Dissertation

INFLUENCE OF THREE VITAMIN D ASSOCIATED POLYMORPHISMS ON BONE, PROSTATE CANCER AND MORTALITY

submitted by

Olivia Trummer, MSc.

for the Academic Degree of

Doctor of Medical Science (Dr. scient. med.)

at the

Medical University of Graz

Department of Internal Medicine

under the Supervision of

Prof. Dr. Barbara Obermayer-Pietsch

Univ.- Doz. Prof. Dr. Wilfried Renner

Priv.- Doz. DDr. Elisabeth Lerchbaum

2013

Declaration

I hereby declare that this dissertation is my own original work and that I have fully acknowledged by name all of those individuals and organisations that have contributed to the research for this dissertation. Due acknowledgement has been made in the text to all other material used. Throughout this dissertation and in all related publications I followed the guidelines of “Good Scientific Practice”.

Date: ..... Signature: .....

Acknowledgement

Lieber Clemens,

Vielen Dank für deine einzigartige Unterstützung auf meinem Weg bis hierher.

Ich freue mich auf die gemeinsamen Herausforderungen, die auf uns warten!

Ich danke Univ.-Prof^{in}. Dr^{in}. Barbara Obermayer-Pietsch herzlich für die Möglichkeit zur Umsetzung meiner Dissertation sowie ihrer unermüdlichen, engagierten Unterstützung für jegliche Belänge zu jeder Tages- und Nachtzeit während meines Studiums. Danke auch für die Motivation und die Möglichkeiten die du uns in deinem Team bietest!

Herzlich bedanken möchte ich mich auch bei Univ.- Doz. Prof. Dr. Wilfried Renner für die wertvollen Diskussionen, das wissenschaftliche Fundament sowie die lehrreiche Zusammenarbeit in den letzten Jahren.

Danke Wilfried!

Mein weiterer Dank gilt meiner lieben Familie, der ich diese Dissertation widme!

Table of contents

Abbreviations and Definitions

A	adenine
ACE	angiotensin converting enzyme
ADA	American Diabetes Association
ALDH2	aldehyde dehydrogenase 2
AM	assay mix
ANOVA	univariate analysis of variance
APCR	activated protein C resistance
BMD	bone mineral density
BMI	body mass index
bp	nucleotide base pairs
C	cytosine
CAD	coronary artery disease
calbindin 9K	calcium-binding proteins
Ca SR	calcium-sensing receptors
CI	confidence interval
CV	coefficient of variation
CRP	C-reactive protein
CYP2R1	cytochrome P450 IIR-1 gene
CYP27B1	25-hydroxyvitamin D3-1-alpha-hydroxylase
DHCR7	7-dehydrocholesterol reductase gene
DNA	desoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
DRE	digital rectal examinations
Dr. scient. med.	Doctor of Medical Science
dUTP	deoxyuridine triphosphate
DWP	96-well deep well plate
DXA	dual energy X-ray absorptiometry
e	empty

EMA	European Medicines Agency
FAM	6-carboxyfluorescein
FRET	fluorescent resonance energy transfer
G	guanine
GC	group-specific component gene
GWA study	genome-wide association study
HR	hazard ratio
HWE	Hardy Weinberg equilibrium
IBM	International Business Machines Corporation
LURIC study	LUdwigshafen RIsk and Cardiovascular Health study
MGB	minor groove binder
min	minute
MR	mendelian randomization
mRNA	messenger ribonucleic acid
nm	nanometers
OPG	osteoprotegerin
OR	odds ratio
PASW	Predictive Analysis SoftWare
PCR	polymerase chain reaction
PSA	prostate-specific antigen
PTH	parathyroid hormone
RAAS	renin- angiotensin- aldosterone system
RANK	receptor activator of nuclear factor kB
RANKL	receptor activator of nuclear factor kB ligand
RCT	randomized controlled trials
RFU	Relative fluorescence units
RNA	ribonucleic acid
rpm	rounds per minute
RXR	retinoid X receptor

r^2	coefficient of determination
SNP(s)	single nucleotide polymorphism(s)
T	thymine
Tm	melting temperature
TNM	tumor node metastasis
TRPV6	transient receptor potential cation channel, subfamily V, member 6
UDG	Uracil-DNA Glycosylase
UK	United Kingdom
UMM	Universial Master Mix
UP	Ultra Pure
USA	United States of America
UVB	ultraviolet B
VDBG	vitamin D-binding alpha globulin
VDBP	vitamin D binding protein
VDR	vitamin D receptor
VDREs	vitamin D-responsive elements
VIC®	fluorescent dye by Applied Biosystems
Vitamin D_3	cholecalciferol
1α-hydroxylase	25-hydroxyvitamin D-1α-hydroxylase
1,25 $(OH)_2$ vitamin D	1,25-dihydroxyvitamin D_3
25-OH-vitamin D	25-hydroxyvitamin D
25-hydroxylase	vitamin D-25-hydroxylase

Abstract in German

Hintergrund und Zielsetzung: Neben den essentiellen Wirkungen von Vitamin D auf die Knochengesundheit gibt es zahlreiche Beobachtungsstudien, die für eine wichtige Rolle des 25-Hydroxyvitamin D (25-OH-Vitamin D)-Status bei extraskelettalen Bereichen wie koronarer Herzkrankheit, Krebs oder Mortalität, sprechen. Bis jetzt ist noch unklar, ob erniedrigte Vitamin D-Spiegel eine Ursache dieser klinischen Manifestationen sind oder ob es sich dabei um die Folge des beeinträchtigten Gesundheitszustandes handelt.

Genomweite Assoziationsstudien identifizierten drei genetische Determinanten für erniedrigte 25-OH-Vitamin D-Spiegel. Basierend auf diesen Daten möchte ich in meiner Arbeit (1.) den Einfluß dieser drei Polymorphismen (SNPs, single nucleotide polymorphisms) auf Vitamin D-relevante Endpunkte aus den Bereichen Knochen, Prostatakrebs und Mortalität untersuchen. Weiters soll mit diesen Polymorphismen eine Mendelsche Randomisierungsstudie durchgeführt werden, welche klären soll, ob (2.) niedrige Vitamin D-Spiegel eine kausale Ursache für höhere Mortalitätsraten sind.

Methoden: SNPs in den Genen GC („group-specific component“, rs2282679), DHCR7 („7-dehydrocholesterol reductase“, rs12785878) und CYP2R1 („Cytochrom P450IIR-1“, rs10741657) wurden in unterschiedlichen Kohorten genotypisiert und auf ihre Relation zu 25-OH-Vitamin D-Spiegeln sowie auf eine mögliche Assoziation mit Vitamin D-relevanten Endpunkten untersucht.

Ergebnisse und Schlussfolgerungen: Der GC-Polymorphismus ist eine wichtige genetische Determinante für den Vitamin D-Status und konnte in allen Studien verifiziert werden. Der CYP2R1- und der DHCR7-Polymorphismus konnten nur in der Studie mit der höchsten Fallzahl als relevant bestätigt werden.

Knochen: Keiner der Polymorphismen war mit der Knochendichte (untersucht an 342 StudienteilnehmerInnen einer Querschnittsstudie) oder mit prävalenten Frakturen assoziiert. TrägerInnen des G-Allels des DHCR7 Polymorphismus zeigten jedoch ein erhöhtes Risiko für prospektive Frakturen (untersucht an 1093 AltersheimpatientInnen einer prospektiven Kohortenstudie).

Prostatakrebs: GC Genotypen waren nicht mit Biomarker detektiertem Rezidiv, Entwicklung von Metastasen und Überleben in 702 untersuchten

Prostatakrebspatienten assoziert. Die drei genetischen Determinanten für niedrige Vitamin D-Spiegel spielen keine übergeordnete Rolle in den untersuchten Endpunkten bei Osteoporose und Prostatakrebs.

Mortalität: GC-, CYP2R1- und DHCR7-Genotypen determinierten 25-OH-Vitamin D-Spiegel zuverlässig, sagten jedoch Gesamt-, kardiovaskuläre- und nicht kardiovaskuläre Mortalität nicht vorher. Niedrige Vitamin D-Spiegel sind zwar mit erhöhter Sterblichkeit assoziiert, sind aber wahrscheinlich nicht die kausale Ursache dafür.

Abstract in English

Background and aim: Vitamin D is known to play an important role in bone health. Muscoskeletal effects of inadequate vitamin D status have long been investigated including its impact on osteoporotic fractures. In the last several years vitamin D got popular for a series of benefical consequences. Many associations of low serum 25 hydroxyvitamin D (25-OH-vitamin D) status and extraskeletal effects such as cardiovascular disease, cancer and mortality have been observed. It remains unclear whether vitamin D is the cause of these clinical patterns or the consequence of an impaired health status.

Based on recently reported genetic determinants of hypovitaminosis D I aimed to investigate (1st) the impact of three single nucleotide polymorphisms (SNPs) on Vitamin D and vitamin D related outcomes in the fields of bone, prostate cancer and mortality studies. Using these SNPs as a tool for Mendelian randomization, I wanted to elucidate (2nd) the question if low vitamin D levels are causal for higher mortality rates.

Methods: Genotypes of SNPs in the group-specific component gene (GC, rs2282679), 7-dehydrocholesterol reductase gene (DHCR7, rs12785878) and cytochrome P450 IIR-1 gene (CYP2R1, rs10741657) were determined in different cohorts and investigated for 25-OH-vitamin D and vitamin D related outcomes.

Results and conclusions: The GC polymorphism was the strongest genetic determinant of 25-OH-vitamin D status and was confirmed in all settings. CYP2R1 and the DHCR7 variants could be approved in the largest study with the highest power to detect statistical differences.

Bone: None of the genotypes was linked to bone mineral density (BMD) or past fractures in a BMD screening study (n=342) and a prospective cohort study with 1093 nursing home residents. Carriers of the G-allele of the DHCR7 polymorphism showed an increased risk for prospective fractures.

Prostate cancer: GC genotypes were not associated with biomarker based recurrence, development of metastases or overall survival in 702 prostate cancer patients. The investigated polymorphisms do not play a major role for the investigated bone and prostate cancer related endpoints.

Mortality: GC, CYP2R1 and DHCR7 genotypes predicted 25-OH-vitamin D status reliably but not all-cause-, cardiovascular-, or non-cardiovascular mortality. This suggests that low 25-OH-vitamin D concentrations are associated with, but unlikely to be causal for higher mortality rates.

1. VITAMIN D

1.1 Forms of vitamin D

The hormone vitamin D consists of several vitamers, specified in Table 1. Vitamin D_2 (ergocalciferol – the plant and yeast derived form) and Vitamin D_3 (cholecalciferol – derived mainly from synthesis in the skin and from animal sources) are the two major isoforms of vitamin D. Vitamin D_2 and vitamin D_3 are known collectively as calciferol and vitamin D without a subscript refers to either D_2 or D_3 or both. Vitamin D_2 was chemically characterized in 1932. Four years later the chemical structure of vitamin D_3 was identified and proven to result from ultraviolet irradiation of 7-dehydrocholesterol [1]. Chemically, these vitamers are characterized as secosteroids, meaning steroids in which one of the bonds in the steroid ring is broken. The structural difference between vitamin D_2 and vitamin D_3 can be found in their side chains. The side chain of D_2 contains a double bound between carbon 22 and 23 as well as a methyl group on carbon 24.

Vitamin D_1 is a compound of ergocalciferol and lumisterol at the ratio of 1:1, whereas vitamin D_4 is a saturated form of vitamin D_2, named 22-dihydroergocalciferol. The last vitamer, vitamin D_5 or sitocalciferol, is synthesised via ultraviolet light from 7-dehydrocholesterol in the skin.

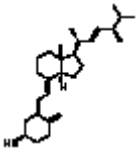
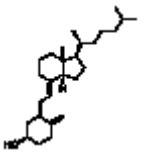
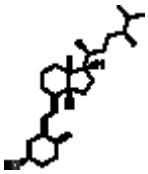
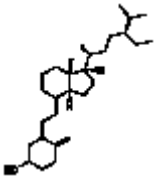
Name	Chemical composition	Structure
Vitamin D1	molecular compound of ergocalciferol with lumisterol, 1:1	
Vitamin D2	ergocalciferol (made from ergosterol)	
Vitamin D3	cholecalciferol (made from 7-dehydrocholesterol in the skin)	
Vitamin D_4	22-dihydroergocalciferol	
Vitamin D_5	sitocalciferol (made from 7-dehydrositosterol)	

Table 1: Vitamers of vitamin D, modified from 2.

1.2 Sources and Metabolism of vitamin D

Humans attain vitamin D mostly from exposure to sunlight, from their diet and from vitamin D-fortified food (e.g. vitamin D fortified milk and orange juice which is common in the US but not in Austria) or vitamin D supplements. During exposure to solar ultraviolet B (UVB) radiation, 7-dehydrocholesterol in the skin is converted to previtamin D_3, which is immediately converted to vitamin D_3 in a heat-dependent process. Excessive exposure to sunlight causes degradation of previtamin D_3 and vitamin D_3 into inactive photoproducts.

Both vitamin D_2 and vitamin D_3, are incorporated into chylomicrons and transported by the lymphatic system into the venous circulation where they can be stored in and, if required, released from fat cells.

Vitamin D as a precursor compound exerts no significant biological activity. Two hydroxylation steps are required to produce the active vitamin D metabolite 1,25-dihydroxyvitamin D_3 (1,25 $(OH)_2$ vitamin D). First, in the circulation vitamin D is bound to vitamin D-binding protein (GC group specific component), which transports the molecule to the liver, where the first hydroxylation takes place [3]. In the liver, vitamin D is converted to 25-hydroxyvitamin D (25-OH-vitamin D) by the enzyme vitamin D-25-hydroxylase (25-hydroxylase). This is the major circulating form of vitamin D that is measured to determine a subject´s vitamin D status. As already mentioned before, this form of vitamin D is biologically inactive and must be converted into the biologically active form 1,25 $(OH)_2$ vitamin D by a second hydroxylation step. The responsible enzyme for this biochemical process is the 25-hydroxyvitamin D-1α-hydroxylase (1α-hydroxylase), which is located mostly in the kidneys, but can be found in many cell systems of the human body. Renal and extrarenal 1,25 $(OH)_2$ vitamin D production supports endocrine, autocrine and paracrine functions (Figure 1). Serum levels of 1,25 $(OH)_2$ vitamin D are significantly determined by renal 1α-hydroxylase activity, which is tightly regulated by several factors related to calcium and phosphorus metabolism [e.g. stimulation by parathyroid hormone (PTH) or inhibition by fibroblast-growth factor 23], whereas dietary calcium can regulate 1α-hydroxylase directly through changes in serum calcium, 1,25 $(OH)_2$ vitamin D hampers its own synthesis through negative feedback. Further, it increases the expression of 24-hydroxylase to turn 1,25 $(OH)_2$ vitamin D into its water-soluble, biologically inactive form calcitroic acid, which is excreted in the bile [4].The regulation of 1α-hydroxylase at extrarenal sites is quite different from that of the renal enzyme. The rates of 1,25 $(OH)_2$ vitamin D synthesis and degradation are under the control of local factors such as cytokines and growth factors, that optimize the levels of 1,25 $(OH)_2$ vitamin D for these cell-specific actions through mechanisms incompletely understood [5].

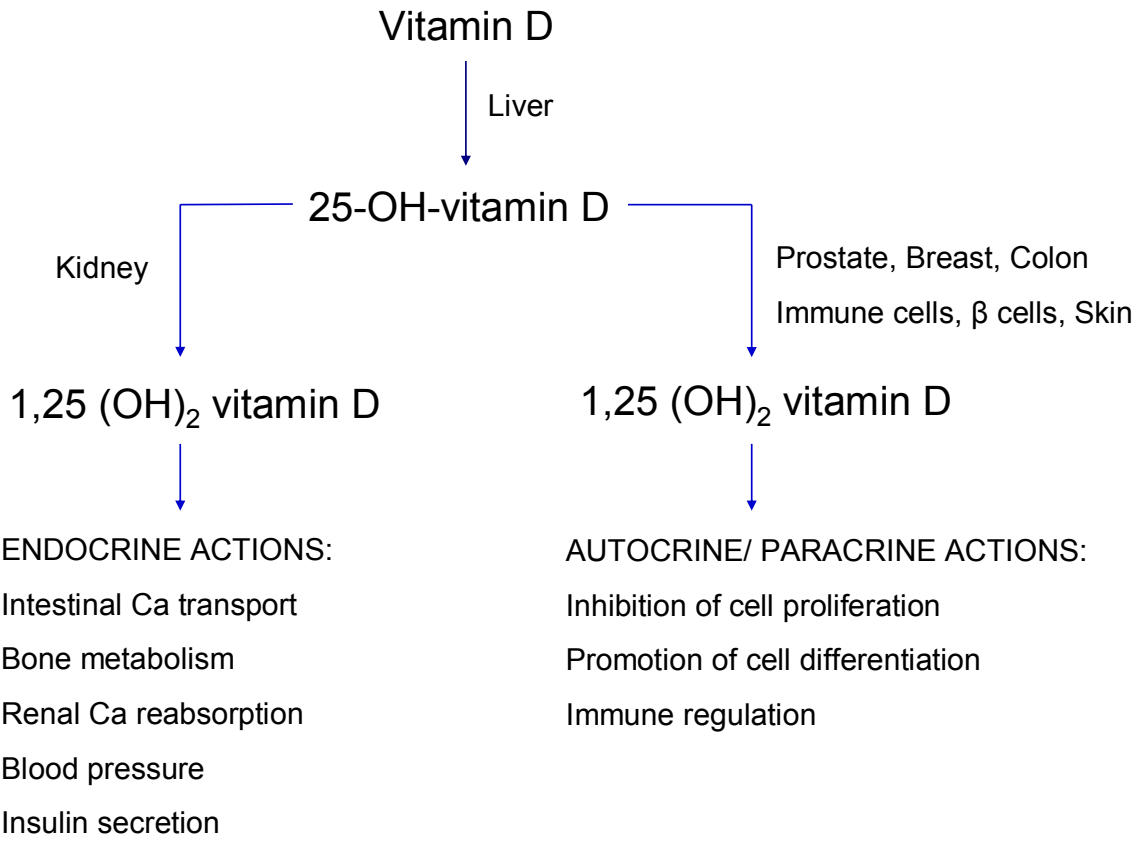


Figure 1: Endocrine, autocrine and paracrine functions from renally and extrarenally produced 1,25 $(OH)_2$ vitamin D, modified from [5].

1.3 Transport of Vitamin D

Vitamin D metabolites are lipophilic molecules with low aqueous solubility that must be bound to plasma proteins in order to be transported in the circulation. The most important of these carrier proteins is vitamin D binding protein GC, which binds the metabolites with high affinity. It binds 25-OH-vitamin D with the highest affinity, then 24.25 $(OH)_2$ vitamin D, and 1,25 $(OH)_2$ vitamin D [6]. Plasma levels of GC are 20 times higher than the total amount of vitamin D metabolites. 99% of circulating vitamin D compounds are therefore protein-bound – mostly to GC, but to a smaller extent also to albumin and lipoproteins. This has a major impact on their pharmacokinetics. GC-bound vitamin D metabolites have limited access to target cells [6] and, therefore, are less susceptible to hepatic metabolism and subsequent biliary excretion, leading to a longer circulating half-life. It is suggested that only the small fraction of unbound metabolites passively enters target cells to be further metabolized or to exert biological activity. For activated vitamin D

compounds (i.e. 1,25 $(OH)_2$ vitamin D and its analogs), biological activity was correlated with the concentration of free hormone [7]. Thus, GC appears to buffer the free levels of active vitamin D compounds, guarding against vitamin D intoxication [8]. GC levels are not regulated by vitamin D but are reduced in liver disease, nephrotic syndrome, and malnutrition and increased during pregnancy and estrogen therapy. The concentration of free 1,25 $(OH)_2$ vitamin D, however, remains constant when GC levels change, an example of the tight self-regulation of vitamin D metabolism [5].

1.4 Mechanism of action

Vitamin D metabolites exert their effects by binding to the cytoplasmatic vitamin D receptor (VDR). After ligand binding, this receptor forms a heterodimer with retinoid X receptor (RXR) and translocates to the nucleus. There, this complex interacts as a transcription factor with specific DNA (desoxyribonucleic acid) regions, called vitamin D-responsive elements (VDREs). By additional interactions with coregulatory proteins, the VDR-RXR complex regulates approximately 3% of the human genome [9].

1.5 Biological functions of vitamin D

1.5.1 Endocrine functions

Calcium balance

Maintenance of serum calcium levels within a narrow range is vital for the normal function of the nervous system, as well as for bone growth and preservation of bone density. The “vitamin D endocrine system” is an essential component of the interactions among kidney, bone, parathyroid gland, and intestine that regulates extracellular calcium levels [5].

The parathyroid gland cells sense the serum calcium level by means of calcium-sensing receptors (CaSR). When calcium levels decrease, e.g. when dietary calcium intake is inadequate, the parathyroid glands secrete parathyroid hormone (PTH) that stimulate 1α-hydroxylase activity in our kidneys and probably other

cells in the body. The increased production of 1,25 $(OH)_2$ vitamin D initiates changes in gene expression that normalize serum calcium in three different ways:

1) by activating the vitamin D-dependent transport system in the small intestine, increasing the absorption of dietary calcium by upregulating the expression of epithelial calcium channels (transient receptor potential cation channel, subfamily V, member 6 [TRPV6]) and calcium-binding proteins (calbindin 9K).

2) by increasing the mobilization of calcium from bone into the circulation when there is insufficient dietary calcium to maintain normal serum calcium levels (see Bone metabolism in the next section) and

3) by increasing the reabsorption of calcium filtered by the kidneys.

The vitamin D endocrine system is further a potent modulator of parathyroid function. Whereas vitamin D deficiency results in parathyroid hyperplasia and increased PTH secretion, 25-OH-vitamin D as well as 1,25 $(OH)_2$ vitamin D administration inhibits PTH synthesis and parathyroid cell growth, thus rendering 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D therapy effective in treating the secondary hyperparathyroidism of chronic kidney disease [10].

Bone metabolism

The renewal of bone (bone remodeling) is responsible for bone strength throughout our life. Old bone is removed by osteoclasts (resorption) and new bone is created by osteoblasts (formation). Vitamin D, when administered for severe vitamin D deficiency or in large doses, tends to increase the recruitment of osteoclasts and plays a part in the mineralization of bone matrix. A severe lack of vitamin D results in osteomalacia. Osteoblasts express the surface ligand RANKL (receptor activator of nuclear factor kB ligand), which can bind either RANK (receptor activator of nuclear factor kB) or osteoprotegerin (OPG), an osteoblast-derived soluble decoy receptor. The binding of RANKL to RANK induces a signaling cascade that results in differentiation and maturation of osteoclasts which dissolve bone matrix and mobilize calcium and other minerals from the skeleton. Adequate calcium and phosphorus levels in the blood promote the mineralization of the skeleton. RANKL expression is stimulated by 1,25 $(OH)_2$ vitamin D as well as by PTH [11]. 1,25 $(OH)_2$ vitamin D further inhibits OPG

production [12]. The regulation of calcium, phosphorus and bone metabolism is displayed in Figure 2.

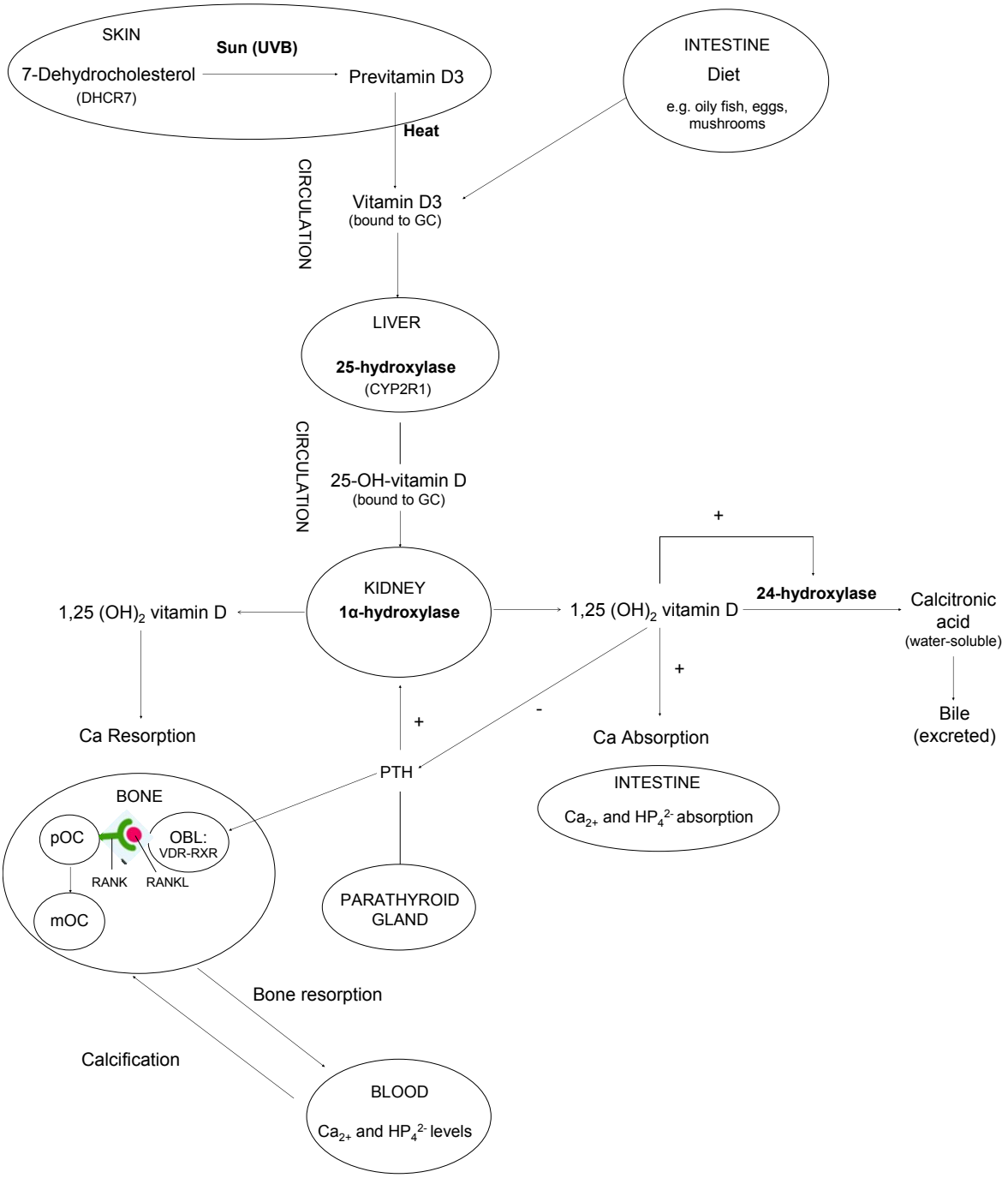


Figure 2: Synthesis and metabolism of vitamin D in the regulation of calcium, phosphorus and bone metabolism, modified from [4].
OBL: osteoblast, pOC: preosteoclast, mOC: mature osteoclast, RANKL: receptor activator of nuclear factor κB ligand, RANK: receptor for RANKL, VDR-RXR: vitamin D retinoid X-receptor complex, Ca2+: calcium, HPO42-: phosphorus, +:increase, -:decrease.

Blood Pressure Regulation

The renin-angiotensin system (RAAS) is a hormone system that plays a major role in blood pressure regulation. The enzyme renin, produced in the kidneys, catalyzes the cleavage of angiotensin I from its precursor molecule angiotensinogen, a serum globulin produced in the liver. In a second step angiotensin converting enzyme (ACE) catalyzes the cleavage of angiotensin I to angiotensin II, a peptide with most potent vasoconstriction properties that causes small arteries to constrict resulting in an increase of blood pressure. Angiotensin II stimulates further the release of the hormone aldosterone from the adrenal cortex. Aldosterone increases reabsorption of sodium and water in the tubules of the kidneys. This increases the volume of fluid in the body and thus also blood pressure.

1,25 $(OH)_2$ vitamin D act as a negative endocrine regulator of the renin-angiotensin system. Mice lacking the VDR gene (VDR-null mice), marked increases in renin expression and plasma angiotensin II production caused hypertension, cardiac hypertrophy, and increased water intake can be found. Since inappropriate activation of the renin-angiotensin system is thought to play a role in some forms of human hypertension, the achievement of adequate vitamin D levels may be important for decreasing the risk of high blood pressure [5, 13].

Insulin Secretion

As the VDR is expressed by insulin-secreting cells of the pancreas 1,25 $(OH)_2$ vitamin D has been shown to increase insulin release from pancreatic islet cells in vitro, and to improve insulin secretion in vitamin D-deficient animals [14]. Insufficient vitamin D levels may have an adverse effect on insulin secretion and glucose tolerance in type 2 diabetes [15].

1.5.2 Autocrine and paracrine functions

Anticancer effects

The local production of 1,25 $(OH)_2$ vitamin D in the breast, the colon and the prostate has been found to interfere with transduction pathways of various growth factor(s)-activated receptors (receptor tyrosine kinases). These receptors modulate transcription and alteration of genomic functions resulting in anticancer

actions such as inhibition of cell proliferation, facilitation of cell differentiation, induction of apoptosis and inhibiton of angiogenesis [16]. Some examples of 1,25 $(OH)_2$ vitamin D up and down regulated genes are given in Table 2 [17].

- **Inhibition of cell proliferation and promotion of cell differentiation**

While cellular proliferation is essential for several physiological processes like growth, tissue renewal or wound healing, uncontrolled proliferation of cells with certain mutations become pathological in all cases of tumor groth whether benign or malignant. Differentiation leads to specialization of cells for specific functions and decreases proliferation. 1,25 $(OH)_2$ vitamin D decreases cellular proliferation of both, normal cells and cancer cells and induces their terminal differentiation [18].

- **Induction of apoptosis and inhibiton of angiogenesis**

Apoptosis is a genetically programmed self-destruct mechanism, where sequence of biochemical events (caspases) lead to a well organized cell death. Beside its physiological function in maintaining constancy of cell numbers in different tissues, apoptosis also prevents the possibility of mutational changes leading to malignancy after DNA damage by removal of such damaged cells. Dysregulated apoptosis has been found to be involved in the pathogenesis of many cancers [16].

Angiogenesis, a physiological and vital process for generating new blood vessels from pre-existing vessels, is in cancerogenesis a fundamental step in the transition of tumors from a dormant state to a malignant one as angiogenesis also provides the blood vessels to tumours for feeding them with nutrients for further growth [19].

Beside cancer prevention and cell death, 1,25 $(OH)_2$ vitamin D can also cause tumor growth retardation as well as tumor regression because of its angiogenesis inhibitory action, which deprives the cancer cells of their nutrients and oxygen for growth and survival. Moreover, the young and immortal cancer cells never grow up, mature and die off [16].

1,25 $(OH)_2$ vitamin D					
positively regulated genes			negatively regulated genes		
Prostate	Colon	Breast	Prostate	Colon	Breast
p21	p21	p21	CDK 2	c-myc	CDK2
p27	p27	p27	MMP-2	PPARδ	CDK4
IGFBP-3	E-cadherin	TGF-β1		Tef-1	Cyclin D1
IGFBP-5	ZO-1	HoxB4		CD44	Cyclin A
E-cadherin	ZO-2	IGFBP-3		Cyclin D1	Cyclin D3
DUSP-10	24-hydroxylase	IGFBP-5		Cyclin E	c-myc
TRR1	CAT1	24-hydroxylase		CDK2	MMP9
SOD2	CAT2			CDK6	
AR	c-Jun				
24-hydroxylase	JunB				
Metallothioneins	JunD				
	FREAC-1/Fox1				
	ZNF-4/KOX7				
	Plectin				
	Filamin				
	K13				
	Kallikrein 10				
	Protease M				

Table 2: Examples of 1,25 $(OH)_2$ vitamin D-regulated genes in prostate-, colon- and breast cancer cells. Positively and negatively expression of these genes were confirmed by immunohistochemistry, Northern or Western blotting techniques, modified from [17].

Immune regulation

Vitamin D in the form of 1,25 $(OH)_2$ vitamin D is a potent immune system modulator. The VDR is expressed by most cells of the immune system, including T cells and antigen-presenting cells, such as dendritic cells and macrophages [20].Macrophages are able to produce 1α-hydroxylase to locally generate 1,25 $(OH)_2$ vitamin D [21]. There is a considerable number of publications reporting a variety of effects of 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D on the immune system function such as enhancing innate immunity and inhibiting the development of autoimmunity [22, 23]

1.6 Vitamin D deficiency

Definition

The Endocrine Society defines vitamin D deficiency as a 25-OH-vitamin D level below 20 ng/ml (50 nmol/l) and vitamin D insufficiency as a 25-OH-vitamin D level of 21-29 ng/ml (52.5-72.5 nmol/l). It should be measured by a reliable assay, to evaluate vitamin D status in subjects who are at risk of vitamin D insufficiency.

Using serum 1,25 $(OH)_2$ vitamin D assays for this purpose is not advisable. It should be used only for monitoring certain conditions such as acquired and inherited disorders of vitamin D and phosphate metabolism.

Further, the Endocrine Society recommends in 2011 prescribing vitamin D supplementation for fall prevention and bone health but not for the purpose of achieving non-skeletal benefits like prevention of cardiovascular disease or death or improving quality of life [24].

Prevalence

Vitamin D insufficiency affects almost 50% of the population worldwide [4]. An estimated 1 billion men worldwide, across all ethnicities and age groups, show a vitamin D deficiency [25]. Vitamin D insufficiency is common in Australia, the Middle East, India, Africa, South America and Europe [4, 26, 27, 28]. 26% of the Austrian population have vitamin D deficiency with serum levels <12 ng/ml [29].

Causes

The major source of vitamin D is exposure to natural sunlight [30]. Very few foods naturally contain or are fortified with vitamin D. Thus, the major cause of vitamin D deficiency is inadequate exposure to sunlight [31, 32]. Wearing a sunscreen with a sun protection factor of 30 reduces vitamin D synthesis in the skin by more than 95% [33]. People with a naturally dark skin tone have natural sun protection and require sun exposure to be at least three to five times longer in order to produce the same amount of vitamin D as a person with a white skin tone [34].

In addition, there is an inverse association of serum 25-OH-vitamin D and body mass index (BMI), and thus, obesity is associated with vitamin D deficiency [35]. Further, decreased hydroxylation of vitamin D to 25-OH-vitamin D (liver diseases, mutations of the 25-hydroxylase gene, drugs) and or decreased 1α-hydroxylation of 25-OH-vitamin D to 1,25 $(OH)_2$ vitamin D (chronic kidney disease, mutations of the 1α-hydroxylase gene, drugs) may be, among others, rare causes of vitamin D insufficiency.

Skeletal consequences

As described in more detail in the previous chapter, vitamin D is essential for bone health. Vitamin D deficiency results in abnormalities in calcium, phosphorus, and bone metabolism. A decrease in the efficiency of intestinal calcium and phosphorus absorption of dietary calcium and phosphorus, resulting in an increase in PTH levels, is the consequence [36, 37]. The so-called secondary hyperparathyroidism maintains serum calcium in the normal range at the expense of mobilizing calcium from the skeleton. An increase in bone resorption is the consequence. The PTH-mediated rise in osteoclastic activity leads to bone weakness and causes a decrease in bone mineral density (BMD), resulting in osteopenia and osteoporosis. Further secondary hyperparathyroidism is accompanied by phosphaturia resulting in low serum phosphorus level, increasing the mineralization defect in the skeleton. In young children who have little mineral in their skeleton, this defect results in a variety of skeletal deformities classically known as rickets. In adults, the epiphyseal plates are closed, and there is enough mineral in the skeleton to prevent skeletal deformities so that this mineralization defect, known as an osteomalacia, often goes undetected. However, osteomalacia causes a decrease in BMD and is associated with isolated or generalized pain in bones and muscles [38, 39]. Vitamin D deficiency also causes muscle weakness; some affected children have difficulties in standing and walking [40, 41], whereas the elderly have increasing sway and more frequent falls [42, 43], thereby increasing their risk of bone fractures.

2. DNA – THE SECRET OF LIFE

The discovery of the DNA – a scientific breakthrough

Figure 3: Francis Crick and James Watson, 1953. Photo: Cold Spring Harbor Laboratory Archives.

The sentence "This structure has novel features which are of considerable biological interest" may be one of science's most famous understatements. It appeared in April 1953 in the scientific paper [44]. where James Watson and Francis Crick presented the structure of the DNA-helix, the molecule that carries genetic information from one generation to the next. Nine years later, in 1962, they shared the Nobel Prize in Physiology or Medicine with Maurice Wilkins for solving one of the most important biological riddles [45.

But in fact it was the work of many scientists that paved the way for the exploration of the DNA. Almost a century before in 1868, a young Swiss physician named Friedrich Mischer isolated an unknown compound from the nuclei of cells that he named “nuclein”, today commonly known as nucleic acid [46].

Two years earlier 1866 the Czech monk Gregor Mendel had finished a series of experiments with peas. His observations turned out to be closely connected to the finding of nuclein. In his studies Mendel showed that certain traits in the peas, such as shape or color, were inherited in different “packages”, which we call genes today [47].

For a long time the connection between nucleic acid and genes was not known. In 1944 the American scientist Oswald Avery managed to prove that genes were made of nucleic acid when transferring the ability to cause disease from one strain of bacteria to another. The previously harmless bacteria could now pass the new trait along to the next generation. What Avery had transferred was nucleic acid.

In the late 1940s the members of the scientific community were aware that DNA was most likely the key molecule to life, even though many were doubtful since it seemed so simple. It was understood that DNA had to contain different amounts of

the four bases adenine (A), thymine (T), guanine (G) and cytosine (C) but nobody had the slightest idea of what the molecule might look like [48].

To solve the puzzle “structure of DNA” Watson and Crick had to combine a couple of distinct pieces of information discovered from different scientists like Rosalind Franklin, Maurice Wilkins, Linus Pauling, and Erwin Chargoff. Franklin discovered that the phosphate backbone of DNA lies on the outside of the molecule not on the inside as was previously thought and that the shape of the molecule was a double helix [49]. It's commonly believed that James Watson and Francis Crick discovered the double helix shape of DNA. But in fact, they based their work on their colleagues´ results at King's College in London. It was also important to understand that two strands run in opposite directions and the specific base pairing inside the molecule [50]. Watson and Crick used stick-and-ball models to test their ideas on the possible structure of the DNA. Finally they set up a valid model containing all the information available at the time. This structure published in their famous Nature paper in April 1953, explained how the DNA molecule could replicate itself during cell division, enabling organisms to reproduce with amazing accuracy except for occasional mutations [51].

The structure of the DNA

The DNA molecule is built up of two chains of several thousands of linked nucleotides, which are twisted like a double helix (Figure 4). Each nucleotide consists of a phosphate group, a deoxyribose and a nitrogenous base. 4 bases occur in the DNA: the purine bases cytosine and thymidine and the pyrimidine bases adenine and guanine. The two strains of the molecule run anti-parallel, meaning they run in opposite directions (5´ to 3´ and 3´ to 5´). The sugar phosphate bone is on the outside of the helix and the bases are on the inside. The backbone looks like a ladder, whereas the bases in the middle form the rungs of the ladder. Each rung of the ladder is made up of two corresponding bases connected by hydrogen bonds: adenine and thymine are connected by two hydrogen bonds, whereas cytosine and guanine build up a (stronger) triple hydrogen bond. The base pairing is thus restricted. This restriction assures the duplication of the DNA in the chromosomes preceding cell division (replication), the transfer of information from the DNA template to RNA (ribonucleic acid) codes

(transcription) which can then be translated into specific proteins on the ribosome in the cytoplasm (translation).

Unfolding the DNA molecule is required for replication. The DNA double helix gets temporarily unzipped in two long stretches of sugar-phosphate backbone with a line of free bases sticking up from it like the teeth of a comb. Each half will then be the template for a new, complementary strand.

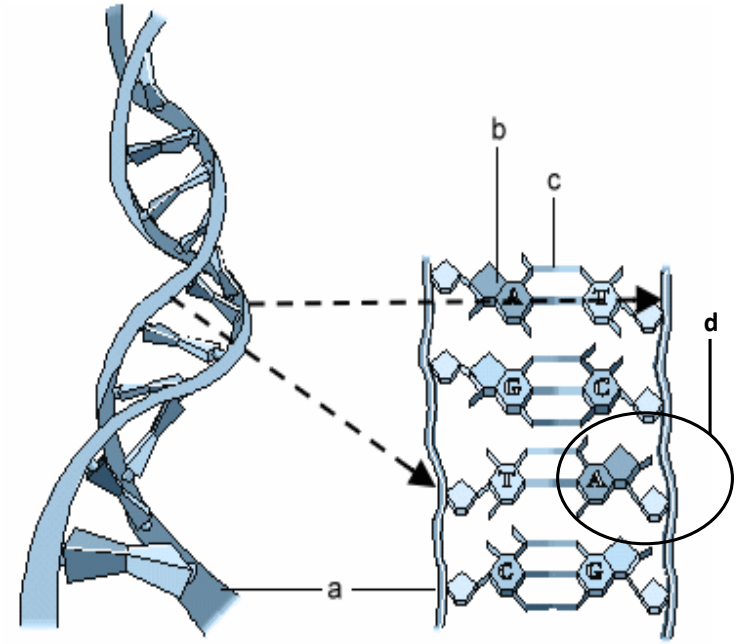


Figure 4: Structure of the DNA molecule, modified from [52].

a: sugar-phosphate backbone, b: position of the bases
c: hydrogen bonds, d: nucleotide

3. GENETIC EPIDEMIOLOGY

Genetic epidemiology focuses on the identification of genetic information (markers) to improve health and prevent disease in families as well as in the general population. As a part of public health, it is influencing health policy decisions and evidence-based medicine by identifying genetic risk factors for diseases and targets for preventive medicine.

3.1 Basic genetic background

This section describes basic genetic background information to understand the principles of genetic epidemiology.

Single nucleotide polymorphism (SNP)

In the last 20 years knowledge of the human genome and polymorphisms (genetic variation in the DNA sequence on a certain locus) increased. In general, DNA variations with a frequency of ≥ 0.01 (1%) are defined as polymorphisms whereas variations with a frequency < 0.01 as mutations. Any two human genomes differ in millions of polymorphisms. These variations in the genomes may cause alterations in an individual's traits, or phenotype, which can be anything from disease risk to physical properties such as height [53]. The most frequent class of polymorphism is the single nucleotide polymorphism. This naturally occurring genetic variability describes a variation in a single base at a certain position (locus) in the DNA sequence. On average, SNPs are believed to be present every 300 to 1000 DNA bases. Thus, several million nucleotide positions are exchanged in the human genome. Larger variations such as deletions (a missing part of a chromosome or DNA sequence), insertions (addition of one or more nucleotide base pairs (bp), short tandem repeat (repeating sequences of 2 to 6 base pairs of DNA) or copy number variations (an abnormal number of copies of one or more sections of DNA due to alterations of the DNA in the cell) are less frequent.

SNPs usually occur more frequently in non-coding regions or intergenic regions than in coding regions. SNPs within a coding sequence do not necessarily change the amino acid sequence of the produced protein, due to degeneration of the genetic code. A SNP, in which both possible bases produce the same polypeptide sequence and therefore an unchanged protein, is called a synonymous or silent

polymorphism. If a different protein sequence is produced the type of polymorphism is a non-synonymous or replacement polymorphism. Replacement polymorphisms may be either missense, resulting in a different amino acid, or nonsense with the consequence of generating a premature stop codon within the coding region of the gene for a protein. Over half of all known disease mutations come from replacement polymorphisms [54]. SNPs that are not in protein-coding regions may affect other gene functions like splicing, transcription factor binding, mRNA (messenger ribonucleic acid) degradation, or the sequence of non-coding RNA.

A SNP is characterized by its genotype, which consists of the two possible bases at a certain locus, so-called alleles, one for each chromosome in the diploid genome. The more frequent allele in a population is called the wild type ("normal") allele, whereas the less frequent version is referred to as the mutant allele or a "variant". Therefore, a genotype can be homozygous (both alleles carry the same nucleotide) for the wild type or the mutant allele whereas a heterozygous (each allele carries a different nucleotide) genotype shows one wild type and one mutant allele.

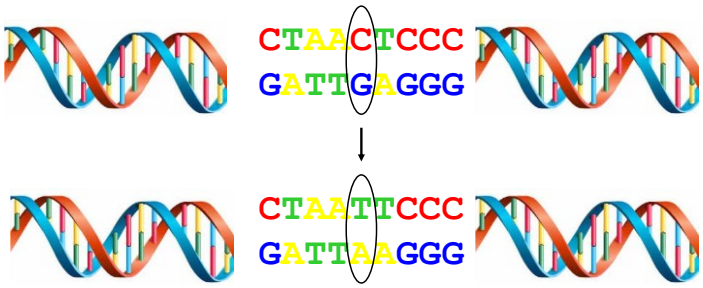


Figure 5: Scheme of a C/T single nucleotide polymorphism, modified from 55.

Mostly, but not always, the mutant variant is disease-related. Some phenotypes (the physical appearance of an organism as a result of the interaction of its genotype and the environment) need both mutated alleles (recessive inheritance) to develop a certain phenotype, whereas in others one mutated allele (dominant inheritance) is sufficient.

The factor V Leiden polymorphism (1601G>A, R506Q), for example, first described in Leiden (Netherlands) by Bertina et al. [56] in 1994, is the most common hereditary hypercoagulability disorder amongst Eurasians [57]. In the DNA sequence at position 1691 (=locus) a guanine (G) is replaced by an

adenosine (A) resulting in an amino acid change from glutamine (Q) to arginine (R) at position 534 in the protein. This amino acid change happens in the gene encoding the coagulation factor V, a member of the coagulation cascade. Normally this protein gets partly cleaved by activated protein C causing an inactivation of this factor for preventing hypercoagulation. The described amino acid exchange leads to resistance of Factor V to activated protein C (APCR, activated protein C resistance), resulting in a higher risk for venous thrombosis [58].

Genotype	Frequency	Commentary
GG	92%	Homozygous for both wild type alleles. No Factor V Leiden variant detectable. Normal coagulation.
GA	8%	Heterozygous for one Factor V Leiden allele. One mutated and one wild type allele. Half of all synthesized Factor V molecules is physiological and half cannot be inactivated. The relative risk of venous thrombosis is increased approximately 3- to 8-fold.
AA	< 0.1%	Homozygous for both factor V Leiden alleles. Two mutated alleles. All synthesized Factor V molecules cannot be inactivated. The relative risk of venous thrombosis is therefore increased approximately 20- to 80-fold.

Table 3: Factor V Leiden genotypes and their risk for venous thrombosis, modified from 59.

Gene dose effect

In general, the gene product in wild types is considered physiological. If a person is heterozygous for a certain polymorphism one allele is “normal” and one may produce a changed amount or impaired gene product (e.g. a protein). Homozygosity for the mutated allele of a functional SNP implies that no or impaired protein is translated. This effect is described as gene dose effect.

The recently discovered LCT-13910 C/T polymorphism [60], one of the genetic causes for adult type of lactose intolerance in Europeans, is an example of a classical gene dose effect. Lactose intolerance (adult type hypolactasia) is the inability to digest and absorb lactose resulting in gastrointestinal symptoms like

abdominal pain, diarrhoea or flatulence when milk or dairy products are ingested. In order for lactose to be absorbed from the intestine it must first be split by the enzyme lactase into glucose and galactose. This enzymatic reaction happens on the surface of the cells that line the small intestine. Lactose intolerance is thus caused by reduced or absent activity of lactase that prevents the splitting of lactose due to the LCT-13910 C/T polymorphism. About 40% of Austrians show a wild type genotype (TT) and produce normal amounts of lactase. Individuals with one effected allele (heterozygous CT genotype with a frequency of 45%) produce half the amount of lactase, which is still sufficient for digesting a normal amount of milk consumption. Patients with two affected alleles (homozygous CC genotype with a frequency of 15%) do not produce any lactase and should therefore avoid milk containing products to circumvent clinical symptoms. This more or less linear relationship between the phenotype and the number of mutated alleles is, as mentioned above, defined as gene dose effect.

SNPs as markers

SNPs have properties and frequencies in the human genome that make them attractive as markers or tools for identification of genes in as yet uncharacterized parts of the genome that may have some relation to specific diseases. There are great expectations that SNPs will be useful in identifying candidate genes that contribute to population-wide, polygenic diseases.

At present several initiatives are ongoing to exploit the information contents of genetic variability. Their purpose is to

- identify genes that contribute to diseases,
- identify gene targets for development of new therapeutic principles,
- identify genes that may predict outcome from therapy [61].

SNPs are used as essential tools in different concepts of genetic epidemiology.

Concepts of genetic epidemiology

Linkage Analysis

Genetic linkage analysis can be used to identify regions of the genome which contain genes that predispose to traits or diseases.

Linkage[1] and linkage disequilibrium[2] are two key concepts in genetic epidemiology. Two loci in linkage disequilibrium will also be linked, but the reverse is not necessarily true. Linkage analysis is often the first stage in the genetic investigation of a trait. It can be used to identify broad genomic regions that might contain a disease gene, even in absence of previous biologically driven hypotheses. If a trait is often passed to offsprings along with a genetic marker (e.g. a SNP), then it can be concluded that gene(s) which are responsible for the trait are located close to these genetic markers on the chromosome. By genotyping these genetic markers and studying their segregation through pedigrees, it is possible to infer their position relative to each other on the genome. This process can be done to map genetic markers or to map disease or trait loci. Nowadays, there exist many sets of linkagemapping markers, in which the markers have been selected to be regularly spaced across the genome (for example, the Marshfield Clinic resource) [62].

Genetic association study

Genetic association studies aim to discover association(s) between one or more genetic polymorphisms and a trait, which might be some quantitative characteristic or a discrete attribute or a disease. An association study differs from a linkage study in that the same allele (or alleles) is (are) associated with the trait in a similar manner across the whole population, while linkage allows different alleles to be associated with the trait in different families. Nevertheless, genetic associations arise because of the fact that human populations share common ancestry and it

1 **Genetic linkage** is the tendency of genes that are located proximal to each other on a chromosome to be inherited together during meiosis. Genes whose loci are nearer to each other are less likely to be separated onto different chromatids during chromosomal crossover, and are therefore said to be genetically linked.

2 Two genetic loci are in **linkage disequilibrium** if, across the population as a whole, are found together on the same haplotype more often or less often than would be expected from a random formation of haplotypes from alleles based on their frequencies (Hardy Weinberg equilibrium).

has been argued that association studies are a special form of linkage study in which the extended family is the wider population.

Genetic susceptibility to common complex disorders probably involves many genes, with smaller effects. This fact, together with the identification of large numbers of SNPs throughout the genome and rapidly decreasing genotyping costs, has led to the importance of association studies in genetic epidemiology.

There are three reasons why association between a genetic marker and a trait might occure in a given population: (1) the polymorphism has a causal role, (2) the polymorphism has no causal role but is associated with a nearby causal variant, or (3) the association is due to some underlying stratification or admixture of the population (bias, see 3.2 The good and the bad fellows of epidemiologic studies).

Direct association:

This type of study is the easiest to analyse and the most powerful, but the difficulty is the identification of candidate polymorphisms. A mutation in a codon which leads to an aminoacid change is a candidate causal variant. However, it is likely that many causal variants responsible for heritability of common complex disorders will be non-coding.

Indirect association:

The polymorphism is a surrogate for the causal locus. This type of association allows us to search for causal genes in an indirect manner. However, indirect associations are weaker than the direct ones and usually it will be necessary to type several surrounding markers to have a high chance of picking up an indirect association.

Confounded association:

The final type of association is that due to confounding by stratification and admixture (substructure) within the population. Confounding, as in the rest of epidemiology, raises the possibility both of generating false findings (positive confounding) or obscuring true causal associations (negative confounding). In the

next chapter the term “confounding” is described in to more detail (see “The good and the bad fellows of epidemiologic studies”).

Genome-wide association studies

GWA studies provide the possibility to analyze millions of genetic variants (typically SNPs) in one analysis in different individuals to investigate if any variant is associated with a specific trait.

In addition to the concept itself other factors made GWA studies possible. One was the advent of biobanks, which are repositories of human genetic material which have greatly reduced the cost and difficulty of collecting sufficient numbers of biological specimens for study [63]. Another was the International HapMap Project which from 2003 onwards has identified a majority of the common SNPs which are interrogated in GWA studies [64]. The haploblock structure[3] identified by the HapMap project allowed the focus on the subset of SNPs that would describe most of the variation [65, 66]. Further, the development of the methods to genotype SNPs using genotyping arrays was an important prearrangement [67].

GWA studies compare DNA of people affected with a disease (cases) with DNA of healthy people (controls). The DNA of all participants is investigated using SNP arrays.

A SNP array or SNP chip is a collection of microscopic DNA spots attached to a solid surface. They are used to genotype multiple regions of the genome simultaneously. Each DNA spot contains picomoles of specific DNA sequence, so called probes (oligos). These short specific sequences, comprising one specific SNP each, are used to hybridize the restricted, amplified and fluorophore labelled target sample (DNA) under high stringency conditions. Probe-target hybridization is quantified by detection of fluorescence to determine relative abundance of nucleic acid sequences in the target [68]. The two most prominent providers of DNA chips are Affymetrix and Illumina.

[3] **Haplotype block:** is a combination of alleles at adjacent loci on the chromosome that are transmitted together. A haplotype may be one locus, several loci, or an entire chromosome depending on the number of recombination events that have occurred between a given set of loci. In context with SNPs a haplotype is a set of SNPs on a single chromosome of a chromosome pair that is statistically associated. It is assumed that these associations, and the identification of a few alleles of a haplotype block, can unambiguously identify all other polymorphic sites in its region. Such information is very valuable for investigating the genetics behind common diseases, and has been investigated in the human species by the International HapMap Project.

The exact number of SNPs per chip depends on the study, but is typically up to a million. Each of these SNPs is genotyped in one single run by loading a chip with the participants DNA. Aim of GWA studies is to investigate if allele frequencies are significantly altered between the case and the control group [69]. Associated SNPs are considered to mark a region of the human genome which influences the risk of disease. One should keep in mind that GWA studies identify SNPs in the DNA, which are associated with a disease but cannot on their own specify which genes or variants are causal [70, 71]. In other words these SNPs may not be the causal ones but rather be in linkage disequilibrium with them. In this analysis the fundamental unit for reporting effect sizes is the odds ratio (OR)[4]. Additionally, a p-value[5] for the significance of the odds OR is typically calculated using a simple chi-squared test[6] [72].

In contrast to methods which specifically test one or a few genetic regions, the GWA studies investigate the entire genome. The approach is therefore non-candidate-driven in contrast to gene-specific candidate-driven studies. GWA studies typically identify hypothesis-free new genetic markers in discovery cohorts, which should be validated in independent replication cohorts.

The results of ORs and p-values can be illustrated in a Manhattan plot. This plot shows the negative logarithm of the p-value as a function of genomic location. Thus the SNPs with the most significant association will stand out on the plot, usually as stacks of points because of haploblock structure. Importantly, the p-value threshold for significance is due to the hypothesis-free approach of GWA studies, corrected for multiple testing issues. The exact thresholds vary by study, but typically p-values must be very low (10 to the power of -7 or -8) to be considered significant in the face of the millions of tested SNPs.

---

4 The **odds ratio** reports the ratio between two proportions: the proportion of individuals in the case group having a specific allele, and the proportions of individuals in the control group having the same allele. When the allele frequency in the case group is much higher than in the control group, the odds ratio will be higher than 1, and vice versa for lower allele frequency.

5 The **p-value** is a marker of the probability in statistical hypothesis testing. This value displays the propability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis (there is no difference between two observed samples) is true. A more detailed description is given in the next chapter (Statistical significance).

6 The **chi-squared test** is any statistical hypothesis test used to compare observed data with data we would expect to obtain according to a specific hypothesis. The null hypothesis, which states that there is no significant difference between the expected and observed result, is valid until the p-value reaches significance (usually $p \leq 0.05$). In other words it displays how much deviation can occur before the investigator must conclude that something other than chance is at work, causing the observed to differ from the expected.

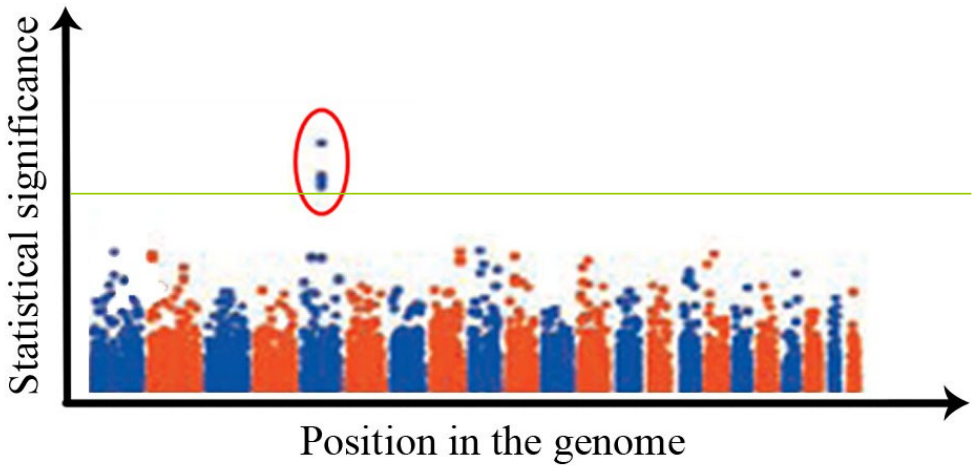


Figure 6: This Manhattan plot shows the statistical association between SNPs and a disease. Dots represent the SNPs and the bands of different tones different chromosomes. In this figure, an association between chromosome 5 SNPs and the disease is presented (red circle). The green line represents the level of statistical significance of the p-value, modified from [73].

Once new genetic associations are identified, researchers can use the information to develop better strategies to detect, treat and prevent the disease, typically common diseases such as asthma, cancer or diabetes [74].

3.2 The good and the bad fellows of epidemiologic studies

When interpreting study results some important definitions and possible errors in measurement should be taken into account, which might influence results drastically.

Causality/Causation

Causality/Causation is the relationship between one or more variables/events (the "causes", e.g. exposure) and a second variable/event (the "effect", e.g. an outcome), where the second event is understood as a consequence of the first. Anything that leads to an effect (directly or indirectly) is a factor of that effect.

Exposure ⟶ Outcome

Table 4: Causality: the relationship of exposure and outcome.

Reverse causation

Reverse causation describes the phenomenon if an assumed cause is not the cause but the consequence of an effect. For example elevated homocystein levels are associtated with coronary artery disease. It seems that the increased biomarker causes the disease but in fact arteriosclerosis leads to elevated levels of homocysteine and not vice versa. In an observation, we often cannot be sure which came first the chicken (exposure or biomarker) or the egg (outcome or disease): did a chicken cause the egg to come into existence or was it the egg that caused the chicken to come into existence? Scienticts have to be aware of reverse causality when interpreting results of associated variables in observational studies.

Association

An association is any relationship between two measured quantities (e.g. variables for exposure and outcome) that renders them statistically dependent [75] whereas a correlation refers to a linear relationship between two quantities. One should be aware that neither the term association nor a correlation implies causality.

A genetic association is the occurrence of two or more traits in a population of individuals, of which at least one is due to a genetic variation. Genetic association studies work with the advantage that the genetic factor can always be seen as direct factor (respectively as a part of the cause) leading to the effect. It should be taken into account, however, that the genetic factor might play an underpart or small role in comparison to other factors like environmental factors or lifestyle in explaining the whole phenotype/ trait.

Exposure ——— Outcome

Table 5: Association: the relationship of exposure and outcome.

The findings of any associations in epidemiological studies may be due to four possibilities:

1. a true effect (causal relationship)
2. chance or random error,

3. bias
4. confounding

While the results may reflect a true effect of an exposure(s) on the development of the outcome under investigation, it should always be considered that the findings might be in fact due to an alternative (2-4) explanation [76]. These errors may not only lead us to conclude the existence of a valid statistical association when it does not exist but also, alternatively, to the absence of an association when one is truly present. Bias and confounding are types of error that must be managed in the collection, analysis, and interpretation of research data [77].

A famous example to demonstrate the difference of associaton and causation is the following: The number of storks per year nesting in small villages of a given country and the number of newborns in these villages are clearly associated. So one may conclude that babies are brought by storks. A closer look reveals that the number of storks as well as the number of newborns reflects the size of a village: a larger village has more families producing more newborns and has more roofs allowing more storks to nest [78].

Statistical significance

When used in statistics, the word significant does not mean important or meaningful, as it does in everydays speech.

Statistical significance refers to whether any difference observed between groups being studied is a real difference or is just arises by chance. The fundamental challenge is that any hypothesis is prone to random error. In statistical testing a result is statistically significant if the observed difference is so extreme that such a result would be expected to occure only due to chance is highly unlikely. Hence, the result provides enough evidence to reject the hypothesis of “no difference” (null hypothesis, H0).

The level of significance is usually depicted by the greek symbol α. The level of significance is selected a priori to statistical analysis of a study. Popular levels of significance are: 10% (0.1), 5% (0.05), 1% (0.01), 0.5% (0.005), and 0.1% (0.001), for many applications, a level of 5% is chosen by convention. The lower the significance level chosen, the stronger the evidence required. If a test of

significance gives a p-value lower than the significance level than the H0 is rejected. Such results are informally referred to as “statistically significant” [79].

Bias

A bias is an error caused by systematically favouring some outcomes over others [80]. They can occur in any research of genetic epidemiology and reflect the potential that the sample studied is not representative of the general population [81]. The best way to avoid bias is a carefully considered study design.Some common types of bias as examples:

Information bias: A systematic distortion or error that arises from the procedures used for classification or measurement of the disease, the exposure or other important variables.

Recall bias: Retrospective cohort studies are prone to recall biases, a subtype of the information bias. People who suffer from an outcome of interest are more likely to rember certain antecedents, or exaggerate or minimise what they consider to be risk factors. For example, investigators studying an adverse health condition of unknown cause may question a population with the condition and a control population without the condition about an exposure plausibly linked with the health effect (case control study). Those who have the condition may be more likely to recall and report an exposure to a suspected agent as a consequence of experiencing, and desiring an explanation for, an adverse health condition. Conversely, the unaffected population may be more likely to forget about exposures of interest. The recall bias becomes especially important in retrospective studies when the interval between the event and the administration of the study questionnaire is long [77].

Selection bias: Because it is not possible to include all individuals with a specific disease or exposure in a research study, a sampling of the population must be selected. Selection bias describes the differential selection of study subjects that may obscure or exaggerate a causal association [82]. These subjects may differ systematically in factors such as health behaviour, socioeconomic status, educational level, age, or ethnicity. A classic example is that people who respond to questionnaires tend to be fitter and more motivated then those who do not. As a

consequence the persons participating in the study tend to be fitter and do not represent the general population.

Publication bias: Publication bias occurs when the decision to publish a study is influenced by the direction of the study results. Although this type of bias does not affect the results of individual studies, it can be apparent in reviews, most notably meta-analyses, and, less formally, in the general understanding of either an exposure effect or treatment effect based on published literature [77].

Confounding

Confounding is an association between two variables of (e.g., an exposure and an illness) that appears to be direct and causal. However, in fact the association only exists because of one or more unconsidered variable(s) and thus leads to spurious interpretation of their relationship [83]. A confounding variable is independently associated with both, the variable of interest pretending to be causal for the outcome and the outcome of interest itself. The effects of confounding can result in:

- An observed difference between study populations when no real difference exists.
- No observed difference between study populations when a true association does exist.
- An underestimate of an effect.
- An overestimate of an effect.

Statistical models should get adjusted for potential confounding factors. For example the outcome weight should be adjusted for sex (men are heavier than women), age (adolescents are heavier that infants) and height (tall people are heavier that small people).

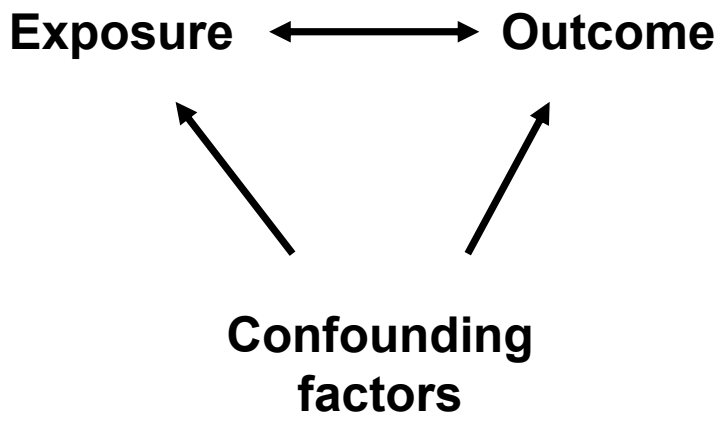


Table 6: Confounding: the relationship of exposure, outcome and confounding factors.

Effects of small magnitude

Biomedical studies sometimes come across small effects of risk factors and interventions, especially in genetic research. GWA studies discover mostly risk variants with ORs <1.2 [84, 85]. Since large cohorts, biobanks, consortia and extensive meta- analyses are increasing the situation where very small effects reach nominal statistical significance become increasingly common. Large effects as the association between smoking and lung cancer in lifestyle epidemiology are rare [86]. Very small effects may still have clinical and public health relevance, when they affect a large segment of the population, or when many small effects act together (composite risk). Simplified interpretation of small effects may be misleading [87], but effects of risk factors and interventions, regardless of their magnitude are important, but have to be interpreted accordingly [86].

3.3 Study types in epidemiology

Epidemiology (Greek; *epi*: on or upon, *demos*: people and *logos*: the study of) is the study of determinants of health-related aspects in specified populations, their distribution and the application to the control of health problems [88]. This basic science of public health is a highly quantitative discipline based on principles of statistics and research methodologies. In other words epidemiologists study the distribution of frequencies and patterns of health events within groups in a population. They use descriptive epidemiology, which characterizes health events in terms of time, place, and person [89]. A further application is the search for causes or factors that are associated with increased risk or probability of disease.

A differentiation between the two major study types “observational” and “experimental” studies is illustrated in Figure 7 [90]. In the following section both types of studies are going to be elucidated. The degree of evidence of the different study types is displayed in Figure 8 [91].

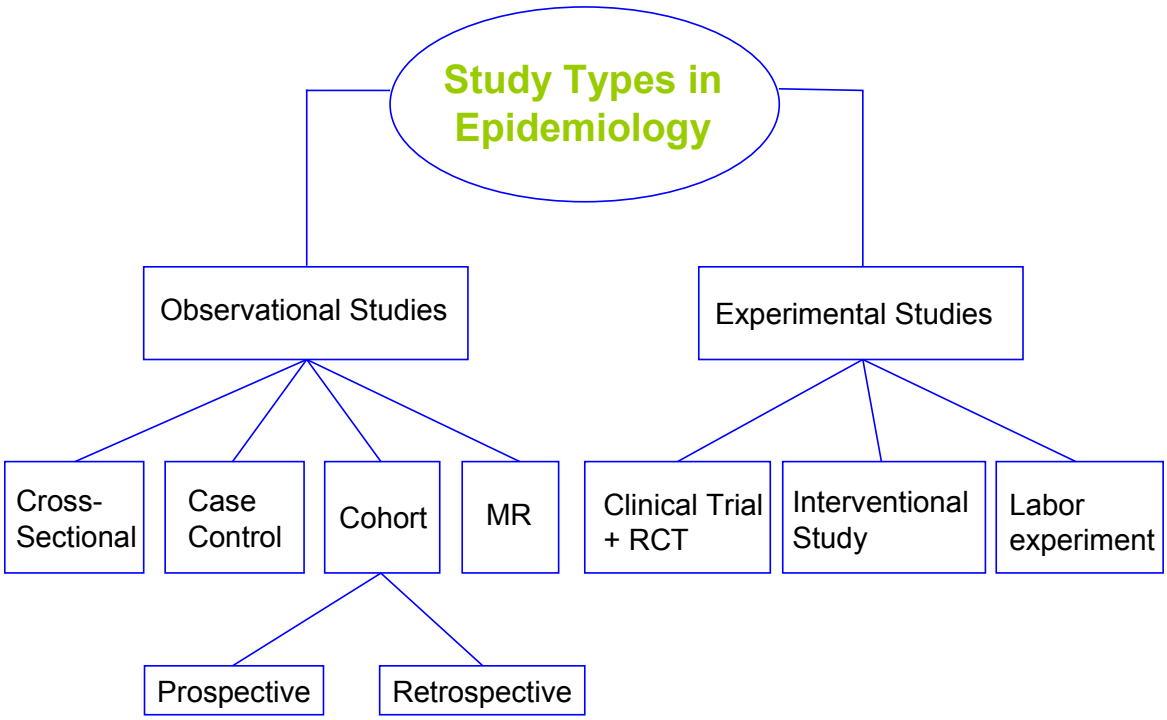


Figure 7: Study types in epidemiology, modified from [90].
MR: Mendelian Randomization, RCT: Randomized Controlled Trial.

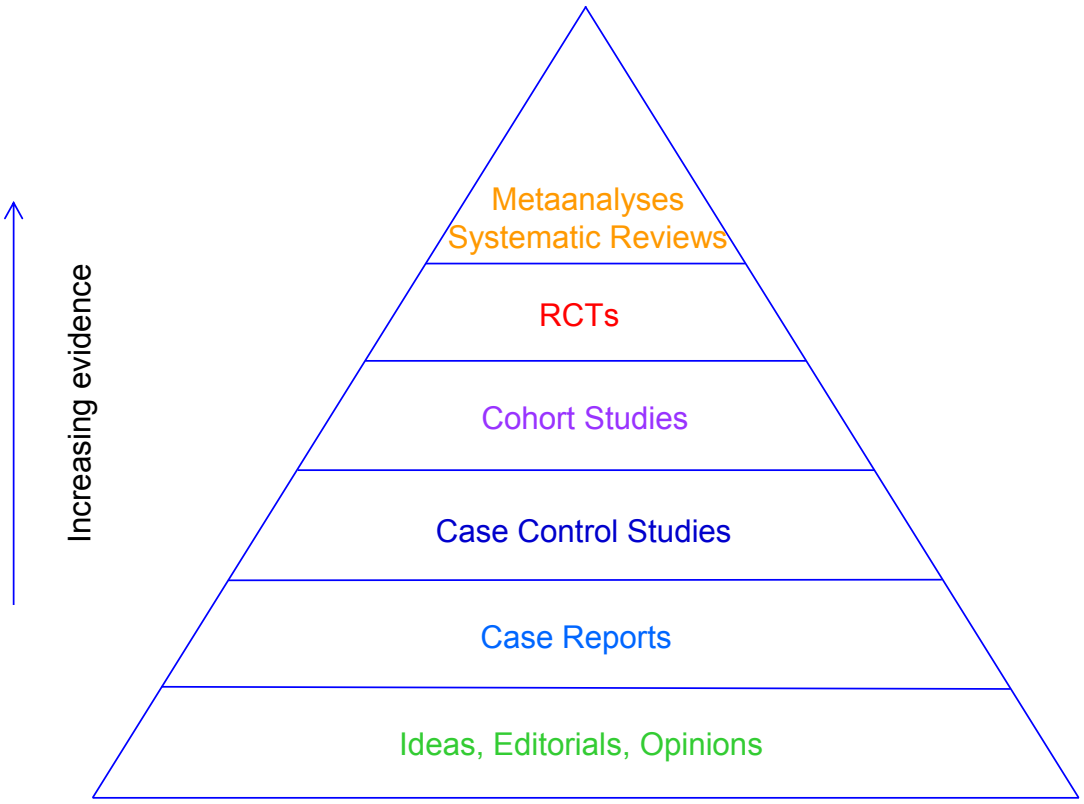


Figure 8: Increasing evidence in epidemiological studies based on the ability to control for bias and to demonstrate cause and effects in humans. The concept of Mendelian Randomization is unaccounted in this illustration, modified from [91]. RCT: randomized controlled trial

3.3.1 Observational Studies

This group of studies is also referred to as non-experimental studies because they are based on observations that take place under natural conditions and should discover relationships between changes that already exist without any manipulation or experimentation. The investigator can only observe the effect of the exposure on the study subjects; he or she plays no role in assigning exposure to the study subjects. This makes observational studies much more vulnerable to methodological problems [92]. A short overview of this study type is given in Table 7 [93]. An example may be the retrospective investigation if playing video games with violent content makes kids more prone to violent behaviour. Parents then may be asked how often their children play video games and whether they play games with a violent background. Afterwards, one might let the children play in a controlled environment with other children. The children should be observed to identify any violent behaviour.

Observational studies have brought important insights into disease aetiology [94]. Appropriate use permits investigation of prevalence, incidence, associations, causes and outcomes. Nevertheless the validity of these studies has been questioned over the past decade due to the fact that the role of selected risk (or protective) factors identified via observational studies could not be confirmed by subsequent large randomized controlled trials (RCT), the gold standard in experimental epidemiology [95]. On the other hand one cannot, for ethical, financial and practical reasons, randomize risk factors using controlled trials in humans for all possible risk factors [96]. Therefore, identification of risk factors still relies on observational studies even if confounding factors, reverse causation, and/ or biases may cause conflicting results [97]. The approach of Mendelian randomization is a possibility to estimate a causal effect between a specific modifiable risk factor (e.g. a biomarker) and a trait or disease of interest (phenotype) by using genetic variants. In this regard, Mendelian randomization can be thought of as a “natural” RCT in the group of observational studies [98].

Study Type	Methodological approach	Advantages	Disadvantages
Case-control study	Sampling conducted with respect to disease status Estimates the odds of having been exposed to a risk factor given the current case-control status	Low costs in relation to efficiency	Disease and exposure status determined simultaneously, making the temporal sequence of events often difficult to establish
Cross-sectional study	Includes all individuals of a population regardless of exposure or disease status	Low costs in relation to efficiency	Disease and exposure status determined simultaneously, making the temporal sequence of events often difficult to establish Overrepresentation of cases with long duration and under representation of cases with short duration of illness, leading to bias
Cohort study	Sampling conducted with respect to exposure status	Exposure status is ascertained before the occurrence of disease, allowing incidence rates of disease to be calculated in people with and without risk factors	Requires the follow up of a large number of individuals until disease development and, hence, is costly Difficult to control confounding variables and maintain high follow-up rates Adverse outcomes may occur before the onset of the disease of interest, leading to survival bias
Mendelian randomization study	Unbiased randomization by randomly assignment of genes during meiosis determining a certain risk factor for an outcome	"Natural RCT", unbiased estimates of effects (associations) can be obtained	Requires fulfilled key assumptions for a genotype to be used as instrument Not unrestricted applicable to binary outcomes

Table 7: Study types of epidemiology: observational studies (non-experimental studies) modified from [93].

Cross-sectional studies

Cross-sectional studies aim to provide data of the entire population at one specific point of time, generated by observation. This descriptive study type is primarily used to determine prevalence, which influences considerably the likelihood of any particular diagnosis and the predictive value of any investigation. If interpreted with

respect to possible confounding, cross-sectional studies can further be used to infer causation. The subjects are assessed to determine whether they were exposed to the relevant agent and whether they have the outcome of interest at one point in time. Some of the subjects will not have been exposed nor have the outcome of interest. This clearly distinguishes this type of study from the other observational studies (cohort and case controlled) where reference to either exposure and/or outcome is made [99]. ORs, absolute and relative risks can be calculated from this type of data.

The use of routinely collected data allows large cross-sectional studies to be quick and cheap, the most important advantage of this study type. In contrast, the most important challenge is differentiating between cause and effect from a discovered association. Further, they often rely on data originally collected for other purposes and are not suitable to study rare diseases. Difficulty in recalling past events may provoke bias.

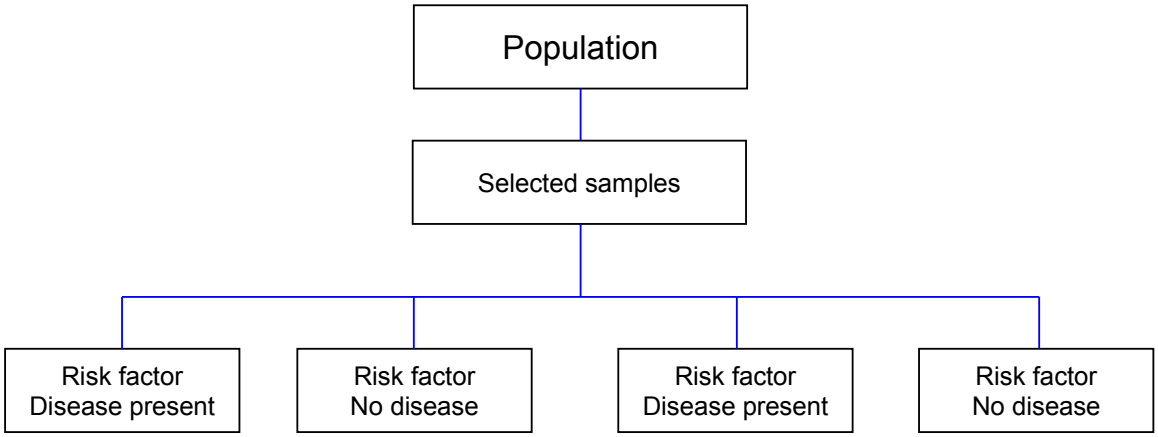


Figure 9: Study design of a cross-sectional study, modified from 99.

Case-control studies

When conditions are rare, case-control studies generate a lot of information from relatively small sample sizes. People with an outcome of interest are matched with a control group and compared retrospectively. ORs, which determine the relative importance of a predictor variable in relation to presence or absence of the outcome and usually approximate the relative risk, can be calculated from this

study type. Case-control studies are very useful for generating hypotheses for further studies [99].

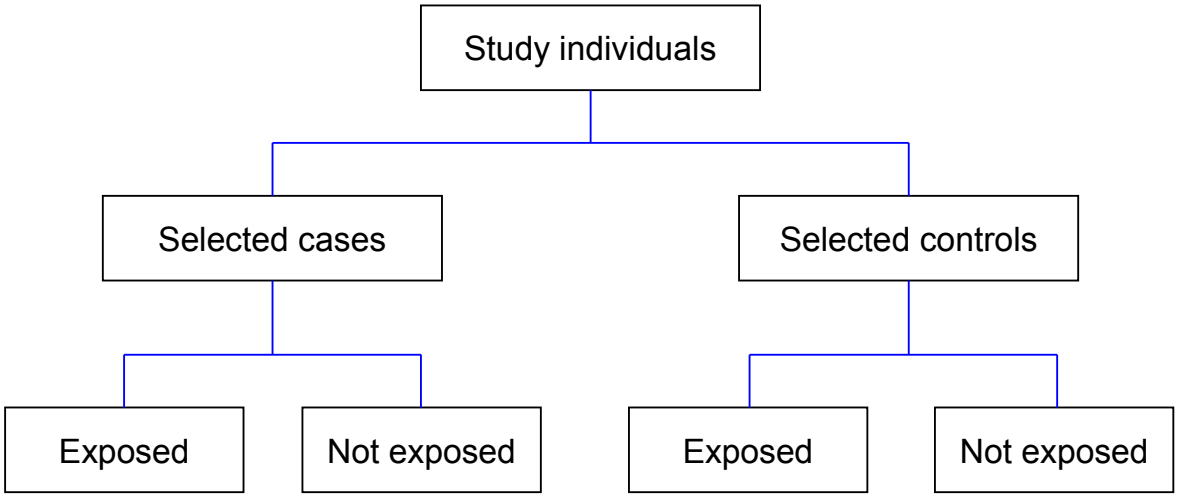


Figure 10: Study design of a case control study, modified from 99.

Cohort studies

Cohort studies are the best method for determining incidence and natural history of a condition if the sample population is representatively selected for the general population. They measure events in temporal sequence, which is why distinguishing causes from effects is possible. Subject selection and loss to follow-up are the major potential causes of bias. Confounding is, as in most observational and experimental studies, a major problem when analyzing cohort studies. Prospective and retrospective cohort studies can be distinguished.

Prospective cohort study

A group of people (cohort) free of the outcome of interest (e.g. type 2 diabetes) is recruited. Each subject must have the potential to develop the outcome of interest (e.g. men should not be included in a cohort study for pre-eclampsia). Investigators then measure a variety of parameters that might be relevant to the development of the condition. All participants are followed up and observed for the occurrence of the outcome. In single cohort studies those people who do not develop the outcome are used as internal controls. If two cohorts are in use one

group is exposed or treated with an agent of interest and the second group acts as an external control.

Retrospective cohort study

For sake of time and money retrospective cohort studies use data already collected for other purposes. The methodology is the same but the study is performed in a post-hoc fashion. The cohort is “followed up” retrospectively. The study period may be many years but the time to complete the study is only as long as it takes to collect and analyze the data [99]. An upcoming disadvantage is for sure that the study was conducted for another purpose and that it is unlikely that all relevant information will have been rigorously collected.

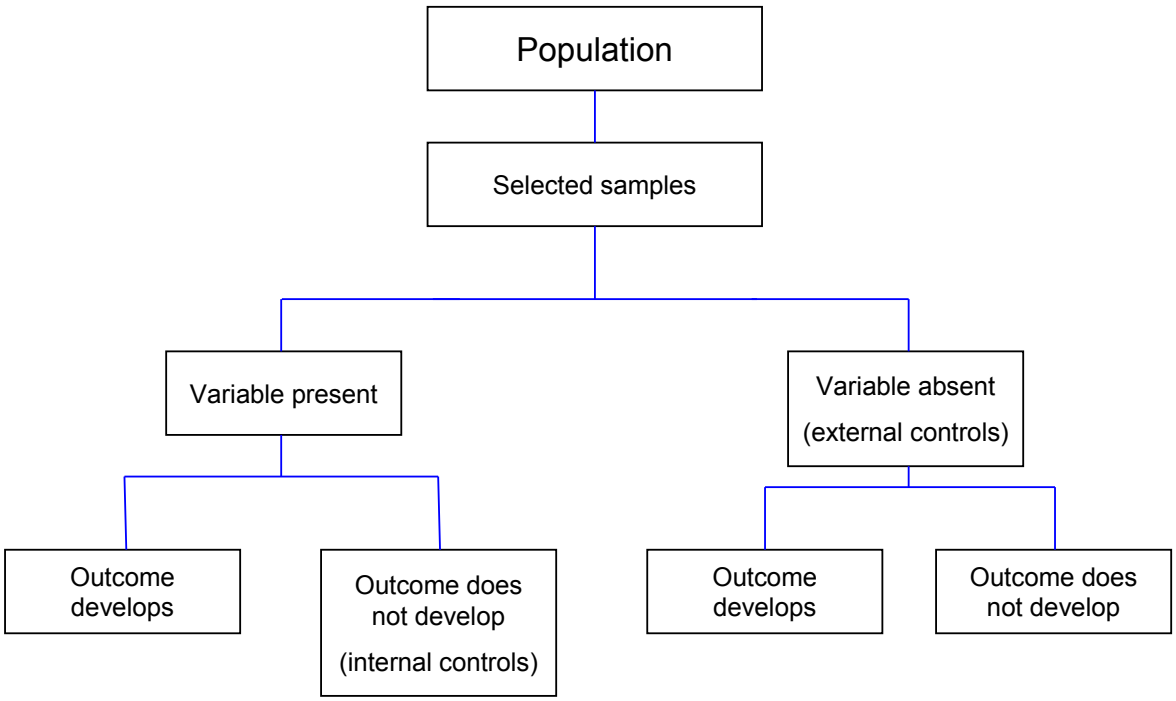


Figure 11: Study design of a cohort study, modified from 99.

Mendelian Randomization

Mendelian randomization (MR) is a method of using measured variation in genes of known function to examine a causal effect of a biomarker on disease. Historically, the first description of the concept of randomization was published 1986 where Katan suggested the use of the Apo E gene to show causality between cholesterol and cancer [100].

The approach of MR utilizes what is sometimes called Mendel's second law within an epidemiological setting. In Mendel's words the second law of random assortment states that *"the behaviour of each pair of differentiating characteristics in hybrid union is independent of the other differences between the two original plants, and, further, the hybrid produces just so many kinds of egg and pollen cells as there are possible constant combination forms"* [101]. Put simply, this suggests that the inheritance of one trait is independent of the inheritance of other traits. As described, conventional genetic epidemiology investigates the genotype-phenotype association within a population by using genetic markers such as SNPs. But it is also possible to exploit the random assignment of genes as a means of reducing confounding in examining exposure–disease associations: this is called Mendelian randomization in the epidemiological context [95].

In other words, Mendelian randomization refers to the random allocation of alleles at the time of gamete formation. A specific genotype carried by a person therefore results from two such randomized transmissions, one from the paternally inherited allele and the other from the maternally inherited allele. A logical consequence of these randomizations is that genotypes are not expected to be associated with known (measurable or not) or unknown confounders for any outcome of interest, except those lying on the causal pathway between the genotype and the outcome. This should hence allow analyzing the genotype-risk factor association and the genotype-outcome association in an unconfounded manner. By combining appropriately the results of these two analyses, one can get an estimate of the unconfounded risk factor-outcome association. The idea to overcome the "confounding" and "reverse causation" problems encountered in observational epidemiology by taking advantage of the natural random allocation of alleles during meiosis is analogous to randomized controlled trials, in which the random allocation of treatment is expected to lead to an even distribution of (known or

unknown) confounding factors across each groups [98, 102]. The genetic effect of an outcome via a phenotype (modified biomarker) is displayed in Figure 12.

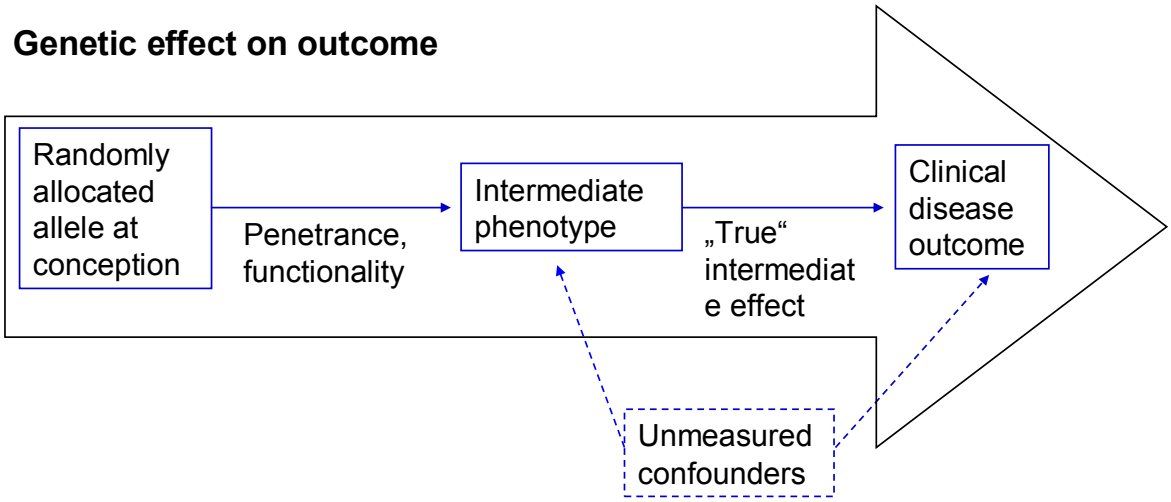


Figure 12: Genetic effect on the outcome, represented by the large open arrow in a Mendelian randomization approach, modified from [102]

To illustrate the approach a classic example is provided: the aldehyde dehydrogenase 2 (ALDH2) gen encodes the enzyme aldehyde dehydrogenase, which catalyzes the chemical transformation from acetaldehyde to acetic acid. Carriers of the ALDH2 *2*2 genotype have reduced alcohol consumption because of adverse reactions (facial flush, headache, nausea and drowsiness) due to acetaldehyde accumulation. This fact has been used to show that alcohol intake increases the risk of esophageal cancer [103] or head and neck cancer [104], which is consistent with the findings from observational studies. Whereas reported alcohol consumption may be subject to measurement errors, ALDH2 genotypes can be measured accurately, are present since birth, result from the random allocation of the paternally and maternally inherited alleles, are strongly associated with alcohol consumption, and therefore provide a unique opportunity to assess, in an unconfounded manner, the risk of disease associated with alcohol consumption [98].

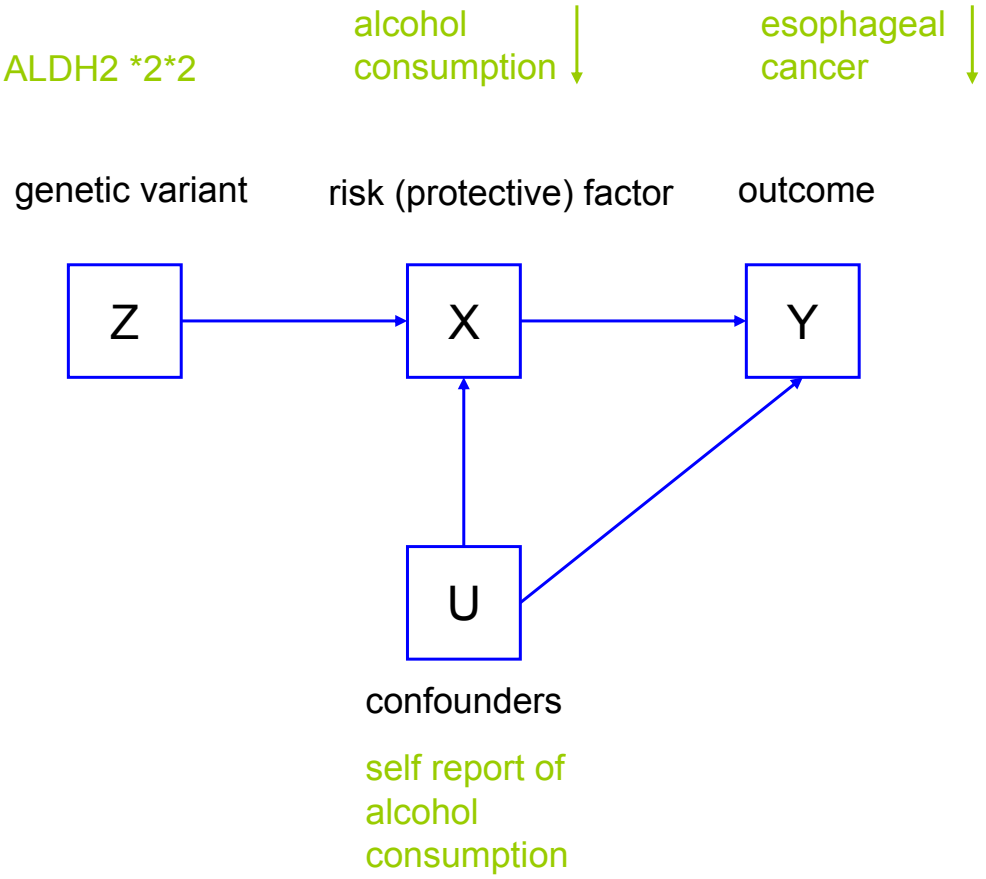


Figure 13: Model of Mendelian randomization representing the causal relationship between the genetic variant (Z), the risk or protective factor (X), the outcome (Y) and the known and unknown confounders (U). To illustrate the approach the example of ALDH2 *2*2 genotypes and alcohol consumption in association with oesophageal cancer is given in green, modified from [98].

3.3.2 Experimental studies

Experimentation is used to test existing theories or to investigate new hypotheses in order to support or to disprove them [105]. Experimental studies provide insights into causality by demonstrating how the outcome changes when a particular factor (variable) is manipulated.

Interventional trial

Interventional studies are considered to provide the most reliable evidence in epidemiological research. This kind of trial can generally be regarded as preventive or therapeutic [106]. Interventions have the intention to improve the condition of an individual of a group of individuals. Examples of intervention studies in public health research are studies that evaluate the impact of a healthier lifestyle (avoiding smoking, reducing alcohol consumption, increasing physical activity) or to start brushing teeth early in babies.

It is important that the target population, i.e. the population to which the findings should be extrapolated, is predefined. In general the intervention (preventive or therapeutic measure) being tested is allocated by the investigator to a group of study subjects. To obtain a clear idea about the effect of the intervention the two (or more) groups should be comparable at the beginning (baseline). Comparability at baseline is achieved best by randomly allocating subjects to the two groups. This is known as simple randomization comparable to flipping a coin. For example, a computer-generated random list could be used. Other randomization methods like block random assignment or stratification are more common. Block randomization, used in smaller trials, ensures that the numbers of participants assigned to each group is equally distributed.

All randomization methods share the goal to ensure that important baseline variables (potential confounding factors) are more evenly distributed between groups than chance alone may assure. However, there are a limited number of baseline variables that can be balanced by stratification because of the potential for small numbers of subjects within each stratum.

It must be considered that randomization can only guarantee balance for large studies and that random imbalance can often occur in small studies. After randomization and execution of the intervention subjects are followed prospectively to compare then intervention versus control group.

An intervention study is able to detect the anticipated effect of the intervention with a high probability. The number of patients enrolled in the study as well as the expected predefined effect size of an intervention has a large impact on the "power" of a study. The sample size required to achieve power is inversely proportional to the treatment effect squared. The smaller the expected effect of an intervention or observation - the larger the sample size needed to be able to conclude, with enough power, that differences are unlikely to be due to chance [96].

Finally, it should be mentioned that any intervention study, as any other study, should be ethically acceptable and that each participant should read, understand and sign an "informed consent form" that includes trial details such as its purpose, duration, required procedures, risks and potential benefits and key contacts.

Controlled clinical trials constitute a separate, important class of intervention. There, the aim is to compare the effectiveness and safety of two (or more) medical treatments or surgical operations or a combination of both. The target population constitutes patients with a specific disease or symptom.

Clinical trial

Clinical trials are sets of tests in medical research and drug development that generate data about safety and efficacy. In a clinical trial, inclusion criteria specify the type of patients who need the treatment under examination. The exclusion criteria exclude patients for which there will be most likely safety concerns or for which the treatment effect might not be clear [107].

While most interventions aim to achieve a change in attitude (a psychological effect), medical treatments need to show their effectiveness apart from any psychological impact (placebo effect). The placebo effect is defined as the pure psychological effect that medical treatments can have on patients. This effect can be detected by applying placebo (inactive medication without active pharmaceutical ingredients with the same taste, texture and look as the "real"

active medication) to patients who are blinded for the fact that they did not receive active treatment. Placebo controlled trials, a trial with a placebo group as control group, are very common. A study is called single-blinded if only the patient does not know whether he or she receives placebo or the active medication. If physician and patient are unaware of the administered treatment then the study is designated as a double-blinded clinical trial. Biases in scoring the effect and safety of the medication can thus be better prevented. In such trials the biological effect can be distinguished from the psychological placebo effect of a drug.

To estimate the effect of an intervention a specific outcome needs to be defined. In context of clinical studies this outcome is designated as endpoint. Multiple testing should therefore in general be avoided. It is more advisable to choose one endpoint, the (primary endpoint). If this is not possible, a correction for multiple testing like Bonferroni adjustment should be applied. The phases of a RCT are illustrated in Figure 14.

The majority of clinical trials are drug trials. It takes many years and billions of Euros to develop and register a new drug. In this context clinical trials are essential, partly because regulatory bodies like the Food and Drug Administration in the United States and the European Medicines Agency (EMA) in Europe have developed stringent criteria on the pharmaceutical industry before a new drug can be registered [107].

Clinical trials developing new drugs are commonly classified into four phases whereas each phase is a separate clinical trial.

- Phase 1: Screening for safety
- Phase 2: Establishing the testing protocol
- Phase 3: Final testing
- Phase 4: Post-approval studies

Each phase has further a different purpose and helps scientists answer a different question.

In Phase 1 trials, researchers test an experimental drug or treatment in a small group of people (20-80) for the first time to evaluate its safety, determine a safe dosage range, and identify side effects.

In Phase 2 trials, the experimental treatment is given to a larger group of people (100-300) to see if it is effective and to further evaluate its safety.

In Phase 3 trials, the treatment is given to large groups of people (1000-3000) to confirm again its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow it to be used safely.

In Phase 4 trials, postmarketing studies delineate additional information, including the treatment's risks, benefits, and optimal use.

Before pharmaceutical companies start clinical trials on a drug extensive preclinical studies are conducted [108].

Study Type	Methodological approach	Advantages	Disadvantages
Clinical trial	Individuals are randomly assigned to intervention and comparison groups Aims to evaluate a cure or a preventive treatment (usually drugs)	Minimizes bias and allows valid statistical testing, as factors that could potentially confound the examined association should occur with roughly equal frequencies in the intervention and comparison groups	Participants studied may not be representative of the general population

Table 8: Study types of epidemiology: experimental studies modified from [93].

The randomized controlled trial – the gold standard in evidence based medicine

The randomized controlled trial, belonging to the group of interventional trials, is a specific type of scientific experiment and the preferred design for a clinical trial. RCTs are often used to test the efficacy of various types of interventions within a patient population. Mostly RCTs are evaluated by an intention-to-treat analysis which is done to avoid the effects of crossover and dropout which may brak the random assignment to the treatment groups in a study [109]. The intention-to-treat effect on an outcome is shown in Figure 15.

This type is considered as the golden standard of determining whether a cause-effect relationship exists between an intervention and outcome [110]. RCTs are not exclusively employed in the field of clinical trials. Other research areas like social or economic sciences are also using this study design. In fact the RCT is a study in which people are allocated at random to receive one of several clinical

interventions [111]. As described previously the act of randomizing patients to receive or not receive an intervention ensures, on average, all other possible causes are equal between the two groups. Thus, any significant differences between groups in the outcome event can be attributed to the intervention and not to some other unidentified factor (chance or confounding). No other study design allows investigators to balance these factors. However, RCTs are not a panacea to answer all clinical questions. Unsuitable are, for example situations in which the outcome is rare and occurs only after long delay. In such instances, observational study designs such as case-control studies or cohort studies are more appropriate. This might also be the case in situations of financial constraints or when low compliance or high drop-out rates are expected RCTs may then not be feasible [96]. It should be kept in mind that random allocation does not protect RCTs against other types of biases [111].

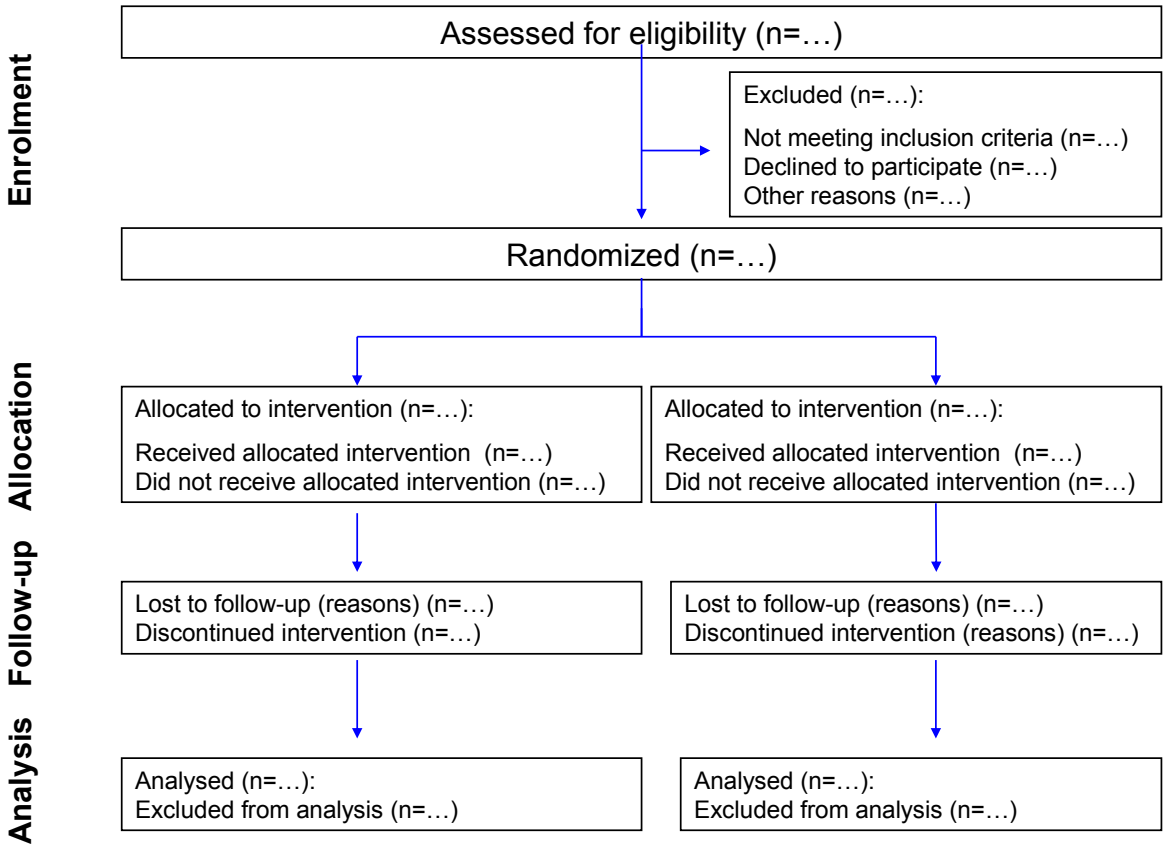


Figure 14: Flow chart: phases (enrolment, intervention allocation, follow-up, and data analysis) of a randomized controlled trial, modified from [112].

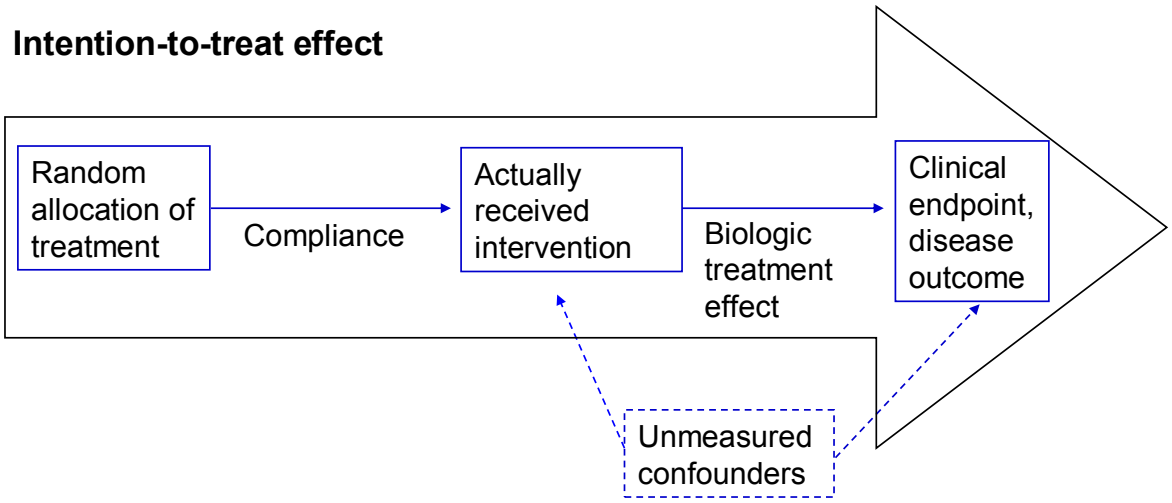


Figure 15: Intention-to-treat effect on the outcome, represented by the large open arrow, modified from [102].

Laboratory experiment

An experiment is a methodical procedure carried out with the goal of verifying, falsifying, or establishing the validity of a hypothesis. Experiments provide insight into cause-and-effect by demonstrating what outcome occurs when a particular factor is manipulated.

Variables in laboratory experiments:

Variables are everything that can be controlled or changed in an experiment (eg. temperature or duration of the experiment) There are three kinds of variables distinguished:

Dependent variable: is what is measured or observed in an experiment and what is affected during the experiment. It is called dependent because it "depends" on the independent variable.

Independent variable: is the changed factor in an experimental setting. To facilate interpretation of the data only one factor should be changed at one time.

Controlled Variables: are kept constant during an experiment

Laboratory experiments are characterized by:

- a dependent and an independent variable

- the independent variable precedes the dependent variable
- data of at least two groups are compared [113].

Hence, laboratory experiments are controlled experiments, where an experimental group is compared with a control group. Ideally, these two groups are identical except for one variable, the independent variable.

3.4 Candidate genes modifying Vitamin D levels

Twin studies suggest a strong heritable component of circulating 25-OH-vitamin D concentrations, with heritability rates ranging from 29% to 77% [114, 115, 116].In 2010, two independent GWA studies identified three common loci of genetic determinants (SNPs) for vitamin D insufficiency: group-specific component (GC, rs2282679), 7-dehydrocholesterol reductase (DHCR7, rs12785878) and cytochrome P450 IIR-1 (CYP2R1, rs10741657). These loci are located nearby genes which are involved in vitamin D transport, cholesterol synthesis and hydroxylation [117, 118]. Wang et al. undertook his investigation in nearly 34000 individuals of 15 cohorts of European ancestry: five epidemiological cohorts were used as discovery cohorts (n = 16125), five as in silico replication cohorts (n = 9367) and five as de novo replication cohorts (n = 8504) [117]. Ahn and his group performed their GWA study in 4501 Europeans from five cohorts. They identified, besides others, the same polymorphisms associated with 25-OH-vitamin D concentrations and replicated them in 2221 additional samples confirming their results concerning these three loci.

These variants near or in genes involved in cholesterol synthesis, hydroxylation and vitamin D transport affect vitamin D status and may therefore identify individuals at increased risk of vitamin D insufficiency.

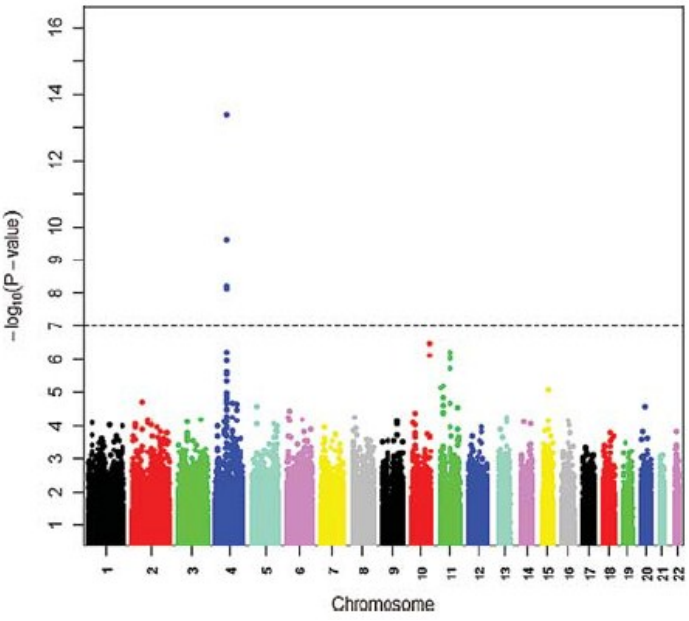


Figure 16: Manhattan Plot of Ahn's GWA study of circulating vitamin D levels: Genome-wide associations of circulating 25-OH-vitamin D graphed by chromosome position and 2log10 p-value based on the five GWA study cohorts, modified from [118].

	Chromosome	Position	Nearest gene(s)	MAF	Combined p value for discovery samples (up to n=16 124)	Combined p value for replication samples (up to n=17 744)	Overall p value
rs3755967	4	72828262	*GC*	0·29	$7·41\times10^{-53}$	$3·00\times10^{-24}$	$2·42\times10^{-75}$
rs17467825	4	72824381	*GC*	0·29	$3·85\times10^{-52}$	$1·61\times10^{-23}$	$6·75\times10^{-74}$
rs1155563	4	72862352	*GC*	0·30	$4·70\times10^{-55}$	$4·26\times10^{-20}$	$2·37\times10^{-73}$
rs2298850	4	72833131	*GC*	0·28	$8·94\times10^{-49}$	$2·12\times10^{-24}$	$2·03\times10^{-71}$
rs7041	4	72837198	*GC*	0·44	$3·74\times10^{-42}$	$1·78\times10^{-18}$	$6·31\times10^{-59}$
rs7944926	11	70843273	*DHCR7/NADSYN1*	0·23	$1·56\times10^{-33}$	$7·57\times10^{-4}$	$8·96\times10^{-36}$
rs12800438	11	70848651	*DHCR7/NADSYN1*	0·23	$5·98\times10^{-33}$	$6·39\times10^{-4}$	$2·54\times10^{-35}$
rs3794060	11	70865327	*DHCR7/NADSYN1*	0·23	$8·09\times10^{-33}$	$6·44\times10^{-4}$	$3·38\times10^{-35}$
rs4945008	11	70898896	*DHCR7/NADSYN1*	0·24	$8·98\times10^{-33}$	$6·11\times10^{-4}$	$4·55\times10^{-35}$
rs4944957	11	70845683	*DHCR7/NADSYN1*	0·23	$1·43\times10^{-32}$	$7·36\times10^{-4}$	$8·70\times10^{-35}$
rs2060793	11	14871886	*CYP2R1*	0·40	$2·69\times10^{-6}$	$2·36\times10^{-7}$	$1·73\times10^{-11}$
rs1993116	11	14866810	*CYP2R1*	0·40	$2·94\times10^{-6}$	$1·28\times10^{-6}$	$6·25\times10^{-11}$
rs12794714	11	14870151	*CYP2R1*	0·43	$6·24\times10^{-5}$	$8·71\times10^{-7}$	$1·84\times10^{-9}$
rs10500804	11	14866849	*CYP2R1*	0·43	$7·43\times10^{-5}$	$1·12\times10^{-6}$	$2·67\times10^{-9}$
rs7116978	11	14838347	*CYP2R1*	0·36	$1·17\times10^{-5}$	$7·59\times10^{-5}$	$4·99\times10^{-9}$

Results within each locus are ordered by strength of association with 25-hydroxyvitamin D concentration. MAF=minor allele frequency.

Table 9: SNPs associated with 25-OH-vitamin D concentrations identified in the GWA study of Wang et al., modified from [117]. Top hit SNPs with the lowest p-values are highlighted in green.

3.4.1 Group specific component (GC) rs2282679

The protein encoded by this gene belongs to the albumin gene family. It is a multifunctional protein found in plasma, ascites, cerebrospinal fluid and on the surface of many cell types. It binds to vitamin D and its plasma metabolites and transports them to target tissues. Alternatively spliced transcript variants encoding different isoforms have been found for this gene [119].

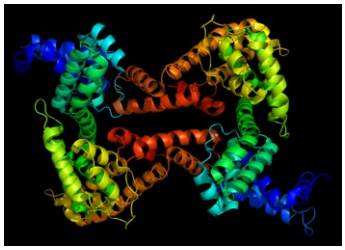


Figure 17: Crystallographic structure of the human vitamin D-binding protein, modified from [121]

Human GC is a glycosylated alpha-globulin with approximately 58 kDa in size. Its 458 amino acids are coded by 1690 nucleotides on chromosome 4 (4q11–q13) and gets synthesized in the liver [120]. The primary structure contains 28 cysteine residues forming multiple disulfide bonds. GC contains 3 domains. Domain 1 is composed of 10 alpha helices, domain 2 of 9 and domain 3 of 4 [121]. The group specific component is also known as vitamin D binding protein (DBP or VDBP) or vitamin D-binding alpha globulin (VDBG).

3.4.2 7-Dehydrocholesterol reductase (DHCR7) rs12785878

This gene encodes an NADPH-dependent enzyme that catalyzes the final step in cholesterol production in many types of cells. It removes the C (7-8)-double bond in the B-ring of sterols and catalyzes the conversion of 7-dehydrocholesterol to cholesterol in our skin (Figure 19). Together with UVB radiation it has therefore major biological relevance for human vitamin D de novo synthesis. DHCR is ubiquitously expressed and its highest concentrations are found in the liver and brain.

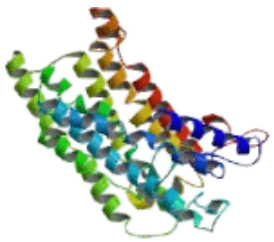


Figure 18: Crystallographic structure of DHCR7, modiefid from [123].

The DHCR7 gen is found on chromosome 11q12–q13 spanning 14 kb and encompassing 9 exons encoding a 425 amino acid protein. Functional DHCR7 protein is a 55.5 kDa integral membrane protein localized in the microsomal membrane [122, 123].

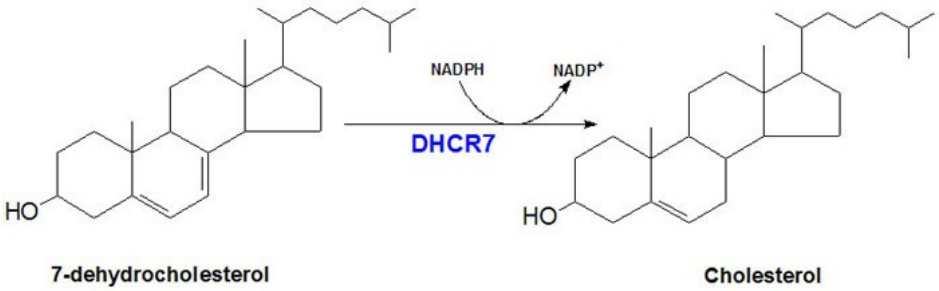


Figure 19: Catalyzation of 7-dehydrocholesterol to cholesterol by DHCR7 with NADPH as cofactor, modified from [122].

3.4.3 Cytochrome P450 IIR-1 (CYP2R1) rs10741657

This gene encodes a member of the cytochrome P450 superfamily of enzymes. The cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids [124, 125]. The synthesis of bioactive vitamin D requires hydroxylation at the 25- and 1-alpha positions by cytochrome P450 enzymes mostly in the liver and kidney, respectively. The enzyme of CYP2R1, the 25-hydroxylase, is responsible for the first hydroxylation of vitamin D (D_2 and D_3) into 25-OH-vitamin D, the major circulatory form of vitamin D [126]. The second hydroxylation is performed by the 25-OH-vitamin D3-1-alpha-hydroxylase encoded by CYP27B1, catalyzing the reaction from 25-OH-vitamin D to 1,25 $(OH)_2$ vitamin D, the biologically active form of vitamin D [127].

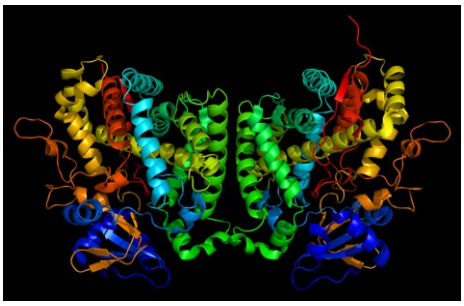


Figure 20: Crystallographic structure of CYP2R1, modified from [125].

CYP2R1 is located on the short arm of chromosome 11 at position 15.2 (11p15.2) [128], contains 5 exons and spans about 15.5 kb [129]. The molecular weight of its protein is 57.4 kD [130].

4. AIMS OF THE THESIS

Vitamin D is known for its crucial role in calcium and bone metabolism. A growing number of extraskeletal diseases including cardiovascular disease [176], prostate cancer [131] as well as mortality rates [180] have been linked to vitamin D concentrations. Until now causal associations have not been clearly confirmed in RCTs indicating that it is unknown whether low vitamin D levels are the cause or the consequence of these extraskeletal associations. Two independent GWA studies have identified three common loci of genetic determinants for vitamin D insufficiency: GC, DHCR7 and CYP2R1 [117, 118]. The aim of the present thesis was to genotype these recently discovered genetic determinants of vitamin D in different settings and to test possible associations with biomarkers and endpoints of bone health and cancer outcomes. Further I wanted to investigate a potential causal role of vitamin D in the often reported association between vitamin D and mortality. According to the aims outlined above this thesis is structured into the following sections:

Vitamin D and bone: Replication of these three allelic determinants of vitamin D insufficiency in two cohorts of different vitamin D supply (vitamin D deficiency and vitamin D levels in the normal range) and to investigate a possible association with BMD and fractures.

Vitamin D and prostate cancer: Investigation of the strongest genetic variant of these three SNPs, the GC polymorphism and its potential to predict 1,25 $(OH)_2$ vitamin D levels and further to address its role in metastatic progression and outcome of prostate cancer.

Vitamin D and mortality: Analyzing the role of these three 25-OH-vitamin D-associated SNPs in predicting all-cause cardiovascular and non-cardiovascular mortality in an unconfounded setting to determine, if low levels of vitamin D are the cause of mortality or the consequence of an impaired health status.

I will start with a detailed description of all marterials and methods applied in this thesis. For each section I will describe respectively a brief background, followed by characterization of the study subjects, procedures, statistics, results and a discussion part.

5. MATERIALS AND METHODS

5.1 Isolation of DNA

5.1.1 Chemical and Reagents

- Buffer B1*
- Buffer B2*
- Wash Buffer BW (ready to use) *
- Wash Buffer B5 (Concentrate)*
- Elution Buffer BE (5 mM Tris/HCl, pH 8.5, ready to use) *
- Proteinase K (lyophilized)*
- Proteinase Buffer PB (ready to use) *
- 96–100 % ethanol (ready to use)
- Aqua dest (DEPC treated water, USB Corporation, UK)

*Contents of the NucleoSpin Blood ® Kit (740899.250) from Machery and Nagel

5.1.2 Equipment

- Automated pipetting system (ep-Motion 5070, Eppendorf Austria)
- Manual dispenser (Multipette ® plus, Eppendorf Austria)
- Pipettes (Eppendorf reference ® 100-1000µl, 10-100µl, 1-10µl, Eppendorf Austria)
- Centrifuge with plate adapter: (Eppendorf 5810R with swing bucket and plate rotor, Eppendorf Austria)
- Shaking incubator (Microplate Shaker Timix5 Control, Edmund Bühler, Germany)
- Incubator (BD53, Germany)
- Personal protection equipment (lab coat, disposable gloves)

5.1.3 Consumables

- 96-well spin column plates (AcroPrep ™ 96 Filter Plate, Pall Life Sciences, Austria)
- 96-well deep well plate (DWP), 2 ml for sample lysis, DNA elution and waste (Nerbe plus 04-072-0020)
- Adhesive PCR film (Thermo Scientific, UK)
- Disposable filter pipet tips (Eppendorf dual filter Tips 0.1-10µl, 2-100 µl, 50-100 µl, Eppendorf Austria)
- Disposable dispenser tips (Combitips plus: 0.5ml, 2.5ml, Eppendorf Austria)
- Tips for the automated pipetting system (epT.I.P.S.® Motion pipete tips with filters 20-300 µl, Eppendorf Austria)

5.1.4 Principle

The 96-well spin column plate isolates genomic DNA from whole blood and other body fluids like serum or plasma as well as from cultred cells. The first step in DNA isolation is to break open the cell and release the cytoplasmic contents including the nucleus. Lysis of the cells is achieved by incubation of whole blood in a solution containing large amounts of chaotropic ions in the presence of proteinase K. Appropriate conditions for the binding of DNA to the silica membrane in the filter column of each well are achieved by addition of ethanol to the lysate. High temperature and mechanical exposure support the initial cell lysis. The binding process is reversible and specific to nucleic acids. Washing steps efficiently remove cell debris and all other components of the fluid except nucleic acids. Pure genomic DNA is finally eluted under low ionic strength conditions in a slightly alkaline elution buffer.

The kit allows purification of highly pure genomic DNA with an A260/A280 ratio between 1.60 and 1.90 and a typical concentration of 40–60 ng per µl. [132]

5.1.5 Procedures

The standard protocol has been modified and adapted for old and/or clotted blood samples.

Preparation of working solutions:

Lysis buffer B3: Transfer of “buffer B1” into “buffer B2”, mix well. The resulting lysis buffer B3 is stable for up to one year at room temperature.

Wash buffer B5: Add 80 ml of ethanol (96–100 %) to wash buffer B5 concentrate. Mark the label of the bottle to indicate that ethanol was added. Store wash buffer B5 at room temperature (18–25°C) up to one year.

Proteinase K: Dissolve lyophilized proteinase K in 3.35 ml proteinase buffer (PB). Proteinase K solution is stable at -20 °C for up to 6 months [132].

Before starting the preparation:

- Define the pattern of DNA how it should be positioned at the 96 DWP. An example of a defined pattern is given below:

Plate 1	1	2	3	4	5	6	7	8	9	10	11	12
A	e	10001	10003	10004	10006	10007	10009	10010	10012	10013	10018	10019
B	10020	10022	10026	10031	10032	10034	10038	10043	10044	10045	10046	10048
C	10050	10051	10053	10054	10055	10056	10058	10060	10062	10063	10064	10066
D	10067	10069	10070	10071	10072	10073	10074	10075	10076	10077	10079	10080
E	10081	10082	10083	10085	10086	10088	10089	10091	10092	10093	10095	10097
F	10098	10099	10101	10102	10103	10105	10110	10111	10112	10113	10115	10116
G	10117	10118	10120	10121	10124	10126	10127	10128	10129	10131	10132	10133
H	10134	10136	10138	10139	10140	10142	10143	10144	10146	10148	10149	10151

Table 10: Example of a DNA pattern on a 96 well plate (e: empty).

- If necessary prepare buffer B3, buffer B5 and proteinase K.
- Preheat Buffer BE up to 55°C (degree Celsius).
- Preheat the incubator up to 70°C.

DNA purification:

1. Lysis of the blood cells

Prepare a DWP with 100 µl blood + 100 µl A. dest. according to the appropriate well position of the predefined pattern.

Add: 15 µl proteinase K

Add: 200 µl buffer B3

Centrifuge: 2000 rpm (rounds per minute), 1 min (minute)

Shaking incubator: 1400 rpm, 10 min

Incubate: 55°C, 20min

2. Adjust DNA binding conditions

Add: 200 µl ethanol (96-100%)

Shaking incubator: 1400 rpm, 1 min

3. Bind DNA to the silica membrane

Ep Motion: Transfer of the lysate from the lysis DWP into the 96-well spin column plates on top of an empty DWP, used as waste in the following centrifuge step, according to the appropriate well position from the predefined pattern.

Centrifuge: 4400 rpm, 1 min

4. Wash silica membrane

Prewash:

Add: 200 µl buffer BW

Centrifuge: 4400 rpm, 1 min

1st wash:

Add: 200 µl buffer B5

Centrifuge: 4400 rpm, 1 min

2nd wash + dry silica membrane:

Add: 200 µl buffer B5

Centrifuge: 4400 rpm, 3 min (residual ethanol is removed during this longer centrifugation step)

5. Elute DNA

Place 96-well spin column plates on top of a newly labelled DWP according to the appropriate well position from the predefined pattern.

Add: 50 µl preheated buffer BE

Incubate: 3 min at room temperature

Centrifuge: 4400 rpm, 1 min

The DNA has now been eluted into the DWP. Discard the filterplate and close the DWP with a PCR film accurately.

Isolated DNA is stored at 4°C for short term storage, at -20 for medium term storage and at -80°C for long time storage.

5.2 SNP Genotyping: 5´3´Exonucleaseassay (TaqMan®)

TaqMan probes were named after the videogame PacMan (Taq Polymerase + PacMan = TaqMan) as its mechanism is based on the PacMan principle.

5.2.1 Chemical and Reagents

- 40x TaqMan® SNP Genotyping Assay for GC, rs2282679 (C___26407519_10, Applied Biosystems by Life Technologies, Invitrogen, Life Tech Austria, Austria), long time storage -20°C.
- 40x TaqMan® SNP Genotyping Assay for CYP2R1, rs10741657 (C___2958430_10, Applied Biosystems by Life Technologies, Invitrogen, Life Tech Austria, Austria), long time storage -20°C.
- 40x TaqMan® SNP Genotyping Assay for DHCR7, rs12785878 (C__32063037_10, Applied Biosystems by Life Technologies, Invitrogen, Life Tech Austria, Austria), long time storage -20°C.

- 2x Universial Master Mix (UMM): TaqMan Gene Expression Master Mix (4369016, Applied Biosystems by Life Technologies, Invitrogen, Life Tech Austria, Austria), store at 2 to 8 °C.
- A. dest (DEPC treated water, USB Corporation, UK)
- Mineral Oil (Sigma-Aldrich Chemie GmbH, Germany)

5.2.2 Equipment

- 12-channel pipet (Research plus pipet, 12-channel, 0.5-10µl, Eppendorf Austria)
- Manual dispenser (Multipette ® plus ,Eppendorf Austria)
- Pipettes (Eppendorf reference ® 100-1000µl, 10-100µl, 0.5-10µl, Eppendorf Austria)
- Microcentrifuge: (Hermle Z233M, Hermle Labortechnik, Germany)
- Centrifuge with plate adapter: (Heraeus Megafuge 1.OR, Heraeus instruments, Germany)
- Vortexer (lab dancer vario, IMLAB, France)
- Personal protection equipment (lab coat, disposable gloves)
- CYCLER (Mastercycler gradient, Eppendorf Austria and Duo Cycler, VWR International Vienna)

5.2.3 Consumables

- 96 well plates (96 well thin wall PCR plates, Simport, Austria)
- Adhesive PCR film (Eppendorf, Austria)
- Disposable filter pipette tips (Eppendorf dual filter Tips 0.1-10µl, 2-100µl, 50-100µl, Eppendorf Austria)
- Disposable dispenser tips (Combitips plus: 0.1ml, 0.5ml, Eppendorf Austria)
- 1.5ml Reaction Tubes (Safe Lock Tubes 1.5ml, Eppendorf Austria)

5.2.4 Principle

TaqMan® 5'3' exonuclease assay chemistry provides a fast and effective way to analyze SNP genotyping results. Each predesigned TaqMan® SNP Genotyping Assay includes two allele-specific TaqMan® MGB (minor groove binder) probes containing distinct fluorescent dyes (VIC® and FAM) and a polymerase chain reaction (PCR) primer pair to detect specific SNP targets. The principle relies on the 5´3´ exonuclease activity of Taq Polymerase to cleave a dual-labeled probe during hybridization to the complementary target sequence and fluorophore-based detection (Figure 21). The resulting fluorescence signal permits quantitative measurements of the accumulation of the product during the exponential stages of the PCR. TaqMan® probes and primer sets (assays) uniquely align with the genome to provide unmatched specificity for the allele of interest.

Assay Mix (AM)

The 40x TaqMan® SNP Genotyping Assay Mix contains assay specific primers and probes.

Primers

Primers are strands of nucleic acid that serve as starting point for the PCR amplification. The polymerase catalyzing this reaction can only add new nucleotides to an existing strand of DNA. The polymerase starts replication at the 3'-end of the primer and copies the opposite strand. Moreover, primers border the amplified PCR product around the location of the SNP.

Probes

Each TaqMan® assay requires two fluorogenic allele-specific probes. One probe matches the wild type sequence and one probe matches the mutant sequence. Each of these probes consists of a fluorophore dye covalently attached to the 5´end of the oligonucleotide, a MGB molecule and a nonfluorescent quencher at the 3´end of the oligo. The MGB molecule increases the melting temperature (Tm) without increasing probe length, thereby allowing the design of probes that are shorter and more specific than traditional dual labeled probes [133]. The reporter dyes used in this study were 6-carboxyfluorescein (FAM, green) for the wild type sequences of GC, CYP2R1 and DHCR7 and VIC® (red) for the mutant sequences. In solution, the proximity of the dyes on each probe allows the quencher to absorb energy from the reporter dye through fluorescent resonance

energy transfer (FRET). In other words, as long as the fluorophore and the quencher are in proximity, quenching inhibits any fluorescence signals. During PCR amplification, only hybridized probes are cleaved by the 5′ to 3′nucleolytic activity of the DNA polymerase. The fragments of the probes are then displaced from the target and polymerization of the strand continues. The degradation of the probe releases the fluorophore and breaks the close proximity of fluorophore and quencher, thus relieving the quenching effect and allowing fluorescence of the fluorophore after excitation. Hence, fluorescence detected in the real-time PCR thermal cycler or as endpoint measurement is directly proportional to the fluorophore released and the amount of DNA template present in the PCR. This process occurs in every cycle and it does not interfere with the exponential accumulation of product. An allele is detected as significantly increased intensity of its reporter dye over background [134].

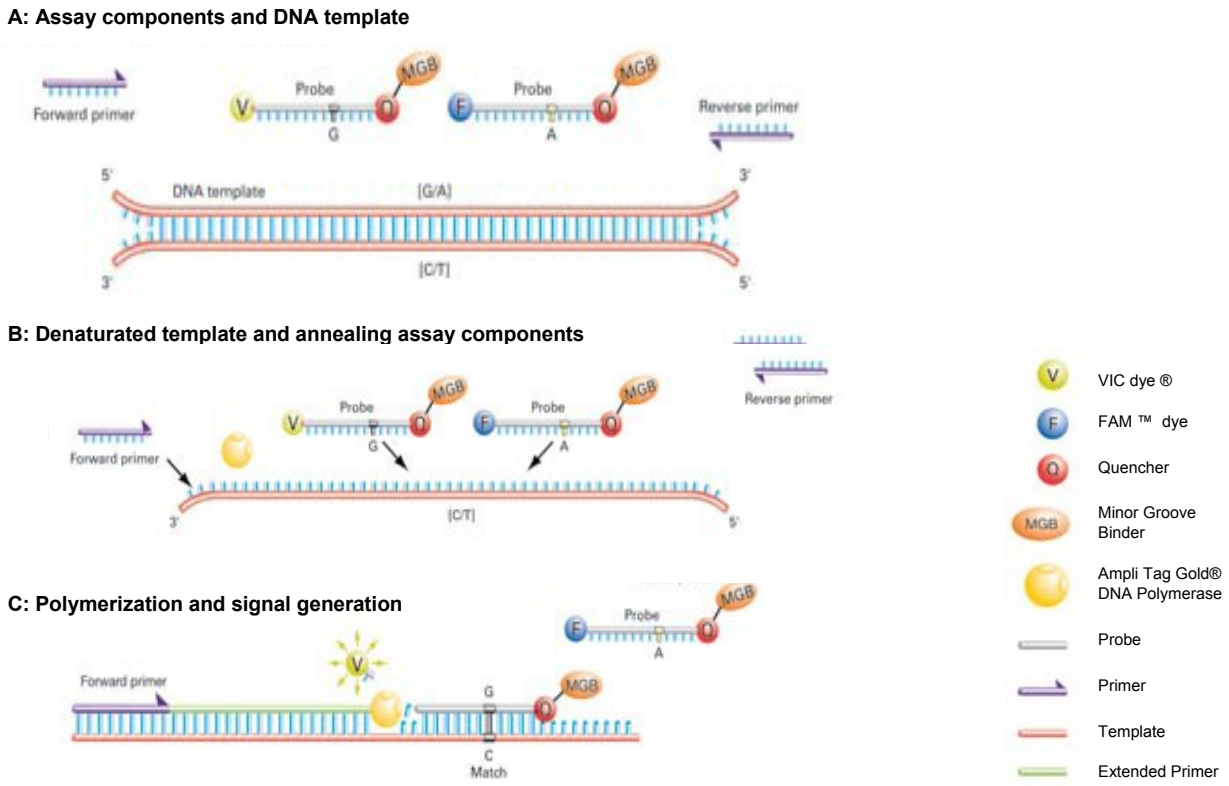


Figure 21: The principle of a 5'3' exonuclease assay, modified from [133]. Allelic discrimination is achieved by the selective annealing of TaqMan® MGB probes. In part C of this figure the DNA polymerase is short before cleaving hybridized probes by its 5´3´exonuclease activity. The probe fragments would then be displaced from the target and polymerization of the strand continues. The

degradation of the probe breaks the close proximity between reporter and quencher, thus allowing the fluorophore to fluoresce.

Universal Master Mix (UMM):

TaqMan Gene Expression Master Mix reagents provide a PCR mix that may be used with any appropriately designed primer and probe to detect any DNA target (including cDNA, genomic, or plasmid DNA). The optimized Mix for TaqMan assays contains AmpliTaq Gold® DNA Polymerase, UP (Ultra Pure), Uracil-DNA Glycosylase (UDG), deoxyribonucleotide triphosphates (dNTPs) with deoxyuridine triphosphate (dUTP), ROX™ Passive Reference, and buffer components optimized for sensitivity, precision, specificity, and duplexing [135].

AmpliTaq Gold® DNA Polymerase: is a key ingredient in Hot Start PCRs. The thermal incubation step required for activation (→"Hot Start") ensures that active enzyme is generated only at temperatures where the DNA is fully denatured. This polymerase is not capable of primer extension at room temperature avoiding that unspecific mispriming events may be enzymatically extended and subsequently amplified.

Uracil-DNA Glycosylase: this enzyme is identical to AmpliTaq Gold DNA polymerase, but is further purified through a proprietary process to reduce bacterial DNA introduced from the host organism. The purification process ensures that non-specific, false-positive DNA products due to bacterial DNA contamination are minimized during PCR.

dNTPs with dUTP: dNTP is a term comprising the four deoxyribonucleotides dATP, dCTP, dGTP and dTTP that are added complementary to the template strand by the polymerase during PCR reaction. dUTP treatment (also know as uracil-N-glycosylase [UNG]), can prevent the reamplification of carryover-PCR products by removing any uracil incorporated into single- or double-stranded amplicons. UDG prevents reamplification of carryover-PCR products in an assay if all previous PCR for that assay was performed using a dUTP-containing Master Mix.

ROX passive reference: provides an internal reference to which the reporter-dye signal can be normalized during data analysis. Normalization is necessary to

correct for fluorescent fluctuations due to changes in concentration or volume [133].

Figure 22: TaqMan Gene Expression Master Mix, modified from [135].

5.2.5 Procedures

The standard protocol has been modified and evaluated for GC rs2282679, DHCR7 rs12785878 and CYP2R1 rs10741657.

Primer and probe sets were designed and manufactured by Applied Biosystems by Life Technologies (Life Tech Austria, Vienna, Austria, Table 13).

All chemicals and reagents are ready for use.

Preparation:

- Defrost DNA 96 DWPs and assay Mixes.

Reaction Mix:

Set up a 5µl reaction per well in a 96 well plate and overlay with 5µl of mineral oil:

- Transfer 2µl of mixed isolated DNA from the 96 DWP into 96 well plate by using a 12-channel pipette.

- Prepare the Master Mix for 96 reactions and calculate 11 reactions as pipetting reserve:

Master Mix in µl	**x1**	**x105**
UMM (2x)	2.5	262.5
A.d.	0.42	44.1
AM (40x)	0.08	8.4
Total	3.0	

Table 11: Components of the Master Mix

- Pipette 3µl of Master Mix into each well of the 96-well plate by using a Multipette.
- Mix DNA and Master Mix.
- Overlay the reaction mix (a total of 5µl) with additional 5 µl of mineral oil.
- Close the 96-well plate with a PCR film.
- Spin down the reaction volume to the bottom of the well by centrifuging up to 1000 rpm.

PCR reaction:

- Perform 40 cycles of PCR using the following temperature profile:

95°C 10 min	1x
92°C 15 sec 60°C 1 min	40x
4°C ∞	1x

Table 12: Cycler conditions of the TaqMan reaction.

Gene	SNP ID	Assay ID	Polymorphism	Context sequence [VIC/FAM]
GC	rs2282679	C___26407519_10	G/T, Transversion Substitution	AGCTAACAATAAAAAATACCTGGCT[G/T]TGTGAGATAATTAAGAGACAGAGAT
CYP2R1	rs10741657	C___2958430_10	A/G, Transition Substitution	TGGTTGGGGAGATACTTTAGCAGGC[A/G]AGGGCTGTCAAGGAAAGTCTTATTA
DHCR7	rs12785878	C__32063037_10	G/T, Transversion Substitution	GGGCTGTCTGATATCACAAAGCTTC[G/T]ATCCTCTCCTGGCCCCGTGGCCGGA

Table 13: Probe sets designed and manufactured by Applied Biosystems by Life Technologies, Invitrogen (Life Tech Austria, Austria).

Figure 23: Eppendorf Mastercycler. Photo: Olivia Trummer

Fluorescence detection:

Endpoint fluorescence of 96 well plates containing cycled PCR reactions is measured in a Fluoroskan Ascent plate reader (Thermo Labsystems, Helsinki, Finland, Figure 24). VIC® dyes are excited with a wavelength of 520 nm (nanometers) and emit light at 520 nm whereas FAM dyes have an excitation of 485 nm and an emission of 520 nm.

Figure 24: Fluoroskan Ascent. Photo: Olivia Trummer.

Measured fluorescence data are exported into Excel format and displayed as a scater plot. Relative fluorescence units (RFU) of VIC® dyes on the y-axis are plotted against RFUs of FAM dyes on the x-axis, as displayed in Figure 25. The right cluster near the x-axis represents the cluster of wildtypes (WT, only probes with FAM dyes match the target sequence of DNA), the cluster in the middle displays individuals with both alleles (heterozygous HE, probes with FAM and VIC® match the DNA sequence). The cluster near the y-axis accordingly shows the homozygous fraction of the investigated samples (HOM, only VIC® dyes match the target sequence. The borders between the clusters are determined by the operating researcher. Visual results are tabulated automatically according to the positions on the predefined 96 well format.

Genotypes of samples in the area under the three clusters are automatically defined as „no genotype available“, coded as “91”.

Genotypes of samples which cannot be definitely assigned to one of the three groups are excluded manually by the operator.

Sample	VIC	FAM
Blank	0.82	3.55
Blank	0.65	2.86
A03	0.65	2.83
A04	0.66	2.85
A05	0.66	2.87
A06	0.83	8.50
A07	1.88	6.32
A08	1.87	5.70
A09	1.79	6.34
A10	2.01	6.94
A11	0.84	7.87
A12	1.46	6.52

Table 14: Raw data of VIC® and FAM fluorescence measurement as an example.

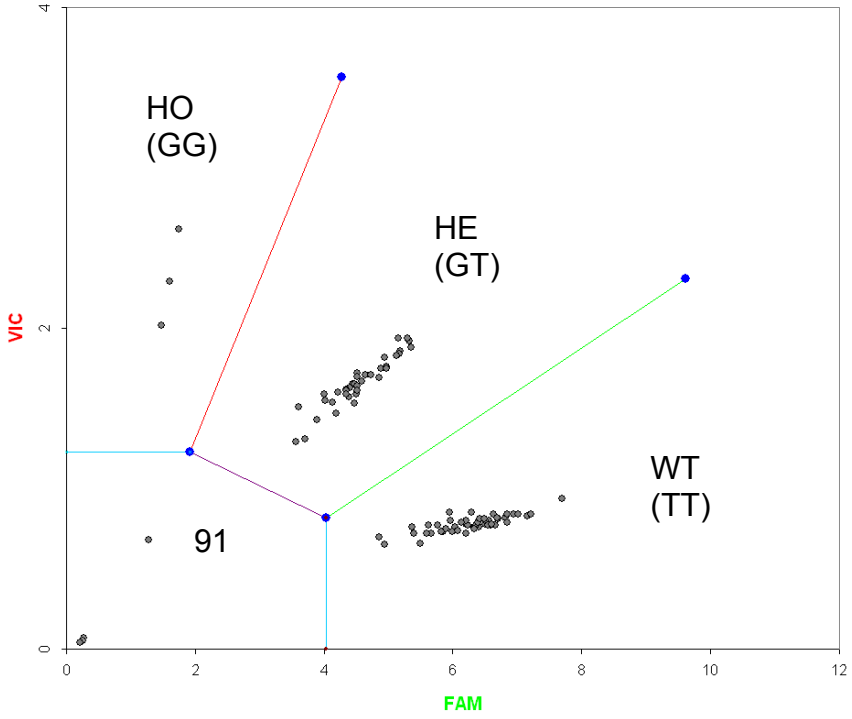


Figure 25: Scatter plot of VIC® (y-axis) and FAM (x-axis) fluorescence data of the GC polymorphism as an example.

Quality controls

- No target controls:

  Each predefined sequence of DNA on a 96 well plate contains a specific pattern of no target controls, specified as empty (e) well positions which are filled up accordingly with A.dest (e.g. Table 10) This unique pattern of no target controls allows to identify contaminations, distortion of the plates when pipetted or measured or sample confusion in the context of a mix-up of whole plates.

- Hardy Weinberg equilibrium[7]:

  Departures from Hardy Weinberg equilibrium (HWE) in investigated cohorts may indicate a real effect of the allele in the population, systematic genotyping errors, or any other kind of bias, like selection bias. Therefore, testing HWE of marker genotype frequencies has been widely recommended as a crucial step in any population-based genetic association study [136]. HWE in these investigations was tested with help of André Rogatkos method to test the hypothesis of equilibrium [137].

- Repetition of Genotyping:

  Of each cohort 41 to at least 96 samples have been genotyped twice to check genotyping quality expressed by number of discrepancies of genotypes between these repetitions.

- Plausibility check:

  Comparison of the frequencies of each genotype group on each plate with the mean of all plates together. Strong aberration of one plate from the mean of all plates may indicate an occurred error (e.g. wrong Assay Mix used for the plate concerned).

7 The **HWE** is a principle stating that genetic variation in a population will remain constant from one generation to the next in the absence of disturbing factors. When mating is random in a large population with no disruptive circumstances, the law predicts that both genotype and allele frequencies will remain constant because they are in equilibrium.

The HWE can be disturbed by a number of forces, including mutations, natural selection, non-random mating, genetic drift, and gene flow. For instance, mutations disrupt the equilibrium of allele frequencies by introducing new alleles into a population. Similarly, natural selection and non-random mating disrupt the HWE because they result in changes in gene frequencies. This occurs because certain alleles help or harm the reproductive success of the organisms that carry them. Another factor that can upset this equilibrium is genetic drift, which occurs when allele frequencies grow higher or lower by chance and typically takes place in small populations. Gene flow, which occurs when breeding between two populations transfers new alleles into a population, can also alter the HWE.

Because all of these disruptive forces commonly occur in nature, the HWE rarely applies in reality. Therefore it describes an idealized state and genetic variations in nature can be measured as changes from this equilibrium state.

After application of all quality controls, data of genotypes and all relevant clinical information were exported into a PASW file (PASW 18.0 software IBM Deutschland GmbH, Ehningen, Germany) to perform statistical analyses.

6.VITAMIN D AND BONE

6.1 Background

The musculoskeletal consequences of inadequate vitamin D levels have intensively been investigated and include accelerated age-related bone loss, as well as morbidity from falls and osteoporotic fractures [138]. Osteoporosis, a complex multifactorial genetic disorder with interaction of environmental and hereditary factors [139], shows numerous different genetic variants contributing to the regulation of this phenotyp. Most are common variants derived from candidate gene approaches or GWA studies with small effect sizes, but rare variants of large effect size may also contribute in certain individuals [140, 141] leading to a high susceptibility to osteoporosis [142].

The GC, DHCR7, and CYP2R1 polymorphism may contribute to BMD and bone fractures, a surrogate and harder endpoint of osteoporosis [117, 118]. Aim of this investigation was to investigate these three allelic determinants of vitamin D insufficiency in two large cohorts of different 25-OH-vitamin D supply and to analyze a possible association with BMD and bone fractures.

6.2 Study subjects

Recruitment, clinical investigation and analysis of laboratory biomarkers of the following cohorts were not carried out in the framework of this thesis nor was my person directly involved in these procedures All conditions are published in detail [143, 144, 145].

Cross-sectional study:

From previous screening studies [143, 144] comprising 628 participants, 286 were excluded due to vitamin D supplementation, missing DNA or missing laboratory data. A flow chart displaying the application of these exclusion criteria is given in Figure 26. The remaining 342 subjects (219 men and 123 women) with a mean age of 55.3 years were investigated for this part of the thesis. All participants were Caucasians and gave their written informed consent before being included in the

study. Clinical examination and questionnaires, anthropometric and bone density measurements by dual energy X-ray absorptiometry (DXA) as well as blood drawing after an overnight fast were performed in all attendants at the Department of Internal Medicine, Division of Endocrinology and Metabolism, at Medical University of Graz. All protocol procedures were approved by the local ethics committee.

Prospective cohort study

A prospective cohort of 1664 participants of 70 years and older was recruited in 95 nursing homes in four provinces in Austria. The study center was located at the Department of Internal Medicine, Division of Endocrinology and Metabolism, at the Medical University of Graz. Patients were screened, enrolled and followed up by five mobile study teams consisting of a physician, a nurse, and a medical student. These teams made day trips to the nursing homes, leaving the study center in the morning and returning in the afternoon [145]. From that cohort, 571 subjects were excluded because of vitamin D supplementation, missing DNA or missing laboratory data and relevant comorbidities such as primary hyperparathyroidism (Figure 26). The remaining 1093 subjects with a mean age of 84.0 years were investigated in this part of the thesis. All subjects were followed from their baseline visit either until death, first hip fracture, or the last follow-up visit, whichever occurred first. At least two staff nurses were responsible for reporting fracture incidents by fax to the study center at the beginning of each month. Additionally, a physician regularly visited all participating centers to search the medical charts for new fracture cases or fatalities. Categories for prospective fractures were hip, radius and other fractures (non vertebral fractures).

All protocol procedures were approved by the local university ethics committee.

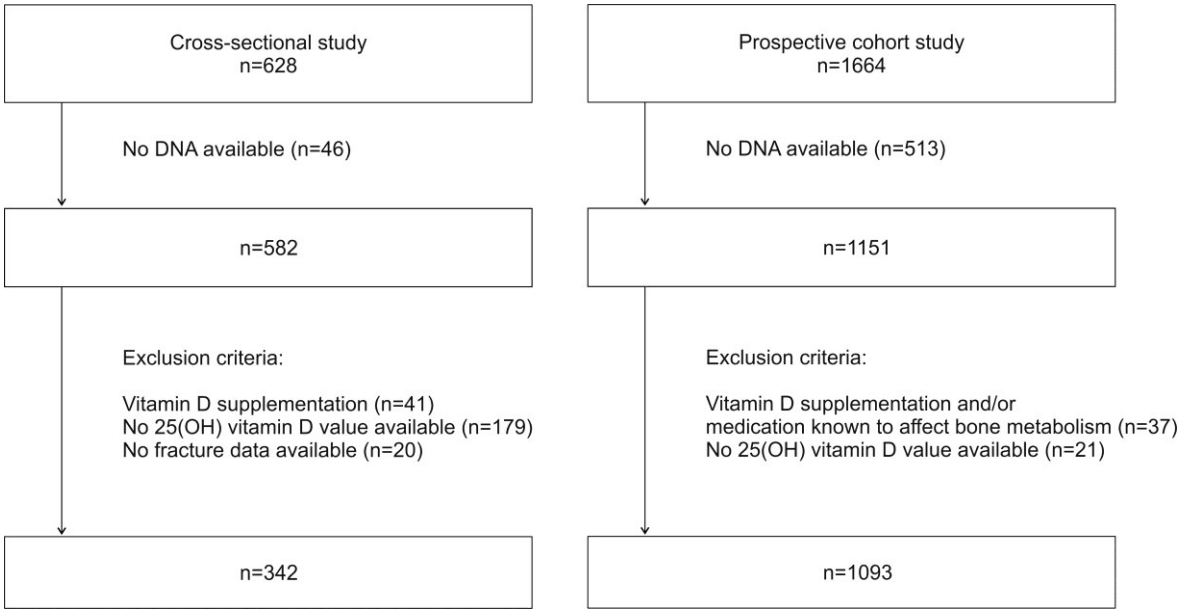


Figure 26: Flow chart of the cross-sectional study and the prospective cohort study, modified from [174].

6.3 Procedures

Clinical investigation and laboratory analysis of biomarkers

Body weight and height were measured in all participants. Blood samples were drawn, sera aliquoted and analyzed for bone-related laboratory parameters on a Hitachi 917 autoanalyzer (Roche, Basel, Switzerland) including albumin-adjusted calcium, phosphate, alkaline phosphatase, and Beta-CrossLaps or stored for later analysis of 25(OH) vitamin D. In the cross-sectional cohort, 25-OH-vitamin D was measured by an enzyme-linked immunosorbent assay (Inkstar, Stillwater, MT USA; intra-assay coefficient of variation [CV] 5.9%, inter-assay CV 6.6%). In the prospective cohort, 25-OH-vitamin D was measured by a radioimmunoassay (Immunodiagnostic Systems, Boldon, UK; intra-assay CV 8.3%; inter-assay CV 12.1%) [146]. In a series of 146 samples measured by both methods, results were in good agreement (slope of relationship = 0.869; y-scale offset 3.7439; R^2 = 0.928). Normal ranges of analyzed parameters are given in Table 15. BMD was measured in the cross-sectional cohort at the lumbar spine (L1-L4), femoral neck and hip by DXA using QDR 4000 plus, Hologic Inc., Erlangen, Germany. The coefficient of variation of repeated measurements in vivo was 2.2% for the lumbar

spine and 2% for the hip. Reference population for age-adjusted comparison (Z-scores) was based on the NHANES database.

In both studies, documentation of prevalent bone fractures was additionally done by current or historical radiographs and medical records [143, 144].

Parameter	Normal range
Calcium	2.20-2.65 mmol/L
Phosphate	2.60-4.50 mg/dL
Alkaline phosphatase	35-105 U/L
Beta-CrossLaps	0.09-0.44 ng/mL
25-OH-vitamin D	30-60 ng/mL

Table 15: Analyzed parameters and their normal range of the cross-sectional and the prospective cohort study.

Genotyping

For genetic analysis, genomic DNA was prepared from EDTA anticoagulated peripheral blood as described in detail in section 5.1 Isolation of DNA. Vitamin D insufficiency polymorphisms GC, DHCR7 and CYP2R1 were analyzed using TaqMan™ fluorogenic 5′-exonuclease assay as depicted in detail in section 5.2 SNP genotyping.

6.4 Statistics

Normal distribution of data was tested with the Kolmogorov-Smirnov test. Categorial variables were compared by the Chi^2 test, multiple group comparisons were carried out by univariate analysis of variance (ANOVA). Data showing unequal variance or lacking normal distribution were analyzed by means of the Kruskal Wallis test. Associations between gene alleles and fractures were calculated by logistic regression analysis including age and sex as potential confounders. Differences were considered significant when a p-value ≤ 0.05 was obtained.

6.5 Results

General results

After application of the exclusion criteria a total of 1093 individuals were available for further analysis in the prospective cohort study (167 men and 926 women) and a total of 342 subjects (219 men and 123 women) in the cross-sectional study. GC, DHCR7 and CYP2R1 genotypes were successfully determined in almost 90% of the individuals of the cross-sectional study and in at least 80% of participants of the prospective cohort. GC, DHCR7 and CYP2R1 genotype frequencies did not deviate from HWE in both cohorts (Table 16).

	Cross-sectional study n = 342 n (%)	HWE p-value	**Prospective cohort study** n = 1093 n (%)	HWE p-value
GC	320 (94)	0.943	950 (87)	0,286
DHCR7	325 (95)	0.844	954 (87)	0,832
CYP2R1	298 (87)	0.356	878 (80)	0,901

Table 16: Overview of successful genotyping rates and Hardy Weinberg equilibrium (HWE) test results in the cross-sectional (n = 342) and the prospective cohort study (n = 1093).

In the cross-sectional cohort, 15% showed a vitamin D deficiency phenotype (n = 51), whereas in the prospective cohort most of the participants were vitamin D deficient (94%, n = 1032), i.e. showed 25(OH) vitamin D levels <20.0 ng/mL. Considering a 25(OH) vitamin D status <30.0 ng/mL as vitamin D deficient, the cross-sectional cohort showed 145 individuals with vitamin D deficiency (42%), whereas again most of the nursing home residents were vitamin D deficient (n = 1076; 98%). General characteristics, anthropometric data, BMD and biochemical measurements are shown in Table 17.

	Cross-sectional study		**Prospective cohort study**	
	men	women	men	women
n	219 (64%)	123 (36%)	167 (15%)	926 (85%)
Age (years)	51.4 ± 11.8 (53.0; 17.0-72.0)	62.1 ± 8.8 (63.0; 40.0-79.0)	84.3 ± 5.8 (85.0; 69.0-99.0)	83.7 ± 6.1 (85.0; 68-100)
Height (cm)	177 ± 7 (177; 160-197)	162 ± 6 (161; 148-176)	166 ± 8 (167; 140-185)	154 ± 7 (154; 128-174)
Weight (kg)	84 ± 12 (83; 60-125)	68 ± 12 (66; 45-98.)	70 ± 12 (70; 44-100)	61 ± 12 (60; 31-110)
25-OH-vitamin D (ng/mL)	33.2 ± 11.6 (33.4; 10.2-70.9)	33.2 ± 15.4 (30.6; 6.5-91.3)	9.7 ± 8.0 (7.2; 0.9-76.3)	8.8 ± 5.9 (7.0; 1.4-54.1)
Vitamin D insufficiency (n/%)				
<20.0 ng/mL	28 (12%)	23 (18%)	158 (94%)	874 (87%)
<30.0 ng/mL	86 (39%)	59 (48%)	164 (98%)	912 (98%)
Calcium [mmol/L]	2.47 ± 0.12 (2.48; 2.02-2.79)	2.36 ± 0.11 (2.35; 2.06-2.06)	2.43 ± 0.12 (2.43; 1.99-2.75)	2.45 ± 0.12 (2.45; 1.80-2.87)
Phosphate [mg/dL]	2.92 ± 0.56 (2.88; 1.50-4.15	3.69 ± 0.58 (3.70; 1.90-5.30)	3.36 ± 0.62 (3.30; 1.70-5.10)	3.76 ± 0.72 (3.70;1.90-7.80)
Alkaline phosphatase [U/L]	69.0 ± 17.5 (66.0; 22.0-140.0)	93.6 ± 29.5 (92.0; 45.0-224.0)	120.5 ± 40.0 (115.0; 60.0-274.0)	120.0 ± 39.5 (113.0; 43.0-283.0)
Beta-CrossLaps (ng/mL)	0.23 ± 0.11 (0.19; 0.04-0.47)	0.22 ± 0.14 (0.19; 0.04-0.80)	0.33 ± 0.19 (0.29; 0.01-0.85)	0.39 ± 0.29 (32.0; 0.10-3.10)

Lumbar BMD (g/cm^2)	1.01 ± 0.15 (0.99; 0.65-1.52)	0.89 ± 0.17 (0.89; 0.53-1.71)	n.a.	n.a.
Femoral neck BMD (g/cm^2)	0.85 ± 0.12 (0.84; 0.59-1.41)	0.70 ± 0.12 (0.70; 0.50-1.00)	n.a.	n.a.
Total hip BMD (g/cm2)	1.02 ± 0.12 (1.01; 0.72-1.44)	0.85 ± 0.14 (0.86; 0.50-1.25)	n.a.	n.a.
Prevalent bone fractures (n)	93	40	65	446

Table 17: Demographic data from the cross-sectional and from the prospective cohort study.
Categorical data are shown as number (percentage) and metric data as mean ± SD (median; range). 25-OH-vitamin D levels were normally distributed in the cross-sectional study, but not in the prospective cohort study. Frequency data are presented as number (percentage), and continuous data as mean ± SD (median; range). n.a.: data are not available, BMD: bone mineral density.

Vitamin D insufficiency genotypes and 25-OH-vitamin D

The GG genotype of the GC polymorphism was associated with lower 25-OH-vitamin D values in the cross-sectional study ($p = 0.001$) as well as in the prospective cohort study ($p = 0.048$). No significant association was found between CYP2R1 or DHCR7 variants and 25-OH-vitamin D values in this investigation (Table 18).

	Cross-sectional cohort		Prospective cohort	
	n (%)	25-OH-vitamin D (ng/mL)	n (%)	25-OH-vitamin D (ng/mL)
CYP2R1				
AA	42 (12.3)	34.2 ± 14.7 (34.5; 10.0-75.2)	147 (13.4)	9.8 ± 7.9 (7.0; 2.8-47.0
AG	155 (45.3)	34.1 ± 12.6 (33.3; 8.7-70.9)	416 (38.1)	8.8 ± 6.4 (7.0; 0.9-76.3)
GG	101 (29.5)	32.2 ± 11.9 (32.0;10.5-64.5)	315 (28.8)	8.8 ± 6.1 (6.9; 3.0-54.1)
p-value		0.460		0.575
DHCR7				
GG	22 (6.4)	31.4±12.8 (29.5; 11.5-62.8)	79 (7.2)	8.9 ± 8.0 (6.6; 3.5-54.1)
GT	117 (34.2)	32.1±13.0 (29.2; 10.5-64.2)	405 (37.1)	8.8 ± 6.1 (6.9; 0.9-47.0)
TT	186 (54.4)	34.7±13.1 (34.2; 10.5-64.2)	470 (43.0)	9.1 ± 6.7 (7.0; 1.4-76.3)
p-value		0.161		0.411
GC				
GG	21 (6.1)	26.1 ± 8.7 (26.2; 10.5-41.5)	73 (6.7)	7.5 ± 4.6 (6.3; 3.0-31.0)
GT	117 (34.2)	31.8 ± 12.9 (30.6; 10.2-65.8)	416 (38.1)	8.7 ± 5.5 (6.9; 2.4;-44.8)
TT	182 (53.2)	35.9 ± 13.3 (34.7; 8.7-75.2)	461 (42.2)	9.4 ± 7.6 (7.1; 0.9-76.3)
p-value		**0.001**		**0.048**

Table 18: 25-OH-vitamin D values according to CYP2R1 (rs10741657), DHCR7 (rs12785878) and GC (rs2282679) genotypes in the cross-sectional and the prospective cohort study.
Data are shown as numbers (%), 25(OH) vitamin D values as mean ± standard deviation (median; range). 25(OH) vitamin D levels were normally distributed in the cross-sectional study, but not in the prospective cohort study.

Vitamin D insufficiency genotypes and BMD

Genotypes of these genetic determinants were further analyzed in association with BMD in the cross-sectional study. Genotype groups did not differ in BMD at any site measured.

Vitamin D insufficiency genotypes and bone fractures

For each gene, the allele associated with lower 25-OH-vitamin D levels was designated as "risk allele" (GC G-allele; CYP2R1 A-allele; DHCR7 G-allele). The potential role of these risk alleles in fracture risk was analyzed in logistic regression analysis including age and sex as confounders. This calculation showed that none of the risk alleles was linked with fractures of the past, neither in the cross-sectional group nor in the prospective cohort (Table 19).

The DHCR7 T-allele, associated with higher 25-OH-vitamin D levels, was significantly associated with a reduced risk of prospective fractures (OR 0.68; 95% confidence interval 0.51 - 0.92; p = 0.011), whereas clearly no association was seen for the GC G-allele or the CYP2R1 A-allele (Table 19).

Cross-sectional study, past fractures:				
	Genotype (n)	No fracture	Any fracture	OR (95% CI); p
CYP2R1	AA (42)	26 (61.9%)	16 (38.1%)	
rs10741657	AG (155)	94 (60.6%)	61 (39.4%)	
	GG (101)	65 (64.4%)	36 (35.6%)	0.826 (0.57 – 1.19); p = 0.307
DHCR7	GG (22)	14 (63.6%)	8 (36.4%)	
rs12785878	GT (117)	68 (58.1%)	49 (41.9%)	
	TT (186)	117 (62.9%)	69 (37.1%)	0.96 (0.66 – 1.38); p = 0.816
GC	GG (21)	11 (52.4%)	10 (47.6%)	
rs2282679	GT (117)	74 (63.2%)	43 (36.8%)	
	TT (182)	112 (61.5%)	70 (38.5%)	0.94 (0.65 – 1.36); p = 0.756
Prospective cohort study, past fractures				
	Genotype (n)	No fracture	Any fracture	OR (95% CI); p
CYP2R1	AA (147)	89 (60.5%)	58 (39.5%)	
rs10741657	AG (416)	222 (53.4%)	194 (46.6%)	
	GG (315)	163 (51.7%)	152 (48.3%)	1.18 (0.97 – 1.43); p = 0.094
DHCR7	GG (79)	41 (51.9%)	38 (48.1%)	
rs12785878	GT (405)	222 (54.8%)	183 (45.2%)	
	TT (470)	254 (54.0%)	216 (46.0%)	0.98 (0.80 – 1.20); p = 0.856
GC	GG (73)	37 (50.7%)	36 (49.3%)	
rs2282679	GT (416)	218 (52.4%)	198 (47.6%)	
	TT (461)	263 (57.0%)	198 (43.0%)	0.86 (0.70 – 1.05); p = 0.141
Prospective cohort study, prospective fractures				
	Genotype (n)	No fracture	Any fracture	OR (95% CI); p
CYP2R1	AA (147)	131 (89.1%)	16 (10.9%)	

rs10741657	AG (416)	371 (89.2%)	45 (10.8%)	
	GG (315)	275 (87.3%)	40 (12.7%)	1.11 (0.82 – 1.50); p = 0.506
DHCR7	GG (79)	67 (84.8%)	12 (15.2%)	
rs12785878	GT (405)	346 (85.4%)	59 (14.6%)	
	TT (470)	428 (91.1%)	42 (8.9%)	**0.68 (0.51 -.0.92); p = 0.011**
GC	GG (73)	62 (84.9%)	11 (15.1%)	
rs2282679	GT (416)	378 (91.9%)	38 (9.1%)	
	TT (461)	392 (85%)	69 (15.0%)	1.29 (0.93 - 1.77); p = 0.125

Table 19: Past and prospective fractures stratified by CYP2R1, DHCR7 and GC genotypes.
Data are presented as number of patients (%) without and with any fractures, odds ratio (OR) and 95% confidence interval (CI), p-value.

6.6 Discussion

This investigation replicates a previously shown association of the GC polymorphism with vitamin D insufficiency. For the first time, this association is visible even in a cohort of nursing home residents with typically very low levels of vitamin D. Moreover, an associaton of the DHCR7 G allele with bone fractures has been discovered. However, we found no association of BMD with any of those single nucleotide polymorphisms.

Further genetic variants in the GC gene (rs4588 and rs7041) have previously been published to be associated with 25-OH-vitamin D levels [139, 147, 148, 149, 150]. We analyzed this vitamin D modulating effect with a new polymorphism (rs2282679) in this gene. All of the previously described data were investigated in populations with normal to moderately decreased 25-OH-vitamin D serum levels. For the first time, we showed that this effect is reproducible down to the lowest 25-OH-vitamin D levels in a very old population of nursing home subjects.

Nevertheless, bone-specific influences of the very small differences of 25-OH-vitamin D levels between genotype groups might have been overwhelmed by the severe overall vitamin D deficiency status in these nursing home residents. Therefore, consequences for bone fractures in diverse genotype groups might have been surpassed.

Lauridsen et al. [147] found a varying number of premenopausal bone fractures between the GC genotypes (rs7041 and rs4588) in a cross-sectional study of almost 600 early postmenopausal Danish women. In our cross-sectional cohort, fracture cases might have been too rare to show any distinction and documentation especially for vertebral fractures might not have been accurate [151]. However, Fang et al. studied GC rs7041 and rs4588 genotypes in 6181 Caucasians with a mean age of almost 70 years. After performing subgroups with vitamin D receptor haplotypes and dietary calcium intake, they identified an increased fracture risk for these variants [150]. It might therefore be possible that in combination with other risk factors CYP2R1, DHCR7 and GC genotypes may as well contribute to an increased risk of fractures despite lack of a detectable direct effect in our prospective nursing home cohort. A risk score with combined “risk alleles” for lower 25-OH-vitamin D levels did not add more information to our findings. Therefore, these data are not shown.

It has to be mentioned that none of the GC genotypes showed an association with BMD in the investigated cross-sectional cohort. This is in line with Lauridsen et al. [147] who also did not find any varying BMD between GC phenotypes of rs4588 and rs7041. Besides an association of the GC polymorphism with 25-OH-vitamin D level, we found consequences for one important endpoint (fractures) but not for a second endpoint (BMD) in osteoporosis. For the prospective cohort (n = 1093), the study had a statistical power of 0.80 or more to detect an OR of 1.5 for each polymorphism for previous fractures. The statistical power was lower in the smaller cross-sectional study (n = 342) and for prospective fractures. Additional studies will be necessary to detect or exclude smaller effects of genetic variants on bone fracture risk.

A relation between fracture risk and vitamin D has long been discussed in literature. Many investigations including meta-analyses from Bischoff-Ferrari [152] have revealed a close correlation between 25-OH-vitamin D serum levels and the dosage of vitamin D applied to the participants. The mechanisms underlying the association between 25-OH-vitamin D levels and fracture risk are currently not fully understood. In the past several years many health aspects have been linked to 25-OH-vitamin D status including muscle weakness [153], impaired body sway and cardiovascular problems [154] such as high blood pressure [155] and cardiac impairment [156]. Almost all of theses effects may play a role in the etiology of

fractures. Further studies, particularly randomized clinical trials, are required to fully elucidate the role of 25-OH-vitamin D in fracture risk.

We can find several examples in literature where single nucleotide polymorphisms clearly associate to clinical biomarkers, known as risk factors for certain diseases. However, no associations between these variants and clinical endpoints have been reported. A classic example is homocysteine which has been suggested as a risk factor for cardiovascular diseases [157, 158]. A functional polymorphism in the methylenetetrahydrofolate reductase gene (MTHFR, 677C>T, rs1801133) which causes reduced enzyme activity is in association with higher homocysteine levels [159]. Remarkably the polymorphism itself was not associated with coronary artery disease [160]. We don´t know yet wheter this phenomenon depends on a “real” genetic loci lying behind the outcomes or on more complex genetic interactions or on epigenetic processes.

Of note, our study has some limitations: Due to technical and ethical reasons it was not possible to measure BMD and vertebral X-rays in the nursing home residents. Furthermore, the present study included three SNPs which had previously been associated with vitamin D levels in two large genome-wide association studies [117, 118]. One should not forget that these SNPs may not be the causal variants, but could rather be in linkage disequilibrium with them. Other yet unknown SNPs within these or other genes may contribute to osteoporosis and/or fracture risk. It has to be mentioned that the rate of past fractures was higher in the cross-sectional study compared to the prospective study, which contained nursing home residents of higher age. The cross-sectional study included patients volunteering for an osteoporosis screening study at the Division of Endocrinology and Metabolism. It is likely that patients with previous fractures have a higher interest in an osteoporosis screening and are more likely to participate in a study than patients without fractures. This possible selection-bias could explain the relatively higher number of previous fractures in the cross-sectional study compared to the prospective study. Further, due to predefined exclusion criteria we had to exclude participants with vitamin D supplementation, missing DNA or missing laboratory data which again can not exclude a selection bias. All this taken into account might have contributed to the not significant but strongly consistent differences in 25-OH-vitamin D levels throughout the groups in both cohorts. Finally, the size and the statistical power (especially for the cross-

sectional study) of the present study were limited and therefore our results need to be replicated in other cohorts.

To the best of our knowledge, this is the first study investigating novel allelic determinants of vitamin D insufficiency and their influence on BMD and bone fractures, showing a significant association of GC variants with 25-OH-vitamin D levels and DHCR7 variants with bone fractures. These findings suggest potential new insights and targets for therapeutic intervention in osteoporosis patients.

7.VITAMIN D AND PROSTATE CANCER

7.1 Background

Prostate cancer is the second most common cancer [161]. Although it is generally a slowly progressing malignancy it shows considerable rates of mortality so that prostate cancer prevention and treatment is an ongoing challenge for clinicians worldwide.

The only established risk factors for prostate cancer are age, family history of prostate cancer and ethnicity. As in the general introduction described, preclinical and epidemiologic data suggest that vitamin D deficiency may play a role in the pathogenesis and progression of prostate cancer [162]. This hypothesis is based on evidence that the vitamin D endocrine system regulates normal prostate growth, cell proliferation, differentiation as well as apoptosis and that black ethnicity and residence in northern latitudes are associated with low circulating 25-OH-vitamin D levels [163, 164, 165].

In the prostate locally produced 1,25 $(OH)_2$ vitamin D [166, 167] may be important for inhibiting cell proliferation and increasing differentiation. In contrast to renal 1α-hydroxylase activity, local prostate production of 1,25 $(OH)_2$ vitamin D seems to be significantly dependent on the substrate availability of 25-OH-vitamin D [168]. It is a widely accepted hypothesis that serum levels of 25-OH-vitamin D are the main determinant of 1,25 $(OH)_2$ vitamin D tissue levels in various organs [176].

Since prostate cancer tissue has reduced 1α-hydroxylase activity [169] and therefore reduced ability to locally convert 25-OH-vitamin D to 1,25-$(OH)_2$-vitamin D, the main source of 1,25 $(OH)_2$ vitamin D to prostate cancer may be from the circulation. Both metabolites may therefore have protective potential [163, 164] and are warrant investigation.

Aim of the present study was to test a possible association between the GC polymorphism and 1,25 $(OH)_2$ vitamin D as well as to replicate the previously reported association between this polymorphism and 25-OH-vitamin D in a large cross-sectional study (LUdwigshafen RIsk and Cardiovascular Health study, LURIC-study).

Further, I aimed to address the role of the GC polymorphism for prostate cancer risk concerning biochemically detected recurrence, development of metastases, and survival in prostate cancer patients from the PROCAGENE study.

7.2 Study subjects

I want to state that recruitment, clinical investigation and laboratory analysis of biomarkers of the following two cohorts was not carried out in the framework of this thesis nor was my person involved in any of these procedures.

LURIC study

The Ludwigshafen Risk and Cardiovascular Health (LURIC) study includes consecutive Caucasian patients hospitalized for coronary angiography between June 1997 and May 2001. The study was approved by the ethics review committee at the “Landesärztekammer Rheinland-Pfalz” (Mainz, Germany) and written informed consent was obtained from all participants. A detailed description of the LURIC study design and baseline characteristics has been published previously [170].From a total of 3316 LURIC participants, only men (n = 2310) were analyzed in the present study.

The PROCAGENE study

The Austrian Prostate Cancer Genetics study comprises 702 prostate cancer patients who were recruited between January 2004 and January 2007. Briefly, PROCAGENE is a prospective study aimed at investigating genetic risk factors, functional relationships between genetic variations and clinical phenotypes, genetically modified response to radiotherapy (radiogenomics), and the prognostic importance of genetic markers for prostate cancer. A more detailed description has been published elsewhere [171]. The study was performed according to the Austrian Gene Technology Act and was approved by the Ethical Committee of the Medical University of Graz. Written informed consent was obtained from all participating subjects and all prostate cancer patients were Caucasians of central–European origin (Austria).

7.3 Procedures

Clinical Investigation and laboratory analysis of biomarkers

LURIC study:

25-OH-vitamin D levels (normal range: 30-60 ng/ml) and 1,25 $(OH)_2$ vitamin D values (normal range: 39-193 pmol/l) were analyzed using radioimmunoassay (25-OH-vitamin D: DiaSorin GmbH, Vienna, Austria, and 1,25 $(OH)_2$ vitamin D: Nichols Institute Diagnostika GmbH, Bad Nauheim, Germany). Intra-assay and inter-assay coefficients of variation were below 10%.

PROCAGENE study:

Clinical properties comprising family history, tumor node metastasis (TNM) stage, histological grade, prostate-specific antigen (PSA), age at time of diagnosis, and means of diagnosis were obtained from medical records. Patients were graded into low- and high-risk groups on the basis of pretreatment prostate-specific antigen (PSA) level, Gleason score, and American Joint Commission on Cancer T-stage. The patients were classified as high risk if they met any of the following criteria: T3/4, grade 3, Gleason score sum 8–10, or PSA>20; otherwise, they were classified as low risk. Follow-up examinations were performed in regular intervals at the Department of Therapeutic Radiology and Oncology in cooperation with the Department of Urology (3 months interval in years 1–3, 6 months interval in years 4–5, and 12 months interval in years 6–15 after diagnosis). PSA measurements and digital rectal examinations (DRE) were performed routinely. The metastatic workup included isotope bone scan, chest X-ray as well as abdominal and pelvic computed tomography. Biochemically detected recurrence was defined according to the Phoenix definition, with absolute nadir +2 ng/ml rise [172] or as the initiation of hormonal therapy due to rising PSA. The primary endpoints were (1) biochemically detected recurrence, (2) development of metastases, and (3) death.

Genotyping

Upon study entry, EDTA-blood was drawn from each participant of the PROCAGENE or LURIC study.and stored at -20°C. Genomic DNA was prepared from EDTA anticoagulated peripherial blood using the described protocol in section 5.1 of the chapter Material and Methods.

Genotypes were analyzed by fluorogenic 5′-exonuclease assays (TaqMan™) as delineated in section 5.2 SNP genotyping.

7.4 Statistics

Groups of continuous variables were compared by ANOVA. The relationship between GC genotypes and the outcome was analyzed by Cox regression analyses. Median follow-up time was calculated according to Schemper and Smith [173]. For this investigation, genotypes were coded as 0 ("wildtype", homozygous for the common T allele), 1 (heterozygous for TG) or 2 (homozygous for the minor allele G). The criterion for statistical significance was $p < 0.05$.

7.5 Results

General results

GC genotypes were successfully analyzed in approximately 95% of both study populations (n = 2179 [94.3 %] male participants of the LURIC study and n = 674 [96.0%] patients of the PROCAGENE study). In the remaining subjects, no DNA was available or genotypes were considered as non-interpretable after two repeats. The frequency of the minor allele (G) was 0.562 in men of the LURIC study and 0.313 in the PROCAGENE study. Genotype frequencies did not deviate from HWE in both cohorts. Demographic data for each study are shown in Table 20 and Table 21.

n		702
Age at diagnosis (years)		68.1 ± 7.2
Stage	T1/T2	362
	T3/T4	293
Risk group	Low risk	457
	High risk	244
GC genotype	TT	318 (47.2%)
	TG	290 (43.0%)
	GG	66 (9.8%)

Table 20: Demographic and genetic data of the PROCAGENE study.

n	2310
Age (years)	61.8 ± 10.7
Height (cm)	174 ± 7
Weight (kg)	83.4 ± 12.9
25-OH-vitamin D (ng/mL)	18.2 ± 9.3
1,25 $(OH)_2$ vitamin D (pmol/l)	35.9 ± 13.9

Table 21: Demographic data on the LURIC study. Data are presented as mean ± standard deviation or n (%).

GC genotypes and vitamin D levels in the LURIC study

In the LURIC study, the G allele of the GC polymorphism was associated with lower 25-OH-vitamin D levels as well as with lower 1,25 $(OH)_2$ vitamin D levels (Table 22).

GC genotype LURIC study	n (%)	1,25 $(OH)_2$ vitamin D [pmol/l]	25-OH-vitamin D [ng/ml]
TT	1131 (51.9)	36.4 ± 14.4	18.6 ± 9.3
TG	870 (39.9)	35.5 ± 13.8	17.9 ± 9.4
GG	178 (8.2)	32.7 ± 11.1	15.1 ± 7.4
p-value (ANOVA)		0.004	<0.001

Table 22: GC genotypes and vitamin D levels (1,25 $(OH)_2$ vitamin D and 25-OH-vitamin D) in the LURIC study. Data are presented as numbers (%), vitamin D values as mean ± standard deviation.

GC genotypes, biochemically detected recurrence, metastasis, and survival in the PROCAGENE study

Due to different event rates and availability of follow-up data, median follow-up times were different for biochemically detected recurrence, development of metastases and death. Follow-up of biochemically detected recurrence was available for 606 (86.3%) patients of the PROCAGENE study. During a median follow-up time of 82 months, biochemically detected recurrence was detected in 179 (29.5%) patients. Follow-up for development of metastases was observed for 701 (99.9%) of patients. During a median follow-up time of 73 months, metastases were detected in 93 (13.3%) patients. Survival follow-up was available for all patients of the PROCAGENE study. During a median follow-up time of 76 months, 123 (17.5%) patients died.

In a univariate Cox regression analysis, GC genotypes were not associated with biochemically detected recurrence (hazard ratio [HR] 089, 95% CI 0.70 - 1.13; p = 0.32), development of metastases (HR 1.20, 95% CI 0.88 - 1.63; p = 0.25) or overall survival (HR 1.10; 95% CI 0.84 - 1.43; p = 0.50. In multivariate Cox regression analysis including the risk-group as potential confounder, similar HR were obtained biochemically detected recurrence: HR 0.92, 95% CI 0.73 – 1.17; p = 0.50; metastasis: HR 1.27, 95% CI 0.93 – 1.72; p = 0.13; overall survival: HR 1.10, 95% CI 0.84 – 1.44; p = 0.49).

7.6 Discussion

The functional GC polymorphism is strongly associated with lower serum 25-OH-vitamin D concentrations [117, 118, 174]. This investigation shows for the first time an additional association of this polymorphism with serum 1,25 $(OH)_2$ vitamin D levels. Both, low levels of 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D have been suggested as risk factors for prostate cancer [2, 4]. Therefore we aimed to investigate the potential role of these vitamin D predictors in three important endpoits of prostate cancer. Our investigation shows that GC genotypes were not associated with biochemically detected recurrence, development of metastases, or overall survival in prostate cancer patients. We therefore conclude that the GC polymorphism does not play a major role in prostate cancer.

A recent meta-analysis by Gilbert and co-workers did not find a significant association of 25-OH-vitamin D levels with total prostate cancer risk or risk of aggressive prostate cancer [175]. These findings confirm our results from this investigation.

A limitation of the study is that 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D levels could not be determined in the PROCAGENE study. We therefore tested the association of the GC polymorphism with 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D levels in 2310 male participants of the cross-sectional LURIC study. We are aware of the fact that a specific cohort of patients referred to coronary angiography has been investigated and that our results may not be generalizable to the PROCAGENE cohort. Despite careful adjustments for possible confounders, we cannot rule out residual confounding.

Due to the confirmed association of the GC polymorphism with 25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D we continued investigating the potential role of the GC polymorphism in prostate cancer. Our study showed non-significant HRs that GC genotypes were not associated with biochemically detected recurrence, development of metastases or overall survival. It should be kept in mind that this investigation had a statistical power of 0.80 to detect an OR $\geq$ 1.3 for carriers of a G allele. We can therefore currently not exclude subtle effects of the GC polymorphism on prostate cancer risk. Our results should be therefore, as it is always the case for genetic association studies, interpreted cautiously until replicated by others.

Nevertheless, according to the results of our investigation, predictors of vitamin D do not play a major role in disease progression and mortality of prostate cancer patients. These results suggest that the GC polymorphism a determinant of vitamin D (25-OH-vitamin D and 1,25 $(OH)_2$ vitamin D) does not play a major role in prostate cancer. The present study was observational and aimed not to investigate the potential benefits of vitamin D supplementation. To investigate causality of vitamin D in prostate cancer Mendelian randomization studies are urgently required.

8.VITAMIN D AND MORTALITY

8.1 Background

Vitamin D is classical known for its functional role in calcium and bone metabolism [4, 9]. Attention has extended to other health effects of 25-OH-vitamin D such as on the cardiovascular system [176], blood pressure [177] or metabolic syndrome [178, 179]. In previous studies, it has been shown that low 25-OH-vitamin D concentrations are associated with all-cause and cardiovascular mortality [180, 181]. It remains unclear whether 25-OH-vitamin D is the cause or the consequence in this association [176].

The principle of Mendelian randomization has been desicribed in detail in section 3.3. Briefly, because genetic polymorphisms are randomly assorted at conception, all covariates between individuals should, theorectically, be perfectly balanced between those with different genotypes thus eliminating all confounding for any association between the genetic polymorphism and disease. Furthermore, if one assumes that the measured genetic polymorphism directly affects the gene product (e.g. a biomarker), then one can also estimate the unconfounded association between the gene product and disease [182].

Genetic determinants for vitamin D insufficiency such as the GC-, DHCR7- and CYP2R1 polymorphism are good candidates to perform a Mendelian randomization study [117, 118]. We therefore analyzed the role of these three 25-OH-vitamin-D-associated SNPs with all-cause, cardiovascular and non-cardiovascular mortality. Aim of the study was to assess a possible causal role of 25-OH-vitamin D in this association.

8.2 Study subjects

I want to state that recruitment, clinical investigation and laboratory analysis of biomarkers of the following cohort was not carried out in the framework of this thesis nor was my person involved directly in the procedures of these concerns. The following cohort is the same which has been used in part to investigate vitamin D and prostate cancer in the previous chapter.

The Ludwigshafen Risk and Cardiovascular Health (LURIC) study is a prospective cohort trial designed to evaluate the effect of genetic polymorphisms and plasma biomarkers on cardiovascular health. Cardiovascular and metabolic phenotypes have been defined or ruled out using standardised methodologies in all study participants. As already mentioned, participants are Caucasian patients hospitalized for coronary angiography between June 1997 and May 2001. The study was approved by the ethics review committee at the “Landesärztekammer Rheinland-Pfalz” (Mainz, Germany). Written informed consent was obtained from each of the participants.

8.3 Procedures

Clinical investigation, laboratory analysis of biomarkers and genotyping

A detailed description of the LURIC study design and baseline characteristics has been published previously [170]. Briefly, the study population comprised 3316 participants. According to the classification of the American Heart Association, coronary artery disease (CAD) was defined as the presence of a visible luminal narrowing (≥20% stenosis) in at least one of 15 coronary segments [183]. Individuals with stenosis <20% were considered as not having CAD. Cardiovascular risk factors such as type 2 diabetes, hypertension and smoking were assessed. Type 2 diabetes mellitus was diagnosed according to the criteria of the American Diabetes Association (ADA). Further, individuals with a history of type 2 diabetes or those receiving oral antidiabetics or insulin were considered diabetic [184]. Hypertension was defined as a systolic and/or diastolic blood pressure exceeding 140 and/or 90 mmHg, respectively, or a history of hypertension documented in medical records. Data on smoking habits were retrieved using questionnaires. To detect “hidden” smokers, plasma cotinine levels were determined using a commercial radioimmunoassay (cotinine RIA; DPC Biermann, Germany). Individuals suffering from acute illnesses other than acute coronary syndromes, chronic non-cardiac diseases and a history of malignancy within the past five years were not eligible.

A fasting blood sample was obtained in the morning before coronary angiography. Selected variables were measured after samples were frozen and stored at -80°C. 25-OH-vitamin D levels were determined using a radioimmunoassay (DiaSorin

GmbH, Vienna, Austria) with intra-assay and inter-assay coefficients of variation of 8.6% and 9.2%, respectively.

Information on mortality rates was collected from local registries. Death certificates were used to classify the deceased into patients who died from cardiovascular versus non-cardiovascular causes. This classification was performed independently by two experienced clinicians who were blinded to any data on the study participants except the information that was required to classify the cause of death.

Upon study entry, each participant of the LURIC study donated a tube of EDTA-blood, which was stored at -20°C. Genomic DNA was prepared from EDTA anticoagulated peripherial blood and genotyping was performed according to the described protocol in section 6.1 of the chapter Material and Methods.

8.4 Statistics

Continuous variables were compared between groups by ANOVA. A linear regression model was used to identify the predictors of 25-OH-vitamin D levels. Cox regression including GC, DHCR7 and CYP2R1 genotypes, age, sex, CAD, smoking and type 2 diabetes mellitus was used to estimate the effect of the variants on mortality (all-cause, cardiovascular and non-cardiovascular). For regression analyses, an allelic model based upon additive gene-dose effects was used and genotypes were coded as 0 ("wildtype", homozygous for common allele), 1 (heterozygous) or 2 (homozygous for minor allele). To take seasonal variations of vitamin D levels into account, categorical variables were adjusted for month of blood taking. The criterion for statistical significance was $p < 0.05$.

8.5 Results

The investigated cohort comprised 3316 persons, including 2310 men (69.7%) and 1006 women (30.3%). GC, DHCR7 and CYP2R1 genotypes were successfully analyzed in 3130 (94.4%), 3109 (93.8%) and 2980 (89.9%) participants. All genotype frequencies did not deviate from HWE. As for the GC and the DHCR7 polymorphism, minor alleles (GC G-allele, DHCR7 G-allele) were associated with lower 25-OH-vitamin D levels. For the CYP2R1 polymorphism, the minor allele (CYP2R1 A-allele) was associated with higher 25-OH-vitamin D levels (Table 24).

Demographic data of the LURIC study are given in Table 23.

Age (years)	62.6 ± 10.6
Male gender	2310 (69.7)
Height (cm)	170 ± 8.7
Weight (kg)	79.7 ± 14.1
25-OH-vitamin D (ng/mL)	17.4 ± 9.7
Vitamin D insufficiency; n (%)	2174 (65.6)
Coronary artery disease; n (%)	2583 (77.9)
Current smoker; n (%)	654 (19.7)
Type 2 diabetes; n (%)	1064 (32.1)
Hypertension; n (%)	2412 (72.7)

Table 23: Demographic data of the LURIC study.
Data are mean ± standard deviation or n (%). Vitamin D insufficiency: 25-OH-vitamin D <=20 ng/ml.

		n (%)	25-OH-vitamin D	p *	p †
GC genotype	TT	1622 (51.8)	18.1 ± 10.2		
	GT	1247 (39.8)	16.9 ± 9.3		
	GG	261 (8.3)	14.3 ± 7.8	<0.001	<0.001
GC G allele frequency		0.283			
DHCR7 genotype	TT	1804 (58.0)	17.9 ± 9.5		
	GT	1110 (35.7)	16.4 ± 10.1		
	GG	261 (8.3)	16.6 ± 9.3	<0.001	<0.001
DHCR7 G allele frequency		0.257			
CYP2R1 genotype	GG	1116 (37.4)	16.74± 9.8		
	AG	1433 (48.1)	17.5 ± 9.4		
	AA	431 (14.5)	17.9 ± 10.3	0.030	0.062
CYP2R1 A allele frequency		0.385			

Table 24: 25-OH-vitamin D-associated genotypes and 25-OH-vitamin D levels. 25-OH-vitamin D values according to GC (rs2282679), DHCR7 (rs12785878) and CYP2R1 (rs10741657) genotypes in the LURIC study. Data are shown as numbers (%), 25-OH-vitamin D values as mean ± standard deviation.* Univariate linear regression, † ANOVA.

25-OH-vitamin D levels were available from 99.5% of the entire study population (n = 3299). In a linear regression model adjusted for month of blood sampling, age and sex, vitamin D levels were predicted by GC genotype (p<0.001), CYP2R1 genotype (p=0.068) and DHCR7 genotype (p<0.001), with a coefficient of determination (r^2) of 0.175.

After a median follow-up of 9.9 years 30.0% of the study population at baseline had died (n = 995). Of these, 619 (62%) deaths were from cardiovascular causes and 355 (36%) were from non-cardiovascular causes. 21 (2%) deaths could not be classified because of insufficient data on the cause of death.

In multivariate Cox regression models including age, sex, type 2 diabetes, CAD, smoking habit, 25-OH-vitamin D levels, GC genotype, CYP2R1 genotype and DHCR7 genotype, none of the genotypes was significantly associated with all-

cause mortality, cardiovascular mortality or non-cardiovascular mortality whereas other classical risk factors for mortality showed a significant association (Table 25). Further, a linear regression model including genotypes of GC, CYP2R1 and DHCR7 to predict a genetically determined 25-OH-vitamin D level for each study participant was calculated. In that model, the genetically determined 25-OH-vitamin D level (ng/ml) was 18,158 - (GC genotype * 1.544) - (DHCR7 genotype * 0.981) + (CYP2R1 genotype * 0.588). The mortality HR per 1 ng/ml genetically determined 25-OH-vitamin D level was 1.015 (95% confidence interval (CI) 0.962 - 1.070; p = 0.59). By contrast, the mortality per 1 ng/ml serum 25-OH-vitamin level was 0.951 (95% CI 0.943 - 0.959; p < 0.001). The point estimate of the effect of the genetically determined (i.e. causal) 25-OH-vitamin D levels was not within the 95% CI of the effect of the 25-OH-vitamin D serum level, suggesting that the two estimates are truly different.

Risk factor	**All-cause mortality**	**Cardiovascular mortality**	**Non-cardiovascular mortality**
	HR (95% CI); p	HR (95% CI); p	HR (95% CI); p
GC (G alleles)	0.95 (0.86 – 1.06); p = 0.35	0.90 (0.78 – 1.02); p = 0.106	1.03 (0.86 – 1.22); p = 0.768
CYP2R1 (A alleles)	1.00 (0.91 - 1.10); p = 0.95	0.94 (1.83 – 1.06); p = 0.313	1.08 (0.92 – 1.27); p = 0.362
DHCR7 (G alleles)	0.94 (0.84 - 1.05); p = 0.26	0.96 (0.84 – 1.11); p = 0.604	0.89 (0.73 – 1.07); p = 0.200
25-OH-vitamin D (ng/ml)	0.96 (0.95 – 0.97); p <0.001	0.96 (0.95 – 0.97); p <0.001	0.97 (0.95 – 0.98); p <0.001
Age (years)	1.07 (1.06-1.08); p <0.001	1.07 (1.06 – 1.08); p <0.001	1.08 (1.06 – 1.09); p <0.001
Sex (female/male)	0.60 (0.51 – 0.71); p <0.001	0.63 (0.51 – 0.76); p <0.001	0.56 (0.43 – 0.74); p <0.001
Smoking (yes/no)	1.49 (1.23 – 1.79); p <0.001	1.33 (1.04 – 1.69); p = 0.023	1.81 (1.35 – 2.42); p <0.001
Type 2 diabetes (yes/no)	1.59 (1.39 – 1.82); p <0.001	1.83 (1.54 – 2.18); p <0.001	1.26 (1.00 – 1.58); p = 0.054
CAD	1.49 (1.20 – 1.85); p = 0.001	1.68 (1.26 – 2.24); p < 0.001	1.17 (0.84 – 1.63); p = 0.351

Table 25: 25-OH-vitamin D-associated genotypes and risk for all-cause, cardiovascular and non-cardiovascular mortality.
A Cox regression model including GC, CYP2R1 and DHCR7 genotypes, age, smoking, type 2 diabetes, hypertension and coronary artery disease to predict mortality. Data are presented as hazard ratio (HR) and 95% confidence interval (CI), p-value.

8.6 Discussion

In the present investigation, we aimed to test the hypothesis that genetically lowered concentrations of 25-OH-vitamin D causally increase mortality rates. Using the study design of a Mendelian randomization, we analyzed SNPs associated with serum 25-OH-vitamin D levels and found that genetically determined differences in 25-OH-vitamin D levels were not associated with mortality rates. This lack of association makes a causal role of 25-OH-vitamin D in mortality unlikely.

The biological mechanism of the preventive power of 25-OH-vitamin D for mortality as well as further pathological conditions such as cardiovascular disease and cancer remains unclear. A possible explanation might be reverse causality, such that specific pathomechanisms involved in cardiovascular disease or cancer could result in a decrease of 25-OH-vitamin D levels. The presented results are in line with other examples. Similar findings have been reported for C-reactive protein (CRP), where CRP plasma concentrations were predictive of specific diseases, Mendelian randomization studies, however, found no evidence for a causal role of this biomarker [185, 186].

It should be taken into account that these data are limited by the fact that a cohort of patients referred to coronary angiography has been investigated so that our results may not be generalizable to patients with other diseases. In all survival analysis the most important factors predicting mortality (age, sex, smoking, type 2 diabetes and CAD) have been considered as possible confounders, nevertheless residual confounding by other factors cannot be ruled out completely. Further, the outcome of this study is binary. The assumption of conventional linearity of the applied mathematical methods is not fully achieved, although using these methods for binary outcome data has been advocated [187, 188].

Strengths of the present study are the in-depth clinical and biochemical characterization of all patients, the high number of participants as well as the long time of follow-up, such that the primary end-point death was reached in almost one third of the study participants.

It should be kept in mind that the present study was observational and not experimental with the aim to investigate the potential benefits of vitamin D supplementation. Interestingly, in a recent large meta-analysis of randomized

clinical trials, only vitamin D3, but not 25-OH-vitamin D, decreased mortality significantly [189]. Nevertheless, we would like to emphasize that the present results do not invalidate the predictive power of 25-OH-vitamin D concentrations for mortality. The biological pathways for the association between 25-OH-vitamin D concentrations and mortality remain still unclear. A causal role for 25-OH-vitamin D in mortality is unlikely.

9.SUMMARY AND FUTURE RESEARCH

The present thesis investigated the potential role of three vitamin D-associated polymorphisms on classical and non-classical actions of vitamin D.

There was no association with any of the SNPs with BMD, but the DHCR7 polymorphism may be a predictor for prospective fractures. Due to failed association with prevalent fractures, this interesting finding needs repliction in larger cohort studies. Further investigations of combined biomarkers as risk profile such as patients with vitamin D deficiency and PTH excess as “dysregulation in the mineral metabolism axis” should be considered.

Data collected from both in vitro and in vivo studies are highly optimistic regarding its potential in prevention and regression of prostate cancer. However, in this investigation carriers of genotypes resulting in lower vitamin D levels of the major metabolite 25-OH-vitamin D and its active form 1,25 $(OH)_2$ vitamin D clear do not have a higher risk concerning biochemically detected recurrence, development of metastases, and survival in prostate cancer. To investigate causality of vitamin D in prostate cancer large Mendelian randomization studies should be performed. Future studies should elucidate further the potential benefits of vitamin D in appropriate doses for prevention of prostate cancer, as an anticancer target and/ or as adjuvants in chemotherapy for the treatment of cancers.

The results of this thesis strongly support that a causal role of vitamin D for mortality is unlikely. The present Mendelian randomization study investigated lifelong decreased vitamin D levels in relation to elevated mortality rates. As already in process, large RCTs should be the focus of future interventions to elucidate if medicinal elevated vitamin D levels act differently than lifelong genetically determined ones.

We know much about the transport, metabolism and gene regulating effectiveness as a hormone of vitamin D and its contribution on numerous pathways. The predictive power of 25-OH-vitamin D levels for many chronic diseases as well as for mortality is beyond question, but the accurate biological role of how vitamin D metabolites effects still remains unclear. Future studies should therefore focus on these pathways extensively to increase knowledge about this concern.

10. REFERENCES

1 History of Vitamin D University of California, Riverside, Vitamin D Workshop.

2 http://en.wikipedia.org/wiki/Vitamin_D, download: 22.08.2012.

3 Gozdzik A, Zhu J, Wong BY, Fu L, Cole DE, Parra EJ. Association of vitamin D binding protein (VDBP) polymorphisms and serum 25(OH)D concentrations in a sample of young Canadian adults of different ancestry. J Steroid Biochem Mol Biol. 2011;127:405-12.

4 Holick MF. Vitamin D deficiency. N Engl J Med. 2007;357:266-81.

5 Dusso AS, Brown AJ, Slatopolsky E. Vitamin D. Am J Physiol Renal Physiol. 2005;289:F8-F28.

6 Cooke NE, Haddad JG. Vitamin D binding protein (Gc-globulin). Endocr Rev. 1989;10: 294-307.

7 Brown AJ, Finch J, Grieff M, Ritter C, Kubodera N, Nishii Y, et al. The mechanism for the disparate actions of calcitriol and 22-oxacalcitriol in the intestine. Endocrinology 1993;133:1158-64.

8 Bouillon R, Van Assche FA, Van Baelen H, Heyns W, De Moor P. Influence of the vitamin D-binding protein on the serum concentration 1,25-dihydroxyvitamin D3. Significance of the free 1,25-dihydroxyvitamin D3 concentration. J Clin Invest. 1981;67: 589-96.

9 Bouillon R, Carmeliet G, Verlinden L, van Etten E, Verstuyf A, Luderer HF, et al. Vitamin D and human health: lessons from vitamin D receptor null mice. Endocrine Reviews. 2008;29:726-76.

10 Dusso AS, Thadhani R, Slatopolsky E. Vitamin D receptor and analogs. Semin Nephrol 2004;24:10-6.

11 Kitazawa S, Kajimoto K, Kondo T, Kitazawa R. Vitamin D3 supports osteoclastogenesis via functional vitamin D response element of human RANKL gene promoter. J Cell Biochem. 2003;89:771-7.

12 Kondo T, Kitazawa R, Maeda S, Kitazawa S. 1,25 Dihydroxyvitamin D3 rapidly regulates the mouse osteoprotegrin gene through dual pathways. J Bone Miner Res. 2004;19:1411-9.

13 http://lpi.oregonstate.edu/infocenter/vitamins/vitaminD/dendocrine.html, download: 17.12.2012.

14 Orwoll E, Riddle M, Prince M. Effects of vitamin D on insulin and glucagon secretion in non-insulin-dependent diabetes mellitus. Am J Clin Nutr. 1994;5:1083-7.

15 Borissova AM, Tankova T, Kirilov G, Dakovska L, Kovacheva R. The effect of vitamin D3 on insulin secretion and peripheral insulin sensitivity in type 2 diabetic patients. Int J Clin Pract. 2003;57:258-61.

16 Chakraborti CK. Vitamin D as a promising anticancer agent. Indian J Pharmacol. 2011;43:113-20.

17 Nagpal S, Na S, Rathnachalam R. Noncalcemic actions of vitamin D receptor ligands. Endocr Rev. 2005;26:662-87.

18 Holick MF. Vitamin D: importance in the prevention of cancers, type 1 diabetes, heart disease, and osteoporosis. Am J Clin Nutr. 2004;79:362-71.

19 Weng YW, Selvaduray KR, Nesaretnam K. The plumbing system in our body-angiogenesis and cancer. Palm Oil Developments 47:10-5.

20 Lin R, White JH. The pleiotropic actions of vitamin D. Bioessays. 2004;26:21-8.

21 Hayes CE, Nashold FE, Spach KM, Pedersen LB. The immunological functions of the vitamin D endocrine system. Cell Mol Biol. 2003;49:277-300.

22 Griffin MD, Xing N, Kumar R. Vitamin D and its analogs as regulators of immune activation and antigen presentation. Annu Rev Nutr. 2003;23:117-45.

23 Prietl B, Pilz S, Wolf M, Tomaschitz A, Obermayer-Pietsch B, Graninger et al. Vitamin D supplementation and regulatory T cells in

apparently healthy subjects: vitamin D treatment for autoimmune diseases? Isr Med Assoc J. 2010;12:136-9.

24 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. J Clin Endocrinol Metab. 2011;96: 1911-30.

25 Gordon CM, DePeter KC, Feldman HA, Grace E, Emans SJ. Prevalence of vitamin D deficiency among healthy adolescents. Arch Pediatr Adolesc Med. 2004;158:531-7.

26 Marwaha RK, Tandon N, Reddy DR, Aggarwal R, Singh R, Sawhney RC, et al. Vitamin D and bone mineral density status of healthy schoolchildren in northern India. Am J Clin Nutr. 2005;82:477-82.

27 Thacher TD, Fischer PR, Strand MA, Pettifor JM. Nutritional rickets around the world: Causes and future directions. Ann Trop Paediatr. 2006;26:1-6.

28 Cutillas-Marco E, Fuertes-Prosper A, Grant WB, Morales-Suárez-Varela M. Vitamin D deficiency in South Europe: effect of smoking and aging. Photodermatol Photoimmunol Photomed. 2012;28:159-61.

29 Kudlacek S, Schneider B, Peterlik M, Leb G, Klaushofer K, Weber K, Woloszczuk W, Willvonseder R. Assessment of vitamin D and calcium status in healthy adult Austrians. Eur J Clin Invest. 2003;33:323-31.

30 Holick MF, Chen TC. Vitamin D deficiency: a worldwide problem with health consequences. Am J Clin Nutr. 2008;87:1080-6.

31 Holick MF, Chen TC, Sauter ER. Vitamin D and skin physiology: a D - lightful story. J Bone Miner Res. 2007;22:28-33.

32 Moan J, Porojnicu AC, Dahlback A, Setlow RB. Addressing the health benefits and risks, involving vitamin D or skin cancer, of increased sun exposure. Proc Natl Acad Sci. USA. 2008;105:668-73.

33 Matsuoka LY, Ide L, Wortsman J, MacLaughlin JA, Holick MF. Sunscreens suppress cutaneous vitamin D3 synthesis. J Clin Endocrinol Metab. 1987;64:1165-8.

34 Clemens TL, Henderson SL, Adams JS, Holick MF. Increased skin pigment reduces the capacity of skin to synthesise vitamin D3. Lancet. 1982;1:74-6.

35 Wortsman J, Matsuoka LY, Chen TC, Lu Z, Holick MF. Decreased bioavailability of vitamin D in obesity. Am J Clin Nutr. 2008;72:690-3.

36 Heaney RP. Functional indices of vitamin D status and ramifications of vitamin D deficiency. Am J Clin Nutr. 2004;80:1706-9.

37 Holick MF, Siris ES, Binkley N, Beard MK, Khan A, Katzer JT, et al. Prevalence of vitamin D inadequacy among postmenopausal North American women receiving osteoporosis therapy. J Clin Endocrinol Metab. 2005;90:3215-24.

38 Malabanan AO, Turner AK, Holick MF. Severe generalized bone pain and osteoporosis in a premenopausal black female: effect of vitamin D replacement. J Clin Densitometr. 1989;1:201-4.

39 Plotnikoff GA, Quigley JM. Prevalence of severe hypovitaminosis D in patients with persistent, nonspecific musculoskeletal pain. Mayo Clin Proc. 2003;78:1463-70.

40 Gordon CM, Williams AL, Feldman HA, May J, Sinclair L, Vasquez A, et al. Treatment of hypovitaminosis D in infants and toddlers. J Clin Endocrinol Metab. 2008;93:2716-21.

41 Holick MF. Resurrection of vitamin D deficiency and rickets. J Clin Invest. 2006;116:2062-72.

42 Bischoff-Ferrari HA, Willett WC, Wong JB, Giovannucci E, Dietrich T, Dawson-Hughes B. Fracture prevention with vitamin D supplementation: a meta-analysis of randomized controlled trials. JAMA. 2005;293:2257-64.

43 Bischoff-Ferrari HA, Dawson-Hughes B, Staehelin HB, Orav JE, Stuck AE, Theiler R, et al. Fall prevention with supplemental and active forms of vitamin D: a meta-analysis of randomised controlled trials. BMJ. 2009;339:b3692.

44 Watson JD, Crick FHC. A Structure for Deoxyribose Nucleic Acid. Nature 1953;171:737-8.

45 http://www.nobelprize.org/educational/medicine/dna_double_helix/readmore.html?print=1, download: 5.10.2012.

46 http://www.ambbea.org/the-rudimentary-development-of-x-ray-technology/, download: 3.10.2012.

47 http://24952031.nhd.weebly.com/scientific-innovations.html, download: 5.10.2012.

48 http://ezinearticles.com/?How-Mere-Human-Minds-Unraveled-the-Ultimate-Complexity-of-Life,-the-DNA-and-the-Double-Helix&id=6002327, download: 5.10.2012.

49 http://www.accessexcellence.org/RC/AB/BC/Rosalind_Franklin.php, download:5.10.2012.

50 http://en.wikipedia.org/wiki/DNA, download: 5.10.2012.

51 http://www.chemheritage.org/discover/online-resources/chemistry-in-history/themes/biomolecules/dna/watson-crick-wilkins-franklin.aspx, download: 5.12.2012.

52 www.JURAJARTNER.COM, download: 5.10.2012.

53 Strachan T, Read A. Human Molecular Genetics. 4th ed. Garland Science. pp 467-95.

54 Stenson PD, Mort M, Ball EV, Howells K, Phillips AD, Thomas NS, Cooper DN. The Human Gene Mutation Database: 2008 update. Genome Med. 2009;1(1):13.

55 http://education.technyou.edu.au/view/91/155/what-does-dna-look, download: 3.10.2012.

56 Bertina RM, Koeleman BP, Koster T,, et al. (May 1994). "Mutation in blood coagulation factor V associated with resistance to activated protein C". Nature 369 (6475): 64-7.

57 Ridker PM, Miletich JP, Hennekens CH, Buring JE. Ethnic distribution of factor V Leiden in 4047 men and women. Implications for venous thromboembolism screening. JAMA. 1997;277:1305-7.

58 Renner W, Köppel H, Hoffmann C, Schallmoser K, Stanger O, Toplak H, et al. Prothrombin G20210A, factor V Leiden, and factor XIII Val34Leu: common mutations of blood coagulation factors and deep vein thrombosis in Austria. Thromb Res. 2000;99:35-9.

59 http://www.labor-renner.at/en/tests/factorV.html, download: 2.10.2012.

60 Obermayer-Pietsch BM, Bonelli CM, Walter DE, Kuhn RJ, Fahrleitner-Pammer A, Berghold A, et al. Genetic predisposition for adult lactose intolerance and relation to diet, bone density, and bone fractures. J Bone Miner Res. 2004;19:42-7.

61 http://cmbi.bjmu.edu.cn/cmbidata/snp/index00.htm, download: 3.10.2012.

62 Teare MD, Barrett JH. Genetic linkage studies. Lancet. 2005;366:1036-44.

63 Greely HT. The uneasy ethical and legal underpinnings of large-scale genomic biobanks. Annu Rev Genomics Hum Genet. 2007;8:343-64.

64 The International HapMap Project, Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu F, Yang H, Ch'Ang L-Y, Huang W. The International HapMap Project. Nature. 2003;426:789-96.

65 The International HapMap Consortium. The International HapMap Project. Nature. 2003;426: 789-96.

66 The International HapMap Consortium. A haplotype map of the human genome. Nature. 2005;437: 1299-1320.

67 Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995;270: 467-70.

68 Müller HJ, Röder T. Der Experimentator: Microarrays. 1st ed. Elsevier Apektrum Akademischer Verlag; 2004.

69 Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nat Protoc. 2011;6:121-33.

70 Manolio TA; Guttmacher, Alan E.; Manolio, Teri A. Genomewide association studies and assessment of the risk of disease. N Engl J Med. 2010;363:166-76.

71 Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008;299:1335-44.

72 Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nat Protoc. 2011;6:121-33.

73 Mannuccio Mannucci P, Lotta LA, Peyvandi F. Genome-Wide Association Studies in Myocardial Infarction and Coronary Artery Disease. J Tehran Heart Cent. 2010;5:116-21.

74 http://www.genome.gov/20019523, download: 31.10.2012.

75 Upton G, Cook I. Oxford Dictionary of Statistics. 2nd ed; 2006

76 Hennekens CH, Buring JE, Epidemiology in Medicine, 1st ed. Philadelphia (USA): Lippincott Williams & Wilkins; 1987.

77 Hsu JL, Banerjee D, Kuschner WG. Understanding and Identifying Bias and Confounding in the Medical Literature. South Med J. 2008;1240-5.

78 Höfer T, Przyrembel H, Verleger S. New evidence for the theory of the stork. Paediatric and Perinatal Epidemiology. 18, 2004,88-92.

79 Stigler S. Fisher and the 5% level. Chance. 2008;21(4):12.

80 Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. Medical epidemiology. 2nd ed. Stamford (CT): Appleton and Lange; 1996. p. 75-85.

81 Attia J, Ioannidis JP, Thakkinstian A, McEvoy M, Scott RJ, Minelli C et al. How to use an article about genetic association: B: Are the results of the study valid? JAMA. 2009;301:191-7.

82 Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002;359:248-52.

83 Greenland S, Morgenstern H. Confounding in health care research. Annu Rev Public Health. 2001;22:189-212.

84 McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP et al. Genomewide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet. 2008;9:356-69.

85 Ioannidis JP, Trikalinos TA, Khoury MJ. Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. Am J Epidemiol 2006;164:609-14.

86 Siontis GC, Ioannidis JP. Risk factors and interventions with statistically significant tiny effects. Int J Epidemiol. 2011;40:1292-307.

87 Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. Clin Trials. 2005;2:282-90.

88 Last JM, editor. Dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001; p 61.

89 http://www.uic.edu/sph/prepare/courses/ph490/resources/epilesson01.pdf, download: 23.10.2012.

90 http://user.meduni-graz.at/andrea.berghold/EBM/EBM_einfuehrung.pdf, download: 24.10.2012.

91 Forrest JL. Introduction to the Basics of Evidence-Based Dentistry: Concepts and Skills. J Evid Based Dent Pract. 2009;9:108-12.

92 Jepsen P, Johnsen SP, Gillman MW, Sørensen HT. Interpretation of observational studies. Heart. 2004;90:956-60.

93 Reitz C, Brayne C, Mayeux R. Epidemiology of Alzheimer Disease. Nat Rev Neurol. 2011;7:137-52.

94 Vandenbroucke JP. When are observational studies as credible as randomised trials? Lancet. 2004;363:1728-31.

95 Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32:1-22.

96 Stolberg HO, Norman G, Trop I. Randomized controlled trials. AJR Am J Roentgenol. 2004;183:1539-44.

97 Ebrahim, S.; Davey, S.G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? Hum. Genet. 2008, 123, 15-33.

98 Bochud M, Rousson V. Usefulness of Mendelian Randomization in Observational Epidemiology. Int J Environ Res Public Health. 2010;7:711-28.

99 Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J. 2003;20:54-60.

100 Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. Lancet. 1986;1:507-8.

101 http://www.mendelweb.org/archive/Mendel.Experiments.txt, downöoad: 17.10.2012.

102 Nitsch D, Molokhia M, Smeeth L, DeStavola BL, Whittaker JC, Leon DA. Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. Am J Epidemiol. 2006;163:397-403.

103 Lewis SJ, Smith GD. Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. Cancer Epidemiol Biomarkers Prev. 2005;14:1967-71.

104 Boccia S, Hashibe M, Galli P, De FE, Asakage T, Hashimoto T, et al. Aldehyde dehydrogenase 2 and head and neck cancer: a meta-analysis implementing a Mendelian randomization approach. Cancer Epidemiol. Biomarkers Prev. 2009, 18, 248-254.

105 Thomas GW. The Physics of Everyday Phenomena: A Conceptual Introduction to Physics. New York: McGraw-Hill Higher Education. 2001; p3.

106 Pocock SJ. Clinical Trials: A practical approach, Chichester, Wiley, 1984.

107 Encyclopedia of Statistics in Behavioral Science. John Wiley & Sons, Ltd.

108 http://www.nlm.nih.gov/services/ctphases.html, download: 11.10.2012.

109 Hollis S, Fiona Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999;319:670-4.

110 Kendall JM. Designing a research project: Randomised Controlled trials and their principles. Emerg Med J. 2003;20:164-8.

111 Jadad AR. Randomised controlled trials: a user's guide. London, England: BMJ Books, 1998.

112 Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.

113 Hager W. Grundlagen einer Versuchsplanung zur Prüfung empirischer Hypo-thesen der Psychologie. Allgemeine Experimentelle Psychologie. Stuttgart: Gustav Fischer Verlag. 1987; 43-253.

114 Orton SM, Morris AP, Herrera BM, Ramagopalan SV, Lincoln MR, Chao MJ, et al. Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis. Am J Clin Nutr. 2008;88:441-7.

115 Hunter D, De Lange M, Snieder H, MacGregor AJ, Swaminathan R, Thakker RV, et al. Genetic contribution to bone metabolism, calcium excretion, and vitamin D and parathyroid hormone regulation. J Bone Miner Res. 2001;16:371-8.

116 Shea MK, Benjamin EJ, Dupuis J, Massaro JM, Jacques PF, D'Agostino RB Sr,, et al. Genetic and non-genetic correlates of vitamins K and D. Eur J Clin Nutr. 2009;63:458-64.

117 Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet. 2010;376:180-8.

118 Ahn J, Yu K, Stolzenberg-Solomon R, Simon KC, McCullough ML, Gallicchio L, et al. Genome-wide association study of circulating vitamin D levels. Hum Mol Genet. 2010;19:2739-45.

119 http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=2638, download: 8.11.2012.

120 Speeckaert M, Huang G, Delanghe JR, Taes YE. Biological and clinical aspects of the vitamin D binding protein (Gc-globulin) and its polymorphism. Clin Chim Acta. 2006;372:33-42.

121 Verboven C, Rabijns A, De Maeyer M, Van Baelen H, Bouillon R, De Ranter C. A structural basis for the unique binding features of the human vitamin D-binding protein. Nat Struct Biol. 2002;9:131-6.

122 http://themedicalbiochemistrypage.org/slos.php, download: 8.11.2012.

123 http://www.uscnk.com/directory/Dehydrocholesterol-reductase-24(DHCR24)-2270.htm, download: 8.11.2012.

124 http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=120227, download: 8.11.2012.

125 http://en.wikipedia.org/wiki/CYP2R1, download: 8.11.2012.

126 Waterham HR, Wanders RJ. Biochemical and genetic aspects of 7-dehydrocholesterol reductase and Smith-Lemli-Opitz syndrome. Biochim Biophys Acta. 2000;1529:340-56.

127 http://omim.org/entry/608713, download: 8.11.2012.

128 http://ghr.nlm.nih.gov/gene/CYP2R1, download: 8.11.2012.

129 Cheng JB, Motola DL, Mangelsdorf DJ, Russell DW. De-orphanization of cytochrome P450 2R1: a microsomal vitamin D 25-hydroxylase. J Biol Chem. 2003;278:38084-93.

130 http://biocyc.org/HUMAN/NEW-IMAGE?type=GENE-IN-PWY&object=HS17721, download: 8.11.2012.

131 Ahonen MH, Tenkanen L, Teppo L, Hakama M, Tuohimaa P. Prostate cancer risk and pre-diagnostic serum 25-hydroxyvitamin D levels (Finland). Cancer Causes Control. 2000;11:847-52.

132 http://www.mn-net.com/Portals/8/attachments/Redakteure_Bio/Protocols/Genomic%20DNA/UM_gDNADxBlood.pdf. Stand: 14.11.2012.

133 http://tools.invitrogen.com/content/sfs/manuals/cms_039284.pdf, download: 14.11.2012.

134 Yuan C, Peterson RJ, Wang C, Goodsaid F, Waters DJ. 59 Nuclease Assays for the Loci CCR5-1/D32. CCR2-V64I, and SDF1-G801A Related to Pathogenesis of AIDS. Clin Chem. 2000;46:124-30.

135 https://products.invitrogen.com/ivgn/product/4369016?ICID=search-4369016#, download: 28.11.2012.

136 Moonesinghe R, Yesupriya A, Chang M, Dowling NF, Khoury MJ, Scott AJ for the CDC/NCI NHANES III Genomics Working Group. A Hardy-Weinberg equilibrium test for analyzing population genetic surveys with complex sample designs. Am J Epidemiol. 2010;171:932-41.

137 Rogatko A, Slifker MJ, Babb JS. Hardy–Weinberg Equilibrium Diagnostics. Theor Popul Biol. 2002;62:251-7.

138 Bischoff-Ferrari HA, Shao A, Dawson-Hughes B, Hathcock J, Giovannucci E, Willett WC. Benefit-risk assessment of vitamin D supplementation. Osteoporos Int. 2010;1:1121-32.

139 van Meurs JB, Trikalinos TA, Ralston SH, Balcells S, Brandi ML, Brixen K, et al. Large-scale analysis of association between LRP5 and LRP6 variants and osteoporosis. JAMA. 2008;299:1277-90.

140 Uitterlinden AG, Ralston SH, Brandi ML, Carey AH, Grinberg D, Langdahl BL, et al. The association between common vitamin D receptor gene variations and osteoporosis: a participant-level meta-analysis. Ann Intern Med. 2006;145:255-64.

141 Rivadeneira F, Styrkarsdottir U, Estrada K, Halldorsson BV, Hsu YH, Richards JB, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. Nat Genet. 2009;41:1199-206.

142 Ralston SH. Genetics of osteoporosis. Ann N Y Acad Sci. 2010;1192:181-9.

143 Obermayer-Pietsch BM, Walter D, Kotschan S, Freigassner-Pritz M, Windhager R, Leb G. Congenital hip dysplasia and bone mineral density of the hip--a new risk factor for osteoporotic fracture? J Bone Miner Res. 2000;15:1678-82.

144 Gugatschka M, Dobnig H, Fahrleitner-Pammer A, Pietschmann P, Kudlacek S, Strele A, et al. Molecularly-defined lactose malabsorption, milk consumption and anthropometric differences in adult males. QJM. 2005;98:857-63.

145 Dobnig H, Piswanger-Solkner JC, Roth M, Obermayer-Pietsch B, Tiran A, Strele A, et al. Type 2 diabetes mellitus in nursing home subjects: effects on bone turnover, bone mass, and fracture risk. J Clin Endocrinol Metab. 2006;91:3355-63.

146 Hollis BW. Comparison of commercially available (125) I-based RIA methods for the determination of circulating 25-hydroxyvitamin D. Clin Chem. 2000;46:1657-61.

147 Lauridsen AL, Vestergaard P, Hermann AP, Brot C, Heickendorff L, Mosekilde L, et al. Plasma concentrations of 25-hydroxy-vitamin D and 1,25-dihydroxy-vitamin D are related to the phenotype of Gc (vitamin D-binding protein): a cross-sectional study on 595 early postmenopausal women. Calcif Tissue Int. 2005;77:15-22.

148 Sinotte M, Diorio C, Berube S, Pollak M, Brisson J. Genetic polymorphisms of the vitamin D binding protein and plasma concentrations of 25-hydroxyvitamin D in premenopausal women. Am J Clin Nutr. 2009;89:634-40.

149 Engelman CD, Fingerlin TE, Langefeld CD, Hicks PJ, Rich SS, Wagenknecht LE, Bowden DW, Norris JM. Genetic and environmental determinants of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D levels in Hispanic and African Americans. J Clin Endocrinol Metab. 2008;93:3381-88.

150 Fang Y, van Meurs JB, Arp P, van Leeuwen JP, Hofman A, Pols HA, et al. Vitamin D binding protein genotype and osteoporosis. Calcif Tissue Int. 2009;85:85-93.

151 Majumdar SR, Kim N, Colman I, Chahal AM, Raymond G, Jen H, Siminoski KG, Hanley DA, Rowe BH. Incidental vertebral fractures discovered with chest radiography in the emergency department: prevalence, recognition, and osteoporosis management in a cohort of elderly patients. Arch Intern Med. 2005;165:905-9.

152 Bischoff-Ferrari HA, Dawson-Hughes B, Staehelin HB, Orav JE, Stuck AE, Theiler R, et al. Fall prevention with supplemental and active forms of vitamin D: a meta-analysis of randomised controlled trials. BMJ. 2009;339:b3692.

153 Pfeifer M, Begerow B, Minne HW. Vitamin D and muscle function. Osteoporos Int. 2002;13:187-94.

154 Pilz S, Tomaschitz A, Drechsler C, Dekker JM, Marz W. Vitamin D deficiency and myocardial diseases. Mol Nutr Food Res. 2010;54:1103-13.

155 Tomaschitz A, Pilz S, Ritz E, Grammer T, Drechsler C, Boehm BO, Marz W. Independent association between 1,25-dihydroxyvitamin D, 25-hydroxyvitamin D and the renin-angiotensin system: The Ludwigshafen Risk and Cardiovascular Health (LURIC) study. Clin Chim Acta. 2010;411:1354-60.

156 Pilz S, Tomaschitz A, Ritz E, Pieber TR. Vitamin D status and arterial hypertension: a systematic review. Nat Rev Cardiol. 2009;6:621-30.

157 Graham IM, Daly LE, Refsum HM, Robinson K, Brattstrom LE, Ueland PM, et al. Plasma homocysteine as a risk factor for vascular disease. The European Concerted Action Project. JAMA. 1997;277:1775-81.

158 Clarke R, Daly L, Robinson K, Naughten E, Cahalane S, Fowler B, et al. Hyperhomocysteinemia: an independent risk factor for vascular disease. N Engl J Med. 1991;324:1149-55.

159 Langsenlehner U, Krippl P, Renner W, Yazdani-Biuki B, Wolf G, Wascher TC, et al. The common 677C>T gene polymorphism of

methylenetetrahydrofolate reductase gene is not associated with breast cancer risk. Breast Cancer Res Treat. 2003;81:169-72.

160 Kadziela J, Janas J, Dzielinska Z, Szperl M, Gazdzik D, Chotkowska E, et al. The C677T mutation in methylenetetrahydrofolate reductase gene, plasma homocysteine concentration and the risk of coronary artery disease. Kardiol Pol. 2003;59:17-26.

161 Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. International Variation in Prostate Cancer Incidence and Mortality Rates. Eur Urol. 2012;61:1079-92.

162 Swami S, Krishnan AV, Feldman D. Vitamin D metabolism and action in the prostate: implications for health and disease. Mol Cell Endocrinol. 2011;347:61-9.

163 Jones G, Strugnell SA, DeLuca HF. Current understanding of the molecular actions of vitamin D. Physiol Rev. 1998;78:1193-231.

164 Chen TC, Holick MF. Vitamin D and prostate cancer prevention and treatment. Trends Endocrinol Metabol. 2003;14:423-30.

165 Giovannucci E. Dietary influences of 1,25(OH)2 vitamin D in relation to prostate cancer: a hypothesis. Cancer Causes Control. 1998;9:567-82.

166 Peehl DM, Feldman D. The role of vitamin D and retinoids in controlling prostate cancer progression. Endocr Relat Cancer. 2003;10:131-40.

167 Young MV, Schwartz GG, Wang L Jamieson DP, Whitlatch LW, Flanagan JN, Lokeshwar BL, Holick MF, Chen TC. The prostate 25-hydroxyvitamin D-1 alpha-hydroxylase is not influenced by parathyroid hormone and calcium: implications for prostate cancer chemoprevention by vitamin D. Carcinogenesis. 2004;25:967-71.

168 Dusso AS, Brown AJ, Slatopolsky E. Vitamin D. Am J Physiol Renal Physiol. 2005;289:F8-F28.

169 Chen TC, Wang L, Whitlatch LW, Flanagan JN, Holick MF. Prostatic 25-hydroxyvitamin D-1alpha-hydroxylase and its implication in prostate cancer. J Cell Biochem. 2003;88:315-22.

170 Winkelmann BR, März W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J; LURIC Study Group (LUdwigshafen RIsk and Cardiovascular Health). Rationale and design of the LURIC study-a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. Pharmacogenomics. 2001;1Suppl1:S1-73.

171 Langsenlehner T, Langsenlehner U, Renner W, Kapp KS, Krippl P, Hofmann G, Clar H, Pummer K, Mayer R. The Glu228Ala polymorphism in the ligand binding domain of death receptor 4 is associated with increased risk for prostate cancer metastases. Prostate. 2008;15;68:264-8.

172 Roach M 3rd, Hanks G, Thames H Jr, Schellhammer P, Shipley WU, Sokol GH, et al. Defining biochemical failure following radiothera¬py with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO phoenix consensus conference. Int J Radiat Oncol Biol Phys 2006;65:965-74.

173 Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Controlled clinical trials. 1996;17:343-6.

174 Trummer O, Schwetz V, Walter-Finell D, Lerchbaum E, Renner W, Gugatschka M, et al. Allelic determinants of vitamin d insufficiency, bone mineral density, and bone fractures.J Clin Endocrinol Metab. 2012 97:1234-40.

175 Gilbert R, Martin RM, Beynon R, Harris R, Savovic J, Zuccolo L, et al. Associations of circulating and dietary vitamin D with prostate cancer risk: a systematic review and dose-response meta-analysis. Cancer Causes Control. 2011;22:319-40.

176 Pilz S, Tomaschitz A, März W, Drechsler C, Ritz E, Zittermann A, et al. Vitamin D, cardiovascular disease and mortality. Clin Endocrinol 2011;75:575-84.

177 Pilz S, Tomaschitz A. Role of vitamin D in arterial hypertension. Expert Rev Cardiovasc Ther. 2010;8:1599-608.

178 Wehr E, Trummer O, Giuliani A, Gruber HJ, Pieber TR, Obermayer-Pietsch B. Vitamin D-associated polymorphisms are related to insulin resistance and vitamin D deficiency in polycystic ovary syndrome. Eur J Endocrinol. 2011;164:741-49.

179 Chiu KC, Chu A, Go VL, Saad MF. Hypovitaminosis D is associated with insulin resistance and beta cell dysfunction. Am J Clin Nutr. 2004;79:820-5.

180 Dobnig H, Pilz S, Scharnagl H, Renner W, Seelhorst U, Wellnitz B, Kinkeldei J, Boehm BO, Weihrauch G, Maerz W. Independent association of low serum 25-hydroxyvitamin d and 1,25-dihydroxyvitamin d levels with all-cause and cardiovascular mortality. Arch Intern Med 2008;168:1340-9.

181 Melamed ML, Michos ED, Post W, Astor B. 25-hydroxyvitamin D levels and the risk of mortality in the general population. Arch Intern Med. 2008;168:1629-37.

182 Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2000;29:722-9.

183 Austen WG, Edwards JE, Frye RL, Gensini GG, Gott VL, Griffith LS, et al. A reporting system on patients evaluated for coronary artery disease. Report of the Ad Hoc Committee for Grading of Coronary Artery Disease, Council on Cardiovascular Surgery, American Heart Association. Circulation. 1975;51:5-40.

184 Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the Expert Committee on the diagnosis and classification of diabetes mellitus. Diabetes Care. 1997;20:1183-97.

185 Marott SC, Nordestgaard BG, Zacho J, Friberg J, Jensen GB, Tybjaerg-Hansen A, et al. Does elevated C-reactive protein increase atrial fibrillation risk? A Mendelian randomization of 47,000 individuals from the general population. J Am Coll Cardiol. 2010;56:789-95.

186 Allin KH, Nordestgaard BG, Zacho J, Tybjaerg-Hansen A, Bojesen SE. C-reactive protein and the risk of cancer: a mendelian randomization study. J Natl Cancer Inst. 2010;102:202-6.

187 Thomas DC, Conti DV. Commentary: The concept of ”Mendelian randomization”. Int J Epidemiol. 2004;33:21-5.

188 Davey Smith G, Lawlor DA, Harbord R, Timpson N, Rumley A, Lowe GD, et al. Association of C-reactive protein with blood pressure and hypertension. life course confounding and Mendelian randomisation tests of causality. Arterioscler Thromb Vasc Biol. 2005;25:1051-6.

189 Bjelakovic G, Gluud LL, Nikolova D, Whitfield K, Wetterslev J, Simonetti RG, Bjelakovic M, Gluud C. Vitamin D supplementation for prevention of mortality in adults. Cochrane Database Syst Rev 2011;7:CD007470.