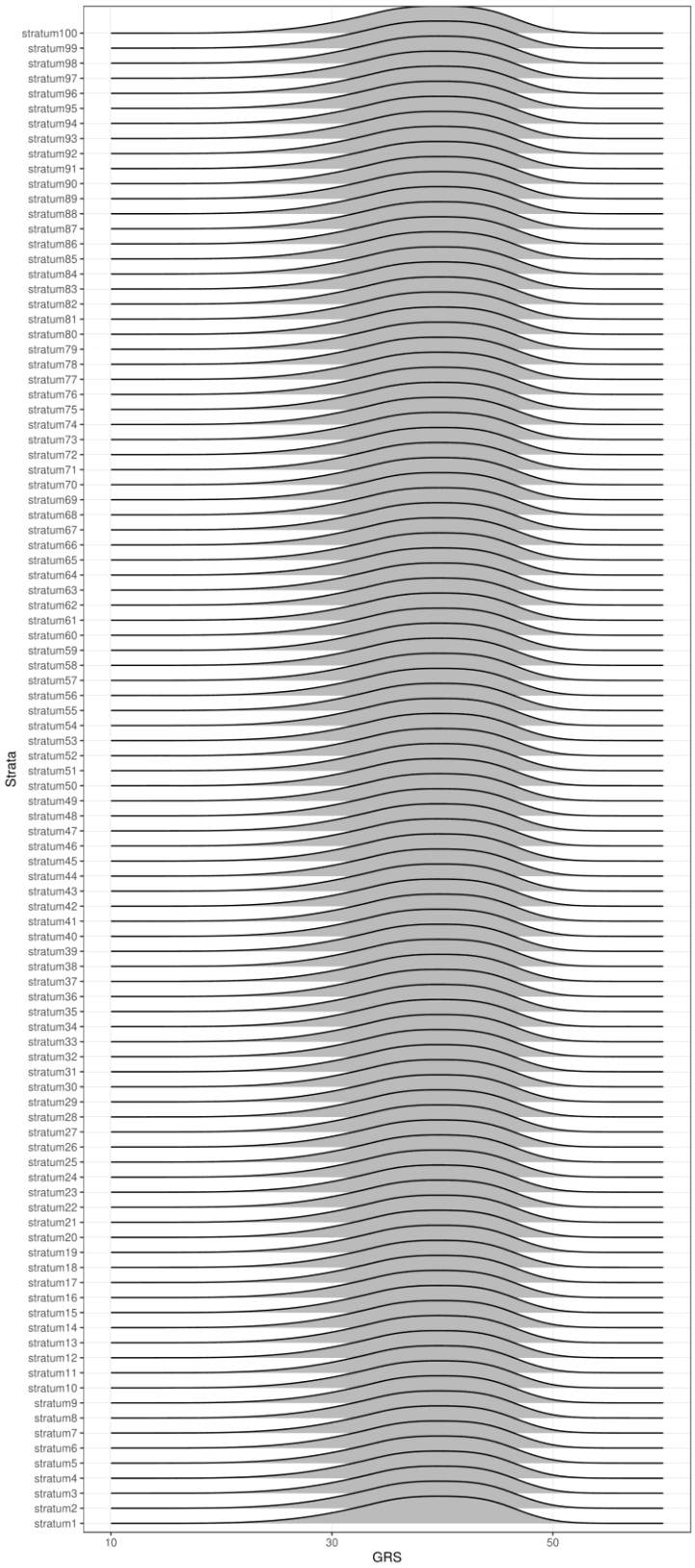
Y-axis stands for the coefficients between nmol/Ls of 25(OH)D against the GRS.



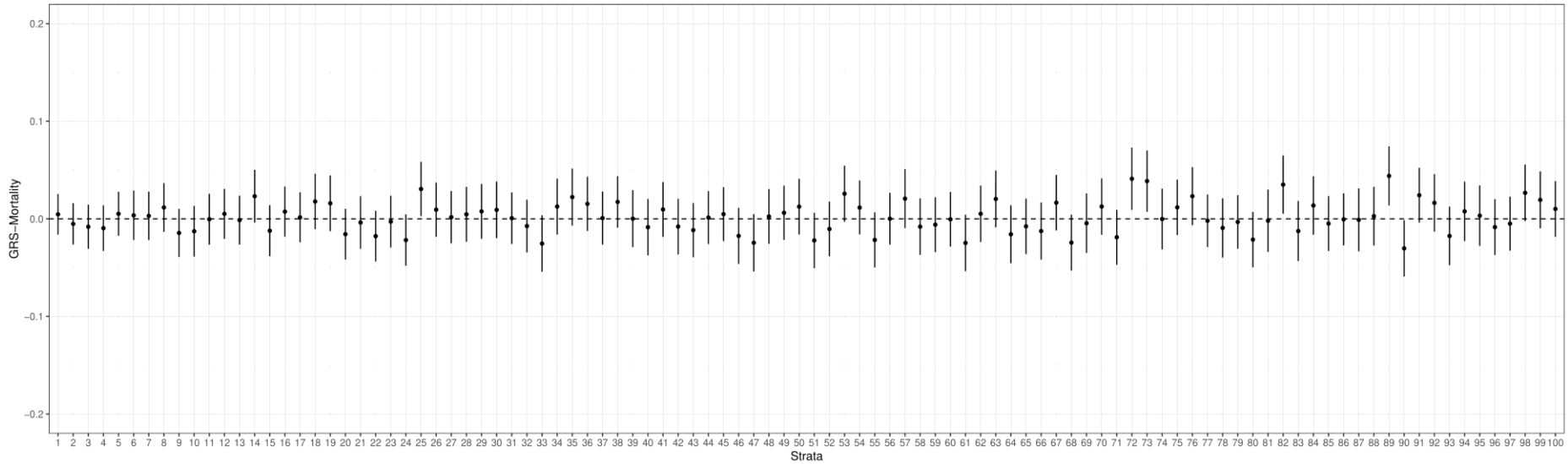
Addendum Figure S02. The distribution of 25(OH)D levels across the 100 strata.



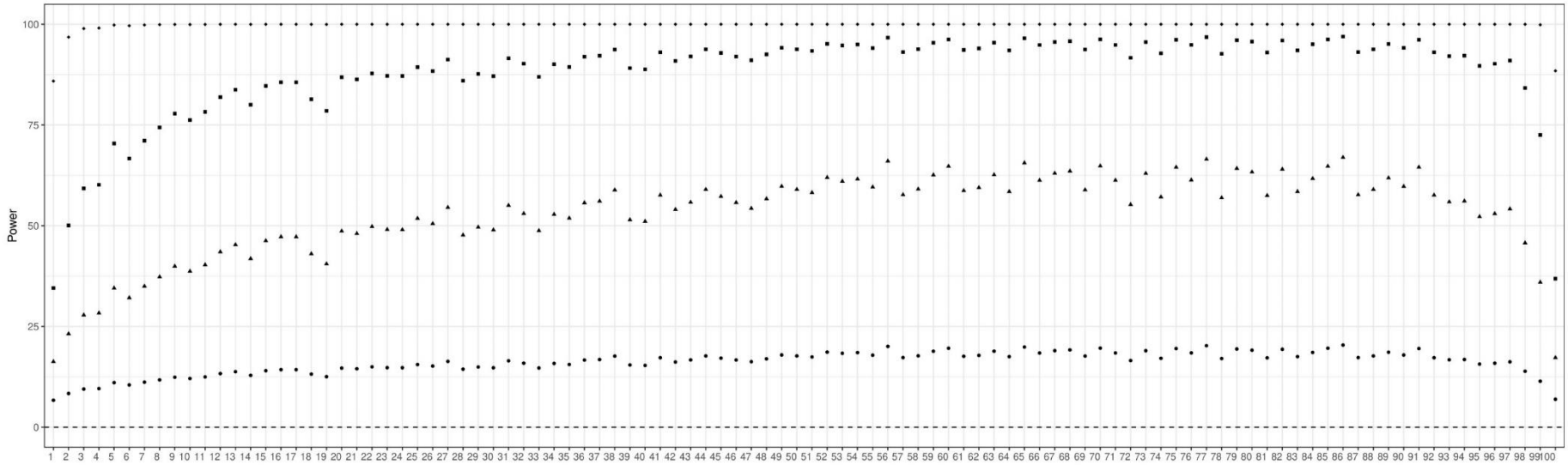
Addendum Figure S03. The distribution of 25(OH)D levels across the 10 strata.



Addendum Figure S04. The distribution of the 25(OH)D GRS across the 100 strata.

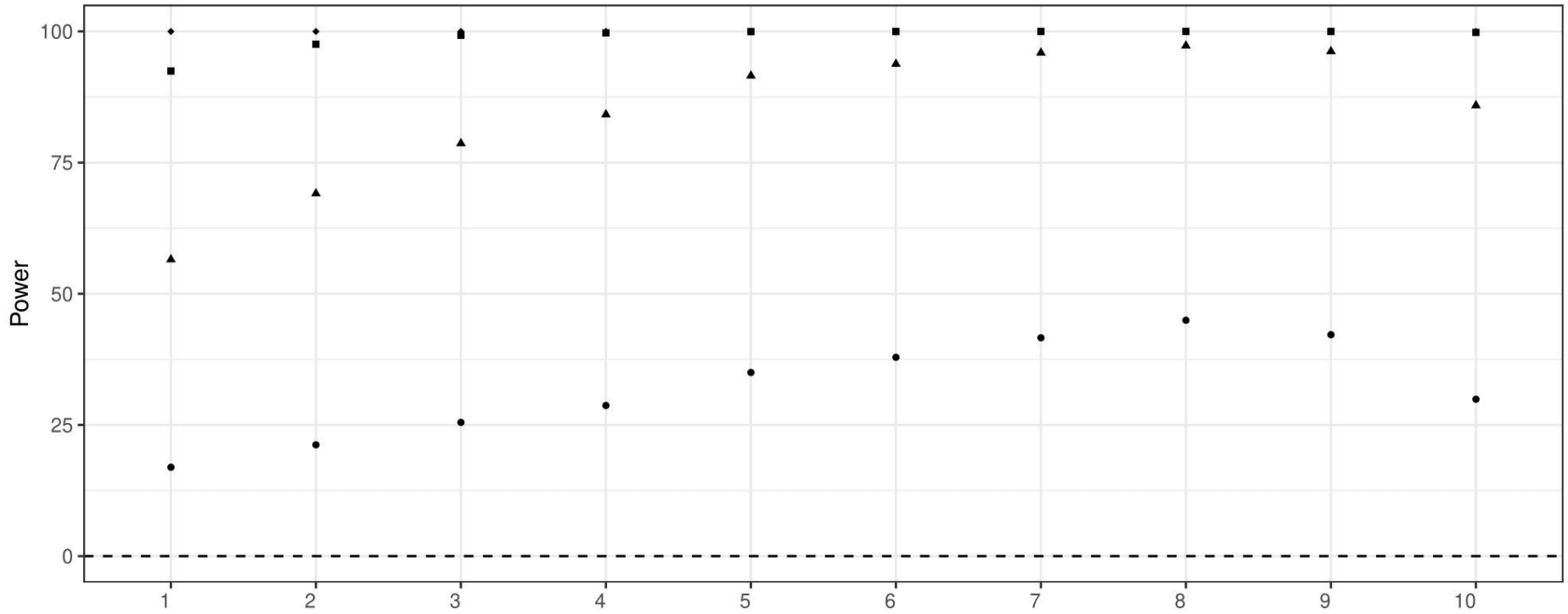


Addendum Figure S5. The association between the GRS and all-cause mortality, presented across the 100 strata.



Addendum Figure S6. Power calculation for all-cause mortality for a range of expected effect sizes (Odds Ratio 0.9, 0.8 0.7, and 0.5) within each strata at alpha (5%), presented across the 100 strata

- ● = power calculation for an expected effect size of 0.9 OR
- ▲ = power calculation for an expected effect size of 0.8 OR
- ■ = power calculation for an expected effect size of 0.7 OR
- ◆ = power calculation for an expected effect size of 0.5 OR



Addendum Figure S7. Power calculations for all-cause mortality for a range of expected effect sizes (Odds Ratio 0.9, 0.8 0.7, and 0.5) within each strata at alpha 5%, presented across the 10 strata.

- ● = power calculation for an expected effect size of 0.9 OR
- ▲ = power calculation for an expected effect size of 0.8 OR
- ■ = power calculation for an expected effect size of 0.7 OR
- ◆ = power calculation for an expected effect size of 0.5 OR

Addendum References:

Argano, C, Mallaci Bocchio, R, Natoli, G, Scibetta, S, Lo Monaco, M & Corrao, S 2023, 'Protective Effect of Vitamin D Supplementation on COVID-19-Related Intensive Care Hospitalization and Mortality: Definitive Evidence from Meta-Analysis and Trial Sequential Analysis', *Pharmaceuticals (Basel),* vol. 16, no. 1, Jan 16.

Bjelakovic, G, Gluud, LL, Nikolova, D, Whitfield, K, Wetterslev, J, Simonetti, RG, Bjelakovic, M & Gluud, C 2014, 'Vitamin D supplementation for prevention of mortality in adults', *Cochrane Database of Systematic Reviews*.

Burgess, S 2023, 'Violation of the Constant Genetic Effect Assumption Can Result in Biased Estimates for Non-Linear Mendelian Randomization', *Hum Hered,* vol. 88, no. 1, pp. 79-90.

Cao, M, He, C, Gong, M, Wu, S & He, J 2023, 'The effects of vitamin D on all-cause mortality in different diseases: an evidence-map and umbrella review of 116 randomized controlled trials', *Frontiers in Nutrition,* vol. 10.

Hamilton, FW, Hughes, DA, Spiller, W, Tilling, K & Smith, GD 2023, 'Non-linear mendelian randomization: evaluation of biases using negative controls with a focus on BMI and Vitamin D', *medRxiv*, p. 2023.08.21.23293658.

Staley, JR & Burgess, S 2017, 'Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization', *Genetic epidemiology,* vol. 41, no. 4, pp. 341-52.

Sutherland, JP, Zhou, A & Hyppönen, E 2022, 'Vitamin D Deficiency Increases Mortality Risk in the UK Biobank: A Nonlinear Mendelian Randomization Study', *Annals of Internal Medicine*.

Tian, H, Mason, AM, Liu, C & Burgess, S 2023, 'Relaxing parametric assumptions for non-linear Mendelian randomization using a doubly-ranked stratification method', *PLoS genetics,* vol. 19, no. 6, p. e1010823.

Zhang, R, Zhang, Y, Liu, Z, Pei, Y, Xu, P, Chong, W, Hai, Y, He, L, He, Y & Yu, J 2022, 'Association between vitamin d supplementation and cancer mortality: A systematic review and meta-analysis', *Cancers,* vol. 14, no. 15, p. 3717.

Zhang, Y, Fang, F, Tang, J, Jia, L, Feng, Y, Xu, P & Faramand, A 2019, 'Association between vitamin D supplementation and mortality: systematic review and meta-analysis', *BMJ*, p. l4673.

Chapter Eight: Study Four

Vitamin D, C-reactive protein, and increased fall risk: A genetic epidemiological study.

The study is unpublished, presented in manuscript form.

Manuscript

Supplementary Material

Addendum

Study 4 Manuscript

List of study 4 manuscript figures:

Study 4 - Figure 1: Phenotypic analyses for the association of measured 25(OH)D and falls in the UK Biobank, with the outcome restricted to falls concurring with C-reactive protein ≥ 20mg/L. 220

Study 4 - Figure 2: Mendelian randomization sensitivity analyses of genetically predicted 25(OH)D and falls concurring with C-reactive protein ≥ 20 mg/L in the UK biobank. 221

List of study 4 manuscript tables:

Study 4 - Table 1: Demographic characteristics of UK Biobank participants.....217

Study 4 - Table 2: Phenotypic analyses for the association between log transformed CRP and categorical 25(OH)D, and the association between 25(OH)D and log transformed CRP, in the UK Biobank.....219

Study 4 - Table 3: Phenotypic analyses for the association of categorical measured 25(OH)D and falls in the UK Biobank, with all falls (adjusted and simple models) and falls concurring with thresholds of C-reactive protein (adjusted and simple models). 220

Vitamin D, C-reactive protein, and increased fall risk: A genetic epidemiological study.

Joshua P Sutherland (BHSc Nut Med, Hons)[1], Ang Zhou (PhD)[1,2,3], Elina Hyppönen (PhD)[1,2]

1. Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, Australia.
2. South Australian Health and Medical Research Institute, Adelaide, Australia
3. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

Joshua P Sutherland: ORCID 0000-0003-4102-3164

Ang Zhou: ORCID 0000-0001-7088-0777

Elina Hyppönen: ORCID 0000-0003-3670-9399

Abstract:

Background: Falls are a major public health concern. Daily vitamin D supplementation is a proposed fall prevention strategy; however, safety concerns have arisen from some clinical trials showing increased fall risk when using higher vitamin D dosing methods. The relationship between vitamin D and falls may be influenced by factors such as inflammation, which can alter the balance of essential nutrients like vitamin D and retinol, potentially affecting motor function. We use a genetic epidemiological approach to explore the association of inflammation, vitamin D and fall risk.

Methods: We included 307,082 UK Biobank participants and conducted observational and Mendelian randomization (MR) analyses to investigate associations between 25-hydroxyvitamin D [25(OH)D] and fall risk, with analyses including restriction to participants who had fallen and had inflammation as defined by CRP $\geq$ 5 mg/L.

Results: In the observational analysis, CRP was associated with a higher (per 5 mg/L CRP increase OR = 1.06, 95% CI 1.05-1.07) and 25(OH)D with a lower odds of falls. The association between 25(OH)D concentrations and falls risk was non-linear ($P < 0.001$), reflecting a plateauing of the association at higher concentrations. There was an interaction between 25(OH)D and CRP on their association with the odds of falls ($P = 0.009$). In participants with CRP $\geq$ 5 mg/L, the association was U-shaped, and fall risk was elevated for both 25(OH)D < 25 nmol/L and $\geq$ 100 nmol/L ($P < 0.004$). The association between high 25(OH)D and falls was most pronounced for participants with CRP $\geq$ 20 mg/L ($\geq$ 100 nmol/L vs. 50-74.99 nmol/L: OR = 2.40, 95% CI, 1.50-3.86). Genetically predicted higher 25(OH)D was not associated with fall risk in the overall population, but a suggestive association with fall risk was seen in participants who had fallen and had CRP $\geq$ 20 mg/L (926 cases; OR = 1.20, 95% CI, 1.00-1.44).

Conclusion: Our study suggests that an inflammatory state can modify fall risk mediated by higher vitamin D. Both low and high 25(OH)D levels associate with more falls in individuals with chronic

inflammation, with supporting evidence seen both in observational and MR analyses. This may provide insight into the increased fall risk following high-dose vitamin D supplementation in clinical trials, warranting further research.

Key words: Falls, Motor Function, Vitamin D, Mendelian Randomization Analysis.

Introduction

Falls represent a significant public health concern, leading to considerable morbidity, mortality, and economic burden on healthcare systems globally [1, 2]. Research suggests that a single fall prevention program could prevent upward of 40,000 falls per year in the US and avoid approximately 100 to 400 million spent annually on medical expenses [3]. In people aged 65 and above, falls are the leading cause of injury [4], and therefore identifying factors that mitigate fall risk in older people is crucial. Daily vitamin D supplementation programs, aimed at redressing vitamin D insufficiency, have been highlighted as one avenue for fall prevention [3]. This is due to the relationship between vitamin D deficiency and increased fall risk [5], as well as the reduction in falls seen in clinical trials using daily dosing of vitamin D3 at 700 to 2000 IU/day, and especially when given to the vitamin D deficient individuals [6, 7].

The vitamin D and falls space, however, is not without controversy. The safety of vitamin D supplementation has been questioned after several clinical trials reported an increased fall risk among the elderly following annual/monthly bolus dosing or high daily doses (> 3200 IU [8, 9]). A study by Sanders et al. (2010) found an increased fall risk after administering a single oral dose of 500,000 IU of vitamin D3 [10]. Similar outcomes, in primary and subgroup analyses, were observed in other studies using monthly doses (60,000 IU to 100,000 IU) [11-13]. However, some notable trials, like the ViDA trial, did not report an increased fall risk despite using similar dosages (100,000 IU monthly) [14], and meta-analyses on vitamin D supplementation have shown conflicting results [8, 15]. Importantly, the Sanders study remains unique, as no other motor function related clinical trial has administered a comparably high single oral dose to an almost entirely vitamin D-replete population [10].

One aspect that might link vitamin D and the risk of falls, and which may explain the observed inconsistent associations, is inflammation. Inflammation (commonly indexed by elevated

C-reactive protein (CRP) concentrations [16]) is associated with a higher rate of falls [17-19], particularly in the elderly [20, 21]. Inflammation can also impact retinol status, which is important for the current context as retinol has a direct competitive relationship with vitamin D [22-24], and it plays an important role in the nervous system in supporting motor function regulation [25, 26]. Elevated CRP concentrations are associated with lower retinol levels, and for example, a modestly elevated CRP ($\geq$ 5 mg/L) has been associated with 25% lower retinol levels compared to those with CRP < 5 mg/L [27]. Given that retinol and vitamin D have a competitive relationship, inflammation could disrupt the balance between these two steroid hormones

Therefore, in this genetic epidemiological study we use data from the UK Biobank to safely explore the interrelationships between higher 25(OH)D, CRP, and the risk of falls.

Methods:

We used information from the UK Biobank which is a prospective cohort that consists of 502,316 individuals aged between 37-73 years, assessed from March 2006 to July 2010 in England, Scotland, and Wales [28]. The participants provided biological samples and participated in survey questionnaires and physical assessments. Analysis in our study was limited to unrelated participants of European ancestry, with MR analyses containing 307,082 individuals, upon restriction to the availability of serum 25(OH)D concentrations. A full overview of exclusion criteria, covariate missingness, composition of models—and the outcome variable case and control numbers therein—can be found in the supplementary appendix **(Figure S1)**. The UK Biobank was approved by the National Information Governance Board for Health and Social Care and the North West Multicenter Research Ethics Committee (11/NW/0382), and informed consent was obtained from all participants. The current study was conducted under UK Biobank project number 20175.

Baseline serum 25(OH)D concentrations were measured using the LIAISON XL 25(OH)D assay from DiaSorin (Stillwater, USA) and the details of the measurement procedure can be found

in the supplementary appendix, page 3. For models using the categorical 25(OH)D we divided the participants to those with 25(OH)D <25 nmol/L, 25-49.9 nmol/L, 50-74.9 nmol/L, 75-99 nmol/L, and ≥ 100 nmol/L, using participants with average/normal levels (50-74.9 nmol/L) as the reference group. Concentrations of serum CRP were measured using high sensitivity analysis via an immunoturbidimetric assay, performed on a Beckman Coulter AU5800 instrument, with an analytical range of 0.08 to 79.96 mg/L [29]. Information for the fall outcome was acquired at baseline via a self-reported questionnaire, containing the question "In the last year have you had any falls", which in our study was coded as a 'yes' or 'no' variable, excluding those who preferred not to answer. We considered 5mg/L as the initial CRP threshold [30], as this has been shown to be associated with ~25% retinol reduction [27]. As retinol deficiency could be expected to be the more severe the stronger the inflammation, we constructed separate outcomes to indicate falls concurring with levels of CRP ≥ 10 mg/L, ≥ 15 mg/L, and ≥ 20 mg/L (with ≥ 20 mg/ retaining n= 940 cases).

Information on potential confounders was obtained from self-reported touchscreen questionnaires (physical activity, smoking, alcohol, age, and sex), physical assessments (height and weight for BMI calculation), and residential address data (Townsend deprivation index). These covariates were used in model adjustments for the phenotypic analyses and are detailed in the supplementary appendix (page 3 to 4, and **Figure S2**).

For the genetic data, the UK Biobank used the Axiom array, which was designed specifically for the UK Biobank and includes approximately 820,000 genetic markers, both common and rare variants [31]. The UK Biobank conducted quality assurance processes to ensure the accuracy and reliability of their genotyping data relating to sample, genotyping, and imputation quality. These processes were described in detail in the genotyping and quality control report and elsewhere. [32, 33]. For our MR analysis, a weighted genetic score was developed by compiling 35 common autosomal SNPs, which were identified from a genome-wide association study conducted

on measured 25(OH)D concentration [34,35] and independently validated in the SUNLIGHT consortium [36]. Further details on the development of the vitamin D genetic score and the selection of the variants utilized can be found in the supplementary appendix (including **Figure S3** and **Table S1**).

Statistical methods:

Phenotypic analyses: We used linear regression to investigate 25(OH)D association with log transformed CRP, and logistic regression to assess the CRP–fall and 25(OH)D–fall associations. We examined if the 25(OH)D–fall association varied by CRP by including an interaction term [CRP x 25(OH)D] in the model. In the presence of interaction, we stratified the models by levels of CRP. To investigate whether the association of 25(OH)D with fall risk is similar for falls which occur in the presence of inflammation, we also conducted analyses defining the outcome of falls based on whether it concurred with elevated CRP concentrations at various thresholds ($\geq$ 5/10/15/20 mg/L). In the simple model we adjusted for age, sex, and assessment center. For analyses with 25(OH)D concentrations as the exposure only, we also controlled for nuisance factors which could affect the measured 25(OH)D concentrations (blood sampling month, fasting time before sample acquisition, and aliquot). In the full model we further adjusted for BMI, physical activity (low, moderate, and high), alcohol (daily or almost daily, once or twice a week, three of four times a week, one to three times a month, special occasions only, and never), smoking (non-smokers, former smokers, and current smokers), education (none, national vocational qualification / certificate of secondary education / A-levels, and degree / professional), and Townsend deprivation index (quartile division of index: least to most deprivation).

Mendelian randomization analyses: We examined genetic evidence for causality in the association of 25(OH)D and falls using the MR approach (**Supplementary Figure S4**). Our primary MR analyses were conducted using the weighted genetic score constructed from the 35 SNPs associated with 25(OH)D, where the weights were based on estimates of SNP-25(OH)D association from the SUNLIGHT Consortium meta-analyses [36]. We conducted sensitivity analyses applying complimentary two sample MR methods, notably IVW, MR-PRESSO, Weighted median, Weighted mode, and MR-Egger. Further details are provided the supplementary methods (pages 4 to 9).

Causal inference in Mendelian Randomization (MR) analysis relies on three assumptions [37] (Figure S2): 1) Genetic score for vitamin D associates with measured 25(OH)D concentrations. 2) Genetic score for vitamin D has no direct impact on falls. 3) Genetic score for vitamin D doesn't associate with confounders of 25(OH)D and falls. To assess the first assumption, we examined the association between the genetic score and 25(OH)D in the UK Biobank (**Figure S5**). Potential horizontal pleiotropy, where the genetic instrument affects outcomes through other pathways, can violate the second and/or third assumptions. To identify horizontal pleiotropy, we assessed the genetic score's association with potential confounders (**Table S2**). We also conducted a leave-block-out analysis, where we excluded variants which had been suggested to have effects also on additional traits (Supplementary appendix page 9, **Tables S3 and S4**) [38]. We also conducted sensitivity analyses using a broader 122 SNP version of the genetic score (Table S1).

R, version 3.6.1 was used for linear MR sensitivity analyses ('TwoSampleMR' and 'MRPRESSO' packages), and STATA, version 14.1 (Stata-Corp LP, College Station, Texas, USA) was used for all other analyses.

Role of funding source.

The study was supported by the National Health and Medical Research Council (NHMRC), grant number 11123603. Joshua P Sutherland studentship is funded by Australian Research Training Program Scholarship. Elina Hypponen is funded by the National Health and Medical Research Council (NHMRC) leadership award (GNT 2025349), Neither funding source had any role in this study's design, conduct or reporting.

Results

As shown in **Table 1** (N = 307,082) the average serum concentration of CRP and 25(OH)D in participants was 1.38 mg/L (geometric mean) and 49.8 nmol/L, respectively. CRP levels were higher in females vs. males, those aged ≥ 60 years vs < 60 years, current smokers and in those with the highest BMI. Compared to others, average levels of 25(OH)D were lower in those with the highest BMI, the lowest physical activity, the northernly located, and current smokers. Falls were more commonly reported by those older than age sixty compared to younger participants, by females compared to males, and by those with the most compared to least socioeconomic deprivation. The prevalence of falls in individuals with high CRP (reported at ≥ 5 mg/L and ≥ 20 mg/L, for the purpose of Table 1) was particularly evident in the obese, current smokers, the most socioeconomically deprived, and the non-educated.

Study 4 - Table 1: Demographic characteristics of UK Biobank participants

		CRP	25(OH)D	CRP ≥ 5 mg/L	CRP ≥ 20 mg/L	Falls All	Falls w/ CRP ≥ 5 mg/L	Falls w/ CRP ≥ 20 mg/L
		N=306,255 Geometric	N=307,082	N=34,961	N=3,519	N=59,825	N=8,992	N=940
	N(%)	Mean (SD)	Mean (SD)	%	%	%	%	%
All	307,082 (100)	1.38 (2.89)	49.8 (21.00)	11.42	1.15	19.48	03.51	0.38
Sex								
Males	144,418 (47.0)	1.33 (2.79)	49.9 (21.0)	09.85	1.15*	15.60	02.40	0.30
Females	162,664 (53.0)	1.42 (2.98)	49.8 (20.9)	12.81	1.15	22.93	04.57	0.46
Age								
< 60	169,475 (55.2)	1.25 (2.93)	48.4 (21.1)	10.41	0.98	17.68	03.01	0.31
≥ 60	137,607 (44.8)	1.55 (2.80)	51.6 (20.6)	12.66	1.36	21.70	04.14	0.46
BMI								
Low 25% 12.1-24.1	76,525 (24.9)	0.79 (2.80)	53.0 (22.0)	05.24	0.86	17.13	01.33	0.24
Mid 50% 24.1-29.8	153,052 (49.8)	1.33 (2.63)	50.9 (20.7)	08.70	0.99	18.17	02.33	0.26
High 25% 29.8-74.7	76,557 (24.9)	2.54 (2.56)	44.4 (19.2)	22.87	1.73	24.18	07.99	0.75
Missing	948 (0.3)	2.25 (3.17)	42.6 (21.3)	23.22	3.39	42.62	16.95	3.20
Location								
South ≤ 51° Lat	102,201 (33.3)	1.31 (2.90)	51.4 (20.5)	10.80	1.07	18.79	03.08	0.32
Mid 52°-53° Lat	144,370 (47.0)	1.41 (2.87)	49.9 (20.9)	11.54	1.18	20.02	03.69	0.41
North 54°- ≥ 55° Lat	60,511 (19.7)	1.43 (2.90)	46.8 (21.6)	12.16	1.20	19.35	03.81	0.41
Smoking								
Non-smokers	167,471 (54.5)	1.26 (2.86)	50.0 (20.6)	09.96	0.98	18.63	02.92	0.29
Ex-smokers	108,973 (35.2)	1.46 (2.87)	50.8 (21.0)	12.08	1.30	19.86	03.78	0.44
Current smokers	30,615 (10.0)	1.79 (2.95)	45.2 (21.8)	16.88	1.53	22.58	05.73	0.66
Missing	1,023 (0.3)	1.74 (2.88)	50.1 (21.8)	16.21	2.06	25.90	06.54	1.30
Alcohol								
Daily	65,476 (21.3)	1.26 (2.82)	51.2 (22.0)	09.57	1.07	18.38	02.60	0.29
1 to 4 times wk	155,379 (50.6)	1.30 (2.84)	50.9 (20.8)	10.11	0.99	17.95	02.80	0.30
1 to 3 times mo	34,047 (11.1)	1.51 (2.90)	48.1 (20.3)	13.07	1.21	20.54	04.09	0.41
Special occasion	32,096 (10.5)	1.74 (2.99)	46.2 (20.4)	16.53	1.63	24.02	06.07	0.66
Never	19,895 (6.5)	1.69 (3.06)	46.0 (21.0)	16.57	1.76	25.83	07.25	0.87
Missing	189 (0.07)	1.79 (2.96)	45.0 (20.5)	16.40	1.59	34.92	08.89	0.00
Physical activity								
Low	91,779 (29.9)	1.61 (2.92)	46.3 (20.2)	14.55	1.48	20.14	04.62	0.53
Moderate	149,059 (48.5)	1.28 (2.85)	50.6 (20.8)	09.89	1.00	18.34	02.81	0.29
High	59,492 (19.4)	1.23 (2.80)	54.0 (21.4)	09.02	0.82	19.69	02.63	0.22
Missing	6,752 (2.2)	2.23 (3.10)	43.3 (21.3)	23.61	2.84	33.84	12.27	1.76
Education								
None	52,009 (17.0)	1.84 (2.83)	50.4 (21.4)	16.29	1.66	23.79	06.02	0.73
NVQ/CSE/A-Levels	108,928 (35.5)	1.43 (2.87)	50.4 (21.2)	11.96	1.14	19.22	03.50	0.34
Degree/professional	143,627 (46.7)	1.20 (2.86)	49.0 (21.0)	09.19	0.96	18.08	02.62	0.29
Missing	2,518 (0.82)	1.65 (2.89)	50.4 (21.0)	14.22	1.31	21.84	04.65	0.35
Townsend index								
Q1 ↓ Deprivation	76,666 (25.0)	1.27 (2.81)	51.9 (20.7)	09.53	0.93	17.10	02.45	0.22
Q2	76,690 (25.0)	1.32 (2.83)	51.5 (20.7)	10.23	1.02	18.07	02.72	0.30
Q3	76,683 (25.0)	1.36 (2.88)	49.9 (20.8)	11.17	1.11	19.47	03.47	0.37
Q4 ↑ Deprivation	76,682 (25.0)	1.57 (3.01)	46.0 (21.1)	14.73	1.53	23.28	05.47	0.64
Missing	361 (0.1)	1.50 (3.12)	50.0 (20.6)	14.40	2.22	22.16	05.07	0.35

CRP, C-reactive protein; NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; Q, quartiles. All models were adjusted for sex, age, assessment centre, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots; P values for all models are ≤ 0.05, except * P = 0.37

Phenotypic analysis:

The association between CRP and 25(OH)D was non-linear (P < 0.001), and in the fully adjusted analyses, log transformed CRP levels were higher at both lower and higher 25(OH)D categories, relative to the 25(OH)D reference category of 50-74.99 nmol/L (**Table 2**). CRP was associated with an increased fall risk (P-non-linearity < 0.001), showing a positive association that plateaued at higher concentrations (**Table S5)**. We observed an inverse cross-sectional association between 25(OH)D concentrations and fall risk, with evidence of non-linearity reflecting a plateauing of the association at higher concentrations (P < 0.001) (**Table 3**). However, we also observed evidence for interaction suggesting that the 25(OH)D-fall relationship differed by CRP levels (P = 0.009). To explore this interaction, we conducted stratified analyses of the 25(OH)D-fall risk relationship above and below the different CRP thresholds ($\geq$ and < 5/10/15/20 mg/L). Below these respective CRP thresholds, we observed an inverse association between 25(OH)D and fall risk (**Table S6**). Above these thresholds, the inverse association remained for CRP $\geq$ 5 mg/L and $\geq$ 10 mg/L, but with greater uncertainty for higher thresholds (e.g., 838 cases and 2,421 controls for CRP $\geq$ 20 mg/L) (Table S6).

We then repeated the initial 25(OH)D-fall analysis, defining cases only as those who fell and had CRP above the thresholds of $\geq$ 5/10/15/20 mg/L (e.g., 838 cases and 239,322 controls for CRP $\geq$ 20 mg/L) (**Table S7**). In these analyses we observed strong evidence for non-linearity (P = 0.00009), and the highest odds of falls co-occurring with inflammation (CRP > 20 mg/L) was seen for those with the highest 25(OH)D levels, while a higher risk was also confirmed for participants with 25(OH)D < 25 nmol/L ($\geq$ 100 nmol/L vs. 50-74.99 nmol/L: OR = 2.40, 95% CI, 1.50-3.86; and < 25nmol/L vs. 50-74.99 nmol/L: OR = 1.29, 95% CI, 1.03-1.61) (Figure 1). The non-linear

association between 25(OH)D and increased fall risk was consistent and significant across the CRP thresholds, being the more pronounced the higher the CRP (Table 3).

As a sensitivity analyses to explore the association between 25(OH)D and fall risk in the context of high inflammation, we repeated the categorical analysis in a rheumatoid arthritis patient population (typically characterized by higher inflammation). In this setting, the association between both lower and higher 25(OH)D and increased fall risk was consistent and significant among those who fell and had rheumatoid arthritis (915 cases, 238,968 controls) (≥ 100 nmol/L vs. 50-74.99 nmol/L: OR = 1.64, 95% CI, 1.00-2.69, P < 0.05; and < 25nmol/L vs. 50-74.99 nmol/L: OR = 1.28, 95% CI, 1.03-1.60, P = 0.03) (see **Table S8**). To explore the association in the context where inflammation is not present, we repeated the categorical analysis restricting to falls occurring below < 5 mg/L CRP, and this reflected an association with fall risk in the deficiency categories (**Table S9**)

Study 4 - Table 2: Phenotypic analyses for the association between log transformed CRP and categorical 25(OH)D, and the association between 25(OH)D and log transformed CRP, in the UK Biobank.

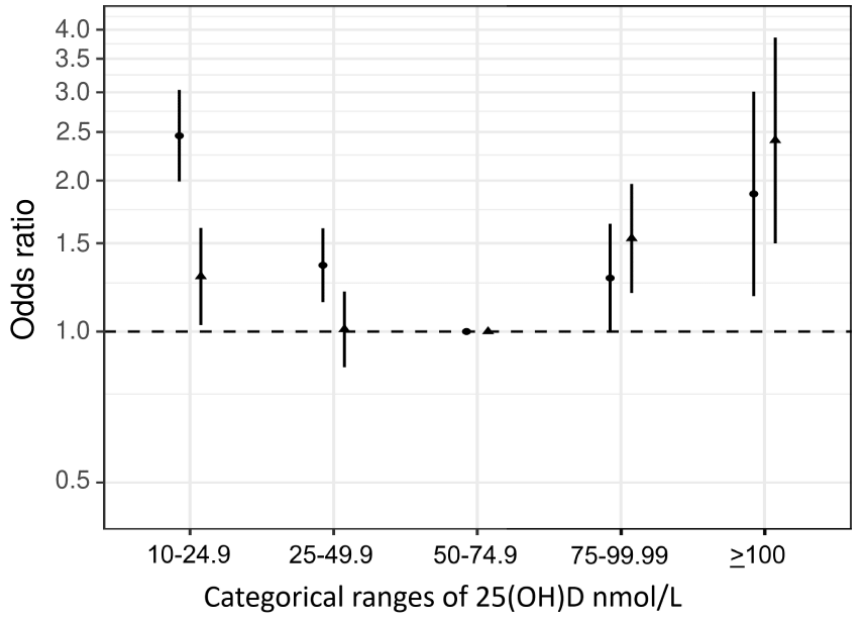
25(OH)D nmol/L	N	CRP Geometric Mean (SD). 95% CIs	Simple Model Beta: 95% CIs	Adjusted Model Beta: 95% CIs
<25 nmol/L	35917	1.63 (3.04). 1.61, 1.65	0.27: 0.26, 0.28	0.03: 0.02, 0.04
25-49.9 nmol/L	126950	1.44 (2.88). 1.43, 1.45	0.13: 0.12, 0.14	0.005: -0.003, 0.01
50-4.99 nmol/L	107450	1.29 (2.82). 1.29, 1.3	*Reference*	*Reference*
75-99.9 nmol/L	31843	1.20 (2.90). 1.19, 1.21	-0.08: -0.09, -0.07	0.007: -0.005, 0.02
100-124.9 nmol/L	4324	1.23 (2.95). 1.19, 1.27	-0.05: -0.09, -0.02	0.08: 0.05, 0.10
≥125 nmol/L	598	1.31 (3.17). 1.19, 1.44	0.05: -0.04, 0.13	0.18: 0.10, 0.26

Fully adjusted non-linearity P value: quadratic non-linearity < 0.001, log ratio test < 0.001. CRP = C-reactive protein. SD = standard deviation. CI = confidence intervals. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Study 4 - Table 3: Phenotypic analyses for the association of categorical measured 25(OH)D and falls in the UK Biobank, with all falls (adjusted and simple models) and falls concurring with thresholds of C-reactive protein (adjusted and simple models).

Fall type	**10-24.9 nmol/L** OR: 95% CI. P	**25-49.9 nmol/L** OR: 95% CI. P	**Reference**	**75-99.99 nmol/L** OR: 95% CI. P	**≥100 nmol/L** OR: 95% CI. P
All falls *Adjusted*	1.18: 1.14, 1.22. <0.001	1.05: 1.03, 1.08. <0.001	*Ref.*	1.00: 0.97, 1.04. 0.83	1.00: 0.93, 1.09. 0.92
All falls *Simple*	1.32: 1.28, 1.37. <0.001	1.10: 1.08, 1.13. <0.001	*Ref.*	0.97: 0.94, 1.01. 0.12	0.97: 0.90, 1.05. 0.40
With CRP *Adjusted*					
≥05 mg/L	1.25: 1.15, 1.34. <0.001	1.07: 1.01, 1.13. 0.02	*Ref.*	1.04: 0.95, 1.14. 0.40	1.31: 1.07, 1.60. 0.009
≥10 mg/L	1.33: 1.17, 1.50. <0.001	1.04: 0.95, 1.14. 0.39	*Ref.*	1.17: 1.01, 1.36. 0.03	1.75: 1.30, 2.35. <0.001
≥15 mg/L	1.39: 1.18, 1.65. <0.001	1.05: 0.93, 1.20. 0.42	*Ref.*	1.30: 1.07, 1.60. 0.01	1.93: 1.30, 2.87. 0.001
≥20 mg/L	1.29: 1.03, 1.61. 0.03	1.01: 0.85, 1.20. 0.91	*Ref.*	1.53: 1.19, 1.96. 0.001	2.40: 1.50, 3.86. <0.001
With CRP *Simple*					
≥05 mg/L	2.28: 2.12, 2.45. <0.001	1.43: 1.35, 1.51. <0.001	*Ref.*	0.86: 0.79, 0.94. 0.001	1.00: 0.82, 1.22. 0.97
≥10 mg/L	2.55: 2.27, 2.86. <0.001	1.41: 1.29, 1.55. <0.001	*Ref.*	0.96: 0.83, 1.11. 0.63	1.34: 1.00, 1.79. 0.05
≥15 mg/L	2.64: 2.25, 3.10. <0.001	1.41: 1.24, 1.61. <0.001	*Ref.*	1.08: 0.89, 1.32. 0.45	1.50: 1.01, 2.22. 0.04
≥20 mg/L	2.45: 1.98, 3.02. <0.001	1.35: 1.14, 1.60. 0.001	*Ref.*	1.28: 1.00, 1.65. 0.05	1.89: 1.18, 3.02. 0.008

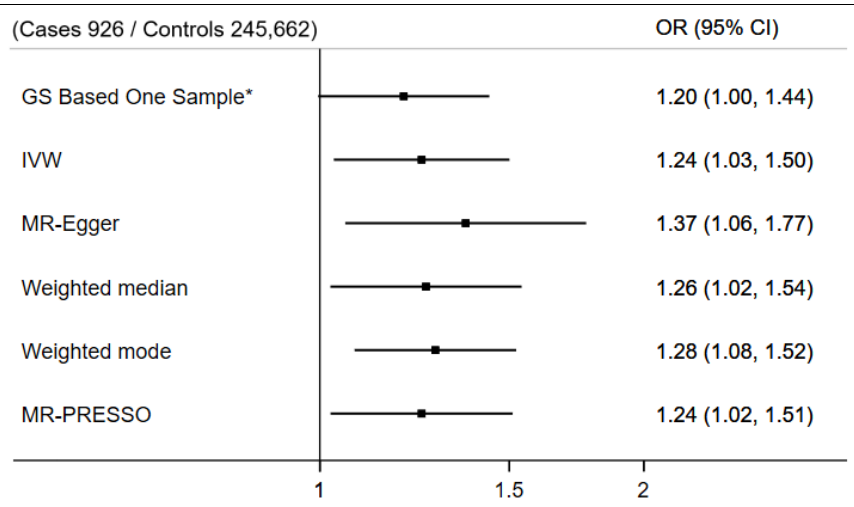
CRP = C-reactive protein. OR = odds ratio. CI = confidence intervals. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.



Study 4 - Figure 1: Phenotypic analyses for the association of measured 25(OH)D and falls in the UK Biobank, with the outcome restricted to falls concurring with C-reactive protein ≥ 20mg/L.

First line = simple ●, second line = adjusted ▲. Reference category: 50-74.99 nmol/L. For P values, see Table 3. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Genetic analysis – Mendelian randomization: We next conducted MR analyses to first assess the association of genetically predicted 25(OH)D and fall outcomes in the general population. In linear MR analyses, there was no evidence for an association between genetically predicted 25(OH)D and falls using the unrestricted outcome sample (fall cases n = 59,303) (per 10 mg/L increase in 25(OH)D, OR = 1.01, 95% CI 0.99-1.04, P = 0.2). However, the same MR analysis but repeated with fall cases reflecting those who had fallen and had CRP above thresholds ≥ 5/10/15/20 mg/L (e.g., 926 cases for CRP ≥ 20 mg/L) provided tentative evidence of a linear association where fall risk increased along with higher 25(OH)D concentrations. This pattern became more pronounced with each threshold increase, with nominal evidence at CRP > 20 mg/L (per 10 mg/L increase in 25(OH)D, OR = 1.20, 95% CI, 1.00-1.44, P = 0.05). Sensitivity analyses using the CRP > 20 mg/L threshold generated similar results when using the 122 SNP version of the genetic score (Table S3), across several pleiotropy robust MR approaches (**Figure 2**), and in the leave-block-out analyses (Table S3).



Study 4 - Figure 2: Mendelian randomization sensitivity analyses of genetically predicted 25(OH)D and falls concurring with C-reactive protein ≥ 20 mg/L in the UK biobank.

*The Genetic Score (GS) Based One Sample approach is presented for comparison purposes. *In the GS Based One Sample approach, the weight of SNPs are based on independent SNP-25(OH)D beta coefficients from the Sunlight Consortium [36]. Genetic analysis adjusted for age, sex, assessment centre, SNP array, principal components 1-40, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. MR PRESSO outlier/s = N/A.

Discussion

In this paper, we explored the relationship between 25(OH)D and fall risk in the context of inflammation, given that inflammation may be a setting where motor function supporting factors, such as retinol [22, 39, 40], could be disrupted by higher vitamin D. We show that the association between vitamin D status measured by 25(OH)D and inflammatory marker CRP is likely to be complex, with more inflammation observed both for participants with the lowest and the highest 25(OH)D concentrations. We also show that the association between 25(OH)D concentrations and the risk of falls appears to be inverse overall, with the highest risks seen for those with the lowest concentrations. However, for participants with the highest 25(OH)D concentrations, risk of falling is higher if they also had evidence for high inflammation, with this finding supported by genetic evidence showing a positive association between higher 25(OH)D and inflammation related fall risk. Overall, our analyses suggests that the association of 25(OH)D and the risk of falls depends on the presence of inflammation, which is compatible with a notion that inflammation suppresses the retinoid system [27]—an important neurological regulator of motor function [25, 26], and a system in competition with vitamin D [22-24].

Previous studies examining vitamin D status and fall risk have generally observed lower 25(OH)D levels (< 50 nmol/L) associating with an increased risk of falls, particularly among older adults [41-45]—and this aligns with vitamin D trials that show a reduction in fall risk arising particularly in those who are deficient [7]. The mechanisms proposed to account for the fall risk and vitamin D insufficiency relationship typically focus on factors including impaired muscle function due to vitamin D receptor-mediated effects on muscle cells [46], reduced bone integrity from secondary hyperparathyroidism leading to increased bone resorption [47], and vitamin D deficiency induced inflammation that may exacerbate neurological an neuromuscular dysfunction [48]. While these proposed mechanisms address fall risk associated with chronic vitamin D insufficiency, our

study extends this consideration to factors that may be relevant in the context of transiently higher vitamin D.

Clinical trials have shown mixed results regarding the protective effects of vitamin D against falls [8, 15, 49], however, findings generally suggest that—where benefits are seen—they arise from smaller daily dosing (< 2000 IU vitamin D3) and do not extend to those with sufficient vitamin D levels [6, 7]. Those studies that have reported increased fall risk (in primary or subgroup analyses) with annual (500,000 IU) [10], monthly (60,000 IU to 100,000 IU) [11-13], or high daily intakes (> 3200 IU) of vitamin D3 [8, 9] suggest the presence of a threshold beyond which additional vitamin D may no longer prevent falls and could even pose a transient increase in fall risk. Inconsistencies notably include the ViDA trial, which used 100,000 IU monthly, and did not report adverse effects [14]. However, the study by Sanders et al. (2010)—which reported an increased risk of falls—is notable, in that no other motor function related study has administered a single oral dose of 500,000 IU vitamin D3 to a largely vitamin D replete population [10]. The study by Latham et al. (2003)—which did not report increased fall risk—is the closest relating trial, having used a single 300,000 IU dose of oral vitamin D3 [50]. However, the prevalence of deficiency was notably higher in the Latham trial [30% under 30 nmol/L 25(OH)D], while in the trial of Sanders et al., only 3% had 25(OH)D under 25 nmol/L [10, 50]. Difference in baseline vitamin D repletion across the studied populations may also be a critical factor that biases meta-analyses of these findings.

One explanation that potentially links fall risk with both lower and higher 25(OH)D levels is inflammation, albeit through distinct mechanisms. Unlike the potential fall risk posed by inadequate VDR stimulation of anti-inflammatory immunomodulatory cells, which can occur at the state of vitamin D deficiency [51], the presence of inflammation in conjunction with higher vitamin D levels may shift the balance of steroid hormone competition [22-24]. This may affect both retinol's

anti-inflammatory role [52], and its neurological-based regulation of motor function [25, 26]. The competitive relationship between vitamin D and retinol is mechanistically explainable at the level of heterodimer synthesis and retinoic acid response elements (RARE) binding. The primary nuclear receptors that RA binds to are the retinoic acid receptor (RAR) and retinoid X receptor (RXR) [53], with vitamin D receptor (VDR)-RXR heterodimers also capable of binding to RARE. In this context, the VDR-RXR is producing a repressive heterodimer species that can suppress RARE mediated gene expression that would otherwise occur from RAR-RXR stimulation [54-56].

Vitamin D's role in inflammation has been widely studied [51], however, the relationship between vitamin D and CRP presents complex and sometimes contradictory findings [57-63]. Several observational studies have reported inverse associations [57-59], supporting the notion that higher vitamin D might reduce systemic inflammation. While a 2018 meta-analysis of trial data showed no overall impact of vitamin D supplementation on CRP, the included studies that did show a lowering of CRP were largely using lower vitamin D3 doses (400 - 7000 IU) or bolus doses given to vitamin D deficient cohorts [63]. Meanwhile, a 2024 meta-analysis found a reduction of CRP levels in post-menopausal women from $\geq$ 3 months of daily vitamin D3 1000 IU [62]. These findings suggest that the vitamin D and CRP relationship may be another complex physiological scenario dependent on repletion status and dosing methods, and where more vitamin D—beyond addressing deficiency—may not necessarily be better.

The null MR findings—derived from analyzing the full sample of falls—suggests that the increased risk of falls associated with higher 25(OH)D is not present in normal physiological conditions, where natural fluctuations of 25(OH)D occur within typical serum level ranges and are likely in balance with other dietary derived steroid hormones. However, given that we saw a genetic association between higher 25(OH)D and increased fall risk in those who fell and had CRP $\geq$ 20 mg/L, this could suggest the potential for clinically relevant/discernible outcomes arising from

disruption to steroid hormone homeostasis—whether that be derived from dietary retinol deficiency, retinol impairment, bolus dosing of vitamin D, or other inflammatory related effects on steroid hormone function. While a mechanistic relationship between retinol and vitamin D was considered as a theoretical basis linking distinct observations related to vitamin D, inflammation and falls, further research will need to elucidate if these proposed mechanisms underlie our findings, and if the phenomenon might occur in response to other forms of interaction between inflammation, vitamin D, and motor function regulation. Further studies using animal models (e.g., murine or drosophila melanogaster [64]) could explore the mechanistic insights raised here, and more effectively explore the potential interactions between inflammation, vitamin D and retinol [27].

Bolus dosing of vitamin D has been safety assessed by multiple government regulators around the world prior to implementation in medical settings [65, 66]. However, increased risk of falling is likely related to factors beyond what safety trials have typically been primed to assess. While the retinol competition could potentially explain the higher risk of falls observed in these trials, there are other possible mechanisms. For example, high dose oral vitamin D is sometimes used to induce vascular calcification in animal models [67]. There are also multiple recorded cases in humans, where a very high to acutely toxic intakes vitamin D supplementation, given by error, have resulted in irreversible tissue calcification [68, 69]. However, there is no reported indication, that the increased falling (associated with bolus dosing of vitamin D) had concurred with calcification tissue injury, which is usually notable, chronic, and irreversible. Other plausible mechanisms could be those which influence vestibular related tissues/function, although to our knowledge no related associations have been reported in the bolus dosing clinical trials [9-13]. While we do not know if the affected clinical trial participants (who reacted poorly to large dose

supplementation [9-13]) had inflammation, our results could suggest that inflammation plays a broader role in high vitamin D related fall risk scenarios.

Strengths and limitations

This study has made pragmatic use of available resources to safely and ethically assess metabolic steps that may underlie a phenomenon that has become challenging to navigate via conventional clinical trial pathways. Our analyses used large scale data obtained from the UK Biobank, with comprehensive analyses to address assess associations both at high and low concentrations. Importantly, results from the MR analyses on 25(OH)D associations and inflammation related risk of falls was consistent across different approaches and supported the findings from observational analyses. While we cannot emulate the same clinically relevant estimations regarding fall risk that have emerged from clinical trials, our study, within the limits of data availability, provides support for potential underlying mechanisms that are worthy of consideration and further exploration. However, our study also has limitations and while the UK Biobank provided the size necessary to detect the phenomenon in a non-clinical trial environment, it lacked measures for some key biomarkers, such as specific inflammatory factors and retinol.

Further limitations with our study relate to restricting our analyses to individuals of white-European descent to minimize potential bias from population stratification, therefore generalizability to other ethnic groups, as a result, may be limited. While the UK Biobank's response rate of 5% may not be representative of the general UK population [70], previous studies using this cohort have replicated expected exposure-disease associations [70]. The use of MR approach is less affected by confounding and reverse causality than other types of observational studies, which strengthens the validity of our findings [71]. However, it's important to note that the

MR approach assumes no horizontal pleiotropy, [72] which may result in biased estimates. To address this issue, we engaged several pleiotropy-robust approaches and found no evidence for an association between the genetic score and potential confounders—with additionally, the genetic instrument being restricted to SNPs with independently replicated evidence for an association with 25(OH)D. [73] MR analyses approximate average effects over the life-course, and the true associations may be more complex than presented here.

Conclusion

Our study suggests that an inflammatory state might increase fall risk mediated by higher vitamin D. Both low and high 25(OH)D levels increased fall risk in individuals with chronic inflammation, with supporting evidence for the association between higher 25(OH)D and increased fall risk in those with chronic inflammation provided via MR analysis. Future research should explore whether the increased fall risk, seen following high-dose vitamin D supplementation in clinical trials, relates to excessive production of VDR-RXR heterodimers that leads to functional retinol deficiency potentially affecting anti-inflammatory and motor function pathways.

Acknowledgements

We thank all UK Biobank participants involved, Dr Anwar Mulugeta (University of South Australia) for support in data management.

Author Contributions

JS developed the research question with EH, wrote the first draft and analyzed the data with guidance from EH and AZ. All authors revised the paper, interpreted results, and approved the final manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Data Availability

This research has been conducted using the UK Biobank resource under application number 20175.

All data and code will be available to approved users upon application to the UK Biobank.

Supplementary Data

Supplementary data to this article can be found online at:

References:

1. Florence CS, Bergen G, Atherly A, Burns E, Stevens J, Drake C. Medical Costs of Fatal and Nonfatal Falls in Older Adults. J Am Geriatr Soc. 2018;66(4):693-8. https://doi.org/10.1111/jgs.15304

2. Mackenzie L, McIntyre A. How Do General Practitioners (GPs) Engage in Falls Prevention With Older People? A Pilot Survey of GPs in NHS England Suggests a Gap in Routine Practice to Address Falls Prevention. Front Public Health. 2019;7:32. https://doi.org/10.3389/fpubh.2019.00032

3. Stevens JA, Lee R. The Potential to Reduce Falls and Avert Costs by Clinically Managing Fall Risk. Am J Prev Med. 2018;55(3):290-7. https://doi.org/10.1016/j.amepre.2018.04.035

4. Appeadu MK, Bordoni B. Falls and fall prevention in the elderly. StatPearls [Internet]: StatPearls Publishing; 2023. https://www.ncbi.nlm.nih.gov/books/NBK560761/

5. Snijder MB, van Schoor NM, Pluijm SMF, van Dam RM, Visser M, Lips P. Vitamin D Status in Relation to One-Year Risk of Recurrent Falling in Older Men and Women. The Journal of Clinical Endocrinology & Metabolism. 2006;91(8):2980-5. https://doi.org/10.1210/jc.2006-0510

6. Wei FL, Li T, Gao QY, Huang Y, Zhou CP, Wang W, et al. Association Between Vitamin D Supplementation and Fall Prevention. Front Endocrinol (Lausanne). 2022;13:919839. https://doi.org/10.3389/fendo.2022.919839

7. Ling Y, Xu F, Xia X, Dai D, Xiong A, Sun R, et al. Vitamin D supplementation reduces the risk of fall in the vitamin D deficient elderly: An updated meta-analysis. Clinical Nutrition. 2021;40(11):5531-7. https://doi.org/10.1016/j.clnu.2021.09.031

8. Zittermann A, Trummer C, Theiler-Schwetz V, Pilz S. Long-term supplementation with 3200 to 4000 IU of vitamin D daily and adverse events: a systematic review and meta-analysis of randomized controlled trials. Eur J Nutr. 2023;62(4):1833-44. https://doi.org/10.1007/s00394-023-03124-w

9. Smith LM, Gallagher JC, Suiter C. Medium doses of daily vitamin D decrease falls and higher doses of daily vitamin D3 increase falls: A randomized clinical trial. The Journal of Steroid Biochemistry and Molecular Biology. 2017;173:317-22. https://doi.org/10.1016/j.jsbmb.2017.03.015

10. Sanders KM, Stuart AL, Williamson EJ, Simpson JA, Kotowicz MA, Young D, et al. Annual high-dose oral vitamin D and falls and fractures in older women: a randomized controlled trial. Jama. 2010;303(18):1815-22. https://doi.org/10.1001/jama.2010.594

11. Bischoff-Ferrari HA, Dawson-Hughes B, Orav EJ, Staehelin HB, Meyer OW, Theiler R, et al. Monthly high-dose vitamin D treatment for the prevention of functional decline: a randomized clinical trial. JAMA internal medicine. 2016;176(2):175-83. https://doi.org/10.1001/jamainternmed.2015.7148

12. Ginde AA, Blatchford P, Breese K, Zarrabi L, Linnebur SA, Wallace JI, et al. High-Dose Monthly Vitamin D for Prevention of Acute Respiratory Infection in Older Long-Term Care Residents: A Randomized Clinical Trial. Journal of the American Geriatrics Society. 2017;65(3):496-503. https://doi.org/10.1111/jgs.14679

13. Waterhouse M, Sanguineti E, Baxter C, Duarte Romero B, McLeod DSA, English DR, et al. Vitamin D supplementation and risk of falling: outcomes from the randomized, placebo-controlled D-Health Trial. Journal of Cachexia, Sarcopenia and Muscle. 2021;12(6):1428-39. https://doi.org/10.1002/jcsm.12759

14. Khaw KT, Stewart AW, Waayer D, Lawes CMM, Toop L, Camargo CA, Jr., et al. Effect of monthly high-dose vitamin D supplementation on falls and non-vertebral fractures: secondary and post-hoc outcomes from the randomised, double-blind, placebo-controlled ViDA trial. Lancet Diabetes Endocrinol. 2017;5(6):438-47. https://doi.org/10.1016/s2213-8587(17)30103-1

15. Bolland MJ, Grey A, Avenell A. Effects of vitamin D supplementation on musculoskeletal health: a systematic review, meta-analysis, and trial sequential analysis. Lancet Diabetes Endocrinol. 2018;6(11):847-58. https://doi.org/10.1016/s2213-8587(18)30265-1

16. Nehring S, Goyal A, Patel B. C Reactive Protein.[Updated 2023 Jul 10]. StatPearls [Internet] Treasure Island (FL): StatPearls Publishing. 2024. https://www.ncbi.nlm.nih.gov/books/NBK441843/

17. Karadağ A, Bingöl M. Loss of balance and fall risk, and their relationship with disease activity, functional status, spinal mobility, quality of life and enthesitis score in patients with ankylosing spondylitis: A cross-sectional study. Cumhuriyet Medical Journal. 2020;42(4):441-8. https://doi.org/10.7197/cmj.819210

18. Ergül EE, Alkan Melikoğlu M. A quantitative assessment of the risk of falls in rheumatoid arthritis patients and determination of the risk factors. Turk J Phys Med Rehabil. 2022;68(2):271-7. https://doi.org/10.5606/tftrd.2022.7868

19. Mojtaba M, Alinaghizadeh H, Rydwik E. Downton Fall Risk Index during hospitalisation is associated with fall-related injuries after discharge: a longitudinal observational study. Journal of Physiotherapy. 2018;64(3):172-7. https://doi.org/10.1016/j.jphys.2018.05.005

20. Britton GB, O'Bryant SE, Johnson LA, Hall JR, Villarreal AE, Oviedo DC, et al. Inflammatory Biomarkers, Depressive Symptoms and Falls Among the elderly in Panama. Current Aging Science. 2019;11(4):236-41. https://doi.org/10.2174/1874609812666190215125104

21. de Amorim JSC, Torres KCL, Carvalho AT, Martins-Filho OA, Lima-Costa MF, Peixoto SV. Inflammatory markers associated with fall recurrence and severity: The Bambuí Cohort Study of Aging. Exp Gerontol. 2020;132:110837. https://doi.org/10.1016/j.exger.2020.110837

22. Metz AL, Walser MM, Olson WG. The interaction of dietary vitamin A and vitamin D related to skeletal development in the turkey poult. The Journal of nutrition. 1985;115(7):929-35.

23. Rohde CM, Manatt M, Clagett-Dame M, DeLuca HF. Vitamin A antagonizes the action of vitamin D in rats. The Journal of nutrition. 1999;129(12):2246-50. https://doi.org/10.1093/jn/115.7.929

24. Conaway HH, Henning P, Lerner UH. Vitamin A Metabolism, Action, and Role in Skeletal Homeostasis. Endocrine Reviews. 2013;34(6):766-97. https://doi.org/10.1210/er.2012-1071

25. Carta M, Stancampiano R, Tronci E, Collu M, Usiello A, Morelli M, et al. Vitamin A deficiency induces motor impairments and striatal cholinergic dysfunction in rats. Neuroscience. 2006;139(4):1163-72. https://doi.org/10.1016/j.neuroscience.2006.01.027

26. Marie A, Darricau M, Touyarot K, Parr-Brownlie LC, Bosch-Bouju C. Role and mechanism of vitamin a metabolism in the pathophysiology of Parkinson's Disease. Journal of Parkinson's disease. 2021;11(3):949-70. https://pubmed.ncbi.nlm.nih.gov/34120916/

27. Thurnham DI, McCabe GP, Northrop-Clewes CA, Nestel P. Effects of subclinical infection on plasma retinol concentrations and assessment of prevalence of vitamin A deficiency: meta-analysis. The Lancet. 2003;362(9401):2052-8. https://doi.org/10.1016/s0140-6736(03)15099-4

28. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine. 2015;12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779

29. Fry D, Almond R, Moffat S, Gordon M, Singh P. UK Biobank Biomarker Project - Companion Document to Accompany Serum Biomarker Data - https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf. 2019. p. 12.

30. Eckschlager C, Schwenoha K, Roth C, Bogner B, Oostingh GJ. Comparative analysis of high CRP-levels in human blood using point-of-care and laboratory-based methods. Practical Laboratory Medicine. 2019;17:e00137. https://doi.org/10.1016/j.plabm.2019.e00137

31. UK Biobank. Genotyping of 500,000 UK Biobank participants Description of sample processing workflow and preparation of DNA for genotyping. UK Biobank Showcase; 2015. https://biobank.ndph.ox.ac.uk/crystal/ukb/docs/ukb_dna_processing.pdf

32. UK Biobank. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource. 2015. https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/genotyping_qc.pdf

33. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. J Nature. 2018;562(7726):203. https://doi.org/10.1038/s41586-018-0579-z

34. Revez JA, Lin T, Qiao Z, Xue A, Holtz Y, Zhu Z, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nature Communications. 2020;11(1). https://doi.org/10.1038/s41467-020-15421-7

35. Zhou A, Selvanayagam JB, Hyppönen E. Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. European Heart Journal. 2022;43(18):1731-9. https://doi.org/10.1093/eurheartj/ehab809

36. Jiang X, O'Reilly PF, Aschard H, Hsu Y-H, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nature Communications. 2018;9(1). https://doi.org/10.1038/s41467-017-02662-2

37. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research. 2020;4:186. https://doi.org/10.12688/wellcomeopenres.15555.3

38. Burgess S, Gill D. Genetic evidence for vitamin D and cardiovascular disease: choice of variants is critical. European Heart Journal. 2022;43(18):1740-2. https://doi.org/10.1093/eurheartj/ehab870

39. Rohde CM, DeLuca HF. All-trans retinoic acid antagonizes the action of calciferol and its active metabolite, 1, 25-dihydroxycholecalciferol, in rats. The Journal of nutrition. 2005;135(7):1647-52. https://pubmed.ncbi.nlm.nih.gov/15987844/

40. Mata-Granados JM, Cuenca-Acevedo R, Luque De Castro MD, Sosa M, Quesada-Gómez JM. Vitamin D deficiency and high serum levels of vitamin A increase the risk of osteoporosis evaluated by Quantitative Ultrasound Measurements (QUS) in postmenopausal Spanish women. Clinical Biochemistry. 2010;43(13-14):1064-8. https://doi.org/10.1016/j.clinbiochem.2010.06.001

41. Arima K, Mizukami S, Nishimura T, Tomita Y, Nakashima H, Abe Y, et al. Epidemiology of the association between serum 25-hydroxyvitamin D levels and musculoskeletal conditions among elderly individuals: a literature review. J Physiol Anthropol. 2020;39(1):38. https://doi.org/10.1186/s40101-020-00249-3

42. Suzuki T, Kwon J, Kim H, Shimada H, Yoshida Y, Iwasa H, et al. Low serum 25-hydroxyvitamin D levels associated with falls among Japanese community-dwelling elderly. J Bone Miner Res. 2008;23(8):1309-17. https://doi.org/10.1359/jbmr.080328

43. Rothenbacher D, Klenk J, Denkinger MD, Herbolsheimer F, Nikolaus T, Peter R, et al. Prospective evaluation of renal function, serum vitamin D level, and risk of fall and fracture in community-dwelling elderly subjects. Osteoporosis International. 2014;25(3):923-32. https://doi.org/10.1007/s00198-013-2565-5

44. Shimizu Y, Kim H, Yoshida H, Shimada H, Suzuki T. Serum 25-hydroxyvitamin D level and risk of falls in Japanese community-dwelling elderly women: a 1-year follow-up study. Osteoporosis International. 2015;26(8):2185-92. https://doi.org/10.1007/s00198-015-3130-1

45. Machado KLLL, Domiciano DS, Machado LG, Lopes JB, Figueiredo CP, Takayama L, et al. Persistent hypovitaminosis D and loss of hip bone mineral density over time as additional risk factors for recurrent falls in a population-based prospective cohort of elderly persons living in the

community. The São Paulo Ageing & Health (SPAH) Study. Osteoporosis International. 2015;26(5):1535-42. https://doi.org/10.1007/s00198-014-3024-7

46. Bollen SE, Bass JJ, Fujita S, Wilkinson D, Hewison M, Atherton PJ. The Vitamin D/Vitamin D receptor (VDR) axis in muscle atrophy and sarcopenia. Cellular Signalling. 2022;96:110355. https://doi.org/10.1016/j.cellsig.2022.110355

47. Lips P. Vitamin D Deficiency and Secondary Hyperparathyroidism in the Elderly: Consequences for Bone Loss and Fractures and Therapeutic Implications. Endocrine Reviews. 2001;22(4):477-501. https://doi.org/10.1210/edrv.22.4.0437

48. Ao T, Kikuta J, Ishii M. The Effects of Vitamin D on Immune System and Inflammatory Diseases. Biomolecules. 2021;11(11). https://doi.org/10.3390/biom11111624

49. Zheng YT, Cui QQ, Hong YM, Yao WG. A meta-analysis of high dose, intermittent vitamin D supplementation among older adults. PLoS One. 2015;10(1):e0115850. https://doi.org/10.1371/journal.pone.0115850

50. Latham NK, Anderson CS, Lee A, Bennett DA, Moseley A, Cameron ID. A randomized, controlled trial of quadriceps resistance exercise and vitamin D in frail older people: the Frailty Interventions Trial in Elderly Subjects (FITNESS). J Am Geriatr Soc. 2003;51(3):291-9. https://doi.org/10.1046/j.1532-5415.2003.51101.x

51. Yin K, Agrawal DK. Vitamin D and inflammatory diseases. J Inflamm Res. 2014;7:69-87. https://doi.org/10.2147/jir.s63898

52. Oliveira LM, Teixeira FME, Sato MN. Impact of Retinoic Acid on Immune Cells and Inflammatory Diseases. Mediators Inflamm. 2018;2018:3067126. https://doi.org/10.1155/2018/3067126

53. Petkovich M, Chambon P. Retinoic acid receptors at 35 years. Journal of Molecular Endocrinology. 2022;69(4):T13-T24. https://doi.org/10.1530/jme-22-0097

54. Jiménez-Lara AM, Aranda A. Vitamin D represses retinoic acid-dependent transactivation of the retinoic acid receptor-β2 promoter: the AF-2 domain of the vitamin D receptor is required for transrepression. Endocrinology. 1999;140(6):2898-907. https://doi.org/10.1210/endo.140.6.6770

55. Carlberg C, Saurat J-H, editors. Vitamin D-retinoid association: molecular basis and clinical applications. The Journal of Investigative dermatology Symposium Proceedings; 1996. https://pubmed.ncbi.nlm.nih.gov/9627699/

56. Rehó B, Fadel L, Brazda P, Benziane A, Hegedüs É, Sen P, et al. Agonist-controlled competition of RAR and VDR nuclear receptors for heterodimerization with RXR is manifested in their DNA binding. Journal of Biological Chemistry. 2023;299(2):102896. https://doi.org/10.1016/j.jbc.2023.102896

57. Tao RX, Zhou QF, Xu ZW, Hao JH, Huang K, Mou Z, et al. Inverse Correlation between Vitamin D and C-Reactive Protein in Newborns. Nutrients. 2015;7(11):9218-28. https://doi.org/10.3390/nu7115468

58. Mokhtari F, Ganjei Z, Yazdanpanah M, Hosseini M. Inverse correlation between vitamin D and CRP levels in alopecia areata: A pilot study. J Cosmet Dermatol. 2023;22(11):3176-80. https://doi.org/10.1111/jocd.15994

59. Kruit A, Zanen P. The association between vitamin D and C-reactive protein levels in patients with inflammatory and non-inflammatory diseases. Clinical Biochemistry.

2016;49(7):534-7. https://doi.org/10.1016/j.clinbiochem.2016.01.002

60. Mellenthin L, Wallaschofski H, Grotevendt A, Völzke H, Nauck M, Hannemann A. Association between serum vitamin D concentrations and inflammatory markers in the general adult population. Metabolism. 2014;63(8):1056-62. https://doi.org/10.1016/j.metabol.2014.05.002

61. Amer M, Qayyum R. Relation Between Serum 25-Hydroxyvitamin D and C-Reactive Protein in Asymptomatic Adults (From the Continuous National Health and Nutrition Examination Survey 2001 to 2006). The American Journal of Cardiology. 2012;109(2):226-30. https://doi.org/10.1016/j.amjcard.2011.08.032

62. Jiang Q, Prabahar K, Saleh SAK, Adly HM, Velu P, Adi AR, et al. The Effects of Vitamin D Supplementation on C-Reactive Protein and Systolic and Diastolic Blood Pressure in Postmenopausal Women: A Meta-Analysis and Systematic Review of Randomized Controlled Trials. Journal of the Academy of Nutrition and Dietetics. 2024;124(3):387-96.e5. https://doi.org/10.1016/j.jand.2023.10.013

63. Mazidi M, Rezaie P, Vatanparast H. Impact of vitamin D supplementation on C-reactive protein; a systematic review and meta-analysis of randomized controlled trials. BMC Nutr. 2018;4:1. https://doi.org/10.1186/s40795-017-0207-6

64. Rebonza E. Analysis of Vitamin D and B12 Function in Drosophila melanogaster. Barrett The Honors College Thesis / Creative Project Collection Arizona State University; 2019. https://hdl.handle.net/2286/R.I.56564

65. Administration DoHaAC-TG. 3.1 Vitamin D Scheduling medicines and poisons 2022 [Available from: https://www.tga.gov.au/resources/publication/scheduling-decisions-interim/scheduling-delegates-interim-decisions-and-invitation-further-comment-accsacms-november-2016/31-vitamin-d.

66. Agency MaHpR. Public Assessment Report UK PAR, SoliCol D3 50,000 IU Tablets, (colecalciferol) 2016 [Available from: https://mhraproducts4853.blob.core.windows.net/docs/c4d72968ccb86a6699e6789718414ce61f43f3a0.

67. Razzaque MS. The dualistic role of vitamin D in vascular calcifications. Kidney Int. 2011;79(7):708-14. https://doi.org/10.1038/ki.2010.432

68. Wang J, Zhou J, Robertson G, Lee V. Vitamin D in Vascular Calcification: A Double-Edged Sword? Nutrients. 2018;10(5):652. https://doi.org/10.3390/nu10050652

69. Nasri H, Mubarak M. Renal injury due to vitamin D intoxication; a case of dispensing error. Journal of Renal Injury Prevention. 2013;2(2):85. https://doi.org/10.12861/jrip.2013.27

70. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. BMJ. 2020:m131. https://doi.org/10.1136/bmj.m131

71. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? International Journal of Epidemiology. 2019;48(3):691-701. https://doi.org/10.1093/ije/dyy202

72. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018:k601. https://doi.org/10.1136/bmj.k601

73. Hyppönen E, Vimaleswaran KS, Zhou A. Genetic Determinants of 25-Hydroxyvitamin D Concentrations and Their Relevance to Public Health. Nutrients. 2022;14(20):4408. https://doi.org/10.3390/nu14204408

Study 4 Supplementary

Supplementary content for: Vitamin D, C-reactive protein, and increased fall risk: A genetic epidemiological study:

Supplementary Methods: Measurement of vitamin D status	**03**
Supplementary Methods: Covariates	**03**
Supplementary Methods: Genetic instruments using 35 variants	**04**
Supplementary Methods: Genetic instrument using 122 vitamin D GWAS variants	**05**
Supplementary Methods: Linear Mendelian randomization	**05**
Supplementary Methods: Leave-block-out analysis	**08**
Figure S1. Study population and participant inclusion criteria.	**10**
Figure S2. Latitudinal zoning of the UK Biobank Assessment Centers.	**11**
Figure S3. Selection of variants for the genetic instrument for measured 25(OH)D concentrations.	**12**
Figure S4. Mendelian randomization assumption diagram.	**13**
Figure S5. Distribution of the vitamin D genetic score (mapped on an X-axis by increasing number of alleles) in the UK Biobank (A), and the association of the vitamin D genetic score (in six groups – lowest to highest number of alleles) with measured 25(OH)D concentrations in the UK Biobank (B).	**14**
Table S1. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.	**15**
Table S2. Association of the vitamin D genetic score with potential confounders in the UK Biobank.	**16**
Table S3. Mendelian randomization analysis for the association of genetically predicted 25(OH)D with falls concurring with C-reactive protein > 20 mg/L in the UK Biobank - 122 SNP genetic score and leave block out analysis.	**17**
Table S4. Functional blocks used in the leave-block-out analyses.	**18**
Table S5. Phenotypic analyses showing the prevalence of falls across CRP ranges, and the association between falls and categorical CRP, in the UK Biobank.	**19**
Table S6. Fully adjusted association between continuous 25(OH)D and falls, stratified above and below CRP thresholds.	**19**
Table S7. Fully adjusted association between continuous 25(OH)D and falls, with the outcome of falls concurring with inflammation across increasing CRP thresholds.	**20**
Table S8. Phenotypic sensitivity analyses for the association of categorical 25(OH)D and falls in the UK Biobank, with the outcome of falls concurring to those with rheumatoid arthritis (adjusted and simple models).	**21**
Table S9. Phenotypic analyses for the association of categorical measured 25(OH)D and falls in the UK Biobank, with falls concurring with CRP <5mg/L.	**21**
References	**22**

Authors:

Joshua P Sutherland (BHSc Nut Med, Hons)[1], Ang Zhou (PhD)[1,2,3], Elina Hyppönen (PhD)[1,2]

4. Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, Australia.
5. South Australian Health and Medical Research Institute, Adelaide, Australia
6. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

Joshua P Sutherland: ORCID 0000-0003-4102-3164

Ang Zhou: ORCID 0000-0001-7088-0777

Elina Hyppönen: ORCID 0000-0003-3670-9399

Corresponding author: Prof Elina Hyppönen; Australian Centre for Precision Health, University of South Australia, c/o South Australian Health and Medical Research Institute, GPO Box 2471, Adelaide SA 5001; +61 8 830 22518; Elina.Hypponen@unisa.edu.au

Supplementary Methods:

Measurement of vitamin D status

The LIAISON XL 25(OH)D assay (DiaSorin, Stillwater, USA) [1] was used to measure baseline serum 25(OH)D concentrations in blood samples. This fully automated, chemiluminescence immunoassay device utilizes magnetic microparticle separation technology.

- The blood samples used had a mean fasting time of 3.8 hours (SD 2.4).
- Assay precision was evaluated using the CLSI EP5-A2 protocol, with 6 samples and 2 controls. The intra-run coefficient of variation (CV) was 2.3%, and the total CV was 7.8%.
- The functional sensitivity of the LIAISON XL assay is ≤ 4.0 ng/ml, in accordance with Clinical Laboratory and Standards Institute (CLSI) EP17-A protocols. [1]

Covariates

The covariates used in adjustments for this study were collected at baseline, with many derived from self-reported touchscreen questionnaires (age, sex, physical activity, smoking, alcohol), residential address data (Townsend deprivation index), [2] and physical assessments (height and weight). The assessment centre at which baseline assessment occurred was itself used as a covariate (with the centers forming categories, as outlined in **Figure S2**). Height data was obtained using a Seca 202 stadiometer, and weight was acquired using the body composition analyser Tanita BC-418MA, or via body weight scales in instances where body composition analysis was not used. BMI was calculated as body weight [in kilograms] divided by height

squared [in meters]) [3] and divided into quartiles (with the two middle quartiles combined into a single 'mid' BMI category)

For Townsend deprivation index, each participant was assigned a score corresponding to socioeconomic data attributed to their postcode; this was derived from the preceding UK census of population and housing. [2, 4] Longitudinal location was determined based on latitudinal zoning of the assessment center (Figure S2). The genetic covariates, and 25(OH)D concentrations were analyzed from samples collected at baseline.

Genetic instruments using 35 variants

Our genetic score for measured 25(OH)D serum concentration was constructed using a weighted approach that involved collating common autosomal SNPs (minor allele frequency $\geq$ 5%) from a GWAS analysis conducted on the UK Biobank dataset, wherein 134 indpendent loci were identified. [5] To ensure the robustness of our findings, we replicated SNP analyses in the previous SUNLIGHT consortium GWAS, [6] and this left us with 35 SNPs that showed consistent direction and a p value of < 0.05 (**Figure S3**).

- So as to avoid bias due to sample overlap, we utilized the SUNLIGHT consortium meta-analyses derived effect estimates for the SNP association with measured 25(OH)D, as the individual SNP weights. [7]
- This enabled us to ensure that the observed GWAS signals were robust.
- We computed the independently derived weighted average of the number of measured 25(OH)D-increasing alleles for an individual, and then multiplied this by the number of available variants.

Genetic instrument using 122 vitamin D GWAS variants

For sensitivity analyses, we developed an alternative genetic instrument by including a broader set of variants that are associated with measured 25(OH)D concentrations. This new instrument comprises of 122 autosomal single nucleotide polymorphisms (SNPs) that were identified from the UK Biobank dataset [5] (Figure S3).

- To avoid any potential bias arising from using internal weights, we used a 10-fold cross-validation approach [7] to construct the score.
- This process involved randomly dividing the samples into 10 equal sub-samples and using the weights from the remaining 9 sub-samples to calculate the scores.
- Information regarding the 122 SNPs used in this instrument is available in **Table S1**.

Mendelian randomization:

In our study we employed a combination of a Genetic Score One Sample method and SNP-based two-sample Mendelian randomization (MR) approaches.

For the Genetic Score One Sample method, we utilized the ratio of coefficients method in computing the MR estimate, [8] where we calculated the vitamin D genetic score – measured 25(OH)D and the vitamin D genetic score – outcome association estimates and then used these to compute the ratio estimator.

Vitamin D genetic score – measured 25(OH)D and the vitamin D genetic score – outcome association estimates were computed by regressing 25(OH)D and outcome on the vitamin D genetic score, respectively. In these regression models, we adjusted for various factors that could affect measured 25(OH)D, including age, sex, assessment centre, SNP array, top 40 genetic principal components, and nuisance factors like the month in which the blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

To address potential horizontal pleiotropy, we also implemented five SNP-based two-sample MR methods, including the inverse-variance weighted method, Mendelian randomization Egger method, weighted median Mendelian randomization method, weighted mode method, and Mendelian randomization pleiotropy residual sum and outlier (MR Presso) test. These methods are complementary to each other and are robust to different patterns of horizontal pleiotropy. Consistency amongst the methods is indicative of robust evidence of causal association. It is important to note that in our analysis, both SNP-25(OH)D and the SNP-falls association were derived from the UK Biobank. To avoid potential bias due to overlapping samples [9], we computed SNP-Fall association estimates in the full sample but took SNP-25(OH)D association estimates from controls only [9].

In contrast to the above outlined Genetic Score One Sample method analysis the following is an overview of the two sample MR approaches:

- SNP-level information was used, and SNP-exposure and SNP-outcome association estimates were taken as inputs.

- The two-sample MR analyses used SNP-25(OH)D and SNP-outcome association estimates from UK Biobank.
- SNP-Fall association estimates were computed in the full sample, but SNP-25(OH)D association estimates were taken from controls only.
- All five two-sample MR methods were conducted using the 35 vitamin D variants used to construct the vitamin D genetic score (Table S1).

Here, we provide a brief overview of each of the employed two-sample MR approaches:

1. Inverse variance-weighted (IVW) method:
    - Assumes that there is no pleiotropy of variants or that the net pleiotropy is zero.
    - The estimator is obtained by the weighted regression of the effect estimates of variant-outcome association on the effect estimates of variant-exposure association, with the weight being the inverse variance of the variant's ratio estimate. [10]
2. Mendelian randomization Egger method:
    - Similar to the IVW method, the Mendelian randomization Egger method assumes that the regression model includes an intercept term, which represents the average pleiotropic effect. [11]
    - This method can account, however, for directional pleiotropy, which introduces additional untestable assumptions stating that the pleiotropic effect must be independent of the variant-exposure association

3. Mendelian randomization Presso method:
    - Mendelian randomization Presso [12] applies an iterative version of the IVW method, where each variant is omitted from the analysis to assess its contribution to heterogeneity.
    - The residual sum of squares (RSS) is calculated at each iteration and used as a measure of heterogeneity.
    - Variants that produce substantially large RSS values are considered outliers and removed from the analysis.
    - The causal estimate in MR Presso is obtained by applying the IVW method with all outlying variants removed.
4. Weighted mode method:
    - Variants are clustered into groups by the similarity of their ratio estimates. [13]
    - Variants are weighted by the inverse variance of their ratio estimates.
    - Clustering is preformed using a normal kernel density function, with its spread depending on a band with parameter.
    - The causal effect is estimated based on the cluster with the largest weight.
    - An unbiased causal effect will be returned if the variants within the cluster with the largest weight are valid instruments.
5. Weighted median-based method:

- Variants are ranked by their ratio estimates, with the causal estimate being the median of the distribution. [14]

- To ensure a consistent estimate from the weighted median method, it is necessary that valid instruments contribute to at least 50% of the total weight.
- Median is taken from a distribution of the ration estimates in which variants with more precise ratio estimates receive more weight.

Leave-block-out analysis

In our leave-block-out analysis, we first grouped variants into functional blocks based on their associated traits, as identified by the PhenoScanner V2 database. [15] To explore sensitivity of the analysis, we then systematically removed one block of variants at a time and repeated the non-linear Mendelian randomization analysis using the vitamin D genetic score constructed from the remaining, non-removed variants. We repeated this process for all functional blocks.

PhenoScanner V2 is a publicly available database that curates results from large-scale genetic association studies. [15] On 09/09/2021, we searched for evidence of variant-disease/trait associations for all 35 variants in this database. Using a GWAS threshold of P < 5E-08, we found that 20 variants were associated with traits other than serum 25(OH)D (**Table S3**). Based on the associated traits, we grouped the variants into four functional blocks: blood, lipids/metabolic, renal traits, and 'unclassified' (with the latter group being for variants whose associated traits did not fall into one of these 3 blocks) (**Table S4**).

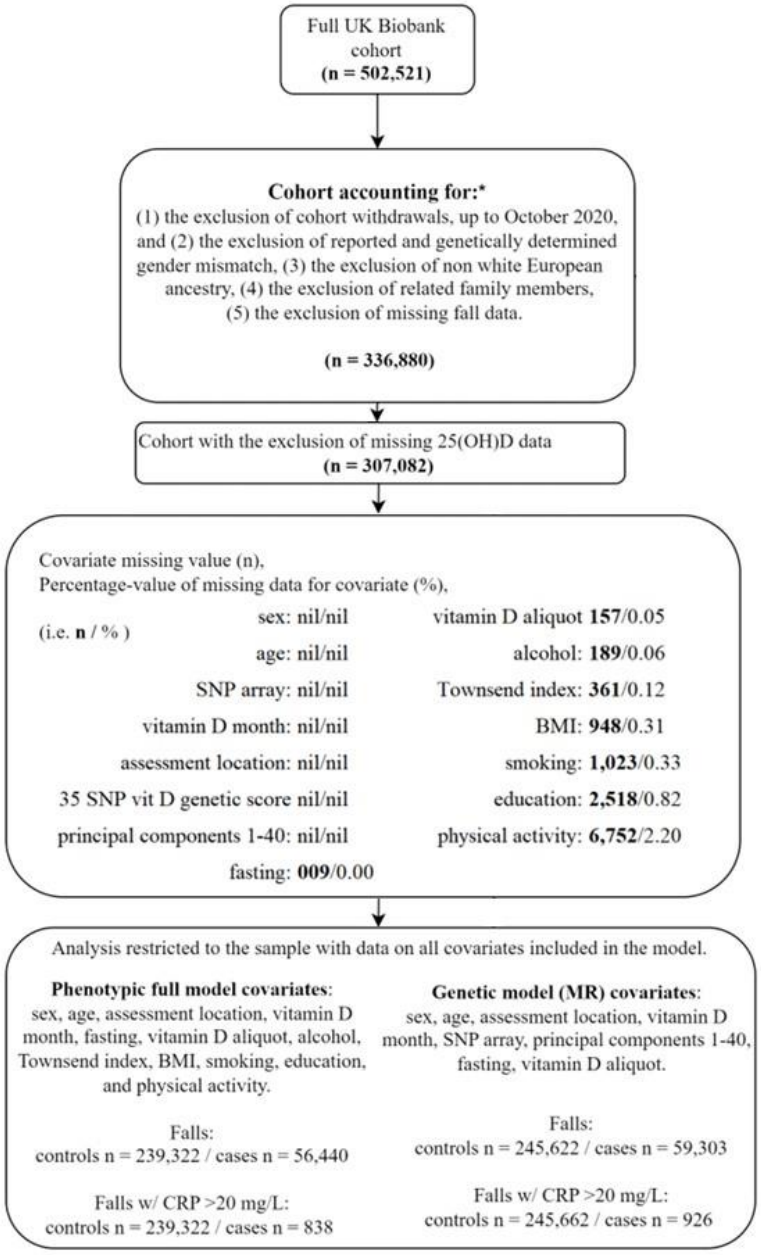


Figure S1. Study population and participant inclusion criteria.

*Exclusion was done in a sequential order. MR = Mendelian randomization.



Figure S2. Latitudinal zoning of the UK Biobank Assessment Centers.

Categorical latitudinal zoning, in ascending order of light to dark: ≤ 51° / 52°-53° / 54°-≥ 55°.
≤ 51° (Croydon, Barts, Hounslow, Reading, Oxford, Bristol, Cardiff, Swansea), 52°-53° (Birmingham, Stoke, Nottingham, Wrexham, Liverpool, Bury, Manchester, Sheffield, Leeds, Stockport), and 54°-≥ 55° (Middlesborough, Newcastle, Glasgow, Edinburgh).

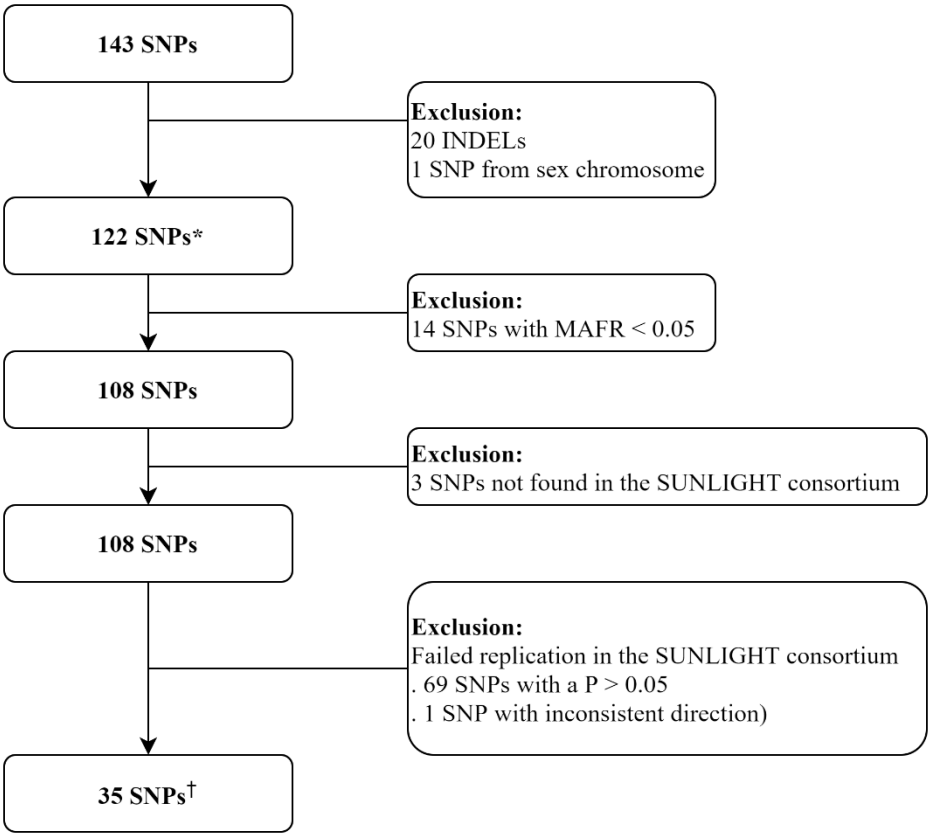


Figure S3. Selection of variants for the genetic instrument for measured 25(OH)D concentrations.

SNP: single nucleotide polymorphism. *SNPs used for the vitamin D genetic score in the primary analysis; †SNPs used for the 122 SNP version of the vitamin D genetic score in the sensitivity analysis. INDEL: insertion and deletion; MAF: minor allele frequency.

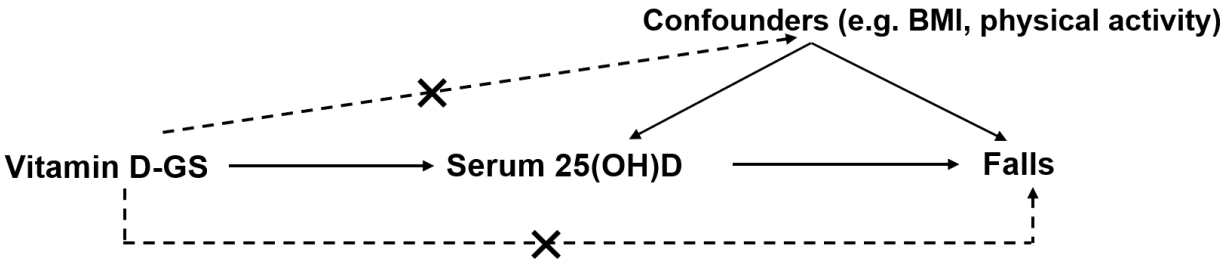


Figure S4. Mendelian randomization assumption diagram.

Valid causal inference from the Mendelian randomization analysis relies on 3 key assumptions: [16] 1) vitamin D genetic score associates with measured 25(OH)D concentrations; 2) vitamin D genetic score has no direct effect on outcome; 3) vitamin D genetic score does not associate with confounders of measured 25(OH)D and outcome. Vitamin D-GS: vitamin D genetic score.

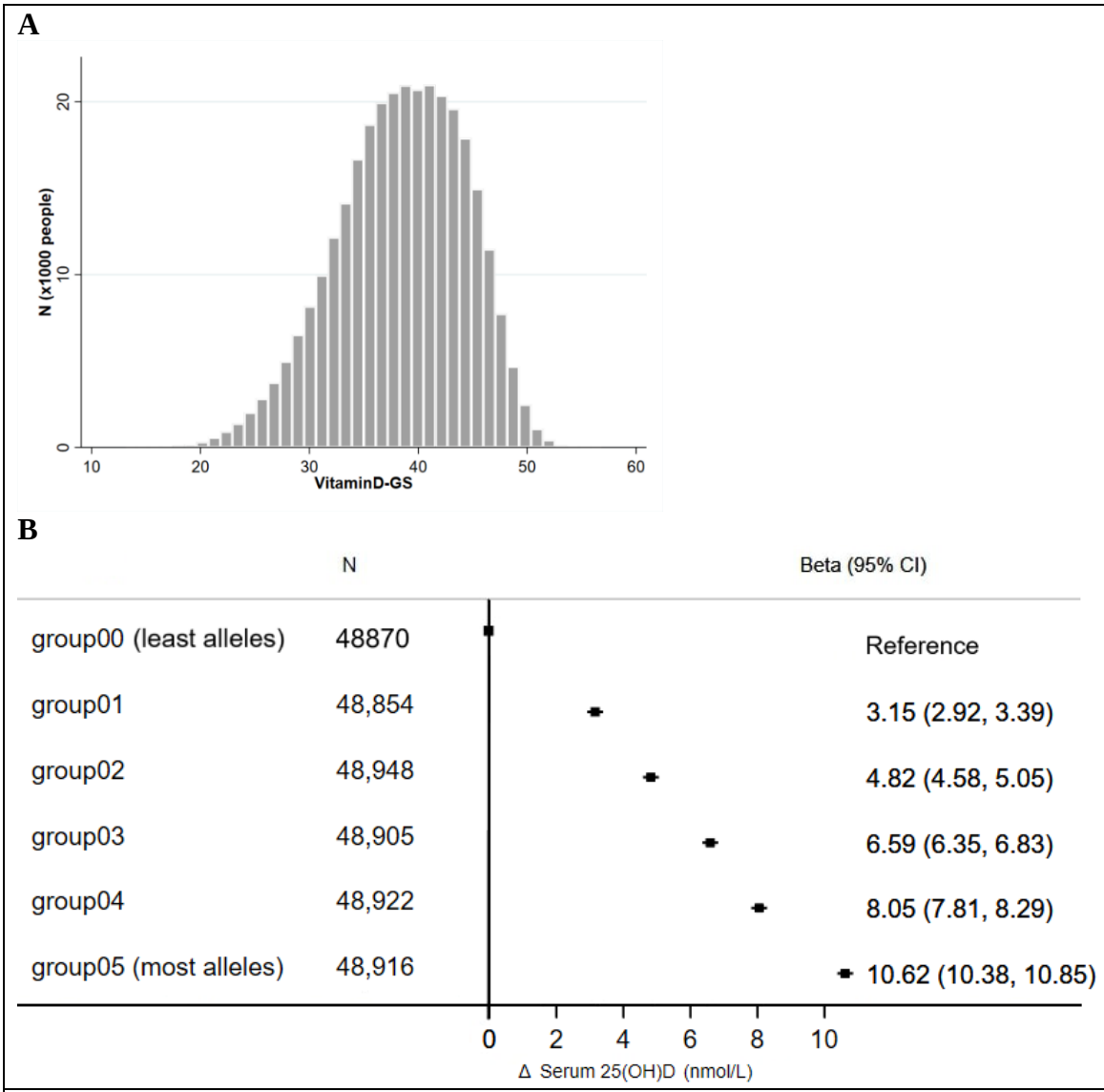


Figure S5. Distribution of the vitamin D genetic score (mapped on an X-axis by increasing number of alleles) in the UK Biobank (A), and the association of the vitamin D genetic score (in six groups – lowest to highest number of alleles) with measured 25(OH)D concentrations in the UK Biobank (B). The association was tested using the linear regression, with the model adjusted for age, sex, genotyping array, assessment center location, top 40 genetic principal components, birth location, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement. Error bars are 95% confidence intervals. Vitamin D-GS: Vitamin D Genetic Score.

Table S1. Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.

See Supplementary Table 01 of the primary thesis:

Supplementary Table 1: Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations. 321

Table S2. Association of the vitamin D genetic score with potential confounders in the UK Biobank.

		Vitamin D-GS
	N (%)	Mean (S.D.)
BMI		
Lowest 25% - 12.1-24.1	76,640 (24.9)	38.41 (5.87)
Middle 50% - 24.1-29.8	153,321 (50.0)	38.42 (5.88)
Highest 25% - 29.8-74.7	76,675 (25.0)	38.45 (5.86)
P[1]		0.88
Smoking		
Non-smokers	167,703 (54.7)	38.42 (5.86)
Ex-smokers	108,118 (35.3)	38.44 (5.89)
Current smokers	30,719 (10.0)	38.39 (5.89)
P[1]		0.18
Alcohol intake		
Daily	65,542 (21.3)	38.34 (5.88)
1 to 4 times week	155,608 (50.6)	38.45 (5.87)
1 to 3 times month	34,098 (11.1)	38.39 (5.87)
Special occasion	32,179 (10.5)	38.36 (5.90)
Never	19,963 (6.5)	38.51 (5.85)
P[1]		0.02
Physical activity		
Low	92,012 (30.6)	38.44 (5.87)
Moderate	149,205 (49.6)	38.42 (5.88)
High	59,561 (19.8)	38.42 (5.87)
P[1]		0.67
Education		
None	52,193 (17.1)	38.44 (5.86)
NVQ/CSE/A-levels	109,099 (35.8)	38.44 (5.88)
Degree/professional	143,735 (47.1)	38.42 (5.88)
P[1]		0.55
Townsend deprivation index quartiles		
Q1 lowest	76,793 (25.0)	38.41 (5.86)
Q2	76,821 (25.0)	38.46 (5.89)
Q3	76,815 (25.0)	38.42 (5.87)
Q4 highest	76,811 (25.0)	38.43 (5.88)
P[1]		0.38

NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; SD, standard deviation; Q, quartiles. Threshold for statistical significance = 0.05/6(confounders) = 0.008. P[1] Values have been adjusted for age, sex, genotyping array, birth location, and assessment center. Vitamin D-GS: vitamin D genetic score

Table S3. Mendelian randomization analysis for the association of genetically predicted 25(OH)D with falls concurring with C-reactive protein ≥ 20 mg/L in the UK Biobank - 122 SNP genetic score and leave block out analysis.

Vitamin D genetic score / leave-out block	**SNPs excluded from the 35 SNP genetic score**	**Outcome* OR: 95% CIs. P**
122 SNP Genetic score	**N/A**	1.18: 1.02, 1.37. 0.02
35 SNP Genetic score	**N/A**	1.20: 1.00, 1.44. 0.05
Blood traits block	rs1047891, rs1352846, rs6672758, rs72997623, rs7528419,rs78151190	1.20: 0.92, 1.57. 0.18
Lipids / Metabolic block	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs7574138, rs6782190	1.18: 0.96, 1.44. 0.12
Renal block	rs1047891, rs17216707, rs77924615	1.18: 0.99, 1.42. 0.07
Unclassified block	rs6782190, rs61891388, rs12803256, rs1047891, rs212100, rs7522116, rs78151190, rs261291, rs12056768, rs1933064	1.20: 0.99, 1.47. 0.06
Renal or Metabolic block	rs76798800, rs78151190, rs77924615, rs261291, rs727857, rs1047891, rs12794714, rs72997623, rs7528419, rs75741381, rs6782190, rs77924615	1.16: 0.95, 1.42. 0.15

*Per 10 nmol/L increase in 25(OH)D. Adjustment includes age, sex, assessment center, SNP array, top 40 genetic principal components, and nuisance factors which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

Table S4. Functional blocks used in the leave-block-out analyses.

Functional block[a]	SNP	Gene	Traits
Blood traits	rs78151190	CARMIL1 (LRRC16A)	Various, incl. hemoglobin, platelet count, reticulocyte count
	rs1352846	GC	White cell count, granulocyte count
	rs1047891	CPS1	Various, incl. amino acid levels, platelet count, hemoglobin, metabolite levels
	rs72997623	DGAT2	Hemoglobin
	rs7528419	CELSR2	Blood protein levels, progranulin levels
	rs6672758	GALNT2	Platelet count, red cell distribution
Lipids/Metabolic	rs76798800	DCST2	Various, incl. fat-free mass, birth weight, height
	rs78151190	CARMIL1 (LRRC16A)	HbA1c, pulse, blood pressure
	rs77924615	PDILT	Blood pressure
	rs261291	ALDH1A2	Cholesterol, lipid metabolism
	rs727857	LINC01122	Fat mass, BMI, impedance, weight
	rs1047891	CPS1	Various, incl. weight, impedance, fat- free mass cholesterol
	rs12794714	CYP2R1	Hip circumference
	rs72997623	DGAT2	HDL cholesterol
	rs7528419	CELSR2	Angina pectoris, coronary artery disease, cholesterol, statin use
	rs75741381	VGF	Impedance
	rs6782190	CADM2	Fat free mass, BMI, impedance
Renal	rs77924615	PDILT	Glomerular filtration rate
	rs1047891	CPS1	Creatinine, chronic kidney disease
	rs17216707	CYP24A1	Creatinine, glomerular filtration rate
Unclassified	rs6782190	CADM2	Nervous feelings, alcohol intake, smoking, risk taking, number of children fathered
	rs61891388	RP11-867G23.13	Qualifications
	rs12803256	AP002387.1	Population differentiation
	rs1047891	CPS1	Headache
	rs212100	SULT2A1	Cholelithiasis, dehydroepiandrosterone sulphate
	rs7522116	FOXO6	Intelligence, qualifications
	rs78151190	CARMIL1 (LRRC16A)	Disorders of mineral metabolism
	rs261291	ALDH1A2	Age-related macular degeneration
	rs12056768	LINC00536	Hair or balding pattern: pattern 4
	rs1933064	FLG-AS1	Atopic dermatitis

[a]Functional blocks identified using trait associations identified through PhenoScanner V2[16].

Table S5. Phenotypic analyses showing the prevalence of falls across CRP ranges, and the association between falls and categorical CRP, in the UK Biobank.

CRP	Fall % (cases)	Simple Model OR: 95% CIs	Adjusted Model OR: 95% CIs
0-0.99 mg/L	16.57 (20,167)	*Reference*	Reference
1-4.99 mg/L	20.40 (30,532)	1.23: 1.20, 1.25	1.08: 1.06, 1.11
5-9.99 mg/L	25.37 (5,649)	1.54: 1.49, 1.60	1.20: 1.16, 1.25
10-14.9 mg/L	25.77 (1,639)	1.56: 1.47, 1.66	1.18: 1.10, 1.25
15-19.9 mg/L	27.19 (741)	1.70: 1.56, 1.87	1.29: 1.17, 1.41
≥ 20 mg/L	26.72 (940)	1.67: 1.54, 1.81	1.36: 1.25, 1.48

CRP = C-reactive protein. OR = odd ratio. CI = confidence intervals. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Table S6. Fully adjusted association between continuous 25(OH)D and falls, stratified above and below CRP thresholds.

CRP	OR: 95% CI. P	Cases / Controls
≥ 5 mg/L	0.9970: 0.996, 0.999. < 0.001	8,157 / 24,547
< 5 mg/L	0.9987: 0.998, 0.999. < 0.001	48,138 / 214,100
≥ 10 mg/L	0.9970: 0.995, 0.999. 0.005	2,972 / 8,739
< 10 mg/L	0.9990: 0.998, 0.999. < 0.001	53,323 / 229,908
≥ 15 mg/L	0.9980: 0.994, 1.000. 0.078	1,506 / 4,291
< 15 mg/L	0.9985, 0.998, 0.999. < 0.001	54,789 / 234,356
≥ 20 mg/L	1.0000: 0.996, 1.004. 0.723	838 / 2421
< 20 mg/L	0.9980: 0.998, 0.999. < 0.001	55,457 / 236,226

CRP = C-reactive protein. OR = odds ratio. CI = confidence intervals. Fully adjusted model was adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; model was additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Table S7. Fully adjusted association between continuous 25(OH)D and falls, with the outcome of falls concurring with inflammation across increasing CRP thresholds.

CRP	OR: 95% CI. P	Cases / Controls
≥ 5 mg/L	0.98: 0.96, 0.99. < 0.001	8,157 / 239,322
≥ 10 mg/L	0.99: 0.97, 1.01. 0.27	2,972 / 239,322
≥ 15 mg/L	1.00: 0.97, 1.03. 0.91	1,506 / 239,322
≥ 20 mg/L	1.04: 1.001, 1.08. 0.04	838 / 239,322

*Per 10 nmol/L increase in 25(OH)D. CRP = C-reactive protein. OR = odds ratio. CI = confidence intervals. Fully adjusted model was adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; model was additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Table S8. Phenotypic sensitivity analyses for the association of categorical 25(OH)D and falls in the UK Biobank, with the outcome of *falls concurring to those with rheumatoid arthritis* (adjusted and simple models).

Fall type	**10-24.9 nmol/L** OR: 95% CI. P	**25-49.9 nmol/L** OR: 95% CI. P	**Reference**	**75-99.99 nmol/L** OR: 95% CI. P	**≥ 100 nmol/L** OR: 95% CI. P
Adjusted	1.28: 1.03, 1.60. 0.03	1.05: 0.90, 1.24. 0.53	*Ref.*	1.48: 1.18, 1.86. 0.001	1.64: 1.00, 2.70. 0.05
Simple	1.97: 1.59, 2.44. <0.001	1.26: 1.07, 1.47. 0.005	*Ref.*	1.34: 1.07, 1.69. 0.01	1.46: 0.90, 2.40. 0.13

OR = odds ratio. CI = confidence intervals. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Table S9. Phenotypic analyses for the association of categorical measured 25(OH)D and falls in the UK Biobank, with falls concurring with CRP <5mg/L.

10-24.9 nmol/L OR: 95% CI. P	**25-49.9 nmol/L** OR: 95% CI. P	**Reference**	**75-99.99 nmol/L** OR: 95% CI. P	**≥ 100 nmol/L** OR: 95% CI. P
1.12: 1.08, 1.16.< *0.001*	1.03: 1.005, 1.05. *0.02*	*Ref.*	1.002: 0.97, 1.04. *0.9*	0.96: 0.89, 1.05. *0.37*

OR = odds ratio. CI = confidence intervals. Simple models were adjusted for sex, age, assessment center, and nuisance factors that could affect 25(OH)D serum measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; fully adjusted models were additionally adjusted for educational status, Townsend deprivation index, body mass index, physical activity, alcohol, and smoking.

Supplementary References:

1. DiaSorin. LIAISON - 25 OH Vitamin D TOTAL Assay 2019.

2. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North: Routledge; 1988.

3. Nuttall FQ. Body Mass Index. Nutrition Today. 2015;50(3):117-28.

4. Vickers D, Rees P. Creating the UK National Statistics 2001 output area classification. Journal of the Royal Statistical Society. 2007;170(2):379-403.

5. Revez JA, Lin T, Qiao Z, Xue A, Holtz Y, Zhu Z, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nature Communications. 2020;11(1).

6. Jiang X, O'Reilly PF, Aschard H, Hsu Y-H, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nature Communications. 2018;9(1).

7. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. International journal of epidemiology. 2013;42(4):1134-44.

8. Palmer TM, Sterne JAC, Harbord RM, Lawlor DA, Sheehan NA, Meng S, et al. Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses. American Journal of Epidemiology. 2011;173(12):1392-403.

9. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597-608.

10. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. International journal of epidemiology. 2010;39(2):417-20.

11. Dudbridge F, Allen RJ, Sheehan NA, Schmidt AF, Lee JC, Jenkins RG, et al. Adjustment for index event bias in genome-wide association studies of subsequent events. Nature communications. 2019;10(1):1561.

12. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genetic Epidemiology. 2016;40(7):597-608.

13. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 2013;37(7):658-65.

14. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International journal of epidemiology. 2015;44(2):512-25.

15. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. Bioinformatics. 2019;35(22):4851-3.

16. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research. 2020;4:186.

Study 4 Addendum

Addendum to study 4:

Vitamin D, C-reactive protein, and increased fall risk: A genetic epidemiological study.

Title: Investigations to support future research.

Joshua P Sutherland (BHSc Nut Med, Hons)[1], Ang Zhou (PhD)[1,2,3], Elina Hyppönen (PhD)[1,2]

1. Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, Australia.
2. South Australian Health and Medical Research Institute, Adelaide, Australia
3. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

Joshua P Sutherland: ORCID 0000-0003-4102-3164

Ang Zhou: ORCID 0000-0001-7088-0777

Elina Hyppönen: ORCID 0000-0003-3670-9399

Introduction

Our study entitled *Is the vitamin D and fall risk relationship dependent on steroid hormone competition? A genetic epidemiological exploration of a novel hypothesis* used a genetic epidemiological approach to explore the association of inflammation (at levels likely to reflect a reduction of retinol, of at least 25% [1]), vitamin D and fall risk. The investigation showed that both low and high 25(OH)D levels may increase fall risk in individuals with chronic inflammation, with supporting evidence for the association between higher 25(OH)D and increased fall risk in those with chronic inflammation provided via Mendelian randomization analysis [2]. This suggests that the increased fall risk following high-dose vitamin D supplementation observed in clinical trials, may relate to excessive production of vitamin D receptor–retinoid X receptor (VDR-RXR) heterodimers that leads to functional retinol deficiency (via steroid hormone competition [3-5]) affecting motor function. Future work in this area needs to explore the calcitriol and retinoic steroid hormone competition hypothesis via methods that can directly observe these biological pathways and observe the impact of steroid hormone competition on motor function regulating specific gene expression. Investigating gene expression may deepen our understanding of the high vitamin D - increased fall risk phenomenon, and lead to potential pharmaceutical or nutraceutical interventions that leverage these insights [6]. This future work is likely to involve rodent animal models and would likely benefit from building upon existing rodent model research that has already explored retinol deficiency and motor function, and the knocking out of retinoic acid genes involved in regulating motor function.

These previous animal model investigations support a relationship between retinol deficiency and impaired motor function in rats [7], as well as an adverse effect on motor function arising from the knocking out of *BRINP* family genes (Bone Morphogenetic Protein/Retinoic

Acid Inducible Neural-Specific Protein) in mice [8-10]. Future rodent model research investigating the calcitriol and retinoic acid competition hypothesis can incorporate messenger ribonucleic acid (m-RNA) extraction [11] to assess the extent to which genetic expression is affected by retinoic acid restriction / vitamin D interventions in motor function regulating genes. To support future work, in the following I present details for additional genetic epidemiological investigations that examine the relationship between retinoic acid-inducible, motor function associated genes, and the outcome of falls in the human setting. To explore how variations in retinoic acid-inducible genes might relate to motor function pathology in the UK Biobank, we investigated functional variants within genes that are known to respond to retinoic acid and regulate healthy motor function, such as the *BRINP1* gene [8, 9].

Methods:

To identify retinol responsive variants that may have a direct effect on motor function we first compiled a list of genes that had a known responsiveness to retinoic acid and were known to contribute (via experimental animal models or human observational research) to the regulation of healthy motor function. From this list we could then search the GWAS Catalog [12] for single nucleotide polymorphisms (SNPs) that had trait associations with acquired neurological conditions that associate with fall risk. The most noteworthy candidate was the ***BRINP*1** gene, as in mouse models, *BRINP1* knock-out induces dysregulated/hyperactive motor function [8, 10] and as in humans, pathological variants in this gene are associated with Parkinson's disease and attention-deficit/hyperactivity disorder [13, 14]—which are both diseases characterized by motor regulation issues [15, 16] and which both associate with an increased risk of fall related injury [17, 18]. We also identified other potentially relevant retinoic acid responsive and motor function

relating genes, although for these, experimental data to confirm effects was less strong, and thus proposed connections are more hypothetical in nature. **Nurr1/NR4A2** requires retinoic acid for RXR-Nurr1 heterodimers and aids in developing dopamine neurons [19] This gene, through its effect on dopamine [20], is associated with Parkinson's disease [21] and has a proposed role in Alzheimer's disease [22, 23], which is also associated with motor function dysregulation [24]. **GABA subunit genes** are implicated in neurological disorders as well [25-27], with GABA insufficiency also linked to cervical dystonia [28]. Other genes of interest that associate with neurological disorders include **BRINP2**, which is linked to schizophrenia [29] and **NRG1** [30], which is associated with bipolar disorder [31], schizophrenia [32], and depression [33]. These disorders themselves underlie an increased risk of falls [34, 35]. Where present, appropriate variants from these genes were incorporated into analyses using SNP data from the UK Biobank and the outcome of fall (as detailed in the primary study). GWAS catalog was accessed April 2023. The population used for the current SNP-fall analyses: n = 336,517 (**Figure 1**).

Statistical methods for the SNP-fall analyses: For the SNP-fall association analyses, logistic regression was used to assess the association between the risk containing SNP allele groups and the fall outcome (using the no-risk allele group as reference). The model was adjusted for age, sex, assessment centre, SNP array, and top 40 genetic principal components. We also investigated effect modification by the selected SNPs and 25(OH)D on their association with falls, by including an interaction term in the logistic regression model.

Results:

The distribution of the prevalence of falls, as seen for the population used for the SNP-association analyses, are presented in **Table 1**. These Table 1 patterns are similar to those discussed in the primary study, with notably higher prevalence of falls in females vs. males, in those with higher body mass index (BMI) vs. those with lower BMI, and in those experiencing the most socioeconomic deprivation vs. the least.

SNP-fall analysis:

To support future research examining the vitamin D and retinol competitive antagonism aspect of our hypothesis, we investigated the chosen SNPs (from the motor function associating retinol responsive genes) for their associations with falls (**Table 2**). There was suggestive evidence for an association between the *BRINP1* SNP rs4837628 (T allele linked to Parkinson's disease) and rs4837631 (C allele linked to attention deficit/hyperactivity) with falls (TC, OR = 1.02, 95% CI, 1.0001-1.05; and CT, OR = 1.02, 95% CI, 1.001-1.05, respectively), with these SNPs being in close but not perfect linkage disequilibrium (LD) (R^2 =0.89) (**Table 3**).

None of the other variants showed any evidence for an association with fall risk. However, we observed evidence for interaction between 25(OH)D and two *BRINP2* SNPs (rs17659437 and rs6670165) on their association with the risk of falls (likelihood ratio test P = 0.006 and 0.02, respectively). In the analyses stratified by these BRINB2 variants, the interaction appeared to reflect relatively minor differences in the strength of the association between 25(OH)D and the odds of falls, although, for the association between low 25(OH)D concentrations (<25nmol/l vs. 50-75nmol/) (**Table 4**), the odds of falls for rs17659437 was notably stronger in the CC homozygotes than in the other genotype groups.

Discussion:

In this additional investigation, we observed some evidence of a genetic pathway via which retinoic acid inducible gene expression (via *BRINP1*) could manifest in an increased risk of falls, and where (via *BRINP2*) the 25(OH)D-fall association could vary based on these potential steroid hormone balance related factors. The results from our SNP-fall analysis affirm our initial interest the *BRINP* gene family, which was predicated on the established retinoic acid inducible aspect of the gene family, and its overt connection with regulation of healthy motor function [8, 9]. In genome wide association studies (GWAS), *BRINP1* SNP – trait associations have included motor function related conditions (i.e., Parkinson's disease and attention-deficit/hyperactivity disorder [36, 37]), suggesting that this part of the gene on chromosome 9 might be particularly related to the hyper-locomotor/hyperactivity aspects seen in the gene knock out studies [8, 10], while with *BRPN2* (GWAS link to schizophrenia [38,39]) the chromosome 1 location also appears to influence hyperactivity [9]. These *BRINP1* SNPs [rs4837628 (Parkinson's disease) and rs4837631 (attention-deficit/hyperactivity disorder)] were the only variants to provide evidence of a connection with falls, while a connection with falls via *BRINP2* [rs17659437 and rs6670165 (schizophrenia)] were perhaps masked by more complex interactions with 25(OH)D, as evidenced by the strong interaction between the 25(OH)D – fall association and *BRINP2* variants.

We note that both Parkinson's disease and attention-deficit/hyperactivity disorder have complex inheritance patterns [40, 41]. Interestingly, however, both Parkinson's disease (nonhereditary and hereditary forms) and attention-deficit/hyperactivity disorder have an established observed connection with deficiency of retinoic acid/retinol, respectively [42, 43]—with this insight being relevant given the connection between these disorders and *BRINP1*.

Further interest (especially considering the tight LD correlation between the SNPs) is recent research associating attention-deficit/hyperactivity disorder and Parkinson's disease together, with the former now noted as increasing the risk for the latter [44].

Conclusions:

Our SNP-fall analyses support existing insights that affirm a BRINP–motor function relationship, with our investigation linking this relationship to fall risk in humans and highlighting likely 25(OH)D interaction. This insight likely warrants the inclusion of BRINP genes analysis in research that directly examines the higher vitamin D and increased fall risk phenomenon via a steroid hormone competition lens. Further studies can explore the mechanistic insights raised here and more effectively seek for interactions between vitamin D, retinol, and dysregulation of *BRINP1* gene expression, with such research ascertaining if motor instability can be induced via *BRINP1* suppression arising from a disruption to steroid hormone activity homeostasis.

Acknowledgements:

We acknowledge Dr Amanda Lumsden for general support on matters of genetics, and Natasha Radcliffe for providing insights pertaining to retinol/*BRINP* related factors.

Table 1. Demographic characteristics of UK Biobank participants

	N(%)	Falls N=66,177 %
All	336,517 (100)	19.65
Sex		
Males	155,763 (46.3)	15.70
Females	180,754 (53.7)	23.05
Age		
<60	184,579 (54.9)	17.85
≥60	151,938 (45.2)	21.83
BMI		
Low 25% 12.1-24.1	83,836 (24.9)	17.27
Mid 50% 24.1-29.8	167,778 (49.9)	18.32
High 25% 29.8-74.7	83,830 (24.9)	24.37
Missing	1,073 (0.3)	42.68
Location		
South ≤51° Lat	111,245 (33.1)	18.93
Mid 52°-53° Lat	159,153 (47.3)	20.21
North 54°- ≥55° Lat	66,119 (19.7)	19.51
Smoking		
Non-smokers	183,379 (54.5)	18.76
Ex-smokers	118,189 (35.1)	20.01
Current smokers	33,817 (10.0)	23.00
Missing	1,132 (0.4)	25.53
Alcohol		
Daily	72,119 (21.4)	18.51
1 to 4 times wk	169,776 (50.5)	18.10
1 to 3 times mo	37,262 (11.1)	20.73
Special occasion	35,276 (10.5)	24.19
Never	21,871 (6.5)	26.06
Missing	213 (0.06)	35.21
Physical activity		
Low	100,937 (30.0)	20.32
Moderate	162,883 (48.4)	18.49
High	64,931 (19.3)	19.85
Missing	7,766 (2.3)	33.47
Education		
None	57,399 (17.1)	23.91
NVQ/CSE/A-Levels	119,298 (35.5)	19.42
Degree/professional	157,066 (46.7)	18.21
Missing	2,754 (0.8)	22.55
Townsend index		
Q1 Deprivation – lowest	84,073 (25.0)	17.26
Q2	84,111 (25.0)	18.20
Q3	84,051 (25.0)	19.58
Q4 Deprivation – highest	84,886 (25.0)	23.54
Missing	396 (0.1)	22.22

CRP, C-reactive protein; NVQ, National Vocational Qualification; CSE, Certificate of Secondary Education; A-levels, Advanced levels; Q, quartiles. All models were adjusted for sex, age, assessment centre; all models in table P value ≤0.002).

Table 2. Association between motor function disorder relating SNPs – from retinol responsive genes – with falls in the UK Biobank.

Mapped Genes	SNP	Trait	Risk Allele	MAF	Trait GWAS P Value	Association of risk allele with fall OR: CI 95%[a]		N[b]	P value
BRINP1	rs4837631	ADHD	C	0.4347	2×10^{-11}	CT CC	1.02: 1.001, 1.05 1.02: 0.99, 1.04	165,982 101,552	0.05 0.20
	rs4837628	Parkinson's disease	T	0.4449	$^{*}1 \times 10^{-6}$	TC TT	1.02: 1.0001, 1.05 1.01: 0.99, 1.04	165,753 105,088	0.05 0.27
	rs1331621	Bipolar disorder	T	0.4299	$^{*}4 \times 10^{-7}$	TC TT	1.01: 0.99, 1.03 1.02: 0.99, 1.05	155,239 43,693	0.22 0.22
	rs10818399	Insomnia	C	0.4066	1×10^{-11}	CG CC	1.00: 0.98, 1.02 0.99: 0.96, 1.01	161,540 54,950	0.87 0.26
	rs942516	Bipolar disorder	A	0.1026	$^{*}5 \times 10^{-6}$	GA AA	1.00: 0.98, 1.02 0.98: 0.93, 1.04	82,729 6,893	0.40 0.60
	rs7847153	Schizophrenia	C	0.2127	$^{*}9 \times 10^{-7}$	CG CC	0.98: 0.95, 1.02 0.98: 0.95, 1.01	124,063 131,361	0.31 0.16
BRINP 2	rs17659437	Schizophrenia	T	0.1607	2×10^{-8}	TC TT	1.04: 0.96, 1.12 1.02: 0.95, 1.10	67,545 239,762	0.33 0.61
	rs16851048	Schizophrenia	T	0.2666	5×10^{-14}	TC TT	1.02: 0.97, 1.07 1.01: 0.96, 1.05	103,368 219,896	0.80 0.26
	rs6670165	Schizophrenia	T	0.1885	4×10^{-11}	CT TT	1.01: 0.99, 1.03 0.99: 0.95, 1.04	102,331 12,029	0.38 0.67
NRG1	rs1383890	Insomnia	T	0.3385	4×10^{-8}	CT TT	1.06: 0.98, 1.03 1.08: 0.99, 1.04	164,128 110,275	0.61 0.19
	rs17673138	Schizophrenia, Bipolar, depression	C	0.0711	$^{*}6 \times 10^{-8}$	AC CC	1.00: 0.97, 1.03 0.98: 0.88, 1.09	49,365 2,184	0.86 0.75
	rs11771523 6	MAD	A	0.0072	$^{*}4 \times 10^{-6}$	TA AA	0.98: 0.93, 1.02 0.81: 0.56, 1.17	16,180 208	0.18 0.26

*Trait association above 5×10^{-8}. MAF = minor allele frequency, ADHD = attention-deficit/hyper-activity disorder, MAD = major depressive disorder. [a] Relative to double non-risk allele group. [b] Total n = 337,466. Logistic regresssion adjusted for age, sex, assessment centre, SNP array, and principal components 1-40.

Table 3. Pairwise linkage disequilibrium statistical matrix: R2 values for the linkage disequilibrium relationships between the BRINP1 SNPs.

rs4837628	rs4837631	rs942516	rs1331621	rs7847153	rs10818399
1	0.894	0.083	0.023	0.003	0.027
0.894	1	0.077	0.011	0.001	0.036
0.083	0.077	1	0.107	0	0.006
0.023	0.011	0.107	1	0.008	0.024
0.003	0.001	0	0.008	1	0
0.027	0.036	0.006	0.024	0	1

Calculated in April of 2023 using LD Matrix: https://ldlink.nci.nih.gov/

Table 4. 25(OH)D-fall association stratification by BRINP2 variant allele groups.

Variant	Allele group	Fall cases in model	25(OH)D Continuous (OR 95% CIs)	25(OH)D <25 nmol/L OR: 95% CIs. P	25(OH)D >75 nmol/L OR: 95% CIs. P	25(OH)D >100 nmol/L OR: 95% CIs. P	P-int* LRT
rs17659437	CC	810 / 18.9%	0.99: 0.9875, 0.9960. P 0.000	1.66: 1.32, 2.10. 0.000 case n = 136 / 16.8%	1.16: 0.89, 1.51. 0.20 case n = 77 / 9.5%	0.65: 0.30, 1.38. 0.30 Case n = 8 / 0.99%	
	TC	11,489 / 19.4%	0.998: 0.997, 0.999. P 0.003	1.11: 1.04, 1.19. 0.001 case n = 1,513 / 13.2%	1.05: 0.98, 1.12. 0.17 case n = 1,243 / 10.8%	1.05: 0.89, 1.24. 0.58 Case n = 172 / 1.50%	
	TT	39,855 / 19.0%	0.998: 0.9978, 0.999. P 0.000	1.14: 1.10 1.18. 0.000 case n = 5,262 / 13.2%	1.01: 0.98, 1.05. 0.58 case n = 4,541 / 11.4%	1.022: 1.021, 1.024. 0.000 Case n = 582 / 1.46%	0.006
rs6670165	CC	37,055 / 19.0%	0.998: 0.9977, 0.999. P 0.000	1.14: 1.10, 1.18. 0.000 case n = 4,968 / 13.4%	1.01: 0.97 1.04. 0.76 case n = 4,214 / 11.4%	0.97: 0.88, 1.07. 0.51 Case n = 535 / 1.44%	
	CT	17,235 / 19.2%	0.999: 0.998, 0.9995. P 0.002	1.10: 1.05, 1.16. 0.000 case n = 2,253 / 13.1%	1.05: 0.99, 1.11. 0.08 case n = 1,888 / 11.0%	1.05: 0.99, 1.11. 0.37 Case n = 267 / 1.55%	
	TT	2,011 / 19.1%	0.995: 0.992, 0.997. P 0.000	1.34: 1.16, 1.57. 0.000 case n = 305 / 15.2%	1.07: 0.91, 1.25. 0.44 case n = 218 / 10.8%	1.07: 0.91, 1.25. 0.09 Case n = 22 / 1.09%	0.02

* LRT - SNP-Fall association, interaction with continuous 25(OH)D

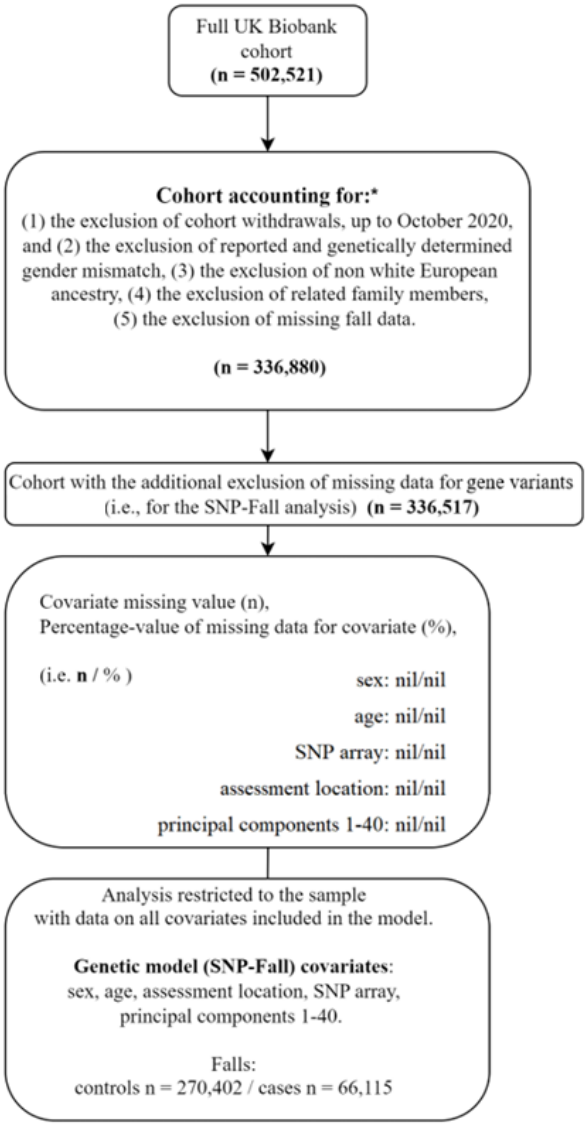


Figure 1. Study population and participant inclusion criteria.

*Exclusion was done in a sequential order.

References:

1. Thurnham DI, McCabe GP, Northrop-Clewes CA, Nestel P. Effects of subclinical infection on plasma retinol concentrations and assessment of prevalence of vitamin A deficiency: meta-analysis. The Lancet. 2003;362(9401):2052-8.

2. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research. 2020;4:186.

3. Rohde CM, DeLuca HF. All-trans retinoic acid antagonizes the action of calciferol and its active metabolite, 1, 25-dihydroxycholecalciferol, in rats. The Journal of nutrition. 2005;135(7):1647-52.

4. Metz AL, Walser MM, Olson WG. The interaction of dietary vitamin A and vitamin D related to skeletal development in the turkey poult. The Journal of nutrition. 1985;115(7):929-35.

5. Mata-Granados JM, Cuenca-Acevedo R, Luque De Castro MD, Sosa M, Quesada-Gómez JM. Vitamin D deficiency and high serum levels of vitamin A increase the risk of osteoporosis evaluated by Quantitative Ultrasound Measurements (QUS) in postmenopausal Spanish women. Clinical Biochemistry. 2010;43(13-14):1064-8.

6. O'Shea R, Ma AS, Jamieson RV, Rankin NM. Precision medicine in Australia: now is the time to get it right. Med J Aust. 2022;217(11):559-63.

7. Carta M, Stancampiano R, Tronci E, Collu M, Usiello A, Morelli M, et al. Vitamin A deficiency induces motor impairments and striatal cholinergic dysfunction in rats. Neuroscience. 2006;139(4):1163-72.

8. Berkowicz SR, Featherby TJ, Qu Z, Giousoh A, Borg NA, Heng JI, et al. Brinp1 −/− mice exhibit autism-like behaviour, altered memory, hyperactivity and increased parvalbumin-positive cortical interneuron density. Molecular Autism. 2016;7(1).

9. Berkowicz SR, Featherby TJ, Whisstock JC, Bird PI. Mice lacking Brinp2 or Brinp3, or both, exhibit behaviors consistent with neurodevelopmental disorders. Frontiers in behavioral neuroscience. 2016;10:196.

10. Kobayashi M, Nakatani T, Koda T, Matsumoto K-I, Ozaki R, Mochida N, et al. Absence of BRINP1 in mice causes increase of hippocampal neurogenesis and behavioral alterations relevant to human psychiatric disorders. Molecular Brain. 2014;7(1):12.

11. Bravo-Ferrer I, Khakh BS, Díaz-Castro B. Cell-specific RNA purification to study translatomes of mouse central nervous system. STAR Protoc. 2022;3(2):101397.

12. Buniello A, Jacqueline A, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research. 2019;47(D1):D1005-D12.

13. Chairta PP, Hadjisavvas A, Georgiou AN, Loizidou MA, Yiangou K, Demetriou CA, et al. Prediction of Parkinson's Disease Risk Based on Genetic Profile and Established Risk Factors. Genes. 2021;12(8):1278.

14. Pol-Fuster J, Cañellas F, Ruiz-Guerra L, Medina-Dols A, Bisbal-Carrió B, Ortega-Vila B, et al. The conserved ASTN2/BRINP1 locus at 9q33.1–33.2 is associated with major psychiatric disorders in a large pedigree from Southern Spain. Scientific Reports. 2021;11(1).

15. Stray LL, Kristensen Ø, Lomeland M, Skorstad M, Stray T, Tønnessen FE. Motor regulation problems and pain in adults diagnosed with ADHD. Behavioral and Brain Functions. 2013;9(1):18.

16. Mazzoni P, Shabbott B, Cortes JC. Motor Control Abnormalities in Parkinson's Disease. Cold Spring Harbor Perspectives in Medicine. 2012;2(6):a009282-a.

17. Pelicioni PHS, Menant JC, Latt MD, Lord SR. Falls in Parkinson's Disease Subtypes: Risk Factors, Locations and Circumstances. International Journal of Environmental Research and Public Health. 2019;16(12):2216.

18. Amiri S, Sadeghi-Bazargani H, Nazari S, Ranjbar F, Abdi S. Attention deficit/hyperactivity disorder and risk of injuries: a systematic review and meta-analysis. Journal of injury and violence research. 2017;9(2):95.

19. Zetterstrom RH, Solomin L, Jansson L, Hoffer BJ, Olson L, Perlmann T. Dopamine neuron agenesis in Nurr1-deficient mice. Science. 1997;276(5310):248-50.

20. Sacchetti P, Carpentier R, Ségard P, Olivé-Cren C, Lefebvre P. Multiple signaling pathways regulate the transcriptional activity of the orphan nuclear receptor NURR1. Nucleic Acids Research. 2006;34(19):5515-27.

21. Al-Nusaif M, Yang Y, Li S, Cheng C, Le W. The role of NURR1 in metabolic abnormalities of Parkinson's disease. Molecular Neurodegeneration. 2022;17(1).

22. Moon M, Jung ES, Jeon SG, Cha M-Y, Jang Y, Kim W, et al. Nurr1 (NR4A2) regulates Alzheimer's disease-related pathogenesis and cognitive function in the 5XFAD mouse model. Aging Cell. 2019;18(1):e12866.

23. Pan X, Kaminga AC, Wen SW, Wu X, Acheampong K, Liu A. Dopamine and dopamine receptors in Alzheimer's disease: a systematic review and network meta-analysis. Frontiers in aging neuroscience. 2019;11:175.

24. Stark SL, Roe CM, Grant EA, Hollingsworth H, Benzinger TL, Fagan AM, et al. Preclinical Alzheimer disease and risk of falls. Neurology. 2013;81(5):437-43.

25. Neelands TR, Zhang J, Macdonald RL. GABA_A Receptors Expressed in Undifferentiated Human Teratocarcinoma NT2 Cells Differ from Those Expressed by Differentiated NT2-N Cells. The Journal of Neuroscience. 1999;19(16):7057-65.

26. Addae C, Yi X, Gernapudi R, Cheng H, Musto A, Martinez-Ceballos E. All-trans-retinoid acid induces the differentiation of encapsulated mouse embryonic stem cells into GABAergic neurons. Differentiation. 2012;83(5):233-41.

27. Song X-Q, Meng F, Ramsey D, Ripps H, Qian H. The GABA ρ1 subunit interacts with a cellular retinoic acid binding protein in mammalian retina. Neuroscience. 2005;136(2):467-75.

28. Berman BD, Pollard RT, Shelton E, Karki R, Smith-Jones PM, Miao Y. GABAA receptor availability changes underlie symptoms in isolated cervical dystonia. Frontiers in Neurology. 2018;9:188.

29. Smeland OB, Wang Y, Lo M-T, Li W, Frei O, Witoelar A, et al. Identification of genetic loci shared between schizophrenia and the Big Five personality traits. Scientific Reports. 2017;7(1).

30. Chapman K, Medrano G, Chaudhary J, Hamra F. NRG1 and KITL signal downstream of retinoic acid in the germline to support soma-free syncytial growth of differentiating spermatogonia. Cell Death Discovery. 2015;1(1):15018.

31. Prata DP, Breen G, Osborne S, Munro J, Clair DS, Collier DA. An association study of the neuregulin 1 gene, bipolar affective disorder and psychosis. Psychiatric genetics. 2009;19(3):113-6.

32. Stefansson H, Petursson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigmundsson T, et al. Neuregulin 1 and Susceptibility to Schizophrenia. The American Journal of Human Genetics. 2002;71(4):877-92.

33. Levchenko A, Vyalova NM, Nurgaliev T, Pozhidaev IV, Simutkin GG, Bokhan NA, et al. NRG1, PIP4K2A, and HTR2C as potential candidate biomarker genes for several clinical subphenotypes of depression and bipolar disorder. Frontiers in genetics. 2020;11:936.

34. Hausdorff JM, Peng C-K, Goldberger AL, Stoll AL. Gait unsteadiness and fall risk in two affective disorders: a preliminary study. BMC psychiatry. 2004;4(1):1-7.

35. Tsuji Y, Akezaki Y, Mori K, Yuri Y, Katsumura H, Hara T, et al. Factors inducing falling in schizophrenia patients. Journal of Physical Therapy Science. 2017;29(3):448-51.

36. Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nature Neuroscience. 2021;24(10):1367-76.

37. Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, et al. Genome-Wide Association Study Confirms SNPs inSNCAand theMAPTRegion as Common Risk Factors for Parkinson Disease. Annals of Human Genetics. 2010;74(2):97-109.

38. Yao X, Glessner JT, Li J, Qi X, Hou X, Zhu C, et al. Integrative analysis of genome-wide association studies identifies novel loci associated with neuropsychiatric disorders. Translational psychiatry. 2021;11(1):69.

39. Ikeda M, Takahashi A, Kamatani Y, Momozawa Y, Saito T, Kondo K, et al. Genome-wide association study detected novel susceptibility genes for schizophrenia and shared trans-populations/diseases genetic effect. Schizophrenia bulletin. 2019;45(4):824-34.

40. Klein C, Westenberger A. Genetics of Parkinson's Disease. Cold Spring Harbor Perspectives in Medicine. 2012;2(1):a008888-a.

41. Wallis D, Russell HF, Muenke M. Genetics of attention deficit/hyperactivity disorder. Journal of Pediatric Psychology. 2008;33(10):1085-99.

42. Marie A, Darricau M, Touyarot K, Parr-Brownlie LC, Bosch-Bouju C. Role and mechanism of vitamin a metabolism in the pathophysiology of Parkinson's Disease. Journal of Parkinson's disease. 2021;11(3):949-70.

43. Li H-H, Yue X-J, Wang C-X, Feng J-Y, Wang B, Jia F-Y. Serum levels of vitamin A and vitamin D and their association with symptoms in children with attention deficit hyperactivity disorder. Frontiers in Psychiatry. 2020;11:599958.

44. Baumeister AA. Is Attention-Deficit/Hyperactivity Disorder a Risk Syndrome for Parkinson's Disease? Harvard Review of Psychiatry. 2021;29(2):142-58.

Chapter Nine: Discussion

This thesis aims to explore the multifaceted role of vitamin D in degenerative health, shedding light on its association with various health outcomes while addressing pertinent safety concerns. The investigation initially examined determinants of vitamin D deficiency, importantly highlighting health vulnerabilities, with high rates of deficiency among specific ethnic groups and socioeconomically deprived individuals. Subsequent analyses revealed a causal relationship between vitamin D and skeletal muscle health, indicating potential implications for age-related muscle loss. However, the benefits of vitamin D on muscle may be limited by obesity-related factors. Links between higher vitamin D status and reduced mortality from respiratory-related illness are also documented. Additionally, the nonlinear MR findings contribute to research efforts continuing to elucidate threshold effects between 25(OH)D status and mortality using genetic epidemiology. By leveraging theoretical understanding of steroid hormone competition, research conducted in this thesis was able to, for the first time, decern in a cross sectional, population-based setting the pattern of increased fall risk associating with a higher 25(OH)D status. Genetic evidence was then provided that could support future research efforts in exploring the initially proposed steroid hormone competition hypothesis.

Fundamentally, it is timely to explore in this thesis the contemporary dichotomy of optimism and concern for vitamin D, as such dichotomy has persisted since the inception of vitamin D study, and appreciating this historical context is crucial for navigating the challenges that lie ahead. A deep comprehension of vitamin D metabolism, developed over recent decades, underscores the essential nature of vitamin D, and the balance it requires. The intricate interplay between inflammation, immunomodulation, metabolism, and its impact on the musculoskeletal system, coupled with the genomic mechanisms driving many of these effects, shapes both vitamin D's profound influence on health and the potential for imbalance to lead to pathology. While existing research has explored this dichotomy in relation to cardiovascular health (Wang et al. 2018), or steroid hormone balance [in terms of cancer proliferation (Jiménez-Lara & Aranda 1999) and bone remodelling (Mata-Granados et al. 2010)], research documented in this thesis investigates how these intersecting mechanisms of vitamin D, in conjunction with inflammation, may impact degenerative health from a largely motor function related perspective. While there can be tendencies to compartmentalise exploration into cardiovascular or skeletal

health [focusing primarily on the systems being examined], exploring degenerative health through the lens of motor function offers unique advantages. Motor function related pathology serves as pivotal points in pathological cascades, with muscular decline often preceding deterioration of respiratory and metabolic health, which are central aspects of overall health decline (Benz et al. 2019; Joo & Kim 2023; Kim et al. 2021). Moreover, the deterioration of motor function impacts movement and the ability to engage in healthy activities, which are essential components of cardiovascular/respiratory and skeletal health. Thus, a motor function focus can extend considerations, revealing crucial connections to broader aspects of health and mortality risk.

While the findings from this thesis are intrinsically linked to the available 25(OH)D data, inference can be made regarding underlying vitamin D metabolism, which serves to place these findings into broader and helpful context. For instance, the effects of adverse metabolic status on vitamin D metabolism are areas of ongoing investigation and, as is known, obesity downregulates CYP2R1 mediated 25(OH)D synthesis and is likely interlinked to compensatory effects, such as leptin-induced increases of calcitriol synthesis (Elkhwanky et al. 2020; Hou et al. 2020). The findings from this thesis—indicating limited benefits of vitamin D on sarcopenia risk in obesity [Study II]—suggest further ways in which adverse metabolic status potentially affect the vitamin D system through a potential shifting of metabolic demand. Mediated by the modulation of proinflammatory cytokines like TNF-α and interleukin-6, alongside the release of anti-inflammatory cytokines such as interleukin-10, the immune system demonstrates a pronounced need for 25(OH)D in the presence of inflammation (Ao, Kikuta & Ishii 2021; Currò et al. 2020; Leal et al. 2020). As is also known, CYP27B1 activity exists throughout macrophages, dendritic cells, and T helper cells (Khashim Alswailmi et al. 2021; Lopez et al. 2021) which, upon activation by inflammation, could constitute a competing draw of already reduced 25(OH)D in obese individuals and could see vitamin D utilisation shifting to immune modulation from muscle support [where it is needed for crucial steps in myogenesis and contractile actions (Romeu Montenegro et al. 2019)]. Understanding such metabolic influences is crucial for developing personalized nutritional strategies to optimize treatment plans for obesity and metabolic disorders, simultaneously addressing immune modulation, inflammation, and muscle loss.

Further examination of vitamin D metabolism elucidates its connection to mortality. To briefly recap, clinical trial evidence suggests vitamin D reduces mortality risk, especially in the 25(OH)D deficient critically ill (Amrein et al. 2014; Hasanloei et al. 2020), those with concurrent COVID-19 and 25(OH)D deficiency (D'Ecclesiis et al. 2022; Meng et al. 2023), and certain cancers (Cao et al. 2023). A connection between vitamin D deficiency and all-cause mortality in a phenotypic context is shown in this thesis and establishes a linear MR association with respiratory-related mortality. Drawing from the research presented in this thesis a fundamental concept is proposed: that the observed reduction in mortality risk attributed to vitamin D is rooted in its role in maintaining the minimal reserves necessary to sustain life. While this concept can be examined from various perspectives, the exploration of its genetic and epigenetic implications emerges as particularly helpful. At the deepest layer of analysis, genetic and epigenetic mechanisms not only directly influence the ageing process but also likely form the foundational framework for which higher-level factors affect mortality risk (Pal & Tyler 2016). Ageing is intricately linked with epigenetic processes, particularly modifications to histones. Histone modifications—including acetylation and methylation—regulate cellular functions like genetic expression, stress response, and DNA repair, ultimately modulating hormetic effects to promote resilience and longevity (Pal & Tyler 2016; Peleg et al. 2016). The interaction between VDR/RXR heterodimers and histones help to illustrate this connection. VDR/RXR heterodimers recruit histone acetyltransferases to acetylate histones, loosening chromatin and enhancing gene transcription, vital for cellular health (Karlic & Varga 2011). These epigenetic mechanisms, highlighted as important in longevity and ageing, suggest that understanding the interplay between VDR/RXR heterodimers, histone modifications, and ageing processes could provide further mechanistic insights into mortality risk and age-related diseases.

The crucial role of calcitriol in preserving respiratory health has previously been emphasized in this thesis, primarily through the production of antimicrobial peptides and the modulation of immune-system cell proliferation (White 2022). This focus gains importance considering Study III's linear MR respiratory related mortality findings, and especially given the shared pathological progression, seen often in degenerative health scenarios, between respiratory illness and loss of muscular integrity (Benz et al. 2019). It is imperative to acknowledge that even in these scenarios, epigenetic mechanisms—and disruptions therein that could potentially exacerbate from inadequate vitamin D levels—retain relevance, with these fundamental

epigenetic processes exerting impact on physiological factors needed to sustain life and likely mitigate degenerative decline.

When assessing the risk of vitamin D, it is crucial to acknowledge that the UK Biobank data lacks sufficient instances of very high levels of 25(OH)D to directly evaluate the impact of excessive 25(OH)D on fall risk. However, this limitation provides an opportunity to investigate whether the increased fall risk observed in some primary and subgroup clinical trial analyses, involving D3 bolus and higher daily dosing (Bischoff-Ferrari et al. 2016; Ginde et al. 2017; Sanders et al. 2010; Waterhouse et al. 2021; Smith, Gallagher & Suiter 2017; Zittermann et al. 2023), can occur in other scenarios and with different 25(OH)D conditions. Fundamentally, it is of public health relevance to ascertain if this phenomenon might extend beyond bolus dosing scenarios, serving to clarify if the phenomenon is solely due to the dosing method, or if it is in part attributable to 25(OH)D level fluctuation. When considering the use of a population without notable 25(OH)D excess, the question arose whether specific physiological states, such as inflammation, might reveal this effect at higher, yet non-excessive 25(OH)D levels. This line of inquiry proved valuable, as it allowed for a discerning of the pattern [higher 25(OH)D associating with increased fall risk] to occur for the first time outside of a clinical trial setting, specifically in individuals with chronic inflammation. Discerning the pattern outside a trial-based scenario underscores the need to investigate mechanisms beyond those typically associated with bolus dosing [such as the activation of counteracting factors like 24-hydroxylase (Mazess, Bischoff-Ferrari & Dawson-Hughes 2021)], with a logical focus for further research being potential competition between VDR/RXR and RAR/RXR in motor function regulating genes, or other as yet fully understood inflammation-25(OH)D interactions.

9.1 The interconnectedness of vitamin D in degenerative health pathology.

Vitamin D-related pathologies are interlinked through similarities in underlying biochemistry, which can helpfully be categorized for clinicians into immunomodulatory/inflammatory and metabolic domains. Understanding these biochemical connections highlights how disruptions in metabolic integrity and chronic inflammation contribute to degenerative health. For example, similar inflammatory profiles are observed in both sarcopenia (Assyov et al. 2023) and COPD (Huang et al. 2021). Chronic inflammatory responses, whether due to repeated exposure to allergens, infection, disrupted feedback signalling, acute injury, or chronic injurious behaviour

such as sedentary lifestyles (Henson et al. 2013), are particularly problematic. Vitamin D has been shown to modulate the inflammasome (Orrù et al. 2020), suggesting a meaningful connection to muscular and respiratory pathologies through immunoinflammatory pathways.

From a metabolic perspective, disruptions in endocrinological signalling, particularly insulin, can result in impaired energy utilisation, affecting all body tissues (Rahman et al. 2021). Vitamin D influences insulin through both genomic and nongenomic mechanisms (Szymczak-Pajor & Śliwińska 2019), and insulin system degradation is independently associated with vitamin D deficiency (Sung et al. 2012). Impaired insulin function is linked to conditions that also correlate with sarcopenia and COPD risk, such as fatty liver disease and reduced muscle strength (Fraser et al. 2022; Viglino et al. 2017). Fundamentally, metabolic dysfunction overlaps with immunomodulatory and inflammatory processes, underscoring the potentially pivotal role dysregulated vitamin D metabolism has in the pathogenesis of comorbid degenerative diseases (Han et al. 2020).

The examples provided above do not suggest that all chronic and comorbid conditions can be straightforwardly remedied by vitamin D repletion, as clinically important as that is, nor do they imply a specific causal sequence in the complex cascade of compounding pathologies. Rather, they illustrate that degenerative health conditions share common biochemical threads and pathological origins rooted in the breakdown of immunoinflammatory modulation and metabolic regulation, which may follow the path determined by that individual's weakest susceptibilities. Vitamin D's relevance in these scenarios is consistent, highlighting the fundamental need to eradicate vitamin D deficiency, particularly in individuals at risk for, or approaching, degenerative health age.

9.2 The complexity of vitamin D in degenerative healthcare settings

Degenerative health concerns are evolving matters in healthcare, with conditions such as sarcopenic obesity (Ji, Li & Ma 2022) becoming an increasing concern—with this condition exemplifying the complexity of vitamin D in a 21st-century context. While addressing rickets once offered a straightforward public health narrative for vitamin D (Rajakumar 2003), modern issues like the obesity epidemic (Temple 2022), widespread sedentary behaviour and increased indoor time (Nakshine et al. 2022) create cycles of pathology for which vitamin D plays a

crucial, but not solitary, role. The ‘vicious cycle’ here is a paradox between lifestyle and biology. Sedentary behaviour and indoor time, which perpetuate sarcopenia (Hämäläinen et al. 2024; Mo et al. 2023) and which are further promoted in its presence (Liu et al. 2014), are linked to increasing obesity and vitamin D deficiency (Solis-Urra et al. 2019). Additionally, obesity itself increases the risk of vitamin D deficiency (Vranić, Mikolašević & Milić 2019), while vitamin D deficiency initially drives sarcopenia. This dysfunctional and cyclical interplay accelerates and perpetuates the primary pathology.

While conditions like rickets were historically also characterized by these kinds of paradoxes [i.e., affected children being less able to engage outdoor play and other activities], such developmental issues were/can be remediated through correction of nutritional intake and corrective devices (Rajakumar 2003). Degenerative diseases, however, occurring in adults, have multiple challenges to overcome, such as potential disease-promoting lifestyle and cultural habits established over a lifetime, in conjunction with relatively lessened overall metabolic health. It is from this point of disadvantage that clinicians and public health proponents are faced with an ever growing and compounding challenge regarding degenerative health conditions in healthcare.

The degenerative health pandemic facing older peoples can be characterized by a complex of pathologized molecular interplay (Cannavo 2023), which underpins why one degenerative health concern can be seen as a risk factor for another and vice versa. A consistent picture is painted throughout this thesis of vitamin D being both at the core of maintaining health and servicing the minimum reserves necessary to stave off pathogenesis across distinct [but ultimately interconnected] degenerative health conditions. Critically, this demonstrates that there is an optimal scenario of adequate vitamin D status supporting overall health and normal physiology, but this is easily disrupted by multiple environmental and lifestyle factors characteristic of contemporary times.

9.3 Considerations for clinical practice and public health

In its current form, the western healthcare landscape largely specialises out the assessment and treatment of chronic conditions across a range of medical specialists and allied health professionals (Britnell 2011). In such a system, a potential challenge for individuals with

degenerative health condition lies with the increasingly acknowledged ‘silo mentality’ (McCartney 2016), where condition-specific healthcare can become isolated across a range of distinct institutions and professionals. Given the importance of vitamin D in a range of degenerative health conditions, and given that the care of degenerative conditions is prone to segregated forms of assessment and intervention, the onus is put on those healthcare professionals that deal consistently at the centre of patient care [i.e., the General Practitioner and other primary care clinicians (Kang et al. 2020)] to ensure that where warranted, any potentiality for vitamin D deficiency has been investigated/addressed, so as to assure that this easily remediated factor is not unnecessarily hindering progress amongst broader rehabilitation efforts.

The vitamin D narrative put forward in this thesis—with the ‘vitamin D character’ being consistent throughout both health and disease—is perhaps a particularly helpful lens through which public health proponents, clinicians, and patients can view complex co-morbidity degenerative health issues. Education is a key focus in contemporary public healthcare, and given the same fundamental public health advice for supporting vitamin D health [i.e., physical activity, healthy outdoor exposure, nutrient rich eating habits], is synonymous with the advice given in general for preventing degenerative health progression (Di Renzo, Gualtieri & De Lorenzo 2021; Santiago & Potashkin 2023), this provides opportunity consolidate educational efforts. It is from this perspective that existing health advice, in conjunction with an expanded form of vitamin D education, could provide a more tangible and cohesive public health narrative.

The same vitamin D lens for clinicians could also provide an educational tool that helps patients to contextualise the steps needed to start undoing the vicious cycles of compounded degenerative health—that is ‘*promoting vitamin D health is promoting a healthier you!*’ Fundamentally, while there remains a need in public health to not venture into radicalization/over enthusiasm territory with vitamin D (Kilbane et al. 2014), the public could perhaps benefit from an understanding that vitamin D is more than just ‘the sunshine vitamin’ and, indeed, understand vitamin D to be a pro-hormone that is connected to healthy ageing in the same way that other forms of endocrinology are connected to and important for healthy ageing.

Given the evolving concerns of degenerative health, it is crucial to ensure public health programs prioritise addressing vitamin D deficiency, particularly for those approaching or already at risk of degenerative health conditions. This encompasses both the ageing and

disability sectors, and particularly includes rural and remote health, where degenerative health conditions can manifest at younger ages (Thomas, Wakerman & Humphreys 2015). The importance of maintaining adequate 25(OH)D levels has been underscored in this thesis, and the need for multifaceted approaches to achieve this has been highlighted. Firstly, consideration of vitamin D status in routine health assessments for at-risk populations should be mandatory where not already implemented, especially given that such individuals may be less likely to achieve healthy sun exposure. For these populations, it could therefore be logical to universally adopt public health measures demonstrated to work in low sun exposure environments—such as Finland's fortification program—and aim for individuals to meet a minimum 400 IU/day intake (Cashman & Kiely 2024), serving to minimally mitigate severe deficiency. Supplementation should be considered where necessary, especially for institutionalised individuals, those whose employment limits healthy sun exposure, and those not living in countries with mandatory food fortification. Programs should also ensure that individuals in institutional settings, such as care homes, have regular access to outdoor light if they are able, facilitating adequate sun exposure. This comprehensive approach aims to create a public health directive that integrates vitamin D maintenance into broader strategies for preventing and managing degenerative health conditions, ensuring a proactive stance in promoting overall health and longevity.

9.4 Concerns and health equity

There does, however, remain a need to balance the promise and potential of vitamin D in the degenerative health space with vitamin D related concerns and health equity issues raised in this project and elsewhere (Mazess, Bischoff-Ferrari & Dawson-Hughes 2021; Shaw 2002). In the time following this thesis's first study, the COVID-19 pandemic drew attention to health outcome disparities that largely manifested according to socioeconomic and ethnic background lines (Magesh et al. 2021; Zaccardi et al. 2023)—with these COVID-19 health outcome disparities, in this sense, running recognisably similarly to those vitamin D deficiency disparities raised in this thesis. Given the role vitamin D sufficiency likely played in supporting survival in the pandemic (Argano et al. 2023), findings documented in this thesis would posit that the health disparity findings [outlined in Study I] are now more relevant than ever, and can serve to inform public health proponents as they seek to ensure that the most vulnerable are better supported, in any such similar repeating of history.

Further consideration also must be given to the contemporary role of bolus dosing. The COVID-19 pandemic made extensive use of vitamin D bolus dosing; and the inclusion of bolus dosing within metanalyses that saw benefit from vitamin D on COVID-19 [particularly in the most deficient] ought to be acknowledged (D'Ecclesiis et al. 2022; Hosseini, El Abd & Ducharme 2022; Meng et al. 2023; Petrelli et al. 2023). However, with the safety concerns raised elsewhere (Mazess, Bischoff-Ferrari & Dawson-Hughes 2021) and explored in this thesis, it raises the question as to whether the 'improved compliancy argument' for bolus dosing (Kearns et al. 2015) remains appropriate for at least the elderly, who are most at risk for falls, and who—in the advent of increased use of individualised blister pack medication provision (Shah et al. 2021)—could now potentially otherwise be in receipt of compliance-assured daily dosing of vitamin D. Indeed, the recent Endocrine Society guideline update has specified daily intake of 800 IU vitamin D3 for those over age 75, from food or supplements (Demay et al. 2024).

9.5 Considerations for future research

The research undertaken in this thesis contributes to the justification of several future research directions, spanning potential for clinical trials and molecular investigations [including possible animal modelling] aiming to elucidate underlying biochemical pathways, specifically pertaining to sarcopenic obesity, and steroid hormone competition and its effects on the motor function regulating areas of the nervous system. Potentially, the most straightforward direction for future research would be investigating the proposed mechanisms for the findings presented in Study II [Sarcopenia], where a potential shift in metabolic demand [driven by obesity-related inflammatory factors] was put forward as a possible reason for why the beneficial effect of 25(OH)D on reducing sarcopenia risk appeared to diminish in the presence of obesity. Factors proposed to metabolically shift 25(OH)D utilisation were the increased inflammatory responses common to obesity, and to which vitamin D is proposed to play immunomodulatory effects against, specifically interleukin 6 (Silberstein 2021).

Fundamentally, there is an axis between sarcopenia, obesity, vitamin D, and interleukin 6, in that adipose tissue is one of the main sources of interleukin 6 (Eder et al. 2009), interleukin 6 is known to contribute to the pathogenesis of sarcopenia (Dalbeni et al. 2023), and interleukin 6 is known to be modulated by vitamin D (Silberstein 2020). Future vitamin D-sarcopenia intervention trials could incorporate interleukin 6 measurements, alongside anthropometric

measures, to gain insight into the interplay between these factors. Interleukin 6 is a generally accessible laboratory test, and several recent advancements (Pan et al. 2022)—including a recent Australian developed lateral flow assay (Rahbar et al. 2021)—have produced considerably more affordable options for including assessment of interleukin 6 in clinical trials.

Given that vitamin D and sarcopenia have garnered increased attention among clinical trialists in recent years, with a noted need for more targeted obesity focused studies (Di Filippo et al. 2022), it is posited that emerging projects focused in this area could incorporate, with relative straightforwardness, combination therapy arms that monitor interleukin 6 and partner anti-inflammatory agents along with a vitamin D intervention, in an effort to determine if benefits from vitamin D are extended into sarcopenic-obese subjects [relative to vitamin D alone]. In exploring potential synergies with vitamin D, several anti-inflammatory nutraceutical interventions stand out, including omega-3 fatty acids (Li et al. 2014), resveratrol (Omraninava et al. 2021), curcumin (Derosa et al. 2016), and epigallocatechin 3-gallate [i.e., green tea extract] (Ahmed 2010). However, questions arise about their efficacy in obese individuals, requiring further investigation into optimal dosing levels for co-administration. Interestingly, recent clinical data draw attention to similarities between vitamin D and resveratrol as anti-inflammatory therapies, particularly in the context of SARS-CoV-2 infection, highlighting their potential synergistic effects (Russo et al. 2023). Indeed, preliminary studies suggest promising outcomes from combining vitamin D with a phytochemical blend containing resveratrol in models of macular degeneration (Hernandez et al. 2021).

Importantly, elevated interleukin 6 levels in severe COVID-19 cases underscore the relevance of these nutraceutical agents in addressing degenerative issues associated with obesity (Coomes & Haghbayan 2020), however, it is with curcumin that growing evidence supports a specific interleukin 6 modulating, anti-inflammatory role specifically in obese individuals, with potential benefits from curcumin intervention seen for non-alcoholic fatty liver disease and sarcopenia (Jazayeri-Tehrani et al. 2019; Joo & Kim 2023; Saud Gany et al. 2023; Shimizu et al. 2019). Potentially relevant for sarcopenic obesity—where increased inflammation may shift metabolic demand of vitamin D—human and animal studies collectively indicate that curcumin may enhance vitamin D's skeletal protective effects and in doing so potentially spare the vitamin’s use and optimise its metabolic demand (Patnaik et al. 2023; Xin et al. 2015).

Encouragingly, early investigations into concurrent curcumin and vitamin D therapy also show promise in mitigating inflammatory pathways associated with osteoarthritis (Patnaik et al. 2023). These findings suggest that combining vitamin D with curcumin and/or resveratrol could offer a promising therapeutic approach for sarcopenic-obese individuals, leveraging their shared anti-inflammatory properties and biochemical mechanisms while potentially bypassing the complex regulatory systems affecting fat-soluble hormones.

The next path forward for research would be further exploring the higher vitamin D, inflammation, and increased fall risk relationship, with a particular focus put on the potential for a steroid hormone competition hypothesis, to which indirect support has been given from the work undertaken in this thesis. In essence, fundamental prerequisite groundwork for an animal model driven approach to explore steroid hormone competition has already been developed in way of the vitamins-A-D-antagonism work done in turkey poult (Metz, Walser & Olson 1985) and rodent studies (Rohde & DeLuca 2005; Rohde et al. 1999). There is a limitation, however, in that the turkey poult study appears to be the only study to use a reciprocal design [i.e., wherein the antagonism effect was explicitly studied in either direction], while the rat studies have primarily looked at vitamin A antagonizing vitamin D, with a focus on calcium regulation. Experimentation in mice has indirectly demonstrated the antagonism effect in action (Hetland et al. 2009), however, the focus again has been on the antagonizing effect of vitamin A. The literature does not appear to explicitly offer a vitamin-A-D-antagonism blueprint for rodents, in the same way that exists for turkey poult. Establishing this experimental blueprint in rodents would constitute an ideal first step in exploring steroid hormone competition, as the mice model sector already offers what would be an advantageous comparison group in way of BRINP knock-out mice models, (Berkowicz et al. 2016 a,b; Kobayashi et al. 2014), which result in mice with motor function dysregulation due to the deletion of a retinoic-acid inducible gene.

Assuming a vitamin-A-D-antagonism blueprint in a mice model can be established [i.e., one that explicitly focuses on motor function outcomes] then comparing the behavioural characteristics of wild-type mice [with and without retinol deprivation, and with and without high vitamin D exposure] to the BRINP knock-out mice, could be a logical place to begin elucidating the steroid hormone antagonistic interplay effects in a controlled environment. In terms of establishing a vitamin-A-D-antagonism blueprint in a mice model, there is precedent to

assume that such attempts would yield the expected results, as the effects of vitamin A deficiency has, at least, been explored in a mice model resulting in motor function impairment (Carta et al. 2006). Interestingly, Carta and colleagues also reported hyperactivity in mice starved of vitamin A and, although the authors did not associate this with BRINP, it would appear to be a logical connection, given the subsequent understating that emerged in the BRINP knockout space in the years following their experiment [i.e., wherein it was demonstrated that knockout of a retinoic acid inducible, motor function regulating gene could produce motor function dysregulation and hyperactivity (Berkowicz et al. 2016 a,b)].

Given that motor function issues have already been demonstrated in vitamin A starved mice (Carta et al. 2006), it's posited here that the induction of motor function dysregulation would likely be intensified in the presence of high [but not acutely toxic] vitamin D status. Exploring this assumption is a logical first step in investigating the steroid hormone competition hypothesis in an animal setting. BRINP 1 m-RNA extraction could also be incorporated into the initial study to assess the extent to which genetic expression is affected by vitamin A restrictions/vitamin D interventions.

9.6 Strengths and limitations

The research documented in this thesis leverages on its breadth and depth, encompassing diverse methodologies to provide a nuanced understanding of the relationship between vitamin D and important degenerative health conditions. By incorporating observational analyses (Mann 2003), MR (Thanassoulis & O'Donnell 2009), and broader forms of genetic epidemiology (Sahebi et al. 2013), associations are identified, but also potential causal pathways have been explored and put forward for consideration. The inclusion of a large and diverse sample from the UK Biobank (Bycroft et al. 2018) strengthens the statistical power of analyses undertaken in this thesis and has also enabled exploration of ethnic disparities in vitamin D health outcomes. Our inclusion of conventional linear MR methodology represents a pivotal strength of this thesis, in way of bringing deeper insight to respiratory and muscular related health and pathology, while allowing for considerations therein that may have been challenging to elucidate in traditional observational studies. Research conducted using the genetic epidemiological approach, documented in this thesis, not only helps to provide a deeper understanding of the mechanisms underlying observed associations, but does so in a manner where conventional RCTs might have

otherwise been impractical or unethical (Frame et al. 2018; Vanclay, Baines & Taylor 2013). Furthermore, by keeping complex biological interactions at the forefront of consideration and discussion, this thesis contributes to a richer understanding of how vitamin D influences health outcomes, thereby supporting further research in this area.

Despite these strengths, this research is not without limitations. Foremost is the restriction to 25(OH)D data in the UK Biobank, with no availability of serum calcitriol data. While 25(OH)D is the generally recognised indicator of vitamin D status, deeper levels of inference—and insights on specific mechanisms—would be possible from analyses using broader vitamin D metabolites (Amrein et al. 2020). While the inclusion of diverse ethnic groups enhances the external validity of the phenotypic study on deficiency determinants, the restricting of analyses to individuals of white-European descent [to facilitate the genetic approach in the UK Biobank] may limit the generalisability of findings to other populations—and considerations therefore are warranted when extrapolating findings from this thesis to more ethnically diverse cohorts. The UK Biobank's low participation response rate may introduce selection biases, potentially impacting the external validity of our findings, while healthy volunteer bias may present confounding issues beyond those readily considered (Watts 2007). Despite employing robust statistical methods to mitigate these biases, residual confounding and breaching of core MR assumptions cannot be entirely ruled out (Davies, Holmes & Davey Smith 2018). While we employed sensitivity analyses and robust methods to address these assumptions, biases may still be introduced, compromising the validity of our estimates. Availability of data was another limitation, with the most notable being the absence of serum retinol data. The concurrent absence of broader inflammatory markers [such as interleukin 6], further limited the granularity to which the metabolic interplay in these analyses could be viewed, and increased reliance on inferring the possible underlying mechanisms.

As is possible when engaging any cutting-edge methodology, the evolving matters of nonlinear MR have provided challenges, which have been noted and now further contextualized elsewhere in this project [chapter 7 supplementary and addendum], and any extrapolation of nonlinear MR results need to be considered accordingly (Burgess, Wood & Butterworth 2023; Tian et al. 2023). While MR offers valuable insights, it cannot fully replicate the controlled conditions of clinical trials. Despite inherent limitations, our comprehensive methodological

approach and innovative techniques offer valuable insights that can inform future research endeavours and public health interventions aimed at optimising vitamin D status and improving health outcomes.

Chapter Ten: Conclusions

In conclusion, the research documented in this thesis presents observational and genetic evidence in support of a relationship between 25(OH)D and degenerative health related factors, including grip strength, skeletal muscle composition, and sarcopenia—while further insight was provided in way of a relationship between 25(OH)D and mortality, specifically via respiratory-related mortality, which held a robust linear MR association. The studies documented in this thesis also demonstrated alarming rates of deficiency in the UK population, a situation made more complex by socioeconomic deprivation and cross ethnic factors. Important concerns surrounding high dose vitamin D supplementation were also explored, helping to support the notion that supplementation strategies ought to consider the vitamin not in isolation, but rather as part of a broader steroid hormone system. Future studies should aim to explore the vitamin D, inflammation, and motor function relationship—and the proposed steroid hormone competition mechanisms therein—via methods better suited to explore genetic expression. Meanwhile, insights from this thesis collectively suggest that public health proponents consider prioritising vitamin D repletion in the degenerative health and healthy ageing space.

The research documented in this thesis reveals the multifaceted role of vitamin D in degenerative health, while underscoring the importance of adopting a holistic approach that considers the complex interplay between vitamin D, genetic factors, environmental influences, and broader health determinants in shaping health outcomes across the lifespan. Through continued research efforts and evidence-based interventions, we can strive towards promoting optimal vitamin D status and, in doing so, help mitigate the burden of degenerative health conditions, ultimately improving the quality of life of individuals and populations worldwide.

References:

Abiri, B & Vafa, M 2020, 'Vitamin D and muscle sarcopenia in aging', in P Guest (ed), *Clinical and Preclinical Models for Maximizing Healthspan: Methods and Protocols*, Humana, New York (NY), USA, pp. 29–47.

Acar, S & Özkan, B 2021, 'Vitamin D metabolism', *Vitamin D*, vol. 2021, pp. 1–23.

Addae, C, Yi, X, Gernapudi, R, Cheng, H, Musto, A & Martinez-Ceballos, E 2012, 'All-trans-retinoid acid induces the differentiation of encapsulated mouse embryonic stem cells into GABAergic neurons', *Differentiation*, vol. 83, no. 5, pp. 233–241.

Afzal, S, Brøndum-Jacobsen, P, Bojesen, SE & Nordestgaard, BG 2014, 'Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts', *BMJ*, vol. 349, p. g6330.

Ahmed, S 2010, 'Green tea polyphenol epigallocatechin 3-gallate in arthritis: progress and promise', *Arthritis Research & Therapy*, vol. 12, no. 2, p. 208.

Al-Nusaif, M, Yang, Y, Li, S, Cheng, C & Le, W 2022, 'The role of NURR1 in metabolic abnormalities of Parkinson's disease', *Molecular Neurodegeneration*, vol. 17, no. 1, p. 46.

Alia, E & Kerr, PE 2021, 'Vitamin D: Skin, sunshine, and beyond', *Clinics in dermatology*, vol. 39, no. 5, pp. 840–846.

Allen, NE, Arnold, M, Parish, S, Hill, M, Sheard, S, Callen, H, Fry, D, Moffat, S, Gordon, M, Welsh, S, Elliott, P & Collins, R 2020, 'Approaches to minimising the epidemiological impact of sources of systematic and random variation that may affect biochemistry assay data in UK Biobank', *Wellcome Open Research*, vol. 5, p. 222.

Alonso, N, Zelzer, S, Eibinger, G & Herrmann, M 2023, 'Vitamin D metabolites: analytical challenges and clinical relevance', *Calcified Tissue International*, vol. 112, no. 2, pp. 158–177.

Amiri, S, Sadeghi-Bazargani, H, Nazari, S, Ranjbar, F & Abdi, S 2017, 'Attention deficit/hyperactivity disorder and risk of injuries: a systematic review and meta-analysis', *Journal of injury and violence research*, vol. 9, no. 2, p. 95.

Amrein, K, Schnedl, C, Holl, A, Riedl, R, Christopher, KB, Pachler, C, Urbanic Purkart, T, Waltensdorfer, A, Münch, A, Warnkross, H, Stojakovic, T, Bisping, E, Toller, W, Smolle, K-H, Berghold, A, Pieber, TR & Dobnig, H 2014, 'Effect of High-Dose Vitamin D3 on Hospital Length of Stay in Critically Ill Patients With Vitamin D Deficiency', *JAMA*, vol. 312, no. 15, p. 1520.

Amrein, K, Scherkl, M, Hoffmann, M, Neuwersch-Sommeregger, S, Köstenberger, M, Tmava Berisha, A, Martucci, G, Pilz, S & Malle, O 2020, 'Vitamin D deficiency 2.0: an update on the current status worldwide', *European Journal of Clinical Nutrition*, vol. 74, no. 11, pp. 1498–1513.

Ao, T, Kikuta, J & Ishii, M 2021, 'The Effects of Vitamin D on Immune System and Inflammatory Diseases', *Biomolecules*, vol. 11, no. 11, p. 1624.

Argano, C, Mallaci Bocchio, R, Natoli, G, Scibetta, S, Lo Monaco, M & Corrao, S 2023, 'Protective Effect of Vitamin D Supplementation on COVID-19-Related Intensive Care Hospitalization and Mortality: Definitive Evidence from Meta-Analysis and Trial Sequential Analysis', *Pharmaceuticals (Basel)*, vol. 16, no. 1, p. 130.

Aspelund, T, Grübler, MR, Smith, AV, Gudmundsson, EF, Keppel, M, Cotch, MF, Harris, TB, Jorde, R, Grimnes, G, Joakimsen, R, Schirmer, H, Wilsgaard, T, Mathiesen, EB, Njølstad, I, Løchen, M-L, März, W, Kleber, ME, Tomaschitz, A, Grove-Laugesen, D, Rejnmark, L, Swart, KMA, Brouwer, IA, Lips, P, Van Schoor, NM, Sempos, CT, Durazo-Arvizu, RA, Škrabáková, Z, Dowling, KG, Cashman, KD, Kiely, M, Pilz, S, Gudnason, V & Eiriksdottir, G 2019, 'Effect of Genetically Low 25-Hydroxyvitamin D on Mortality Risk: Mendelian Randomization Analysis in 3 Large European Cohorts', *Nutrients*, vol. 11, no. 1, p. 74.

Assyov, Y, Ganeva, I, Ikonomov, S, Nedeva, I, Velikov, T, Kamenov, Z & Velikova, T 2023, ' Interleukin-6: Unravelling its role in sarcopenia pathogenesis and exploring therapeutic avenues', *Pharmacia*, vol. 70, no. 4, pp. 1493–1498.

Avenell, A, Mak, JC & O'Connell, D 2014, 'Vitamin D and vitamin D analogues for preventing fractures in post-menopausal women and older men', *Cochrane Database of Systematic Reviews*, vol. 2014, no. 4, p. CD000227.

Balvers, MG, Brouwer-Brolsma, EM, Endenburg, S, de Groot, LC, Kok, FJ & Gunnewiek, JK 2015, 'Recommended intakes of vitamin D to optimise health, associated circulating 25-hydroxyvitamin D concentrations, and dosing regimens to treat deficiency: workshop report and overview of current literature', *Journal of Nutritional Science*, vol. 4, p. e23.

Batman, A, Saygili, ES, Yildiz, D, Sen, EC, Erol, RS, Canat, MM, Ozturk, FY & Altuntas, Y 2021, 'Risk of hypercalcemia in patients with very high serum 25-OH vitamin D levels', *International Journal of Clinical Practice*, vol. 75, no. 7, p. e14181.

Begum, S & Mirghani, HO 2024, 'The Role of Vitamin D in COVID-19 Survival and Prevention: A Meta-analysis', *Sudan Journal of Medical Sciences (SJMS)*, vol. 19, no. 1, pp. 72–83.

Benedik, E 2021, 'Sources of vitamin D for humans', *International Journal for Vitamin and Nutrition Research*, vol. 92, no. 2, pp. 118–125.

Benz, E, Trajanoska, K, Lahousse, L, Schoufour, JD, Terzikhan, N, De Roos, E, de Jonge, GB, Williams, R, Franco, OH & Brusselle, G 2019, 'Sarcopenia in COPD: a systematic review and meta-analysis', *European Respiratory Review*, vol. 28, no. 154.

Berkowicz, SR, Featherby, TJ, Qu, Z, Giousoh, A, Borg, NA, Heng, JI, Whisstock, JC & Bird, PI 2016, '*Brinp1*$^{-/-}$ mice exhibit autism-like behaviour, altered memory, hyperactivity and increased parvalbumin-positive cortical interneuron density', *Molecular Autism*, vol. 7, no. 1, p. 22.

Berkowicz, SR, Featherby, Whisstock, JC & Bird, PI 2016, ' Mice lacking *Brinp2* or *Brinp3*, or both, exhibit behaviors consistent with neurodevelopmental disorders', *Molecular Autism*, vol. 7, no. 1, p. 22.

Berman, BD, Pollard, RT, Shelton, E, Karki, R, Smith-Jones, PM & Miao, Y 2018, 'GABAA receptor availability changes underlie symptoms in isolated cervical dystonia', *Frontiers in Neurology*, vol. 9, p. 188.

Berry, DJ, Vimaleswaran, KS, Whittaker, JC, Hingorani, AD & Hyppönen, E 2012, 'Evaluation of genetic markers as instruments for Mendelian randomization studies on vitamin D', *PLoS One*, vol. 7, no. 5, p. e37465.

Berry, DJ, Hyppönen, E & Cortina-Borja, M 2013, 'Investigating the association of vitamin D seasonality on inflammatory and hemostatic markers', *Chronobiology International*, vol. 30, no. 6, pp. 786–95.

Bi, CS, Li, X, Qu, HL, Sun, LJ, An, Y, Hong, YL, Tian, BM & Chen, FM 2020, 'Calcitriol inhibits osteoclastogenesis in an inflammatory environment by changing the proportion and function of T helper cell subsets (Th2/Th17)', Cell Proliferation, vol. 53, no. 6, p. e12827.

Bikle, DD & Schwartz, J 2019, 'Vitamin D Binding Protein, Total and Free Vitamin D Levels in Different Physiological and Pathophysiological Conditions', *Frontiers in Endocrinology*, vol. 10, p. 317.

Bikle, DD 2020a, 'Vitamin D: Newer Concepts of Its Metabolism and Function at the Basic and Clinical Level', *Journal of the Endocrine Society*, vol. 4, no. 2, p. bvz038.

Bikle, DD 2020b, 'The Free Hormone Hypothesis: When, Why, and How to Measure the Free Hormone Levels to Assess Vitamin D, Thyroid, Sex Hormone, and Cortisol Status', *JBMR Plus*, vol. 5, no. 1, p. e10418.

Binkley, N, Krueger, DC, Morgan, S & Wiebe, D 2010, 'Current status of clinical 25-hydroxyvitamin D measurement: an assessment of between-laboratory agreement', *JClinica Chimica Acta*, vol. 411, no. 23–24, pp. 1976–1982.

Bischoff-Ferrari, HA, Dawson-Hughes, B, Orav, EJ, Staehelin, HB, Meyer, OW, Theiler, R, Dick, W, Willett, WC & Egli, A 2016, 'Monthly high-dose vitamin D treatment for the prevention of functional decline: a randomized clinical trial', *JAMA Internal Medicine*, vol. 176, no. 2, pp. 175–183.

Bjelakovic, G, Gluud, LL, Nikolova, D, Whitfield, K, Wetterslev, J, Simonetti, RG, Bjelakovic, M & Gluud, C 2014, 'Vitamin D supplementation for prevention of mortality in adults', *Cochrane Database of Systematic Reviews*, vol. 10, no. 1, p. CD007470.

Bolland, MJ, Grey, A, Gamble, GD & Reid, IR 2014, 'The effect of vitamin D supplementation on skeletal, vascular, or cancer outcomes: a trial sequential meta-analysis', *The Lancet Diabetes & Endocrinology*, vol. 2, no. 4, pp. 307–320.

Bolland, MJ, Grey, A & Avenell, A 2018, 'Effects of vitamin D supplementation on musculoskeletal health: a systematic review, meta-analysis, and trial sequential analysis', *The Lancet Diabetes & Endocrinology*, vol. 6, no. 11, pp. 847–858.

Bolland, MJ, Grey, A & Avenell, A 2018, 'Assessment of research waste part 2: wrong study populations- an exemplar of baseline vitamin D status of participants in trials of vitamin D supplementation', *BMC Medical Research Methodology*, vol. 18, no. 1, p. 103.

Bollen, SE & Atherton, PJ 2021, 'Myogenic, genomic and non-genomic influences of the vitamin D axis in skeletal muscle', *Cell Biochemistry Function*, vol. 39, no. 1, pp. 48–59.

Bone, AE, Hepgul, N, Kon, S & Maddocks, M 2017, 'Sarcopenia and frailty in chronic respiratory disease: Lessons from gerontology', *Chronic respiratory disease*, vol. 14, no. 1, pp. 85–99.

Bornstedt, ME, Gjerlaugsen, N, Pepaj, M, Bredahl, MKL & Thorsby, PM 2019, 'Vitamin D Increases Glucose Stimulated Insulin Secretion from Insulin Producing Beta Cells (INS1E)', *International Journal of Endocrinolology and Metabolism*, vol. 17, no. 1, Jan, p. e74255.

Bouillon, R, Carmeliet, G, Verlinden, L, van Etten, E, Verstuyf, A, Luderer, HF, Lieben, L, Mathieu, C & Demay, M 2008, 'Vitamin D and human health: lessons from vitamin D receptor null mice', *Endocrine Reviews*, vol. 29, no. 6, pp. 726–776.

Bouillon, R, Manousaki, D, Rosen, C, Trajanoska, K, Rivadeneira, F & Richards, JB 2022, 'The health effects of vitamin D supplementation: evidence from human studies', *Nature Reviews Endocrinology*, vol. 18, no. 2, pp. 96–110.

Bowden, J, Davey Smith, G & Burgess, S 2015, 'Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression', *International Journal of Epidemiology*, vol. 44, no. 2, pp. 512–525.

Bowden, J, Davey Smith, G, Haycock, PC & Burgess, S 2016, 'Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator', *Genetic Epidemiology*, vol. 40, no. 4, pp. 304–314.

Britnell, M 2011, 'The role of the 'specialist' in healthcare', *Clinical Medicine Journal*, vol. 11, no. 4, pp. 329–331.

Britton, GB, O'Bryant, SE, Johnson, LA, Hall, JR, Villarreal, AE, Oviedo, DC, Pérez-Lao, AR & Carreira, MB 2019, 'Inflammatory Biomarkers, Depressive Symptoms and Falls Among the elderly in Panama', *Current Aging Science*, vol. 11, no. 4, pp. 236–241.

Brożyna, AA, Jozwicki, W, Janjetovic, Z & Slominski, AT 2011, 'Expression of vitamin D receptor decreases during progression of pigmented skin lesions', *Human Pathology*, vol. 42, no. 5, pp. 618–631.

Brunner, J 2016, *Data analysis with SAS: An Open Textbook*, Department of Statistical Sciences, University of Toronto, Toronto, Canada.

Buniello, A, Jacqueline, A, Cerezo, M, Harris, LW, Hayhurst, J, Malangone, C, McMahon, A, Morales, J, Mountjoy, E, Sollis, E, Suveges, D, Vrousgou, O, Whetzel, PL, Amode, R, Guillen, JA, Riat, HS, Trevanion, SJ, Hall, P, Junkins, H, Flicek, P, Burdett, T, Hindorff, LA, Cunningham, F & Parkinson, H 2019, 'The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019', *Nucleic Acids Research,* vol. 47, no. D1, pp. D1005–D1012.

Burgess, S, Butterworth, A & Thompson, SG 2013, 'Mendelian randomization analysis with multiple genetic variants using summarized data', *Genetic Epidemiology,* vol. 37, no. 7, pp. 658–665.

Burgess, S, Davey Smith, G, Davies, NM, Dudbridge, F, Gill, D, Glymour, MM, Hartwig, FP, Holmes, MV, Minelli, C, Relton, CL & Theodoratou, E 2020, 'Guidelines for performing Mendelian randomization investigations', *Wellcome Open Research,* vol. 4, p. 186.

Burgess, S, Wood, AM & Butterworth, AS 2023, 'Mendelian randomisation and vitamin D: the importance of model assumptions–Authors' reply', *The Lancet Diabetes & Endocrinology,* vol. 11, no. 1, pp. 15–16.

Butera, A, Smirnova, L, Ferrando-May, E, Hartung, T, Brunner, T, Leist, M & Amelio, I 2023, 'Deconvoluting gene and environment interactions to develop an "epigenetic score meter" of disease', *EMBO Molecular Medicine,* vol. 15, no. 9, p. e18208.

Bycroft, C, Freeman, C, Petkova, D, Band, G, T. Elliott, L, Sharp, K, Motyer, A, Damjan, V, Delaneau, O, O'Connell, J, Cortes, A, Welsh, S, McVean, G, Leslie, S, Donnelly, P & Marchini, J 2017, 'Genome-wide genetic data on ~500,000 UK Biobank participants', *bioRxiv,* vol. Preprint available at https://www.biorxiv.org/content/10.1101/166298v1.full.pdf.

Bycroft, C, Freeman, C, Petkova, D, Band, G, Elliott, LT, Sharp, K, Motyer, A, Vukcevic, D, Delaneau, O & O'Connell, J 2018, 'The UK Biobank resource with deep phenotyping and genomic data', *Nature,* vol. 562, no. 7726, pp. 203–209.

Caire-Juvera, G, Ritenbaugh, C, Wactawski-Wende, J, Snetselaar, LG & Chen, Z 2009, 'Vitamin A and retinol intakes and the risk of fractures among participants of the Women's Health Initiative Observational Study', *The American Journal of Clinical Nutrition,* vol. 89, no. 1, pp. 323–330.

Cairncross, CT, Stonehouse, W, Conlon, CA, Grant, CC, McDonald, B, Houghton, LA, Eyles, D, Camargo, CA, Coad, J & Von Hurst, PR 2017, 'Predictors of vitamin D status in New Zealand preschool children', *Maternal & Child Nutrition,* vol. 13, no. 3, p. e12340.

Camargo, CA, Sluyter, J, Stewart, AW, Khaw, KT, Lawes, CMM, Toop, L, Waayer, D & Scragg, R 2020, 'Effect of monthly high-dose vitamin D supplementation on acute respiratory infections in older adults: a randomized controlled trial', *Clinical Infectious Diseases*, vol. 71, no. 2, pp. 311–317, doi: 10.1093/cid/ciz801. PMID: 31420647.

Camargo Jr, CA, Toop, L, Sluyter, J, Lawes, CM, Waayer, D, Khaw, K-T, Martineau, AR & Scragg, R 2021, 'Effect of monthly vitamin D supplementation on preventing exacerbations of asthma or chronic obstructive pulmonary disease in older adults: post hoc analysis of a randomized controlled trial', *Nutrients,* vol. 13, no. 2, p. 521.

Cannavo, A 2023, 'Molecular Mechanisms Underlying Chronic and Degenerative Diseases', *International Journal of Molecular Sciences,* vol. 24, no. 15, Aug 7, p. 12507.

Canpolat, U, Özcan, F, Özeke, Ö, Turak, O, Yayla, Ç, Açıkgöz, SK, Cay, S, Topaloğlu, S, Aras, D & Aydoğdu, S 2015, 'Impaired cardiac autonomic functions in apparently healthy subjects with vitamin D deficiency', *Annals of Noninvasive Electrocardiology,* vol. 20, no. 4, pp. 378–385.

Cao, M, He, C, Gong, M, Wu, S & He, J 2023, 'The effects of vitamin D on all-cause mortality in different diseases: an evidence-map and umbrella review of 116 randomized controlled trials', *Frontiers in Nutrition,* vol. 10, p. 1132528.

Carlberg, C & Saurat, J-H 1996, 'Vitamin D-retinoid association: molecular basis and clinical applications', *The Journal of Investigative Dermatology. Symposium Proceedings*, vol. 1, no. 1, pp. 82–86.

Carlberg, C 2014, 'Genome-wide (over)view on the actions of vitamin D', *Frontiers in Physiology*, vol. 5, p. 167.

Carlberg, C 2018, 'Vitamin D Genomics: From In Vitro to In Vivo', *Frontiers in Endocrinology*, vol. 9, p. 250.

Carlberg, C & Muñoz, A 2020, 'An update on vitamin D signaling and cancer', *Seminars in Cancer Biology*, vol. 79, pp. 217–230.

Carlberg, C 2022b, ' Vitamin D and Its Target Genes', *Nutrients*, vol. 14, no. 7, p. 1354.

Carlberg, C 2022a, 'Vitamin D in the Context of Evolution', *Nutrients*, vol. 14, no. 15, p. 3018.

Carlberg, C 2024, 'Vitamin D, chromatin, and epigenetics', in M Hewison et al (eds), *Feldman and Pike's Vitamin D (Fifth Edition)*, Academic Press, Cambridge (MA). USA, pp. 229-239.

Carta, M, Stancampiano, R, Tronci, E, Collu, M, Usiello, A, Morelli, M & Fadda, F 2006, 'Vitamin A deficiency induces motor impairments and striatal cholinergic dysfunction in rats', *Neuroscience*, vol. 139, no. 4, pp. 1163–1172.

Cashman, KD, O'Sullivan, SM, Galvin, K & Ryan, M 2020, 'Contribution of Vitamin D2 and D3 and Their Respective 25-Hydroxy Metabolites to the Total Vitamin D Content of Beef and Lamb', *Current Developments in Nutrition*, vol. 4, no. 7, p. nzaa112.

Cashman, KD & Kiely, M 2024, 'Chapter 58 - Vitamin D and food fortification', in M Hewison et al (eds), *Feldman and Pike's Vitamin D (Fifth Edition)*, Academic Press, Cambridge (MA), USA, pp. 135–160.

Castro, HM & Ferreira, JC 2023, 'Linear and logistic regression models: when to use and how to interpret them?', *Jornal Brasileiro de Pneumologia*, vol. 48, no. 6, p. e20220439.

Català-Moll, F, Ferreté-Bonastre, AG, Godoy-Tena, G, Morante-Palacios, O, Ciudad, L, Barberà, L, Fondelli, F, Martínez-Cáceres, EM, Rodríguez-Ubreva, J, Li, T & Ballestar, E 2022, 'Vitamin D receptor, STAT3, and TET2 cooperate to establish tolerogenesis', *Cell Reports*, vol. 38, no. 3, p. 110244.

Cataldi, S, Stanley, AT, Miniaci, MC & Sulzer, D 2022, 'Interpreting the role of the striatum during multiple phases of motor learning', *The FEBS Journal*, vol. 289, no. 8, pp. 2263–2281.

Cavalier, E, Huyghebaert, L, Rousselle, O, Bekaert, AC, Kovacs, S, Vranken, L, Peeters, S, Le Goff, C & Ladang, A 2020, 'Simultaneous measurement of 25(OH)-vitamin D and 24,25(OH)2-vitamin D to define cut-offs for CYP24A1 mutation and vitamin D deficiency in a population of 1200 young subjects', *Clinical Chemistry and Laboratory Medicine*, vol. 58, no. 2, pp. 197–201.

Chairta, PP, Hadjisavvas, A, Georgiou, AN, Loizidou, MA, Yiangou, K, Demetriou, CA, Christou, YP, Pantziaris, M, Michailidou, K & Zamba-Papanicolaou, E 2021, 'Prediction of Parkinson's Disease Risk Based on Genetic Profile and Established Risk Factors', *Genes*, vol. 12, no. 8, p. 1278.

Chakraborti, CK 2011, 'Vitamin D as a promising anticancer agent', *Indian Journal of Pharmacology*, vol. 43, no. 2, p. 113.

Chapman, K, Medrano, G, Chaudhary, J & Hamra, F 2015, 'NRG1 and KITL signal downstream of retinoic acid in the germline to support soma-free syncytial growth of differentiating spermatogonia', *Cell Death Discovery*, vol. 1, no. 1, p. 15018.

Cheema, MR & Chaudhry, AY 2016, 'Quality-of-life indicators and falls due to vitamin D deficiency', *International Journal of General Medicine*, vol. 22, no. 9, pp. 21–25.

Chehsmazar, E, Zarrati, M, Yazdani, B, Razmpoosh, E, Hosseini, AF & Shidfar, F 2021, 'The effect of vitamin D supplementation on serum concentrations of dehydroepiandrosterone, paraoxonase 1, apolipoproteins, free fatty acid and insulin in vitamin D deficient obese and overweight individuals under a low-calorie diet program: a randomized controlled trial', *Nutrition & Food Science*, vol. 51, no. 5, pp. 765–780.

Cheng, JB, Motola, DL, Mangelsdorf, DJ & Russell, DW 2003, 'De-orphanization of cytochrome P450 2R1: a microsomal vitamin D 25-hydroxilase', *Journal of Biological Chemistry*, vol. 278, no. 39, pp. 38084–38093.

Cheng, S, So, WY, Zhang, D, Cheng, Q, Boucher, BJ & Leung, PS 2016, 'Calcitriol Reduces Hepatic Triglyceride Accumulation and Glucose Output Through Ca2+/CaMKKβ/AMPK Activation Under Insulin-Resistant Conditions in Type 2 Diabetes Mellitus', *Current Molecular Medicine*, vol. 16, no. 8, pp. 747–758.

Choi, SW, Mak, TS & O'Reilly, PF 2020, 'Tutorial: a guide to performing polygenic risk score analyses', *Nature Protocols*, vol. 15, no. 9, pp. 2759–2772.

Christakos, S, Li, S, DeLa Cruz, J, Verlinden, L & Carmeliet, G 2020, 'Vitamin D and Bone', in PH Stern (ed), *Bone Regulators and Osteoporosis Therapy. Handbook of Experimental Pharmacology*, 2019/12/04 edn, vol. 262, Springer, Cham, Switzerland, pp. 47–63.

Chun, RF, Hernandez, I, Pereira, R, Swinkles, L, Huijs, T, Zhou, R, Liu, NQ, Shieh, A, Guemes, M & Mallya, S 2016, 'Differential responses to vitamin D2 and vitamin D3 are associated with variations in free 25-hydroxyvitamin D', *Journal of Endocrinology*, vol. 157, no. 9, pp. 3420–3430.

Coomes, EA & Haghbayan, H 2020, 'Interleukin-6 in COVID-19: a systematic review and meta-analysis', *Reviews in Medical Virology*, vol. 30, no. 6, pp. 1–9.

Cooper, C, Harvey, NC, Bishop, NJ, Kennedy, S, Papageorghiou, AT, Schoenmakers, I, Fraser, R, Gandhi, SV, Carr, A, D'Angelo, S, Crozier, SR, Moon, RJ, Arden, NK, Dennison, EM, Godfrey, KM, Inskip, HM, Prentice, A, Mughal, MZ, Eastell, R, Reid, DM & Javaid, MK 2016, 'Maternal gestational vitamin D supplementation and offspring bone health (MAVIDOS): a multicentre, double-blind, randomised placebo-controlled trial', *The Lancet Diabetes & Endocrinology*, vol. 4, no. 5, pp. 393–402.

Courbebaisse, M & Cavalier, E 2020, 'Vitamin D in 2020: An old pro-hormone with potential effects beyond mineral metabolism', *Nutrients*, vol. 12, no. 11, p. 3378.

Cruz-Jentoft, AJ, Bahat, G, Bauer, J, Boirie, Y, Bruyère, O, Cederholm, T, Cooper, C, Landi, F, Rolland, Y, Sayer, AA, Schneider, SM, Sieber, CC, Topinkova, E, Vandewoude, M, Visser, M, Zamboni, M, Bautmans, I, Baeyens, J-P, Cesari, M, Cherubini, A, Kanis, J, Maggio, M, Martin, F, Michel, J-P, Pitkala, K, Reginster, J-Y, Rizzoli, R, Sánchez-Rodríguez, D & Schols, J 2019, 'Sarcopenia: revised European consensus on definition and diagnosis', *Age and Ageing*, vol. 48, no. 1, pp. 16–31.

Cummings, SR, Kiel, DP & Black, DM 2016, 'Vitamin D Supplementation and Increased Risk of Falling', *JAMA Internal Medicine*, vol. 176, no. 2, p. 171.

Currò, M, Ferlazzo, N, Costanzo, MG, Caccamo, D & Ientile, R 2020, 'Vitamin D status influences transcriptional levels of RANKL and inflammatory biomarkers which are associated with activation of PBMC', *Clinica Chimica Acta*, vol. 507, pp. 219–223.

D'Ecclesiis, O, Gavioli, C, Martinoli, C, Raimondi, S, Chiocca, S, Miccolo, C, Bossi, P, Cortinovis, D, Chiaradonna, F & Palorini, R 2022, 'Vitamin D and SARS-CoV2 infection, severity and mortality: A systematic review and meta-analysis', *PLoS One*, vol. 17, no. 7, p. e0268396.

Dalbeni, A, Natola, LA, Garbin, M, Zoncapè, M, Cattazzo, F, Mantovani, A, Vella, A, Canè, S, Kassem, J & Bevilacqua, M 2023, 'Interleukin-6: A New Marker of Advanced-Sarcopenic HCC Cirrhotic Patients', *Cancers*, vol. 15, no. 9, p. 2406.

Danese, VC, Pepe, J, Ferrone, F, Colangelo, L, De Martino, V, Nieddu, L, Ferrazza, G, Panzini, E, Pascone, R, Blocki, F, Minisola, S & Cipriani, C 2023, 'The Mutual Interplay between Bone, Glucose and Lipid Metabolism: The Role of Vitamin D and PTH', *Nutrients*, vol. 15, no. 13, p. 2998.

Darling, AL & Lanham-New, SA 2024, 'Chapter 51 - Defining thresholds for vitamin D I: scientific rationale for serum 25-hydroxyvitamin D cutoffs of 25 and 50nmol/L', in M Hewison et al (eds), *Feldman and Pike's Vitamin D (Fifth Edition)*, Academic Press, Cambridge (MA), USA, pp. 1–9.

Daugaard, S, Garde, AH, Hansen, ÅM, Vistisen, HT, Rejnmark, L & Kolstad, HA 2018, 'Indoor, outdoor, and night work and blood concentrations of vitamin D and parathyroid hormone', *Scandinavian Journal of Work, Environment & Health*, vol. 44, no. 6, pp. 647–657.

Davey Smith, G & Hemani, G 2014, 'Mendelian randomization: genetic anchors for causal inference in epidemiological studies', *Human Molecular Genetics*, vol. 23, no. R1, pp. R89–R98.

Davies, NM, Holmes, MV & Davey Smith, G 2018, 'Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians', *BMJ*, vol. 362, p. k601.

Del Valle, HB, Yaktine, AL, Taylor, CL & Ross, AC (eds) 2011, *Dietary Reference Intakes for Calcium and Vitamin D*, Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium, National Academies Press (US), Washington (DC), USA. PMID: 21796828.

Delmonico, MJ & Beck, DT 2017, 'The Current Understanding of Sarcopenia: Emerging Tools and Interventional Possibilities', *American Journal of Lifestyle Medicine*, vol. 11, no. 2, pp. 167–181.

DeLuca, HF & Clagett-Dame, M 2013, 'Vitamin D', in WJ Lennarz & MD Lane (eds), *Encyclopedia of Biological Chemistry (Second Edition)*, Academic Press, Cambridge (MA), USA, pp. 535–539.

Demay, MB, Pittas, AG, Bikle, DD, Diab, DL, Kiely, ME, Lazaretti-Castro, M, Lips, P, Mitchell, DM, Murad, MH, Powers, S, Rao, SD, Scragg, R, Tayek, JA, Valent, AM, Walsh, JME & McCartney, CR 2024, 'Vitamin D for the prevention of disease: An Endocrine Society clinical practice guideline', Journal of Clinical Endocrinology & Metabolism, vol. 109, no. 8, pp. 1907-1947.

Dent, CE 1964, 'Dangers of Vitamin-D Intoxication', *British Medical Journal* vol. 1, no. 5386, p. 834.

Department of Health and Aged Care, Therapeutic Goods Administration 2022, *3.1 Vitamin D Scheduling medicines and poisons*, viewed 10/12/2022, <https://www.tga.gov.au/resources/publication/scheduling-decisions-interim/scheduling-delegates-interim-decisions-and-invitation-further-comment-accsacms-november-2016/31-vitamin-d>.

Derosa, G, Maffioli, P, Simental-Mendía, LE, Bo, S & Sahebkar, A 2016, 'Effect of curcumin on circulating interleukin-6 concentrations: A systematic review and meta-analysis of randomized controlled trials', *Pharmacological Research*, vol. 111, pp. 394–404.

Deurenberg, P, Yap, M & Van Staveren, W 1998, 'Body mass index and percent body fat: a meta analysis among different ethnic groups', *International Journal of Obesity*, vol. 22, no. 12, pp. 1164–1171.

Di Filippo, L, De Lorenzo, R, Giustina, A, Rovere-Querini, P & Conte, C 2022, 'Vitamin D in Osteosarcopenic Obesity', *Nutrients*, vol. 14, no. 9, p. 1816.

Di Renzo, L, Gualtieri, P & De Lorenzo, A 2021, 'Diet, Nutrition and Chronic Degenerative Diseases', *Nutrients*, vol. 13, no. 4, p. 1372.

Diasorin 2019, *LIAISON - 25 OH Vitamin D TOTAL Assay*, Diasorin S.p.A, Saluggia, Italy. <https://www.diasorin.com/sites/default/files/allegati_prodotti/ese_brochure_25_oh_vit._d_total_low.pdf>.

Dieckmann, K-P, Andura, O, Pichlmeier, U, Otte, KM, Isbarn, H & Wülfing, C 2021, 'Are serum levels of 25-hydroxy vitamin D reduced following orchiectomy in testicular cancer patients?', *Basic and Clinical Andrology*, vol. 31, no. 1, p. 14.

Ding, C, Gao, D, Wilding, J, Trayhurn, P & Bing, C 2012, 'Vitamin D signalling in adipose tissue', *British Journal of Nutrition*, vol. 108, no. 11, pp. 1915–1923.

Dodds, RM, Granic, A, Robinson, SM & Sayer, AA 2020, 'Sarcopenia, long-term conditions, and multimorbidity: findings from UK Biobank participants', *Journal of Cachexia, Sarcopenia and Muscle*, vol. 11, no. 1, pp. 62–68.

Duester, G 1994, 'Are Ethanol-Induced Birth Defects Caused by Functional Retinoic Acid Deficiency?', in R Blomhoff (ed), *Vitamin A in Health and Disease*, CRC Press, Boca Raton (FL). USA, p. 363.

Dugger, BN & Dickson, DW 2017, 'Pathology of Neurodegenerative Diseases', *Cold Spring Harbor Perspectives in Biology*, vol. 9, no. 7, pp. 1–22.

Duguid, J 1930, 'The Toxicity of Vitamin D', *The Lancet*, vol. 216, no. 5592, pp. 983–985.

Dunlop, E, James, AP, Cunningham, J, Strobel, N, Lucas, RM, Kiely, M, Nowson, CA, Rangan, A, Adorno, P, Atyeo, P & Black, LJ 2021, 'Vitamin D composition of Australian foods', *Food Chemistry*, vol. 358, p. 129836.

Durup, D, Jørgensen, HL, Christensen, J, Tjønneland, A, Olsen, A, Halkjær, J, Lind, B, Heegaard, A-M & Schwarz, P 2015, 'A Reverse J-Shaped Association Between Serum 25-Hydroxyvitamin D and Cardiovascular Disease Mortality: The CopD Study', *The Journal of Clinical Endocrinology & Metabolism*, vol. 100, no. 6, pp. 2339–2346.

Dzik, KP & Kaczor, JJ 2019, 'Mechanisms of vitamin D on skeletal muscle function: oxidative stress, energy metabolism and anabolic state', *European Journal of Applied Physiology*, vol. 119, no. 4, pp. 825–839.

Ebrahimzadeh, A, Mohseni, S, Narimani, B, Ebrahimzadeh, A, Kazemi, S, Keshavarz, F, Yaghoubi, MJ & Milajerdi, A 2023, 'Association between vitamin D status and risk of covid-19 in-hospital mortality: A systematic review and meta-analysis of observational studies', *Critical Reviews in Food Science and Nutrition*, vol. 63, no. 21, pp. 5033–5043.

Eder, K, Baffy, N, Falus, A & Fulop, AK 2009, 'The major inflammatory mediator interleukin-6 and obesity', *Inflammation Research*, vol. 58, pp. 727–736.

El-Sharkawy, A & Malki, A 2020, 'Vitamin D signaling in inflammation and cancer: Molecular mechanisms and therapeutic implications', *Molecules*, vol. 25, no. 14, p. 3219.

Elíes, J, Yáñez, M, Pereira, TM, Gil-Longo, J, MacDougall, DA & Campos-Toimil, M 2020, 'An update to calcium binding proteins', in Md. Shahidul Islam (ed), *Calcium Signaling*, Springer, Cham, Switzerland, pp. 183–213.

Elkhwanky, MS, Kummu, O, Piltonen, TT, Laru, J, Morin Papunen, L, Mutikainen, M, Tavi, P & Hakkola, J 2020, 'Obesity Represses CYP2R1, the Vitamin D 25 Hydroxylase, in the Liver and Extrahepatic Tissues', *Journal of Bone and Mineral Research Plus*, vol. 4, no. 11, p. e10397.

Fallin, MD, Duggal, P & Beaty, TH 2016, 'Genetic Epidemiology and Public Health: The Evolution From Theory to Technology', *American Journal of Epidemiology*, vol. 183, no. 5, pp. 387–393.

Feng, Q, Zhang, H, Dong, Z, Zhou, Y & Ma, J 2017, 'Circulating 25-hydroxyvitamin D and lung cancer risk and survival: A dose-response meta-analysis of prospective cohort studies', *Medicine (Baltimore)*, vol. 96, no. 45, p. e8613.

Fleet, JC 2022, 'Vitamin D-mediated regulation of intestinal calcium absorption', *Nutrients*, vol. 14, no. 16, p. 3351.

Frame, L, Fischer, J, Geller, G & Cheskin, L 2018, 'Use of Placebo in Supplementation Studies—Vitamin D Research Illustrates an Ethical Quandary', *Nutrients*, vol. 10, no. 3, p. 347.

Francés, MB, Larrea, L, Depiaggio, M, Duque, R, Vázquez, G, Oset, MM, Cardenal, R, Ferrero, VT & Amr, O 2017, 'Vitamin D levels in blood and survival in glioblastoma', *International Journal of Radiation Oncology, Biology, Physics*, vol. 99, no. 2, p. S188.

Franssen, FM, Rutten, EP, Groenen, MT, Vanfleteren, LE, Wouters, EF & Spruit, MA 2014, 'New reference values for body composition by bioelectrical impedance analysis in the general population: results from the UK Biobank', *Journal of the American Medical Directors Association*, vol. 15, no. 6, pp. 448.e1–448.e6.

Fraser, BJ, Blizzard, L, Buscot, M-J, Schmidt, MD, Dwyer, T, Venn, AJ & Magnussen, CG 2022, 'Muscular strength measured across the life-course and the metabolic syndrome', *Nutrition, Metabolism and Cardiovascular Diseases*, vol. 32, no. 5, pp. 1131–1137.

Fraser, WD, Tang, JCY, Dutton, JJ & Schoenmakers, I 2020, 'Vitamin D Measurement, the Debates Continue, New Analytes Have Emerged, Developments Have Variable Outcomes', *Calcified Tissue International*, vol. 106, no. 1, pp. 3–13.

Frérot, M, Lefebvre, A, Aho, S, Callier, P, Astruc, K & Aho Glélé, LS 2018, 'What is epidemiology? Changing definitions of epidemiology 1978-2017', *PLoS One*, vol. 13, no. 12, p. e0208442.

Fry, D, Almond, R, Moffat, S, Gordon, M & Singh, P 2019, *UK Biobank Biomarker Project - Companion Document to Accompany Serum Biomarker Data*, UK Biobank, Stockport, UK. <https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf>.

Gagliano Taliun, SA & Evans, DM 2021, 'Ten simple rules for conducting a mendelian randomization study', *PLoS Computational Biology*, vol. 17, no. 8, p. e1009238.

Galior, K, Grebe, S & Singh, R 2018, 'Development of Vitamin D Toxicity from Overcorrection of Vitamin D Deficiency: A Review of Case Reports', *Nutrients*, vol. 10, no. 8, p. 953.

Gill, BD & Indyk, HE 2018, 'Analysis of Vitamin D_2 and Vitamin D_3 in Infant and Adult Nutritional Formulas by Liquid Chromatography-Tandem Mass Spectrometry: A Multilaboratory Testing Study', *Journal of AOAC International*, vol. 101, no. 1, pp. 256–263.

Ginde, AA, Blatchford, P, Breese, K, Zarrabi, L, Linnebur, SA, Wallace, JI & Schwartz, RS 2017, 'High-Dose Monthly Vitamin D for Prevention of Acute Respiratory Infection in Older Long-Term Care Residents: A Randomized Clinical Trial', *Journal of the American Geriatrics Society*, vol. 65, no. 3, pp. 496–503.

Girgis, CM & Brennan-Speranza, TC 2021, 'Vitamin D and Skeletal Muscle: Current Concepts From Preclinical Studies', *JBMR Plus*, vol. 5, no. 12, p. e10575.

Giustina, A, Bouillon, R, Dawson-Hughes, B, Ebeling, PR, Lazaretti-Castro, M, Lips, P, Marcocci, C & Bilezikian, JP 2023, 'Vitamin D in the older population: a consensus statement', *Endocrine*, vol. 79, no. 1, pp. 31–44.

Gogulothu, R, Nagar, D, Gopalakrishnan, S, Garlapati, VR, Kallamadi, PR & Ismail, A 2020, 'Disrupted expression of genes essential for skeletal muscle fibre integrity and energy metabolism in Vitamin D deficient rats', *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 197, p. 105525.

Grant, AH 1930, 'The Effect of Rachitic Diets on Experimental Tuberculosis: IV. Effects of Disturbing Optimal Ratio between Calcium, Vitamin C and Vitamin D', *American Review of Tuberculosis*, vol. 21, no. 1, pp. 115–126.

Grant, WB, Bhattoa, HP & Boucher, BJ 2017, 'Seasonal variations of U.S. mortality rates: Roles of solar ultraviolet-B doses, vitamin D, gene exp ression, and infections', *Journal of Steroid Biochemistry and Molecular Biology*, vol. 173, pp. 5–12.

Haghighat, N, Sohrabi, Z, Bagheri, R, Akbarzadeh, M, Esmaeilnezhad, Z, Ashtary-Larky, D, Barati-Boldaji, R, Zare, M, Amini, M, Hosseini, SV, Wong, A & Foroutan, H 2023, 'A Systematic Review and Meta-Analysis of Vitamin D Status of Patients with Severe Obesity in Various Regions Worldwide', *Obesity Facts*, vol. 16, no. 6, pp. 519–539.

Halfon, M, Phan, O & Teta, D 2015, 'Vitamin D: a review on its effects on muscle strength, the risk of fall, and frailty', *BioMed Research International*, vol. 2015, p. 953241.

Hallemans, A, Verbeque, E & Van de Walle, P 2020, 'Motor functions', *Handbook of Clinical Neurology*, vol. 173, pp. 157–170.

Hämäläinen, O, Tirkkonen, A, Savikangas, T, Alén, M, Sipilä, S & Hautala, A 2024, 'Low physical activity is a risk factor for sarcopenia: a cross-sectional analysis of two exercise trials on community-dwelling older adults', *BMC Geriatrics*, vol. 24, no. 1, pp. 1–13.

Han, MS, White, A, Perry, RJ, Camporez, JP, Hidalgo, J, Shulman, GI & Davis, RJ 2020, 'Regulation of adipose tissue inflammation by interleukin 6', *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 117, no. 6, pp. 2751–2760.

Han, SS & Chatterjee, N 2018, 'Review of Statistical Methods for Gene-Environment Interaction Analysis', *Current Epidemiology Reports*, vol. 5, no. 1, pp. 39–45.

Harris, LJ & Innes, JRM 1931, 'The mode of action of vitamin D', *Biochemical Journal*, vol. 25, no. 1, pp. 367–390.5.

Hartono, AC, Sidharta, VM, Astiarani, Y & Regina, R 2023, 'Association between melanin and vitamin D: A systematic review', *JKKI: Jurnal Kedokteran dan Kesehatan Indonesia*, vol. 14, no. 1, pp. 95–103.

Hartwig, FP, Davey Smith, G & Bowden, J 2017, 'Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption', *International Journal of Epidemiology*, vol. 46, no. 6, pp. 1985–1998.

Hasanloei, MAV, Rahimlou, M, Eivazloo, A, Sane, S, Ayremlou, P & Hashemi, R 2020, 'Effect of Oral Versus Intramuscular Vitamin D Replacement on Oxidative Stress and Outcomes in Traumatic Mechanical Ventilated Patients Admitted to Intensive Care Unit', *Nutrition in Clinical Practice*, vol. 35, no. 3, pp. 548–558.

Hausdorff, JM, Peng, C-K, Goldberger, AL & Stoll, AL 2004, 'Gait unsteadiness and fall risk in two affective disorders: a preliminary study', *BMC Psychiatry*, vol. 4, no. 1, pp. 1–7.

Hauta-Alus, HH, Holmlund-Suila, EM, Rita, HJ, Enlund-Cerullo, M, Rosendahl, J, Valkama, SM, Helve, OM, Hytinantti, TK, Surcel, H-M, Mäkitie, OM, Andersson, S & Viljakainen, HT 2018, 'Season, dietary factors, and physical activity modify 25-hydroxyvitamin D concentration during pregnancy', *European Journal of Nutrition*, vol. 57, no. 4, pp. 1369–1379.

Henson, J, Yates, T, Edwardson, CL, Khunti, K, Talbot, D, Gray, LJ, Leigh, TM, Carter, P & Davies, MJ 2013, 'Sedentary time and markers of chronic low-grade inflammation in a high risk population', *PLoS One*, vol. 8, no. 10, p. e78350.

Hernandez, M, Recalde, S, González-Zamora, J, Bilbao-Malavé, V, Sáenz de Viteri, M, Bezunartea, J, Moreno-Orduña, M, Belza, I, Barrio-Barrio, J & Fernandez-Robredo, P 2021, 'Anti-inflammatory and anti-oxidative synergistic effect of vitamin D and nutritional complex on retinal pigment epithelial and endothelial cell lines against age-related macular degeneration', *Nutrients*, vol. 13, no. 5, p. 1423.

Hess, AF & Unger, LJ 1921, 'The cure of infantile rickets by artificial light and by sunlight', *Experimental Biology and Medicine*, vol. 18, no. 8, pp. 298–298.

Hess, AF, Weinstock, M & Helman, F 1925, 'The development of antirachitic potency in phytosterol and cholesterol following irradiation', *Proceedings of the Society for Experimental Biology and Medicine*, vol. 22, no. 4, pp. 227–228.

Hetland, RB, Alexander, J, Berg, JP, Svendsen, C & Paulsen, JE 2009, 'Retinol-induced intestinal tumorigenesis in Min/+ mice and importance of vitamin D status', *Anticancer Research*, vol. 29, no. 11, pp. 4353–4360.

Hewison, M & Adams, JS 2011, 'Chapter 45 - Extrarenal 1α-Hydroxylase', in D Feldman, JW Pike & JS Adams (eds), *Vitamin D (Third Edition)*, Academic Press, San Diego (CA), USA, pp. 777–804.

Hewison, M 2012, 'Vitamin D and immune function: autocrine, paracrine or endocrine?', *Scandinavian Journal of Clinical and Laboratory Investigation*, vol. 72, no. sup243, pp. 92–102.

Holick, MF, Binkley, NC, Bischoff-Ferrari, HA, Gordon, CM, Hanley, DA, Heaney, RP, Murad, MH & Weaver, CM 2011, 'Evaluation, Treatment, and Prevention of Vitamin D Deficiency: an Endocrine Society Clinical Practice Guideline', *The Journal of Clinical Endocrinology & Metabolism*, vol. 96, no. 7, pp. 1911–1930.

Holick, MF & Slominski, AT 2024, 'Chapter 3 - Photobiology of vitamin D', in M Hewison et al (eds), *Feldman and Pike' s Vitamin D (Fifth Edition)*, Academic Press, Cambridge (MA), USA, pp. 27–45.

Holmes, MV, Dale, CE, Zuccolo, L, Silverwood, RJ, Guo, Y, Ye, Z, Prieto-Merino, D, Dehghan, A, Trompet, S, Wong, A, Cavadino, A, Drogan, D, Padmanabhan, S, Li, S, Yesupriya, A, Leusink, M, Sundstrom, J, Hubacek, JA, Pikhart, H, Swerdlow, DI, Panayiotou, AG, Borinskaya, SA, Finan, C, Shah, S, Kuchenbaecker, KB, Shah, T, Engmann, J, Folkersen, L, Eriksson, P, Ricceri, F, Melander, O, Sacerdote, C, Gamble, DM, Rayaprolu, S, Ross, OA, McLachlan, S, Vikhireva, O, Sluijs, I, Scott, RA, Adamkova, V, Flicker, L, Bockxmeer, FM, Power, C, Marques-Vidal, P, Meade, T, Marmot, MG, Ferro, JM, Paulos-Pinheiro, S, Humphries, SE, Talmud, PJ, Mateo Leach, I, Verweij, N, Linneberg, A, Skaaby, T, Doevendans, PA, Cramer, MJ, van der Harst, P, Klungel, OH, Dowling, NF, Dominiczak, AF, Kumari, M, Nicolaides, AN, Weikert, C, Boeing, H, Ebrahim, S, Gaunt, TR, Price, JF, Lannfelt, L, Peasey, A, Kubinova, R, Pajak, A, Malyutina, S, Voevoda, MI, Tamosiunas, A, Maitland-van der Zee, AH, Norman, PE, Hankey, GJ, Bergmann, MM, Hofman, A, Franco, OH, Cooper, J, Palmen, J, Spiering, W, de Jong, PA, Kuh, D, Hardy, R, Uitterlinden, AG, Ikram, MA, Ford, I, Hyppönen, E, Almeida, OP, Wareham, NJ, Khaw, KT, Hamsten, A, Husemoen, LL, Tjønneland, A, Tolstrup, JS, Rimm, E, Beulens, JW, Verschuren, WM, Onland-Moret, NC, Hofker, MH, Wannamethee, SG, Whincup, PH, Morris, R, Vicente, AM, Watkins, H, Farrall, M, Jukema, JW, Meschia, J, Cupples, LA, Sharp, SJ, Fornage, M, Kooperberg, C, LaCroix, AZ, Dai, JY, Lanktree, MB, Siscovick, DS, Jorgenson, E, Spring, B, Coresh, J, Li, YR, Buxbaum, SG, Schreiner, PJ, Ellison, RC, Tsai, MY, Patel, SR, Redline, S, Johnson, AD, Hoogeveen, RC, Hakonarson, H, Rotter, JI, Boerwinkle, E, de Bakker, PI, Kivimaki, M, Asselbergs, FW, Sattar, N, Lawlor, DA, Whittaker, J, Davey Smith, G, Mukamal, K, Psaty, BM, Wilson, JG, Lange, LA, Hamidovic, A, Hingorani, AD, Nordestgaard, BG, Bobak, M, Leon, DA, Langenberg, C, Palmer, TM, Reiner, AP, Keating, BJ, Dudbridge, F & Casas, JP 2014, 'Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data', *BMJ*, vol. 349, p. g4164.

Horas, K, van Herck, U, Maier, GS, Maus, U, Harrasser, N, Jakob, F, Weissenberger, M, Arnholdt, J, Holzapfel, BM & Rudert, M 2020, 'Does vitamin D deficiency predict tumour malignancy in patients with bone tumours? Data from a multi-center cohort analysis', *Journal of Bone Oncology*, vol. 25, p. 100329.

Hosseini, B, El Abd, A & Ducharme, FM 2022, 'Effects of Vitamin D Supplementation on COVID-19 Related Outcomes: A Systematic Review and Meta-Analysis', *Nutrients*, vol. 14, no. 10, May 20, p. 2134.

Hou, J, He, C, He, W, Yang, M, Luo, X & Li, C 2020, 'Obesity and Bone Health: A Complex Link', *Frontiers in Cell Development Biology*, vol. 8, p. 600181.

Huang, H, Huang, X, Zeng, K, Deng, F, Lin, C & Huang, W 2021, 'Interleukin-6 is a strong predictor of the frequency of COPD exacerbation within 1 year', *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 16, pp. 2945–2951.

Hunter, DJ 2005, 'Gene–environment interactions in human diseases', *Nature Reviews Genetics*, vol. 6, no. 4, pp. 287–298.

Hyppönen, E, Vimaleswaran, KS & Zhou, A 2022, 'Genetic Determinants of 25-Hydroxyvitamin D Concentrations and Their Relevance to Public Health', *Nutrients*, vol. 14, no. 20, p. 4408.

Irfan, SA, Ali, AA, Shabbir, N, Altaf, H, Ahmed, A, Kunnath, JT, Boorle, NVLD, Miguel, AK, Loh, CC & Gandrakota, N 2022, 'Effects of vitamin d on systemic lupus erythematosus disease activity and autoimmunity: a systematic review and meta-analysis', *Cureus*, vol. 14, no. 6, p. e25896.

Jaiswal, V, Ishak, A, Peng Ang, S, Babu Pokhrel, N, Shama, N, Lnu, K, Susan Varghese, J, Storozhenko, T, Ee Chia, J, Naz, S, Sharma, P & Jaiswal, A 2022, 'Hypovitaminosis D and cardiovascular outcomes: A systematic review and meta-analysis', *IJC Heart & Vasculature*, vol. 40, p. 101019.

Jakobsen, J & Christensen, T 2021, 'Natural Vitamin D in Food: To What Degree Does 25 Hydroxyvitamin D Contribute to the Vitamin D Activity in Food?', *Journal of Bone and Mineral Research Plus*, vol. 5, no. 1, p. e10453.

Jazayeri-Tehrani, SA, Rezayat, SM, Mansouri, S, Qorbani, M, Alavian, SM, Daneshi-Maskooni, M & Hosseinzadeh-Attar, M-J 2019, 'Nano-curcumin improves glucose indices, lipids, inflammation, and Nesfatin in overweight and obese patients with non-alcoholic fatty liver disease (NAFLD): a double-blind randomized placebo-controlled clinical trial', *Nutrition & Metabolism*, vol. 16, no. 1, pp. 1–13.

Jefferson, GE, Schnell, DM, Thomas, DT & Bollinger, LM 2017, 'Calcitriol concomitantly enhances insulin sensitivity and alters myocellular lipid partitioning in high fat-treated skeletal muscle cells', *Journal of Physiology and Biochemistry*, vol. 73, no. 4, pp. 613–621.

Jensen, MB 2021, 'Vitamin D and male reproduction', *Hormone Research in Paediatrics*, vol. 94, no. Suppl 1, p. 8.

Ji, T, Li, Y & Ma, L 2022, 'Sarcopenic obesity: an emerging public health problem', *Aging and Disease*, vol. 13, no. 2, p. 379.

Jiang, X, O'Reilly, PF, Aschard, H, Hsu, Y-H, Richards, JB, Dupuis, J, Ingelsson, E, Karasik, D, Pilz, S, Berry, D, Kestenbaum, B, Zheng, J, Luan, J, Sofianopoulou, E, Streeten, EA, Albanes, D, Lutsey, PL, Yao, L, Tang, W, Econs, MJ, Wallaschofski, H, Völzke, H, Zhou, A, Power, C, McCarthy, MI, Michos, ED, Boerwinkle, E, Weinstein, SJ, Freedman, ND, Huang, W-Y, Van Schoor, NM, Van Der Velde, N, Groot, LCPGMD, Enneman, A, Cupples, LA, Booth, SL, Vasan, RS, Liu, C-T, Zhou, Y, Ripatti, S, Ohlsson, C, Vandenput, L, Lorentzon, M, Eriksson, JG, Shea, MK, Houston, DK, Kritchevsky, SB, Liu, Y, Lohman, KK, Ferrucci, L, Peacock, M, Gieger, C, Beekman, M, Slagboom, E, Deelen, J, Heemst, DV, Kleber, ME, März, W, De Boer, IH, Wood, AC, Rotter, JI, Rich, SS, Robinson-Cohen, C, Den Heijer, M, Jarvelin, M-R, Cavadino, A, Joshi, PK, Wilson, JF, Hayward, C, Lind, L, Michaëlsson, K, Trompet, S, Zillikens, MC, Uitterlinden, AG, Rivadeneira, F, Broer, L, Zgaga, L, Campbell, H, Theodoratou, E, Farrington, SM, Timofeeva, M, Dunlop, MG, Valdes, AM, Tikkanen, E, Lehtimäki, T, Lyytikäinen, L-P, Kähönen, M, Raitakari, OT, Mikkilä, V, Ikram, MA, Sattar, N, Jukema, JW, Wareham, NJ, Langenberg, C, Forouhi, NG, Gundersen, TE, Khaw, K-T, Butterworth, AS, Danesh, J, Spector, T, Wang, TJ, Hyppönen, E, Kraft, P & Kiel, DP 2018, 'Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels', *Nature Communications*, vol. 9, no. 1, p. 260.

Jiang, X, Kiel, DP & Kraft, P 2019, 'The genetics of vitamin D', *Bone*, vol. 126, pp. 59–77.

Jiménez-Lara, AM & Aranda, A 1999, 'Vitamin D represses retinoic acid-dependent transactivation of the retinoic acid receptor-β2 promoter: the AF-2 domain of the vitamin D receptor is required for transrepression', *Endocrinology*, vol. 140, no. 6, pp. 2898–2907.

Johnson, JB, Ekanayake, C, Caravani, F, Mani, JS, Lal, P, Calgaro, SJ, Prasad, SS, Warner, RD & Naiker, M 2023, 'A review of vitamin D and its precursors in plants and their translation to active metabolites in meat', *Food Reviews International*, vol. 39, no. 4, pp. 1770–1798.

Jolliffe, DA, Greenberg, L, Hooper, RL, Mathyssen, C, Rafiq, R, de Jongh, RT, Camargo, CA, Griffiths, CJ, Janssens, W & Martineau, AR 2019, 'Vitamin D to prevent exacerbations of COPD: systematic review and meta-analysis of individual participant data from randomised controlled trials', *Thorax*, vol. 74, no. 4, pp. 337–345.

Jolliffe, DA, Camargo, CA, Sluyter, JD, Aglipay, M, Aloia, JF, Ganmaa, D, Bergman, P, Bischoff-Ferrari, HA, Borzutzky, A & Damsgaard, CT 2021, 'Vitamin D supplementation to prevent acute respiratory infections: a

systematic review and meta-analysis of aggregate data from randomised controlled trials', *The Lancet Diabetes & Endocrinology*, vol. 9, no. 5, pp. 276–292.

Joo, SK & Kim, W 2023, 'Interaction between sarcopenia and nonalcoholic fatty liver disease', *Clinical and Molecular Hepatology*, vol. 29, no. Suppl, p. S68.

Jorde, R, Sneve, M, Figenschau, Y, Svartberg, J & Waterloo, K 2008, 'Effects of vitamin D supplementation on symptoms of depression in overweight and obese subjects: randomized double blind trial', *Journal of Internal Medicine*, vol. 264, no. 6, pp. 599–609.

Kamiya, S, Nakamori, Y, Takasawa, A, Takasawa, K, Kyuno, D, Ono, Y, Magara, K & Osanai, M 2023, 'Suppression of the vitamin D metabolizing enzyme CYP24A1 provides increased sensitivity to chemotherapeutic drugs in breast cancer', *Oncology Reports*, vol. 49, no. 5, pp. 1–10.

Kang, M, Robards, F, Luscombe, G, Sanci, L & Usherwood, T 2020, 'The relationship between having a regular general practitioner (GP) and the experience of healthcare barriers: a cross-sectional study among young people in NSW, Australia, with oversampling from marginalised groups', *BMC Family Practice*, vol. 21, no. 1, p. 220.

Kaplan, MH, Hufford, MM & Olson, MR 2015, 'The development and in vivo function of T helper 9 cells', *Nature Reviews Immunology*, vol. 15, no. 5, pp. 295–307.

Karlic, H & Varga, F 2011, 'Impact of vitamin D metabolism on clinical epigenetics', *Clinical Epigenetics*, vol. 2, no. 1, pp. 55–61.

Kearns, MD, Binongo, JN, Watson, D, Alvarez, JA, Lodin, D, Ziegler, TR & Tangpricha, V 2015, 'The effect of a single, large bolus of vitamin D in healthy adults over the winter and following year: a randomized, double-blind, placebo-controlled trial', *European Journal of Clinical Nutrition*, vol. 69, no. 2, pp. 193–197.

Khashim Alswailmi, F, Shah, SIA, Nawaz, H & Al-Mazaideh, GM 2021, 'Molecular Mechanisms of Vitamin D-Mediated Immunomodulation', *Galen Medical Journal*, vol. 10, p. e2097.

Khaw, KT, Stewart, AW, Waayer, D, Lawes, CMM, Toop, L, Camargo, CA Jr, et al. 2017, 'Effect of monthly high-dose vitamin D supplementation on falls and non-vertebral fractures: secondary and post-hoc outcomes from the randomised, double-blind, placebo-controlled ViDA trial', *The Lancet Diabetes & Endocrinology*, vol. 5, no. 6, pp. 438–447.

Kilbane, MT, O'Keane, M, Morrin, M, Flynn, M & McKenna, MJ 2014, 'The double-edged sword of vitamin D in Ireland: the need for public health awareness about too much as well as too little', *Irish Journal of Medical Science*, vol. 183, no. 3, pp. 485–487.

Kim, IM, Norris, KC & Artaza, JN 2016, 'Chapter Twelve - Vitamin D and Cardiac Differentiation', in G Litwack (ed), *Vitamins & Hormones*, vol. 100, Academic Press, Cambridge (MA), USA, pp. 299–320.

Kim, SH, Jeong, JB, Kang, J, Ahn, D-W, Kim, JW, Kim, BG, Lee, KL, Oh, S, Yoon, SH, Park, SJ & Lee, DH 2021, 'Association between sarcopenia level and metabolic syndrome', *PLoS One*, vol. 16, no. 3, p. e0248856.

Kobayashi, M, Nakatani, T, Koda, T, Matsumoto, K-I, Ozaki, R, Mochida, N, Takao, K, Miyakawa, T & Matsuoka, I 2014, 'Absence of BRINP1 in mice causes increase of hippocampal neurogenesis and behavioral alterations relevant to human psychiatric disorders', *Molecular Brain*, vol. 7, no. 1, p. 12.

Kumar, M, Parchani, A, Kant, R & Das, A 2023, 'Relationship Between Vitamin D Deficiency and Non-alcoholic Fatty Liver Disease: A Cross-Sectional Study From a Tertiary Care Center in Northern India', *Cureus*, vol. 15, no. 2, Feb, p. e34921.

Kweder, H & Eidi, H 2018, 'Vitamin D deficiency in elderly: Risk factors and drugs impact on vitamin D status', *Avicenna Journal of Medicine*, vol. 8, no. 04, pp. 139–146.

Lan, S-H, Lai, C-C, Chang, S-P, Lu, L-C, Hung, S-H & Lin, W-T 2020, 'Vitamin D supplementation and the outcomes of critically ill adult patients: a systematic review and meta-analysis of randomized controlled trials', Scientific Reports, vol. 10, no. 1.

Latham, A & Cunniff, C 2021, 'Smith Lemli Opitz Syndrome', in JC Carey et al (eds), *Cassidy and Allanson's Management of Genetic Syndromes*, Wiley, Hoboken NJ, USA, pp. 851–862.

Latham, NK, Anderson, CS, Lee, A, Bennett, DA, Moseley, A & Cameron, ID 2003, 'A randomized, controlled trial of quadriceps resistance exercise and vitamin D in frail older people: the Frailty Interventions Trial in Elderly Subjects (FITNESS)', *Journal of the American Geriatrics Society*, vol. 51, no. 3, pp. 291–299.

Leal, LKAM, Lima, LA, Alexandre de Aquino, PE, Costa de Sousa, JA, Jataí Gadelha, CV, Felício Calou, IB, Pereira Lopes, MJ, Viana Lima, FA, Tavares Neves, KR, Matos de Andrade, G & Socorro de Barros Viana, G 2020, 'Vitamin D (VD3) antioxidative and anti-inflammatory activities: Peripheral and central effects', *European Journal of Pharmacology*, vol. 879, p. 173099.

Leary, PF, Zamfirova, I, Au, J & McCracken, WH 2017, 'Effect of latitude on vitamin D levels', *Journal of Osteopathic Medicine*, vol. 117, no. 7, pp. 433–439.

Lenherr-Taube, N, Young, EJ, Furman, M, Elia, Y, Assor, E, Chitayat, D, Uster, T, Kirwin, S, Robbins, K, Vinette, KMB, Daneman, A, Marshall, CR, Collins, C, Thummel, K, Sochett, E & Levine, MA 2021, 'Mild Idiopathic Infantile Hypercalcemia-Part 1: Biochemical and Genetic Findings', *Journal of Clinical Endocrinology and Metabolism*, vol. 106, no. 10, pp. 2915–2937.

Leung, PS 2019, 'The Modulatory Action of Vitamin D on the Renin-Angiotensin System and the Determination of Hepatic Insulin Resistance', *Molecules*, vol. 24, no. 13, p. 2479.

Levchenko, A, Vyalova, NM, Nurgaliev, T, Pozhidaev, IV, Simutkin, GG, Bokhan, NA & Ivanova, SA 2020, 'NRG1, PIP4K2A, and HTR2C as potential candidate biomarker genes for several clinical subphenotypes of depression and bipolar disorder', *Frontiers in Genetics*, vol. 11, p. 936.

Li, K, Huang, T, Zheng, J, Wu, K & Li, D 2014, 'Effect of marine-derived n-3 polyunsaturated fatty acids on C-reactive protein, interleukin 6 and tumor necrosis factor α: a meta-analysis', *PLoS One*, vol. 9, no. 2, p. e88103.

Li, YC 2024, 'Vitamin D and the renin–angiotensin system', in M Hewison et al (eds), *Feldman and Pike's Vitamin D (Fifth Edition)*, Academic Press, Cambridge (MA), USA, pp. 925–951.

Liu, CK, Leng, X, Hsu, F-C, Kritchevsky, S, Ding, J, Earnest, C, Ferrucci, L, Goodpaster, B, Guralnik, J & Lenchik, L 2014, 'The impact of sarcopenia on a physical activity intervention: the Lifestyle Interventions and Independence for Elders Pilot Study (LIFE-P)', *The Journal of Nutrition, Health & Aging*, vol. 18, pp. 59–64.

Liu, N, Su, H, Zhang, Y & Kong, J 2020, 'The protective effect of 1,25(OH)2D3 against cardiac hypertrophy is mediated by the cyclin-dependent kinase inhibitor p21', *European Journal of Pharmacology*, vol. 888, p. 173510.

Lopez, DV, Al-Jaberi, FA, Woetmann, A, Ødum, N, Bonefeld, CM, Kongsbak-Wismann, M & Geisler, C 2021, 'Macrophages control the bioavailability of vitamin D and vitamin D-regulated T cell responses', *Frontiers in Immunology*, vol. 12, p. 722806.

Louw, JA, Werbeck, A, Louw, M, Kotze, T, Cooper, R & Labadarios, D 1992, 'Blood vitamin concentrations during the acute-phase response', *Critical Care Medicine*, vol. 20, no. 7, pp. 934–941.

Lovely, C, Rampersad, M, Fernandes, Y & Eberhart, J 2017, 'Gene-environment interactions in development and disease', *Wiley Interdisciplinary Reviews: Developmental Biology*, vol. 6, no. 1, p. e247.

Lung, B, Mowery, M & Komatsu, D 2022, 'Calcitriol ', in StatPearls (ed), *StatPearls*, StatPearls Publishing, Treasure Island (FL), USA, Available at https://www.ncbi.nlm.nih.gov/books/NBK526025/.

Luo, X, Xiong, J, Cai, H, Zou, R, Li, F, Wang, Y & Wang, C 2022, 'Effects of Vitamin D Deficiency on the Function of the Cardiac Autonomic Nervous System in Rats', *Cardiovascular Therapeutics*, vol. 2022, p. 4366948.

Lyall, DM, Celis-Morales, C, Ward, J, Iliodromiti, S, Anderson, JJ, Gill, JMR, Smith, DJ, Ntuk, UE, Mackay, DF, Holmes, MV, Sattar, N & Pell, JP 2017, 'Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study', *JAMA Cardiology*, vol. 2, no. 8, pp. 882–889.

Macdonald, HM, Wood, AD, Tang, JC & Fraser, WD 2012, 'Comparison of vitamin D2 and vitamin D3 supplementation in increasing serum 25-hydroxyvitamin D status: a systematic review and meta-analysis', *The American Journal of Clinical Nutrition*, vol. 96, no. 5, pp. 1152–1153.

Mady, LJ, Zhong, Y, Dhawan, P & Christakos, S 2023, 'Role of Coactivator Associated Arginine Methyltransferase 1 (CARM1) in the Regulation of the Biological Function of 1,25-Dihydroxyvitamin D(3)', *Cells*, vol. 12, no. 10, p. 1407.

Magesh, S, John, D, Li, WT, Li, Y, Mattingly-App, A, Jain, S, Chang, EY & Ongkeko, WM 2021, 'Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic review and meta-analysis', *JAMA Network Open*, vol. 4, no. 11, p. e2134147.

Makris, K, Sempos, C & Cavalier, E 2020, 'The measurement of vitamin D metabolites: part I—metabolism of vitamin D and the measurement of 25-hydroxyvitamin D', *Hormones*, vol. 19, no. 2, pp. 81-96.

Mann, C 2003, 'Observational research methods. Research design II: cohort, cross sectional, and case-control studies', *Emergency Medicine Journal*, vol. 20, no. 1, pp. 54-60.

Manna, P & Jain, SK 2012, 'Vitamin D up-regulates glucose transporter 4 (GLUT4) translocation and glucose utilization mediated by cystathionine-γ-lyase (CSE) activation and H2S formation in 3T3L1 adipocytes', *Journal of Biological Chemistry*, vol. 287, no. 50, pp. 42324–42332.

Manousaki, D, Mitchell, R, Dudding, T, Haworth, S, Harroud, A, Forgetta, V, Shah, RL, Luan, JA, Langenberg, C, Timpson, NJ & Richards, JB 2020, 'Genome-wide Association Study for Vitamin D Levels Reveals 69 Independent Loci', *The American Journal of Human Genetics*, vol. 106, no. 3, pp. 327–337.

Martens, P-J, Gysemans, C, Verstuyf, A & Mathieu, C 2020, 'Vitamin D's effect on immune function', *Nutrients*, vol. 12, no. 5, p. 1248.

Martineau, AR, Timms, PM, Bothamley, GH, Hanifa, Y, Islam, K, Claxton, AP, Packe, GE, Moore-Gillon, JC, Darmalingam, M & Davidson, RN 2011, 'High-dose vitamin D3 during intensive-phase antimicrobial treatment of pulmonary tuberculosis: a double-blind randomised controlled trial', *The Lancet*, vol. 377, no. 9761, pp. 242–250.

Martineau, AR, Jolliffe, DA, Hooper, RL, Greenberg, L, Aloia, JF, Bergman, P, Dubnov-Raz, G, Esposito, S, Ganmaa, D, Ginde, AA, Goodall, EC, Grant, CC, Griffiths, CJ, Janssens, W, Laaksi, I, Manaseki-Holland, S, Mauger, D, Murdoch, DR, Neale, R, Rees, JR, Simpson, S, Stelmach, I, Kumar, GT, Urashima, M & Camargo, CA 2017, 'Vitamin D supplementation to prevent acute respiratory tract infections: systematic review and meta-analysis of individual participant data', *BMJ*, vol. 356, p. i6583.

Mata-Granados, JM, Cuenca-Acevedo, R, Luque De Castro, MD, Sosa, M & Quesada-Gómez, JM 2010, 'Vitamin D deficiency and high serum levels of vitamin A increase the risk of osteoporosis evaluated by Quantitative Ultrasound Measurements (QUS) in postmenopausal Spanish women', *Clinical Biochemistry*, vol. 43, no. 13–14, pp. 1064–1068.

Mazess, RB, Bischoff-Ferrari, HA & Dawson-Hughes, B 2021, 'Vitamin D: Bolus Is Bogus—A Narrative Review', *JBMR Plus*, vol. 5, no. 12, p. e10567.

Mazzoni, P, Shabbott, B & Cortes, JC 2012, 'Motor Control Abnormalities in Parkinson's Disease', *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 6, p. a009282.

Mba, CM, Koulman, A, Forouhi, NG, Sharp, SJ, Imamura, F, Jones, K, Meadows, SR, Assah, F, Mbanya, JC & Wareham, NJ 2022, 'Association between circulating 25-hydroxyvitamin D and cardiometabolic risk factors in adults in rural and urban settings', *Nutrition & Diabetes*, vol. 12, no. 1, p. 34.

McCartney, M 2016, 'Margaret McCartney: Breaking down the silo walls', *BMJ*, vol. 354.

McCourt, AF & O'Sullivan, AM 2022, 'Using food fortification to improve vitamin D bioaccessibility and intakes', *Proceedings of the Nutrition Society*, vol. 81, no. 1, pp. 99–107.

Medicines and Healthcare products Regulatory Agency (MHRA) 2016, *Public Assessment Report UK PAR, SoliCol D3 50,000 IU Tablets, (colecalciferol)*, MHRA, London, UK. <https://mhraproducts4853.blob.core.windows.net/docs/c4d72968ccb86a6699e6789718414ce61f43f3a0>.

Meng, J, Li, X, Liu, W, Xiao, Y, Tang, H, Wu, Y, Xiong, Y & Gao, S 2023, 'The role of vitamin D in the prevention and treatment of SARS-CoV-2 infection: A meta-analysis of randomized controlled trials', *Clinical Nutrition*, vol. 42, no. 11, pp. 2198–2206.

Menon, JC, Kumari, A, Sajjan, SM & Dabadghao, P 2024, 'Novel Mutation in CYP2R1 Causing Vitamin D-Dependent Rickets Type 1b', *JCEM Case Reports*, vol. 2, no. 3, pp. 1–5.

Metz, AL, Walser, MM & Olson, WG 1985, 'The interaction of dietary vitamin A and vitamin D related to skeletal development in the turkey poult', *The Journal of Nutrition*, vol. 115, no. 7, pp. 929–935.

Meyer, MB & Pike, JW 2020, 'Mechanistic homeostasis of vitamin D metabolism in the kidney through reciprocal modulation of Cyp27b1 and Cyp24a1 expression', *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 196, p. 105500.

Meza-Meza, MR, Ruiz-Ballesteros, AI & de la Cruz-Mosso, U 2022, 'Functional effects of vitamin D: From nutrient to immunomodulator', *Critical reviews in food science and nutrition*, vol. 62, no. 11, pp. 3042–3062.

Mezquita & Mezquita 2019, 'Two Opposing Faces of Retinoic Acid: Induction of Stemness or Induction of Differentiation Depending on Cell-Type', *Biomolecules*, vol. 9, no. 10, p. 567.

Mo, Y, Zhou, Y, Chan, H, Evans, C & Maddocks, M 2023, 'The association between sedentary behaviour and sarcopenia in older adults: a systematic review and meta-analysis', *BMC Geriatrics*, vol. 23, no. 1, p. 877.

Modica, R, Liccardi, A, Minotta, R, Benevento, E, Cannavale, G & Colao, A 2023, 'Parathyroid diseases and metabolic syndrome', *Journal of Endocrinological Investigation*, vol. 46, no. 6, pp. 1065–1077.

Moon, M, Jung, ES, Jeon, SG, Cha, M-Y, Jang, Y, Kim, W, Lopes, C, Mook-Jung, I & Kim, K-S 2019, 'Nurr1 (NR4A2) regulates Alzheimer's disease-related pathogenesis and cognitive function in the 5XFAD mouse model', *Aging Cell*, vol. 18, no. 1, p. e12866.

Moonesinghe, R, Ioannidis, JP, Flanders, WD, Yang, Q, Truman, BI & Khoury, MJ 2012, 'Estimating the contribution of genetic variants to difference in incidence of disease between population groups', *European Journal of Human Genetics*, vol. 20, no. 8, pp. 831–836.

Morrow, EL, Duff, MC & Mayberry, LS 2022, 'Mediators, Moderators, and Covariates: Matching Analysis Approach for Improved Precision in Cognitive-Communication Rehabilitation Research', *Journal of Speech, Language, and Hearing Research*, vol. 65, no. 11, pp. 4159–4171.

Movassaghi, M, Bianconi, S, Feinn, R, Wassif, CA & Porter, FD 2017, 'Vitamin D levels in Smith-Lemli-Opitz syndrome', *American Journal of Medical Genetics Part A*, vol. 173, no. 10, pp. 2577–2583.

Murcray, CE, Lewinger, JP & Gauderman, WJ 2008, 'Gene-Environment Interaction in Genome-Wide Association Studies', *American Journal of Epidemiology*, vol. 169, no. 2, pp. 219–226.

Na, SY, Kim, KB, Lim, YJ & Song, HJ 2022, 'Vitamin D and Colorectal Cancer: Current Perspectives and Future Directions', Journal of Cancer Prevention, vol. 27, no. 3, pp. 147-56.

Nair, R & Maseeh, A 2012, 'Vitamin D: The “sunshine” vitamin', *Journal of Pharmacology & Pharmacotherapeutics*, vol. 3, no. 2, pp. 118–126.

Nakatani, S, Nakatani, A, Mori, K, Emoto, M, Inaba, M & Razzaque, MS 2022, 'Fibroblast growth factor 23 as regulator of vitamin D metabolism', in MS Razzaque (ed), *Phosphate Metabolism. Advances in Experimental Medicine and Biology, vol 1362.*, Springer, Cham, Switzerland, pp. 47–54.

Nakshine, VS, Thute, P, Khatib, MN & Sarkar, B 2022, 'Increased screen time as a cause of declining physical, psychological health, and sleep patterns: a literary review', *Cureus*, vol. 14, no. 10, p. e30051.

Navarro-Valverde, C, Caballero-Villarraso, J, Mata-Granados, JM, Casado-Díaz, A, Sosa-Henríquez, M, Malouf-Sierra, J, Nogués-Solán, X, Rodríguez-Mañas, L, Cortés-Gil, X & Delgadillo-Duarte, J 2018, 'High serum retinol as a relevant contributor to low bone mineral density in postmenopausal osteoporotic women', *Calcified Tissue International*, vol. 102, pp. 651–656.

Neale, RE, Baxter, C, Romero, BD, McLeod, DS, English, DR, Armstrong, BK, Ebeling, PR, Hartel, G, Kimlin, MG & O'Connell, R 2022, 'The D-Health Trial: a randomised controlled trial of the effect of vitamin D on mortality', *The Lancet Diabetes & Endocrinology*, vol. 10, no. 2, pp. 120–128.

Neelands, TR, Zhang, J & Macdonald, RL 1999, 'GABA(A) Receptors Expressed in Undifferentiated Human Teratocarcinoma NT2 Cells Differ from Those Expressed by Differentiated NT2-N Cells', *The Journal of Neuroscience*, vol. 19, no. 16, pp. 7057–7065.

Nguyen, C, Wagner, C, Marshall, P & Jurutka, P 2021, 'The Vitamin D Receptor and RXR Heteropartner Cooperate via Synergistic Activation to Control Target Gene Expression', *The FASEB Journal*, vol. 35, no. S1, p. 01848.

Noordzij, M, Dekker, FW, Zoccali, C & Jager, KJ 2010, 'Measures of disease frequency: prevalence and incidence', *Nephron Clinical Practice*, vol. 115, no. 1, pp. c17–c20.

Nudel, R, Zetterberg, R, Hemager, N, Christiani, CAJ, Ohland, J, Burton, BK, Greve, AN, Spang, KS, Ellersgaard, D, Gantriis, DL, Bybjerg-Grauholm, J, Plessen, KJ, Jepsen, JRM, Thorup, AAE, Werge, T, Mors, O & Nordentoft, M 2022, 'A family-based study of genetic and epigenetic effects across multiple neurocognitive, motor, social-cognitive and social-behavioral functions', *Behavioral and Brain Functions*, vol. 18, no. 1, p. 14.

Nuttall, FQ 2015, 'Body Mass Index', *Nutrition Today*, vol. 50, no. 3, pp. 117–128.

Omraninava, M, Razi, B, Aslani, S, Imani, D, Jamialahmadi, T & Sahebkar, A 2021, 'Effect of resveratrol on inflammatory cytokines: A meta-analysis of randomized controlled trials', *European Journal of Pharmacology*, vol. 908, p. 174380.

Orrù, B, Szekeres-Bartho, J, Bizzarri, M, Spiga, A & Unfer, V 2020, 'Inhibitory effects of Vitamin D on inflammation and IL-6 release. A further support for COVID-19 management?', *European Review for Medical & Pharmacological Sciences*, vol. 24, no. 15, pp. 8187–8193.

Osz, J, McEwen, AG, Bourguet, M, Przybilla, F, Peluso-Iltis, C, Poussin-Courmontagne, P, Mély, Y, Cianférani, S, Jeffries, CM, Svergun, DI & Rochel, N 2020, 'Structural basis for DNA recognition and allosteric control of the retinoic acid receptors RAR–RXR', *Nucleic Acids Research*, vol. 48, no. 17, pp. 9969–9985.

Pal, S & Tyler, JK 2016, 'Epigenetics and aging', *Science Advances*, vol. 2, no. 7, p. e1600584.

Palmer, D, Soule, S, Gaddam, RR, Elder, P, Chambers, S & Doogue, M 2022, 'Unbound vitamin D concentrations are not decreased in critically ill patients', *Internal Medicine Journal*, vol. 52, no. 1, pp. 89–94.

Palmer, TM, Sterne, JAC, Harbord, RM, Lawlor, DA, Sheehan, NA, Meng, S, Granell, R, Smith, GD & Didelez, V 2011, 'Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses', *American Journal of Epidemiology*, vol. 173, no. 12, pp. 1392–1403.

Pan, S-C, Wu, Y-F, Lin, Y-C, Lin, S-W & Cheng, C-M 2022, 'Paper-based interleukin-6 test strip for early detection of wound infection', *Biomedicines*, vol. 10, no. 7, p. 1585.

Pan, X, Kaminga, AC, Wen, SW, Wu, X, Acheampong, K & Liu, A 2019, 'Dopamine and dopamine receptors in Alzheimer's disease: a systematic review and network meta-analysis', *Frontiers in Aging Neuroscience*, vol. 11, p. 175.

Park, CY & Han, SN 2021, 'The Role of Vitamin D in Adipose Tissue Biology: Adipocyte Differentiation, Energy Metabolism, and Inflammation', *Journal of Lipid and Atherosclerosis*, vol. 10, no. 2, p. 130.

Patnaik, R, Riaz, S, Sivani, BM, Faisal, S, Naidoo, N, Rizzo, M & Banerjee, Y 2023, 'Evaluating the potential of Vitamin D and curcumin to alleviate inflammation and mitigate the progression of osteoarthritis through their effects on human chondrocytes: A proof-of-concept investigation', *PLoS One*, vol. 18, no. 12, p. e0290739.

Pei, YY, Zhang, Y, Peng, XC, Liu, ZR, Xu, P & Fang, F 2022, 'Association of Vitamin D Supplementation with Cardiovascular Events: A Systematic Review and Meta-Analysis', *Nutrients*, vol. 14, no. 15, p. 3158.

Peleg, S, Feller, C, Ladurner, AG & Imhof, A 2016, 'The Metabolic Impact on Histone Acetylation and Transcription in Ageing', *Trends in Biochemical Sciences*, vol. 41, no. 8, pp. 700–711.

Pelicioni, PHS, Menant, JC, Latt, MD & Lord, SR 2019, 'Falls in Parkinson's Disease Subtypes: Risk Factors, Locations and Circumstances', *International Journal of Environmental Research and Public Health*, vol. 16, no. 12, p. 2216.

Petkovich, M & Chambon, P 2022, 'Retinoic acid receptors at 35 years', *Journal of Molecular Endocrinology*, vol. 69, no. 4, pp. T13-T24.

Petrelli, F, Oldani, S, Borgonovo, K, Cabiddu, M, Dognini, G, Ghilardi, M, Parati, MC, Petro', D, Dottorini, L & Rea, C 2023, 'Vitamin D3 and COVID-19 outcomes: an umbrella review of systematic reviews and meta-analyses', *Antioxidants*, vol. 12, no. 2, p. 247.

Pol-Fuster, J, Cañellas, F, Ruiz-Guerra, L, Medina-Dols, A, Bisbal-Carrió, B, Ortega-Vila, B, Llinàs, J, Hernandez-Rodriguez, J, Lladó, J, Olmos, G, Strauch, K, Heine-Suñer, D, Vives-Bauzà, C & Flaquer, A 2021, 'The conserved ASTN2/BRINP1 locus at 9q33.1–33.2 is associated with major psychiatric disorders in a large pedigree from Southern Spain', *Scientific Reports*, vol. 11, no. 1, p. 14529

Prata, DP, Breen, G, Osborne, S, Munro, J, Clair, DS & Collier, DA 2009, 'An association study of the neuregulin 1 gene, bipolar affective disorder and psychosis', *Psychiatric Genetics*, vol. 19, no. 3, pp. 113–116.

Rahbar, M, Wu, Y, Subramony, JA & Liu, G 2021, 'Sensitive colorimetric detection of interleukin-6 via lateral flow assay incorporated silver amplification method', *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 778269.

Rahman, MS, Hossain, KS, Das, S, Kundu, S, Adegoke, EO, Rahman, MA, Hannan, MA, Uddin, MJ & Pang, MG 2021, 'Role of Insulin in Health and Disease: An Update', *International Journal of Molecular Sciences*, vol. 22, no. 12, p. 6403.

Rajakumar, K 2003, 'Vitamin D, cod-liver oil, sunlight, and rickets: a historical perspective', *Pediatrics*, vol. 112, no. 2, pp. e132–e135.

Ramanarayanan, P, Heine, G & Worm, M 2023, 'Vitamin A and vitamin D induced nuclear hormone receptor activation and its impact on B cell differentiation and immunoglobulin production', *Immunology Letters*, vol. 23, pp. 80-86.

Rehó, B, Fadel, L, Brazda, P, Benziane, A, Hegedüs, É, Sen, P, Gadella, TWJ, Tóth, K, Nagy, L & Vámosi, G 2023, 'Agonist-controlled competition of RAR and VDR nuclear receptors for heterodimerization with RXR is manifested in their DNA binding', *Journal of Biological Chemistry*, vol. 299, no. 2, p. 102896.

Rejnmark, L 2011, 'Effects of vitamin D on muscle function and performance: a review of evidence from randomized controlled trials', *Therapeutic Advances in Chronic Disease*, vol. 2, no. 1, pp. 25–37.

Remelli, F, Vitali, A, Zurlo, A & Volpato, S 2019, 'Vitamin D Deficiency and Sarcopenia in Older Persons', *Nutrients*, vol. 11, no. 12, p. 2861.

Revez, JA, Lin, T, Qiao, Z, Xue, A, Holtz, Y, Zhu, Z, Zeng, J, Wang, H, Sidorenko, J, Kemper, KE, Vinkhuyzen, AAE, Frater, J, Eyles, D, Burne, THJ, Mitchell, B, Martin, NG, Zhu, G, Visscher, PM, Yang, J, Wray, NR & McGrath, JJ 2020, 'Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration', *Nature Communications*, vol. 11, no. 1, p. 1647.

Rezuș, E, Cardoneanu, A, Burlui, A, Luca, A, Codreanu, C, Tamba, BI, Stanciu, GD, Dima, N, Bădescu, C & Rezuș, C 2019, 'The Link Between Inflammaging and Degenerative Joint Diseases', *International Journal of Molecular Sciences*, vol. 20, no. 3, p. 614.

Riemerschmid, G 1921, 'Some aspects on the effect of solar radiation and other factors of fhe climatic environment on human beings (Continued): B: solar radiation in South Africa and Europe and its effect on the human body (1,039)', *South African Medical Journal*, vol. 19, no. 16, pp. 282–287.

Rivera-Paredez, B, Hidalgo-Bravo, A, León-Reyes, G, León-Maldonado, LS, Aquino-Gálvez, A, Castillejos-López, M, Denova-Gutiérrez, E, Flores, YN, Salmerón, J & Velázquez-Cruz, R 2021, 'Total, Bioavailable, and Free 25-Hydroxyvitamin D Equally Associate with Adiposity Markers and Metabolic Traits in Mexican Adults', *Nutrients*, vol. 13, no. 10, p. 3320.

Rizzoli, R 2021, 'Vitamin D supplementation: upper limit for safety revisited?', *Aging Clinical and Experimental Research*, vol. 33, no. 1, pp. 19–24.

Rochel, N 2022, 'Vitamin D and Its Receptor from a Structural Perspective', *Nutrients*, vol. 14, no. 14, p. 2847.

Rohde, CM, Manatt, M, Clagett-Dame, M & DeLuca, HF 1999, 'Vitamin A antagonizes the action of vitamin D in rats', *The Journal of Nutrition*, vol. 129, no. 12, pp. 2246–2250.

Rohde, CM & DeLuca, HF 2005, 'All-trans retinoic acid antagonizes the action of calciferol and its active metabolite, 1, 25-dihydroxycholecalciferol, in rats', *The Journal of Nutrition*, vol. 135, no. 7, pp. 1647–1652.

Roizen, JD, Casella, A, Lai, M, Long, C, Tara, Z, Caplan, I, O'Lear, L & Levine, MA 2018, 'Decreased Serum 25-Hydroxyvitamin D in Aging Male Mice Is Associated With Reduced Hepatic Cyp2r1 Abundance', *Endocrinology*, vol. 159, no. 8, pp. 3083–3089.

Romeu Montenegro, K, Carlessi, R, Cruzat, V & Newsholme, P 2019, 'Effects of vitamin D on primary human skeletal muscle cell proliferation, differentiation, protein synthesis and bioenergetics', *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 193, p. 105423.

Royston, P, Ambler, G & Sauerbrei, W 1999, 'The use of fractional polynomials to model continuous risk variables in epidemiology', *International Journal of Epidemiology*, vol. 28, no. 5, pp. 964–974.

Ruiz-García, A, Pallarés-Carratalá, V, Turégano-Yedro, M, Torres, F, Sapena, V, Martin-Gorgojo, A & Martin-Moreno, JM 2023, 'Vitamin D supplementation and its impact on mortality and cardiovascular outcomes: systematic review and meta-analysis of 80 randomized clinical trials', *Nutrients*, vol. 15, no. 8, p. 1810.

Russo, C, Valle, MS, Malaguarnera, L, Romano, IR & Malaguarnera, L 2023, 'Comparison of Vitamin D and Resveratrol Performances in COVID-19', *Nutrients*, vol. 15, no. 11, p. 2639.

Sacchetti, P, Carpentier, R, Ségard, P, Olivé-Cren, C & Lefebvre, P 2006, 'Multiple signaling pathways regulate the transcriptional activity of the orphan nuclear receptor NURR1', *Nucleic Acids Research*, vol. 34, no. 19, pp. 5515–5527.

Sahebi, L, Dastgiri, S, Ansarin, K, Sahebi, R & Mohammadi, SA 2013, 'Study designs in genetic epidemiology', *ISRN Genetics*, vol. 2013, p. 952518

Sahni, SS, Kakkar, S, Kumar, R & Goraya, JS 2021, 'Osteomalacic myopathy in children and adolescents with vitamin-D deficiency', *Neurology India*, vol. 69, no. 6, p. 1650.

Sambrook, PN, Cameron, ID, Cumming, RG, Lord, SR, Schwarz, JM, Trube, A & March, LM 2002, 'Vitamin D deficiency is common in frail institutionalised older people in northern Sydney', *Medical Journal of Australia*, vol. 176, no. 11, p. 560.

Sanada, F, Taniyama, Y, Muratsu, J, Otsu, R, Shimizu, H, Rakugi, H & Morishita, R 2018, 'Source of Chronic Inflammation in Aging', *Frontiers in Cardiovascular Medicine*, vol. 5, p. 12.

Sanders, KM, Stuart, AL, Williamson, EJ, Simpson, JA, Kotowicz, MA, Young, D & Nicholson, GC 2010, 'Annual high-dose oral vitamin D and falls and fractures in older women: a randomized controlled trial', *JAMA*, vol. 303, no. 18, pp. 1815–1822.

Santiago, JA & Potashkin, JA 2023, 'Physical activity and lifestyle modifications in the treatment of neurodegenerative diseases', *Frontiers in Aging Neuroscience*, vol. 15, p. 1185671.

Saponaro, F, Saba, A & Zucchi, R 2020, 'An Update on Vitamin D Metabolism', *International Journal of Molecular Sciences*, vol. 21, no. 18, p. 6573.

Saud Gany, SL, Chin, K-Y, Tan, JK, Aminuddin, A & Makpol, S 2023, 'Curcumin as a Therapeutic Agent for Sarcopenia', *Nutrients*, vol. 15, no. 11, p. 2526.

Schett, G & Neurath, MF 2018, 'Resolution of chronic inflammatory disease: universal and tissue-specific concepts', *Nature Communications*, vol. 9, no. 1, p. 3261.

Scragg, R 1981, 'Seasonality of Cardiovascular Disease Mortality and the Possible Protective Effect of Ultra-violet Radiation', *International Journal of Epidemiology*, vol. 10, no. 4, pp. 337–341.

Sears, B & Perry, M 2015, 'The role of fatty acids in insulin resistance', *Lipids Health Dis*, vol. 14, Sep 29, p. 121.

Sempos, CT, Durazo-Arvizu, RA, Dawson-Hughes, B, Yetley, EA, Looker, AC, Schleicher, RL, Cao, G, Burt, V, Kramer, H, Bailey, RL, Dwyer, JT, Zhang, X, Gahche, J, Coates, PM & Picciano, MF 2013, 'Is There a Reverse J-Shaped Association Between 25-Hydroxyvitamin D and All-Cause Mortality? Results from the U.S. Nationally Representative NHANES', *The Journal of Clinical Endocrinology & Metabolism*, vol. 98, no. 7, pp. 3001–3009.

Seraphin, G, Rieger, S, Hewison, M, Capobianco, E & Lisse, TS 2023, 'The impact of vitamin D on cancer: A mini review', *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 231, p. 106308.

Shah, G, Erickson, JL, Luxenburg, J, Hu, B, Reali-Sorrell, M, Lovelace, R, Pfoh, E, Kobaivanova, N & Brateanu, A 2021, 'Impact on patients' compliance with medication using prepacked blisters for multidrug medical therapy: I-COMPLY Study', *American Journal of Health-System Pharmacy*, vol. 78, no. 18, pp. 1713–1719.

Shah, K, Seeley, S, Schulz, C, Fisher, J & Gururaja Rao, S 2022, 'Calcium Channels in the Heart: Disease States and Drugs', *Cells*, vol. 11, no. 6, p. 943.

Shakoor, H, Feehan, J, Al Dhaheri, AS, Ali, HI, Platat, C, Ismail, LC, Apostolopoulos, V & Stojanovska, L 2021, 'Immune-boosting role of vitamins D, C, E, zinc, selenium and omega-3 fatty acids: Could they help against COVID-19?', *Maturitas*, vol. 143, pp. 1–9.

Shaurova, T, Seshadri, M & Hershberger, PA 2018, 'Chapter 101 - Vitamin D and Lung Cancer', in D Feldman (ed), *Vitamin D (Fourth Edition)*, Academic Press, Cambridge (MA), USA, pp. 875–890.

Shaw, NJ 2002, 'Vitamin D deficiency in UK Asian families: activating a new concern', *Archives of Disease in Childhood*, vol. 86, no. 3, pp. 147–149.

Shea, RL & Berg, JD 2017, 'Self-administration of vitamin D supplements in the general public may be associated with high 25-hydroxyvitamin D concentrations', *Annals of Clinical Biochemistry*, vol. 54, no. 3, pp. 355–361.

Shimizu, K, Funamoto, M, Sunagawa, Y, Shimizu, S, Katanasaka, Y, Miyazaki, Y, Wada, H, Hasegawa, K & Morimoto, T 2019, 'Anti-inflammatory action of curcumin and its use in the treatment of lifestyle-related diseases', *European Cardiology Review*, vol. 14, no. 2, p. 117.

Silberstein, M 2020, 'Correlation between premorbid IL-6 levels and COVID-19 mortality: Potential role for Vitamin D', *International Immunopharmacology*, vol. 88, p. 106995.

Silberstein, M 2021, 'COVID-19 and IL-6: Why vitamin D (probably) helps but tocilizumab might not', *European Journal of Pharmacology*, vol. 899, p. 174031.

Simpson, RU 2011, 'Selective knockout of the vitamin d receptor in the heart results in cardiac hypertrophy: is the heart a drugable target for vitamin D receptor agonists?', *Circulation*, vol. 124, no. 17, pp. 1808–1810.

Sinnott-Armstrong, N, Tanigawa, Y, Amar, D, Mars, N, Benner, C, Aguirre, M, Venkataraman, GR, Wainberg, M, Ollila, HM, Kiiskinen, T, Havulinna, AS, Pirruccello, JP, Qian, J, Shcherbina, A, Rodriguez, F, Assimes, TL, Agarwala, V, Tibshirani, R, Hastie, T, Ripatti, S, Pritchard, JK, Daly, MJ & Rivas, MA 2021, 'Genetics of 35 blood and urine biomarkers in the UK Biobank', *Nature Genetics*, vol. 53, no. 2, pp. 185–194.

Sîrbe, C, Rednic, S, Grama, A & Pop, TL 2022, 'An Update on the Effects of Vitamin D on the Immune System and Autoimmune Diseases', *International Journal of Molecular Sciences*, vol. 23, no. 17, p. 9784.

Sîrbu, AC, Sabin, O, Bocșan, IC, Vesa, ȘC & Buzoianu, AD 2023, 'The Effect of Vitamin D Supplementation on the Length of Hospitalisation, Intensive Care Unit Admission, and Mortality in COVID-19—A Systematic Review and Meta-Analysis', *Nutrients*, vol. 15, no. 15, p. 3470.

Sizar, O, Khare, S, Goyal, A & Givler, A 2024, 'Vitamin D Deficiency', in StatPearls (ed), *StatPearls*, StatPearls Publishing, Treasure Island (FL), USA, Available at https://www.ncbi.nlm.nih.gov/books/NBK532266/.

Small, R 2006, 'Vitamin D and the heart', *The Journal of Lancaster General Hospital*, vol. 1, no. 3, pp. 99–102.

Smeland, OB, Wang, Y, Lo, M-T, Li, W, Frei, O, Witoelar, A, Tesli, M, Hinds, DA, Tung, JY, Djurovic, S, Chen, C-H, Dale, AM & Andreassen, OA 2017, 'Identification of genetic loci shared between schizophrenia and the Big Five personality traits', *Scientific Reports*, vol. 7, no. 1, p. 2222.

Smith, LM, Gallagher, JC & Suiter, C 2017, 'Medium doses of daily vitamin D decrease falls and higher doses of daily vitamin D3 increase falls: a randomized clinical trial', *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 173, pp. 317–322.

Snegarova, V & Naydenova, D 2020, 'Vitamin D: a Review of its Effects on Epigenetics and Gene Regulation', *Folia Medica*, vol. 62, no. 4, pp. 662–667.

Solis-Urra, P, Cristi-Montero, C, Romero-Parra, J, Zavala-Crichton, JP, Saez-Lara, MJ & Plaza-Diaz, J 2019, 'Passive commuting and higher sedentary time is associated with vitamin D deficiency in adult and older women: results from Chilean National Health Survey 2016–2017', *Nutrients*, vol. 11, no. 2, p. 300.

Sommer, A 2008, 'Vitamin A deficiency and clinical disease: an historical overview', *The Journal of Nutrition*, vol. 138, no. 10, pp. 1835–1839.

Song, X-Q, Meng, F, Ramsey, D, Ripps, H & Qian, H 2005, 'The GABA ρ1 subunit interacts with a cellular retinoic acid binding protein in mammalian retina', *Neuroscience*, vol. 136, no. 2, pp. 467–475.

Song, ZY, Yao, Q, Zhuo, Z, Ma, Z & Chen, G 2018, 'Circulating vitamin D level and mortality in prostate cancer patients: a dose-response meta-analysis', *Endocrine Connections*, vol. 7, no. 12, Dec 1, pp. R294–R303.

Spaulding, HR & Yan, Z 2022, 'AMPK and the Adaptation to Exercise', *Annual Review of Physiology*, vol. 84, pp. 209–227.

Srinivasan, A, Wong, FK & Karponis, D 2020, 'Calcitonin: A useful old friend', *Journal of Musculoskeletal & Neuronal Interactions*, vol. 20, no. 4, pp. 600–609.

Staley, JR & Burgess, S 2017, 'Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization', *Genetic Epidemiology*, vol. 41, no. 4, pp. 341–352.

Stark, SL, Roe, CM, Grant, EA, Hollingsworth, H, Benzinger, TL, Fagan, AM, Buckles, VD & Morris, JC 2013, 'Preclinical Alzheimer disease and risk of falls', *Neurology*, vol. 81, no. 5, pp. 437–443.

Stefansson, H, Petursson, H, Sigurdsson, E, Steinthorsdottir, V, Bjornsdottir, S, Sigmundsson, T, Ghosh, S, Brynjolfsson, J, Gunnarsdottir, S, Ivarsson, O, Chou, TT, Hjaltason, O, Birgisdottir, B, Jonsson, H, Gudnadottir, VG, Gudmundsdottir, E, Bjornsson, A, Ingvarsson, B, Ingason, A, Sigfusson, S, Hardardottir, H, Harvey, RP, Lai, D, Zhou, M, Brunner, D, Mutel, V, Gonzalo, A, Lemke, G, Sainz, J, Johannesson, G, Andresson, T, Gudbjartsson, D, Manolescu, A, Frigge, ML, Gurney, ME, Kong, A, Gulcher, JR & Stefansson, K 2002, 'Neuregulin 1 and Susceptibility to Schizophrenia', *The American Journal of Human Genetics*, vol. 71, no. 4, pp. 877–892.

Steindel, SJ 2010, 'International classification of diseases, 10th edition, clinical modification and procedure coding system: descriptive overview of the next generation HIPAA code sets', *Journal of the American Medical Informatics Association*, vol. 17, no. 3, pp. 274–282.

Stokes, CS & Lammert, F 2016, 'Vitamin D supplementation: less controversy, more guidance needed', *F1000Research*, vol. 5, pp. Faculty Rev-2017.

Stray, LL, Kristensen, Ø, Lomeland, M, Skorstad, M, Stray, T & Tønnessen, FE 2013, 'Motor regulation problems and pain in adults diagnosed with ADHD', *Behavioral and Brain Functions*, vol. 9, no. 1, p. 18.

Sudlow, C, Gallacher, J, Allen, N, Beral, V, Burton, P, Danesh, J, Downey, P, Elliott, P, Green, J & Landray, M 2015, 'UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age', *PLoS Medicine*, vol. 12, no. 3, p. e1001779.

Sung, CC, Liao, MT, Lu, KC & Wu, CC 2012, 'Role of vitamin D in insulin resistance', *Journal of Biomedicine & Biotechnology*, vol. 2012, p. 634195.

Suzuki, M & Nishimura, Y 2022, 'The ventral striatum contributes to the activity of the motor cortex and motor outputs in monkeys', *Frontiers in Systems Neuroscience*, vol. 16, p. 979272.

Szymczak-Pajor, I & Śliwińska, A 2019, 'Analysis of association between vitamin D deficiency and insulin resistance', *Nutrients*, vol. 11, no. 4, p. 794.

Taderegew, MM, Woldeamanuel, GG, Wondie, A, Getawey, A, Abegaz, AN & Adane, F 2023, 'Vitamin D deficiency and its associated factors among patients with type 2 diabetes mellitus: a systematic review and meta-analysis', *BMJ Open*, vol. 13, no. 10, p. e075607.

Taylor, PN & Davies, JS 2018, 'A review of the growing risk of vitamin D toxicity from inappropriate practice', *British Journal of Clinical Pharmacology*, vol. 84, no. 6, pp. 1121–1127.

Temple, NJ 2022, 'The Origins of the Obesity Epidemic in the USA–Lessons for Today', *Nutrients*, vol. 14, no. 20, p. 4253.

Tenny, S & Hoffman, MR 2017, 'Relative risk', in Statpearls (ed), *StatPearls*, StatPearls Publishing, Treasure Island (FL), USA, Available at https://pubmed.ncbi.nlm.nih.gov/28613574/.

Thanassoulis, G & O'Donnell, C 2009, 'Mendelian randomization: nature's randomized trial in the post–genome era', *JAMA*, vol. 301, no. 22, pp. 2386–2388.

Thomas, SL, Wakerman, J & Humphreys, JS 2015, 'Ensuring equity of access to primary health care in rural and remote Australia - what core services should be locally available?', *International Journal for Equity in Health*, vol. 14, no. 1, p. 111.

Thurnham, DI, McCabe, GP, Northrop-Clewes, CA & Nestel, P 2003, 'Effects of subclinical infection on plasma retinol concentrations and assessment of prevalence of vitamin A deficiency: meta-analysis', The Lancet, vol. 362, no. 9401, pp. 2052-8.

Thurnham, DI & McCabe, GP 2012, 'Influence of infection and inflammation on biomarkers of nutritional status with an emphasis on vitamin A and iron', in World Health Organization (ed), *Report: Priorities in the assessment of vitamin A and iron status in populations, Panama City, Panama*, World Health Organization, Geneva, pp. 63–80.

Tian, H, Mason, AM, Liu, C & Burgess, S 2023, 'Relaxing parametric assumptions for non-linear Mendelian randomization using a doubly-ranked stratification method', *PLoS Genetics*, vol. 19, no. 6, p. e1010823.

Tintut, Y & Demer, LL 2021, 'Potential impact of the steroid hormone, vitamin D, on the vasculature', *American Heart Journal*, vol. 239, pp. 147–153.

Tobey, J 1936, 'Vitamin D milks', *The American Journal of Nursing*, vol. 36, no. 6, pp. 576–580.

Trimarco, V, Manzi, MV, Mancusi, C, Strisciuglio, T, Fucile, I, Fiordelisi, A, Pilato, E, Izzo, R, Barbato, E & Lembo, M 2022, 'Insulin resistance and vitamin D deficiency: a link beyond the appearances', *Frontiers in Cardiovascular Medicine*, vol. 9, p. 859793.

Tung, KT, Wong, RS, Tsang, HW, Chan, BN, Wong, SY, So, H-K, Tung, JY, Ho, MH, Wong, WH & Ip, P 2021, 'An assessment of risk factors for insufficient levels of vitamin D during early Infancy', *Nutrients*, vol. 13, no. 4, p. 1068.

Turck, D, Bohn, T, Castenmiller, J, De Henauw, S, Hirsch-Ernst, KI, Knutsen, HK, Maciuk, A, Mangelsdorf, I, McArdle, HJ, Pentieva, K, Siani, A, Thies, F, Tsabouri, S, Vinceti, M, Lanham-New, S, Passeri, G, Craciun, I, Fabiani, L, De Sousa, RF, Martino, L, Martínez, SV & Naska, A 2023, 'Scientific opinion on the tolerable upper intake level for vitamin D, including the derivation of a conversion factor for calcidiol monohydrate', *EFSA Journal*, vol. 21, no. 8, p. e08145.

Turner, DP & Deng, H 2020, 'A Conceptual Introduction to Regression', *Headache: The Journal of Head and Face Pain*, vol. 60, no. 6, pp. 1047–1055.

Tyagi, M, Imam, N, Verma, K & Patel, AK 2016, 'Chromatin remodelers: We are the drivers!!', *Nucleus*, vol. 7, no. 4, pp. 388-404.

UK Biobank 2007, *UK Biobank: Protocol for a large-scale prospective epidemilogical resource - Protocol No: UKBB-PROT-09-06*, UK Biobank, Stockport, UK. <https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf>.

UK Biobank 2011a, *Body Composition Measurement Version 1.0*. <https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/body_composition.pdf>.

UK Biobank 2011b, *UK Biobank Grip Strength Measurement Version 1.0*, UK Biobank, Stockport, UK. <https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/Gripstrength.pdf>.

UK Biobank 2015a, *Genotyping of 500,000 UK Biobank participants Description of sample processing workflow and preparation of DNA for genotyping*, UK Biobank, Stockport, UK. <https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/genotyping_sample_workflow.pdf>.

UK Biobank 2015b, *Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource*, UK Biobank, Stockport, UK. <https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/genotyping_qc.pdf>.

UK Biobank 2021, *Mortality data: linkage to death registries* UK Biobank, Stockport, UK, <https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/DeathLinkage.pdf>.

van den Heuvel, EGHM, Lips, P, Schoonmade, LJ, Lanham-New, SA & van Schoor, NM 2024, 'Comparison of the Effect of Daily Vitamin D2 and Vitamin D3 Supplementation on Serum 25-Hydroxyvitamin D Concentration (Total 25(OH)D, 25(OH)D2, and 25(OH)D3) and Importance of Body Mass Index: A Systematic Review and Meta-Analysis', *Advances in Nutrition,* vol. 15, no. 1, p. 100133.

van Driel, M & van Leeuwen, JP 2023, 'Vitamin D and bone: a story of endocrine and auto/paracrine action in osteoblasts', *Nutrients,* vol. 15, no. 3, p. 480.

Vanclay, F, Baines, JT & Taylor, CN 2013, 'Principles for ethical research involving humans: ethical professional practice in impact assessment Part I', *Impact Assessment and Project Appraisal,* vol. 31, no. 4, pp. 243–253.

Vatanparast, H & Longworth, ZL 2023, 'How does Canada's new vitamin D fortification policy affect the high prevalence of inadequate intake of the vitamin?', *Applied Physiology, Nutrition, and Metabolism,* vol. 48, no. 11, pp. 870–875.

Verbanck, M, Chen, C-Y, Neale, B & Do, R 2018, 'Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases', *Nature Genetics,* vol. 50, no. 5, pp. 693–698.

Vieth, R 2020, 'Vitamin D supplementation: cholecalciferol, calcifediol, and calcitriol', *European Journal of Clinical Nutrition,* vol. 74, no. 11, pp. 1493–1497.

Viglino, D, Jullian-Desayes, I, Minoves, M, Aron-Wisnewsky, J, Leroy, V, Zarski, JP, Tamisier, R, Joyeux-Faure, M & Pépin, JL 2017, 'Nonalcoholic fatty liver disease in chronic obstructive pulmonary disease', *European Respiratory Journal,* vol. 49, no. 6, p. 1601923.

Visscher, PM, Wray, NR, Zhang, Q, Sklar, P, McCarthy, MI, Brown, MA & Yang, J 2017, '10 Years of GWAS Discovery: Biology, Function, and Translation', *American Journal of Human Genetics,* vol. 101, no. 1, pp. 5–22.

Vranić, L, Mikolašević, I & Milić, S 2019, 'Vitamin D deficiency: consequence or cause of obesity?', *Medicina,* vol. 55, no. 9, p. 541.

Vyas, SP & Goswami, R 2021, 'Calcitriol and Retinoic acid antagonize each other to suppress the production of IL-9 by Th9 cells', *The Journal of Nutritional Biochemistry,* vol. 96, p. 108788.

Wacker, M & Holick, MF 2013, 'Sunlight and Vitamin D', *Journal of General Internal Medicine,* vol. 5, no. 1, pp. 51–108.

Wallace, A, Gibson, S, De La Hunty, A, Lamberg-Allardt, C & Ashwell, M 2010, 'Measurement of 25-hydroxyvitamin D in the clinical laboratory: current procedures, performance characteristics and limitations', *Steroids,* vol. 75, no. 7, pp. 477–488.

Wang, J, Zhou, J, Robertson, G & Lee, V 2018, 'Vitamin D in Vascular Calcification: A Double-Edged Sword?', *Nutrients,* vol. 10, no. 5, p. 652.

Wang, T, Bengtsson, G, Kärnefelt, I & Björn, LO 2001, 'Provitamins and vitamins $D_2$and D_3 in Cladina spp. over a latitudinal gradient: possible correlation with UV levels', *Journal of Photochemistry & Photobiology*, vol. 62, no. 1-2, pp. 118–122.

Wang, TJ, Zhang, F, Richards, JB, Kestenbaum, B, Van Meurs, JB, Berry, D, Kiel, DP, Streeten, EA, Ohlsson, C & Koller, DL 2010, 'Common genetic determinants of vitamin D insufficiency: a genome-wide association study', *The Lancet*, vol. 376, no. 9736, pp. 180–188.

Wang, Y, Zhu, J & Deluca, HF 2012, 'Where is the vitamin D receptor?', *Archives of Biochemistry and Biophysics*, vol. 523, no. 1, pp. 123–133.

Waterhouse, M, Sanguineti, E, Baxter, C, Duarte Romero, B, McLeod, DSA, English, DR, Armstrong, BK, Ebeling, PR, Hartel, G, Kimlin, MG, O'Connell, RL, Pham, H, Pols, JC, Venn, AJ, Webb, PM, Whiteman, DC & Neale, RE 2021, 'Vitamin D supplementation and risk of falling: outcomes from the randomized, placebo-controlled D-Health Trial', *Journal of Cachexia, Sarcopenia and Muscle*, vol. 12, no. 6, pp. 1428–1439.

Watts, G 2007, 'UK Biobank gets 10% response rate as it starts recruiting volunteers', *BMJ*, vol. 334, no. 7595, p. 659.

Wei, X, Yohannan, S & Richards, JR 2019, 'Physiology, cardiac repolarization dispersion and reserve', in StatPearls (ed), *StatPearls*, StatPearls Publishing, Treasure Island (FL), USA, Available at https://www.ncbi.nlm.nih.gov/books/NBK537194/.

Weiler, HA 2017, 'Vitamin D Supplementation for Infants', in World Health Organization (ed), *e-Library of Evidence for Nutrition Actions (eLENA)*, World Health Organization, Geneva.

Wheeler, BJ, Snoddy, AME, Munns, C, Simm, P, Siafarikas, A & Jefferies, C 2019, 'A Brief History of Nutritional Rickets', *Frontiers in Endocrinology*, vol. 10, p. 795.

White, JH 2022, 'Emerging roles of vitamin D-induced antimicrobial peptides in antiviral innate immunity', *Nutrients*, vol. 14, no. 2, p. 284.

Wicklow, B, Gallo, S, Majnemer, A, Vanstone, C, Comeau, K, Jones, G, L'Abbe, M, Khamessan, A, Sharma, A, Weiler, H & Rodd, C 2016, 'Impact of Vitamin D Supplementation on Gross Motor Development of Healthy Term Infants: A Randomized Dose-Response Trial', *Physical & Occupational Therapy In Pediatrics*, vol. 36, no. 3, pp. 330–342.

Xin, M, Yang, Y, Zhang, D, Wang, J, Chen, S & Zhou, D 2015, 'Attenuation of hind-limb suspension-induced bone loss by curcumin is associated with reduced oxidative stress and increased vitamin D receptor expression', *Osteoporosis International*, vol. 26, pp. 2665–2676.

Yi, Z, Wang, L & Tu, X 2021, 'Effect of Vitamin D Deficiency on Liver Cancer Risk: A Systematic Review and Meta-Analysis', Asian Pacific Journal of Cancer Prevention, vol. 22, no. 4, pp. 991-97.

Young, K, Beggs, MR, Grimbly, C & Alexander, RT 2022, 'Regulation of 1 and 24 hydroxylation of vitamin D metabolites in the proximal tubule', *Experimental Biology and Medicine (Maywood)*, vol. 247, no. 13, pp. 1103–1111.

Zaccardi, F, Tan, PS, Shah, BR, Everett, K, Clift, AK, Patone, M, Saatci, D, Coupland, C, Griffin, SJ & Khunti, K 2023, 'Ethnic disparities in COVID-19 outcomes: a multinational cohort study of 20 million individuals from England and Canada', *BMC Public Health*, vol. 23, no. 1, p. 399.

Zetterstrom, RH, Solomin, L, Jansson, L, Hoffer, BJ, Olson, L & Perlmann, T 1997, 'Dopamine neuron agenesis in Nurr1-deficient mice', *Science*, vol. 276, no. 5310, pp. 248–250.

Zhang, Y-P, Wan, Y-D, Sun, T-W, Kan, Q-C & Wang, L-X 2014, 'Association between vitamin D deficiency and mortality in critically ill adult patients: a meta-analysis of cohort studies', *Critical Care*, vol. 18, no. 6, p. 684.

Zhang, H, Wen, X, Zhang, Y, Wei, X & Liu, T 2015, 'Vitamin D deficiency and increased risk of bladder carcinoma: a meta-analysis', Cellular Physiology and Biochemistry, vol. 37, no. 5, pp. 1686-92.

Zhang, Y, Fang, F, Tang, J, Jia, L, Feng, Y, Xu, P & Faramand, A 2019, 'Association between vitamin D supplementation and mortality: systematic review and meta-analysis', *BMJ*, p. l4673.

Zhang, Y, Li, J, Yang, M & Wang, Q 2023, 'Effect of vitamin D supplementation on COVID-19 patients: A systematic review and meta-analysis', *Frontiers in Nutrition*, vol. 10, p. 1131103.

Zhang, R, Zhang, Y, Liu, Z, Pei, Y, Xu, P, Chong, W, Hai, Y, He, L, He, Y & Yu, J 2022, 'Association between vitamin D supplementation and cancer mortality: A systematic review and meta-analysis', Cancers, vol. 14, no. 15, p. 3717.

Zhong, Y, Armbrecht, HJ & Christakos, S 2009, 'Calcitonin, a regulator of the 25-hydroxyvitamin D3 1alpha-hydroxylase gene', *Journal of Biological Chemistry*, vol. 284, no. 17, Apr 24, pp. 11059–11069.

Zhu, JG, Ochalek, JT, Kaufmann, M, Jones, G & Deluca, HF 2013, 'CYP2R1 is a major, but not exclusive, contributor to 25-hydroxyvitamin D production in vivo', *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 110, no. 39, pp. 15650–15655.

Zhu, M, Wang, T, Wang, C & Ji, Y 2016, 'The association between vitamin D and COPD risk, severity, and exacerbation: an updated systematic review and meta-analysis', *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 11, pp. 2597–2607.

Zittermann, A, Trummer, C, Theiler-Schwetz, V, Lerchbaum, E, März, W & Pilz, S 2021, 'Vitamin D and cardiovascular disease: an updated narrative review', *International Journal of Molecular Sciences*, vol. 22, no. 6, p. 2896.

Zittermann, A, Trummer, C, Theiler-Schwetz, V & Pilz, S 2023, 'Long-term supplementation with 3200 to 4000 IU of vitamin D daily and adverse events: a systematic review and meta-analysis of randomized controlled trials', *European Journal of Nutrition*, vol. 62, no. 4, pp. 1833–1844.

Supplementary Materials

Supplementary Table 1: Genome-wide significant vitamin D variants used for the genetic instruments for measured 25(OH)D concentrations.

No	SNP	CHR	BP	Gene	A1	A2	A1F	UK Biobank*			SUNLIGHT Consortium†			SNPs for 35 SNP genetic score	SNPs for 122 SNP genetic score
								Beta	SE	P	SNPs for 35 SNP score	SE	P		
1	rs6671730[a]	1	2339139	*PEX10*	G	A	0.565714	0.0147881	0.00201077	1.92E-13	0.0061	0.0023	0.006652	Yes	Yes
2	rs35408430	1	17560195	*PADI1*	C	T	0.657806	0.0214952	0.00209979	1.36E-24	0.0236985	0.00564768	0.00002715	Yes	Yes
3	rs7522116	1	41835685	*FOXO6*	C	T	0.433767	0.0134641	0.00202533	2.97E-11	0.0116727	0.00540416	0.03077654	Yes	Yes
4	rs7528419	1	109817192	*CELSR2*	G	A	0.224671	0.0197401	0.00238729	1.35E-16	0.0179046	0.00644566	0.0054732	Yes	Yes
5	rs1933064	1	152301576	*FLG-AS1*	A	G	0.46961	0.015731	0.00203195	9.80E-15	0.0155068	0.00539368	0.00404027	Yes	Yes
6	rs76798800	1	154994978	*DCST2*	G	T	0.733745	0.0121989	0.00225962	6.71E-08	0.0173898	0.00617041	0.00482841	Yes	Yes
7	rs6672758	1	230303512	*GALNT2*	T	C	0.800872	0.0175857	0.00250898	2.40E-12	0.0156301	0.00666121	0.01895423	Yes	Yes
8	rs727857	2	58981967	*LINC01122*	G	A	0.388511	0.0140184	0.00206152	1.05E-11	0.0109131	0.00550772	0.04754487	Yes	Yes
9	rs1047891	2	211540507	*CPS1*	C	A	0.684179	0.0152142	0.00214041	1.18E-12	0.0126572	0.00572581	0.02706743	Yes	Yes
10	rs2012736	2	234622379	*UGT1A5, UGT1A6, UGT1A7, UGT1A8, UGT1A9, UGT1A10*	C	A	0.919186	0.0483073	0.00366555	1.16E-39	0.0384413	0.01038238	0.00021344	Yes	Yes
11	rs6782190	3	85639672	*CADM2*	G	A	0.352488	0.0172156	0.00208415	1.45E-16	0.0206718	0.00562451	0.00023756	Yes	Yes
12	rs705117	4	72608115	*GC*	C	T	0.1477	0.0334179	0.00280601	1.06E-32	0.0269429	0.00744137	0.00029382	Yes	Yes
13	rs1352846	4	72617775	*GC*	A	G	0.708567	0.193471	0.00219074	0	0.2221843	0.00589771	1.40E-310	Yes	Yes
14	rs78151190	6	25619007	*CARMIL1 (LRRC16A)*	A	C	0.871284	0.0168754	0.00297406	1.39E-08	0.0187117	0.00829773	0.02413132	Yes	Yes
15	rs75741381	7	100809458	*VGF*	C	G	0.852362	0.0166065	0.00282521	4.15E-09	0.0214474	0.00736197	0.00357669	Yes	Yes
16	rs12056768	8	116988527	*LINC00536*	T	G	0.417091	0.0234029	0.00202418	6.44E-31	0.0176616	0.00545433	0.00120331	Yes	Yes
17	rs77532868	10	88081438	*GRID1*	T	C	0.054042	0.0265692	0.00440069	1.57E-09	0.0280553	0.01353113	0.03813628	Yes	Yes
18	rs12794714	11	14913575	*CYP2R1*	G	A	0.578197	0.0878964	0.00201629	0	0.0702488	0.00540376	1.22E-38	Yes	Yes
19	rs61891388	11	66079818	*RP11-867G23.13*	G	T	0.455921	0.0125532	0.00200799	4.06E-10	0.0114254	0.00538961	0.03401532	Yes	Yes
20	rs1660839	11	71094232	*AP002387.1*	A	G	0.248849	0.0292665	0.00230557	6.40E-37	0.014173	0.00623639	0.02304867	Yes	Yes
21	rs12803256	11	71132868	*AP002387.1*	G	A	0.776732	0.104243	0.00239998	0	0.0839119	0.00602549	4.39E-44	Yes	Yes
22	rs12798050[b]	11	71223256	*S100A11P3*	T	C	0.830503	0.109998	0.00264849	0	0.0348	0.0024	1.00E-47	Yes	Yes
23	rs72997623	11	75488054	*DGAT2*	A	C	0.084662	0.0276158	0.00358139	1.25E-14	0.0200157	0.00937765	0.03280964	Yes	Yes
24	rs1149605	11	76485216	*RP11-21L23.4*	C	T	0.170397	0.0220166	0.00266133	1.31E-16	0.0209786	0.0072426	0.00377288	Yes	Yes
25	rs10859995	12	96375682	*HAL*	T	C	0.417366	0.0403465	0.0020206	1.05E-88	0.036551	0.00540543	1.36E-11	Yes	Yes
26	rs8018720	14	39556185	*SEC23A*	G	C	0.176673	0.0378247	0.00260904	1.26E-47	0.040852	0.00705183	6.91E-09	Yes	Yes
27	rs261291	15	58680178	*ALDH1A2*	T	C	0.644772	0.0273653	0.00208561	2.50E-39	0.0113468	0.0056366	0.04410853	Yes	Yes
28	rs77924615	16	20392332	*PDILT*	G	A	0.806515	0.0166321	0.00255158	7.11E-11	0.0195535	0.00670691	0.00355194	Yes	Yes
29	rs212100	19	48376995	*SULT2A1*	T	C	0.164001	0.0661522	0.00269018	1.61E-133	0.0193875	0.00719712	0.00706453	Yes	Yes
30	rs10426	19	51517798	*KLK10*	A	G	0.213433	0.0256629	0.00243056	4.64E-26	0.0146449	0.0065379	0.02509092	Yes	Yes
31	rs6123359	20	52714706	*BCAS1*	G	A	0.102225	0.0341831	0.00331429	6.10E-25	0.0373636	0.00940288	0.00007078	Yes	Yes

No	SNP	CHR	BP	Gene	A1	A2	A1F	UK Biobank*			SUNLIGHT Consortium†			SNPs for 35 SNP genetic score	SNPs for 122 SNP genetic score
								Beta	SE	P	SNPs for 35 SNP score	SE	P		
32	rs17216707	20	52732362	*CYP24A1*	T	C	0.817316	0.0376264	0.00263713	3.47E-46	0.0646902	0.0066412	2.02E-22	Yes	Yes
33	rs2585442	20	52737123	*CYP24A1*	G	C	0.240654	0.0356675	0.00237687	6.70E-51	0.0381477	0.00635676	1.96E-09	Yes	Yes
34	rs2762943	20	52790786	*CYP24A1*	G	T	0.923071	0.0457231	0.00373798	2.10E-34	0.032534	0.01442103	0.02406994	Yes	Yes
35	rs2074735	22	31535872	*PLA2G3*	C	G	0.064096	0.0278196	0.00407045	8.23E-12	0.0213678	0.01054517	0.04273241	Yes	Yes
36	rs11591147	1	55505647	*PCSK9*	T	G	0.018041	0.0450903	0.00747745	1.64E-09	-	-	-	No	Yes
37	rs2131925	1	63025942	*DOCK7*	G	T	0.356375	0.0229402	0.0020845	3.61E-28	-	-	-	No	Yes
38	rs140371183	1	152098428	*PUDPP2 (HDHD1P2)*	G	A	0.032179	0.0870147	0.00565335	1.86E-53	-	-	-	No	Yes
39	rs12123821	1	152179152	*FLG-AS1, RP11-107M16.2*	T	C	0.047527	0.0785529	0.00467652	2.55E-63	-	-	-	No	Yes
40	rs61816761	1	152285861	*FLG-AS1,FLG*	A	G	0.015926	0.12315	0.00804767	7.35E-53	-	-	-	No	Yes
41	rs10908419	1	154567699	*ADAR*	G	A	0.510067	0.012342	0.00199206	5.81E-10	-	-	-	No	Yes
42	rs11264322	1	155087933	*Y_RNA*	G	A	0.570235	0.0093848	0.0020223	3.47E-06	-	-	-	No	Yes
43	rs10908465	1	155389688	*ASH1L*	T	C	0.267332	0.0168816	0.00224924	6.12E-14	-	-	-	No	Yes
44	rs867772	1	220972343	*MARC1*	A	G	0.315487	0.0146031	0.00215141	1.14E-11	-	-	-	No	Yes
45	rs7604788	2	21190024	*RP11-116D2.1*	T	C	0.033433	0.0336737	0.0055518	1.32E-09	-	-	-	No	Yes
46	rs541041	2	21294975	*APOB*	G	A	0.180761	0.0154585	0.00258635	2.27E-09	-	-	-	No	Yes
47	rs1260326	2	27730940	*GCKR*	C	T	0.606565	0.0206128	0.00203644	4.41E-24	-	-	-	No	Yes
48	rs11127186	2	28881407	*AC074011.2*	C	T	0.495795	0.0109023	0.00203379	8.30E-08	-	-	-	No	Yes
49	rs2710651	2	63166379	*EHBP1*	G	A	0.471877	0.0114568	0.00199623	9.51E-09	-	-	-	No	Yes
50	rs3849374	2	101443397	*NPAS2, AC092168.2*	G	C	0.821971	0.0161021	0.00261564	7.46E-10	-	-	-	No	Yes
51	rs7569755	2	118648261	*HTR5BP*	A	G	0.29058	0.01425	0.00221206	1.18E-10	-	-	-	No	Yes
52	rs13060130	3	84440527	*AC108696.1*	C	T	0.860311	0.0149893	0.00287224	1.80E-07	-	-	-	No	Yes
53	rs9861009	3	141654685	*TFDP2 (RP11-271K21.11)*	C	T	0.727515	0.0140213	0.00225294	4.86E-10	-	-	-	No	Yes
54	rs78649910	4	3482213	*DOK7*	T	A	0.893821	0.0211949	0.00325203	7.15E-11	-	-	-	No	Yes
55	rs4364259	4	15892159	*RP11-442P12.1*	A	G	0.202148	0.0159119	0.00250595	2.16E-10	-	-	-	No	Yes
56	rs4616820	4	57745481	*REST*	C	T	0.535046	0.012286	0.00201755	1.13E-09	-	-	-	No	Yes
57	rs35057908	4	69372082	*UGT2B29P*	T	A	0.431309	0.0110085	0.00202343	5.31E-08	-	-	-	No	Yes
58	rs13104260	4	70348090	*UGT2B4*	A	G	0.256938	0.0072531	0.00228347	0.00149139	-	-	-	No	Yes
59	rs11732896	4	88287993	*HSD17B11*	G	A	0.701209	0.0160047	0.00217341	1.79E-13	-	-	-	No	Yes
60	rs28364331	4	100201295	*RP11-696N14.1, ADH1A*	G	A	0.018086	0.068614	0.00747169	4.19E-20	-	-	-	No	Yes
61	rs1229984	4	100239319	*ADH1B*	T	C	0.024889	0.0450574	0.00637113	1.53E-12	-	-	-	No	Yes
62	rs10070734	5	87940026	*LINC00461*	C	T	0.709531	0.0132137	0.00219753	1.82E-09	-	-	-	No	Yes
63	rs31612	5	108996643	*AC012603.1*	T	C	0.825562	0.014528	0.00264902	4.15E-08	-	-	-	No	Yes
64	rs72834856	6	22801858	*RP1-209A6.1*	T	G	0.927936	0.0249871	0.00385103	8.67E-11	-	-	-	No	Yes

No	SNP	CHR	BP	Gene	A1	A2	A1F	UK Biobank*			SUNLIGHT Consortium†			SNPs for 35 SNP genetic score	SNPs for 122 SNP genetic score
								Beta	SE	P	SNPs for 35 SNP score	SE	P		
65	rs28374650	6	32623367	*HLA-DQB1*	C	T	0.756438	0.0135623	0.00232502	5.44E-09	-	-	-	No	Yes
66	rs9476310	6	57767576	*RP11-325M4.2*	T	C	0.511363	0.0117571	0.00200093	4.21E-09	-	-	-	No	Yes
67	rs9490317	6	121859499	*RNU4-76P*	C	T	0.445894	0.011051	0.00201182	3.95E-08	-	-	-	No	Yes
68	rs2248551	6	131924689	*MED23*	G	A	0.834778	0.0233623	0.0026823	3.04E-18	-	-	-	No	Yes
69	rs10085881	7	21577960	*DNAH11*	T	C	0.717815	0.0145575	0.00223829	7.83E-11	-	-	-	No	Yes
70	rs7784802	7	64015379	*ZNF680*	T	A	0.360991	0.0138131	0.00207209	2.62E-11	-	-	-	No	Yes
71	rs6966728	7	104618318	*LINC01004*	C	T	0.537368	0.0117532	0.00203758	8.01E-09	-	-	-	No	Yes
72	rs2346264	7	133536351	*EXOC4*	A	C	0.217315	0.0138826	0.00243596	1.21E-08	-	-	-	No	Yes
73	rs34290760	8	9185179	*RP11-115J16.1*	C	G	0.970888	0.0334615	0.00593241	1.70E-08	-	-	-	No	Yes
74	rs804281	8	11611865	*GATA4*	G	A	0.583605	0.0132996	0.00202139	4.72E-11	-	-	-	No	Yes
75	rs28692966	8	25892919	*EBF2*	A	G	0.252936	0.0148311	0.00230018	1.14E-10	-	-	-	No	Yes
76	rs2725371	8	30854033	*PURG*	G	A	0.697691	0.0118392	0.00218287	5.84E-08	-	-	-	No	Yes
77	rs4738684	8	59393273	*CYP7A1*	G	A	0.665522	0.0124102	0.00211491	4.41E-09	-	-	-	No	Yes
78	rs13284054	9	107669073	*ABCA1*	C	T	0.117727	0.0175688	0.00313411	2.07E-08	-	-	-	No	Yes
79	rs10887718	10	82042624	*MAT1A*	C	T	0.471787	0.0111219	0.00199828	2.61E-08	-	-	-	No	Yes
80	rs3925446	10	91495322	*KIF20B*	A	G	0.199129	0.0152166	0.00249607	1.09E-09	-	-	-	No	Yes
81	rs4418728	10	94839724	*CYP26A1*	T	G	0.451688	0.0109611	0.00200006	4.24E-08	-	-	-	No	Yes
82	rs61883501	11	13882754	*RP11-98J9.2, RP11-98J9.3*	A	C	0.966768	0.0017709	0.00553993	0.749232	-	-	-	No	Yes
83	rs116970203	11	14876718	*PDE3B*	G	A	0.972844	0.376873	0.00612043	0	-	-	-	No	Yes
84	rs117576073	11	14912573	*CYP2R1*	G	T	0.987265	0.147177	0.00886081	5.91E-62	-	-	-	No	Yes
85	rs78168201	11	70971149	*SHANK2*	T	C	0.013766	0.0893469	0.00864963	5.18E-25	-	-	-	No	Yes
86	rs964184	11	116648917	*ZPR1 (ZNF259)*	C	G	0.868376	0.0431755	0.00294423	1.09E-48	-	-	-	No	Yes
87	rs613808	11	116710968	*APOA1-AS*	G	A	0.719993	0.0264476	0.00223913	3.40E-32	-	-	-	No	Yes
88	rs2847500	11	120114421	*POU2F3*	G	A	0.876497	0.021925	0.00302751	4.42E-13	-	-	-	No	Yes
89	rs12317268	12	21352541	*SLCO1B1*	A	G	0.848996	0.0208967	0.00278474	6.19E-14	-	-	-	No	Yes
90	rs11182428	12	38526387	*RNA5SP358*	T	C	0.480005	0.0125352	0.00199379	3.23E-10	-	-	-	No	Yes
91	rs1038165	12	68665940	*MDM1*	T	C	0.583349	0.0120567	0.002018	2.31E-09	-	-	-	No	Yes
92	rs11108368	12	96386138	*HAL*	G	A	0.606145	0.0038307	0.00209129	0.0669916	-	-	-	No	Yes
93	rs12372115	12	97982701	*RMST*	G	T	0.929281	0.0217954	0.00387948	1.93E-08	-	-	-	No	Yes
94	rs73413596	12	111582630	*CUX2*	C	T	0.073854	0.0216996	0.00382584	1.41E-08	-	-	-	No	Yes
95	rs7149014	14	29802911	*RP11-562L8.1*	T	C	0.370807	0.0129475	0.00208615	5.42E-10	-	-	-	No	Yes
96	rs12881545	14	101176212	*DLK1*	C	G	0.673439	0.011822	0.00213142	2.91E-08	-	-	-	No	Yes
97	rs1800588	15	58723675	*ALDH1A2, LIPC*	C	T	0.784797	0.0329215	0.00242187	4.38E-42	-	-	-	No	Yes

No	SNP	CHR	BP	Gene	A1	A2	A1F	UK Biobank*			SUNLIGHT Consortium†			SNPs for 35 SNP genetic score	SNPs for 122 SNP genetic score
								Beta	SE	P	SNPs for 35 SNP score	SE	P		
98	rs55829990	15	63790642	*USP3*	T	C	0.655996	0.0186013	0.00210294	9.12E-19	-	-	-	No	Yes
99	rs62007299	15	77711719	*PEAK1*	G	A	0.287463	0.0133407	0.00219977	1.32E-09	-	-	-	No	Yes
100	rs325384	15	100229761	*MEF2A*	C	T	0.715795	0.0141728	0.00221789	1.66E-10	-	-	-	No	Yes
101	rs17231506	16	56994528	*CETP*	C	T	0.676894	0.0184236	0.00213148	5.45E-18	-	-	-	No	Yes
102	rs11076175	16	57006378	*CETP*	G	A	0.178358	0.0230493	0.00260705	9.47E-19	-	-	-	No	Yes
103	rs4327060	16	72807438	*RP5-991G20.1*	C	T	0.945604	0.0243589	0.00439221	2.92E-08	-	-	-	No	Yes
104	rs4575545	16	79755446	*RP11-345M22.1, RP11-345M22.2*	G	A	0.695172	0.0155823	0.00217312	7.47E-13	-	-	-	No	Yes
105	rs11542462	16	82033810	*SDR42E1*	G	A	0.865656	0.023334	0.00291776	1.27E-15	-	-	-	No	Yes
106	rs10454087	17	40735641	*RETREG3 (FAM134C)*	C	T	0.715178	0.0135306	0.00220667	8.70E-10	-	-	-	No	Yes
107	rs2952289	17	66464414	*RP11-120M18.2*	T	C	0.798032	0.017715	0.00249217	1.18E-12	-	-	-	No	Yes
108	rs8091117	18	28919794	*DSG1*	C	A	0.934702	0.0263626	0.0040284	5.98E-11	-	-	-	No	Yes
109	rs4121823	18	47144223	*LIPG*	T	A	0.154667	0.0192879	0.00277797	3.83E-12	-	-	-	No	Yes
110	rs590215	18	57904088	*RP11-795H16.2*	C	T	0.734073	0.0129217	0.00225753	1.04E-08	-	-	-	No	Yes
111	rs2037511	18	61366207	*SERPINB11*	A	G	0.166007	0.0181228	0.00267963	1.35E-11	-	-	-	No	Yes
112	rs142158911	19	11190534	*LDLR*	A	G	0.114608	0.0255317	0.00314553	4.79E-16	-	-	-	No	Yes
113	rs187429064	19	19380513	*AC138430.4, TM6SF2*	G	A	0.011266	0.0648324	0.00947915	7.95E-12	-	-	-	No	Yes
114	rs3814995	19	36342212	*NPHS1*	C	T	0.688405	0.012558	0.00214992	5.18E-09	-	-	-	No	Yes
115	rs7412	19	45412079	*APOE*	T	C	0.082073	0.0300485	0.00363434	1.36E-16	-	-	-	No	Yes
116	rs484195	19	45421877	*APOC1*	A	G	0.384386	0.0155156	0.00209699	1.37E-13	-	-	-	No	Yes
117	rs8113404	19	53065579	*ZNF808, ZNF701*	T	C	0.304586	0.012173	0.00217136	2.07E-08	-	-	-	No	Yes
118	rs11606	19	54658102	*CNOT3*	G	C	0.425162	0.0120363	0.00205119	4.41E-09	-	-	-	No	Yes
119	rs2207132	20	39142516	*MAFB*	G	A	0.96711	0.0345955	0.00557773	5.56E-10	-	-	-	No	Yes
120	rs2229742	21	16339172	*NRIP1*	G	C	0.896549	0.0251483	0.00327069	1.48E-14	-	-	-	No	Yes
121	rs6003456	22	23356100	*AP000362.1*	T	A	0.765336	0.013277	0.00236378	1.95E-08	-	-	-	No	Yes
122	rs115621755	22	50853134	*PPP6R2*	C	T	0.67288	0.0124309	0.00212296	4.76E-09	-	-	-	No	Yes

A1: serum-25(OH)D-increasing allele; A2: alternative allele; A1F: allele frequency for A1; SE: standard error; SNP: single nucleotide polymorphism; CHR: chromosome number; BP: base-pair position, Genome Reference Consortium Human Build 37 (GRCh37); Vitamin D Genetic Score: genetic instrument (using 35 GWAS variants) for measured 25(OH)D concentration used in the primary analysis; 122 version of the Vitamin D Genetic Score: genetic instrument (using 122 GWAS variants) for measured 25(OH)D concentration used in the sensitivity analysis; *obtained from Revez, 2020 et al.; †imputed summary statistics, obtained from Revez, 2020 et al., serum 25(OH)D has been natural-log transformed; [a]SNP proxy in the SUNLIGHT consortium: rs1123571, $r^2 = 0.86806$ (1000 Genome, EUR); [b]SNP proxy in the SUNLIGHT consortium: rs2186777, $r^2 = 1$ (1000 Genome, EUR).

Supplementary Table 1 Reference:

Revez, JA, Lin, T, Qiao, Z, Xue, A, Holtz, Y, Zhu, Z, Zeng, J, Wang, H, Sidorenko, J, Kemper, KE, Vinkhuyzen, AAE, Frater, J, Eyles, D, Burne, THJ, Mitchell, B, Martin, NG, Zhu, G, Visscher, PM, Yang, J, Wray, NR & McGrath, JJ 2020, 'Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration', *Nature Communications,* vol. 11, no. 1, p. 1647.

Appendices – statements of authorship

University of South Australia

Statement of Authorship

Title of Paper	Differences and determinants of vitamin D deficiency among UK Biobank participants: a cross-ethnic and socioeconomic study
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details Relevant reference details	https://www.clinicalnutritionjournal.com/article/S0261-5614(20)30639-7/fulltext
PRINCIPAL AUTHOR	
Name of Principal Author (Candidate)	Joshua Sutherland
Contribution to the Paper Brief description of your work in this publication	JS wrote the first draft and analyzed the data with guidance from EH and AZ. All authors revised the paper, interpreted results and approved the final manuscript.
Overall Percentage (%)	60%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted] Date 20/06/24
CO-AUTHOR CONTRIBUTIONS	
By signing the Statement of Authorship, each author certifies that: I. the candidate's stated contribution to the publication is accurate (as detailed above); II. permission is granted for the candidate in include the publication in the thesis; and III. the sum of all co-author contribution is equal to 100% less the candidate's stated contribution.	
Name of Co-Author 1	Ang Zhou

The University of South Australia
www.unisa.edu.au
CRICOS Provider No. 00121B

Last updated: May 2023

University of South Australia

% Contribution to the Paper	15%	
Signature *Signatures will be redacted by the Library at the time of publication*	Cambridge VPN client certificate Digitally signed by Cambridge VPN client certificate Date: 2024.06.17 09:17:40 +01'00'	Date 17-Jun-24
Name of Co-Author 2	Matthew Leach	
% Contribution to the Paper	5%	
Signature *Signatures will be redacted by the Library at the time of publication*	Associate Professor Matthew Leach Digitally signed by Associate Professor Matthew Leach Date: 2024.06.17 17:42:12 +10'00'	Date 17-Jun-24
Name of Co-Author 3	Elina Hypponen	
% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted]	Date 20/06/24
Name of Co-Author 4		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 5		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 6		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 7		
% Contribution to the Paper		

University of South Australia

Statement of Authorship

Title of Paper	Muscle Traits, Sarcopenia, and Sarcopenic Obesity: A Vitamin D Mendelian Randomization Study
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details Relevant reference details	https://www.mdpi.com/2072-6643/15/12/2703 https://doi.org/10.3390/nu15122703
PRINCIPAL AUTHOR	
Name of Principal Author (Candidate)	Joshua Sutherland
Contribution to the Paper Brief description of your work in this publication	JS wrote the first draft and analysed the data. EH and AZ advised on analyses, revised the paper with JS, and supervised the study. All authors interpreted results and have read and agreed to the published version of the manuscript.
Overall Percentage (%)	60%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted] Date 20/06/24
CO-AUTHOR CONTRIBUTIONS	
By signing the Statement of Authorship, each author certifies that: I. the candidate's stated contribution to the publication is accurate (as detailed above); II. permission is granted for the candidate in include the publication in the thesis; and III. the sum of all co-author contribution is equal to 100% less the candidate's stated contribution.	
Name of Co-Author 1	Ang Zhou

University of South Australia

% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	Cambridge VPN client certificate Digitally signed by Cambridge VPN client certificate Date: 2024.06.17 09:20:30 +01'00'	Date 17-Jun-24
Name of Co-Author 2	Elina Hypponen	
% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted]	Date 20/06/24
Name of Co-Author 3		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 4		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 5		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 6		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 7		
% Contribution to the Paper		

University of South Australia

Statement of Authorship

Title of Paper	Vitamin D Deficiency Increases Mortality Risk in the UK Biobank A Nonlinear Mendelian Randomization Study
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details Relevant reference details	https://www.acpjournals.org/doi/10.7326/M21-3324 https://doi.org/10.7326/M21-3324
PRINCIPAL AUTHOR	
Name of Principal Author (Candidate)	Joshua Sutherland
Contribution to the Paper Brief description of your work in this publication	Analysis and interpretation of the data: JS, AZ, EH. Drafting of the article: JS, AZ, EH. Critical revision for important intellectual content: AZ, EH Final approval of the article: JS, AZ, EH.
Overall Percentage (%)	60%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted] Date 20/06/24
CO-AUTHOR CONTRIBUTIONS	
By signing the Statement of Authorship, each author certifies that: I. the candidate's stated contribution to the publication is accurate (as detailed above); II. permission is granted for the candidate in include the publication in the thesis; and III. the sum of all co-author contribution is equal to 100% less the candidate's stated contribution.	
Name of Co-Author 1	Ang Zhou

University of South Australia

% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	Cambridge VPN client certificate Digitally signed by Cambridge VPN client certificate Date: 2024.06.17 09:21:28 +01'00'	Date 17-Jun-24
Name of Co-Author 2	Elina Hypponen	
% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted]	Date 20/06/24
Name of Co-Author 3		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 4		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 5		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 6		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 7		
% Contribution to the Paper		

University of South Australia

Statement of Authorship

Title of Paper	Is the vitamin D and fall risk relationship dependent on steroid hormone competition? A genetic epidemiological exploration of a novel hypothesis.
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details Relevant reference details	

PRINCIPAL AUTHOR

Name of Principal Author (Candidate)	Joshua Sutherland
Contribution to the Paper Brief description of your work in this publication	JS developed the research question with EH, wrote the first draft and analyzed the data with guidance from EH and AZ. All authors revised the paper, interpreted results, and approved the final manuscript.
Overall Percentage (%)	60%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted] Date 20/06/24

CO-AUTHOR CONTRIBUTIONS

By signing the Statement of Authorship, each author certifies that:

I. the candidate's stated contribution to the publication is accurate (as detailed above);

II. permission is granted for the candidate in include the publication in the thesis; and

III. the sum of all co-author contribution is equal to 100% less the candidate's stated contribution.

Name of Co-Author 1	Ang Zhou

University of South Australia

% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	Cambridge VPN client certificate Digitally signed by Cambridge VPN client certificate Date: 2024.06.17 09:22:56 +01'00'	Date 17-Jun-24
Name of Co-Author 2	Elina Hypponen	
% Contribution to the Paper	20%	
Signature *Signatures will be redacted by the Library at the time of publication*	[redacted]	Date 20/06/24
Name of Co-Author 3		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 4		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 5		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 6		
% Contribution to the Paper		
Signature *Signatures will be redacted by the Library at the time of publication*		Date
Name of Co-Author 7		
% Contribution to the Paper		