New Insights into Genetic Variation and Cancer

Claire Siemes

ACKNOWLEDGEMENTS

The work presented in this thesis was conducted at the Department of Epidemiology & Biostatistics in close collaboration with the Department of Internal Medicine of the Erasmus Medical Center, Rotterdam, The Netherlands.

Most of the studies described in this thesis were performed in The Rotterdam Study that is supported by the Erasmus Medical Center and Erasmus University Rotterdam, the Netherlands Organization for Scientific Research (NWO), the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

The contribution of inhabitants, general practitioners and pharmacists of the Ommoord district to The Rotterdam Study is greatly acknowledged.

Financial support by the Departments of Epidemiology & Biostatistics, Sanquin bloedbank, Jurriaanse stichting, Merck Sharp & Dohme, Eli Lilly, Sanofi-Aventis, Merck, Amgen Oncology, Altana Pharma, and LEO Pharma is gratefully acknowledged.

Design cover : C. Siemes, Capelle aan den IJssel and Optima Grafische Communicatie, Rotterdam
Layout and printed by : Optima Grafische Communicatie, Rotterdam

ISBN : 978-90-8559-283-9

New Insights into Genetic Variation and Cancer

Nieuwe inzichten in genetische variatie en kanker

Proefschrift

ter verkrijging van de graad van doctor

aan de Erasmus Universiteit Rotterdam

op gezag van de rector magnificus

Prof. dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden

op woensdag 9 mei 2007 om 13:45 uur.

door

Claire Siemes

geboren te Leidschendam, Nederland

PROMOTIECOMMISSIE

Promotores:	Prof. dr. B.H.Ch. Stricker
	Prof. dr. H.A.P. Pols
Overige leden:	Prof. dr. J.W.W. Coebergh
	Prof. dr. E.J. Kuipers
	Prof. dr. ir. F.E. van Leeuwen
Co-promotor:	Dr. L.E. Visser

Cover

A crab is the symbol for cancer. The Dutch Cancer Society (KWF – Kankerbestrijding) uses a crab with a sword to symbolize the fight against cancer. Although the meaning of a crab as symbol is not entirely known, some believe it represents the dualty of the disease (dead and fighting) such as the ying-yang symbol, or compare the limbs of the crab with metastases (Hippocrates 460 – 377 A.D.)

Alles wat groot is begon ooit klein,
Je hoeft niet meteen een vlinder te zijn.
(John O. Ewbank)

Voor mijn ouders

CONTENTS

1. General Introduction

1.1. Introduction

BACKGROUND

Cancer has become the leading cause of death in women, surpassing heart diseases.[1] In males it takes a second place. The disease burden is large and grows as a result of aging of the population, especially in western countries. Cancer research in the past has led to a better understanding of treatment and prognosis. Nevertheless, despite substantial efforts in this research field, the pathogenesis of cancer is still largely unclear.

Since long it has been recognized, that cancer is not merely one disease, but has to be considered as a collective term for a diversity of diseases that may differ strongly regarding pathogenesis, clinical appearance, treatment and prognosis. Nevertheless, a joint feature of all tumors is their uncontrolled growth. The withdrawal of a cell from normal growth regulation is a result of cell transformations. These transformations are caused by carcinogens that initiate or promote the formation of cancer (carcinogenesis). Besides the time it takes for one single malignant cell to become a clinically manifest tumor, carcinogenesis is a long-term process for another reason as well. Transformation has a multistep character that results in accumulation of mutations, a process that might take years. Consequently, a person in the middle of such a process might die from other causes before developing a clinical manifest cancer. FIGURE 1 gives an example of the stages in the evolution of colon cancer.

This multistep process can be explained by a theory of Kenneth J. Rothman on disease risks.[2,3] Many different risk factors have been identified for several cancers and in this theory these are

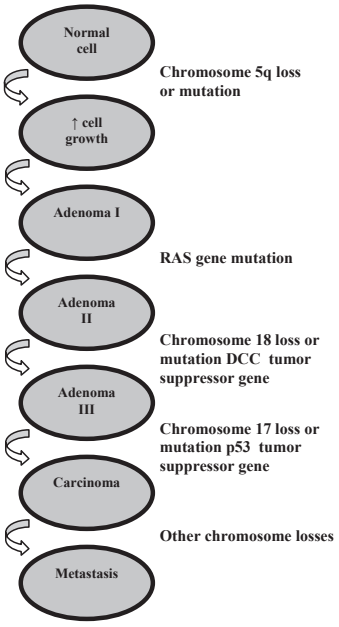

Figure 1. Stages in the evolution of colon cancer

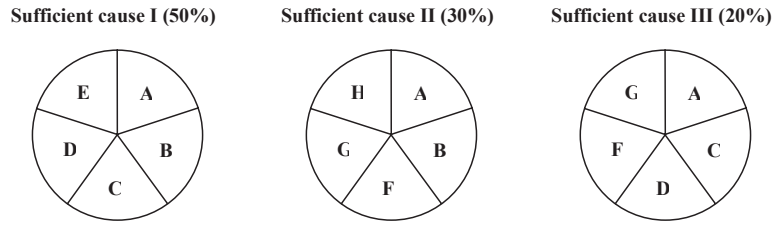

Figure 2. Multistep process of cancer development explained by sufficient causes.

called component causes. A combination of these causes (the so-called sufficient cause) results in cancer development. Nevertheless, it might be that different combinations of component causes lead to the same disease. Therefore, for an individual cancer, different sufficient causes might be present that each lead to a proportion of the total cancer development (FIGURE 2).

As can be seen in FIGURE 2, component cause A is present in all sufficient causes. This component is therefore a necessary cause, since without this factor, the disease of interest will not become manifest. The hope in cancer research is to identify a final common pathway representing a necessary cause for tumors of all types. As a combined feature of all cancers is their uncontrolled growth, necessary or component causes are expected to be involved in the withdrawal of a cell to normal growth regulation.

One of the interests in cancer research nowadays is genetic variation that might be either a component cause or necessary cause. It has been suggested that at least 5 to 10 percent of all risk factors for cancer is genetic.[4] For a long time, the Mendelian method of simple mutations at a single gene (or locus) was the only way of explaining the genetic basis of diseases. This monogenetic approach was associated with cancers with a low incidence. When finding numerous genes as a cause of more common familial diseases, this Mendelian way of thinking was failing.[5] A more subtle approach was necessary to study the etiology of these so-called complex diseases in which a multifactorial scenario, as presented above, was needed to explain cancer development. Since the detection of the human genome sequence, new opportunities for studying the genetic basis of complex diseases have become possible. Specific genes of interest are considered, based upon biological plausibility. These genes are called candidate genes. Variation in such a gene might lead to interindividual susceptibility to cancer (see paragraph 1.4).

AIM AND OUTLINE

The aim of this thesis is to obtain a better understanding of the role which genetic variation in genes encoding for specific enzymes plays in cancer development and prognosis. As for the associatiation with cancer multiple candidate genes can be studied, a selection of two potentially important topics was made;

1. The first part of this thesis will focus on genetic variation in **cytochrome P450** genes, which encode for a group of enzymes with extensive metabolic properties, and their association with cancer.
2. The second part of this thesis will focus on the role genetic variation plays in **inflammatory processes** and their association with cancer.

In the following paragraphs of this introductory chapter, more information will be given on the source population used for analyses (1.2), the demographics of cancer (1.3), and the most important aspects of genetic variation (1.4). Chapter 2 specifically focuses on the association between variation in cytochrome P450 enzyme encoding genes and susceptibility to cancer. First, an overview of the literature on association studies of cytochrome p450 enzyme gene variations and cancer is given in paragraph 2.1. Next, two studies are performed on the association of cytochrome P450 *3A* gene variants, steroid hormone levels and hormone-related cancers. The first describes steroid hormone levels and the association with prostate cancer according to genotype in males (2.2). The second study reports steroid hormone levels and the association with breast cancer according to genotype in females (2.3). Chapter 2 concludes with a study on the combined effect of cytochrome P450 *2C9* gene variant carriage and NSAID use on colorectal cancer risk (2.4). Chapter 3 describes the second topic of this thesis; inflammation and cancer. It starts with a paragraph on cumulative NSAID use and colorectal cancer risk (3.1). This association is further explored in paragraph 3.2, in which a study is presented on cumulative use of different classes of NSAIDs (i.e. non-selective COX, COX-1 selective and COX-2 selective NSAIDs) and their association with cancer alone and stratified for a *COX-2* gene variant. Ultimately, the role of inflammation in the induction and progression of cancer was studied in detail in paragraph 3.3., by investigating the association of baseline C-reactive protein (CRP) levels, variation in the *CRP* gene and cancer. In a general discussion (Chapter 4) the main results of this thesis are briefly reported and placed in perspective, strengths and limitations of genetic association studies are discussed, clinical implications are considered and finally, recommendations for future research are proposed.

1.2. The Rotterdam Study

DESIGN

For all analyses presented in this thesis, data from The Rotterdam Study were used.[6] The Rotterdam Study is a single center population-based prospective cohort study that started in 1990. All inhabitants of the Rotterdam suburb Ommoord, aged 55 years and over, were invited to participate. Of the 10,275 eligible subjects, 7983 (4878 women and 3105 men) agreed to participate (78%), including 897 individuals living in one of the six homes for the elderly. Almost all (99,8%) participants are of Caucasian descent. The main objective of The Rotterdam Study is to investigate the prevalence and incidence of and risk factors for chronic disabling diseases in the elderly (hence its national name ERGO; Erasmus Rotterdam Gezondheid en Ouderen. This name later changed in Erasmus Rotterdam Gezondheids Onderzoek, as it started to include younger individuals as well). At the start of The Rotterdam Study, chronic cardiovascular, neurological, locomotor and ophthalmologic diseases were the main topic of interest, but during the years, these were expanded with other interests like psychiatric diseases and malignancies. The aim of The Rotterdam Study is to contribute to a better understanding, prevention and treatment of chronic diseases in the elderly. The Medical Ethics Committee of the Erasmus Medical Center approved the study and written informed consent was obtained from all participants.

HOME INTERVIEW

All participants were visited at home at the start of the study for a standardised interview on current health status, medical history, smoking habits, socio-economic status, current drug use (ATC-classification), use of medical facilities, dietary habits, alcohol consumption, and physical activity. Dietary data were collected with a 170-item semi-quantitative food frequency questionnaire (SFFQ) adapted for use in the elderly. The two-step dietary assessment comprised a simple self-administered questionnaire (20 min) followed by a structured interview with trained dieticians (20 min) based on the complete questionnaire. This questionnaire has been validated for use in the elderly.[7]

RESEARCH CENTER VISITS

After the home interview was performed, each participant visited the research center two times. During these visits, several measurements were performed: cognitive function, indicators for Parkinson's disease, Dual-Energy X-ray Absorptiometry (bone mineral density), X-rays of hands, thoraco-lumbar spine, hips and knees, an extensive ophthalmologic examination, ultrasound assessment of cardiac dimensions, diameter of the abdominal aorta, carotic arterial wall thickness and plaque thickness, a computerized ECG, blood pressure readings

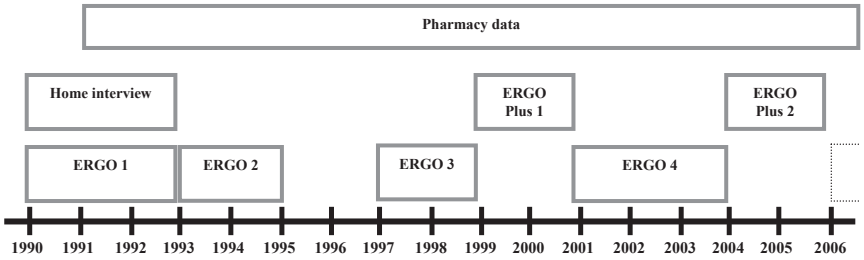

Figure 1. Cross-sectional surveys in The Rotterdam Study

(brachial artery, posterior tibial artery), anthropometrics and a limited physical examination. Additionally, 10 non-fasting blood samples were collected and stored at -20 ºC before and after blood sampling, until a requested measurement. DNA was extracted from these samples for all participants using the standard salting out procedure.[8]

PHARMACY DATA

Besides the self-reported information on medication use at baseline, detailed information on day-to-day medication use is available for all participants since January 1st 1991 (see FIGURE 1). Seven computerized pharmacies cover the research area and are all linked to one network. In this way, the date of prescription, the total amount of drug units per prescription and the prescribed defined daily dosage (DDD, this is the recommended dose for an adult given for the most important indication) [9] are available per drug defined by an Anatomical Therapeutical Chemical (ATC) code.[10]

FOLLOW-UP

Baseline data collection was performed from October 1990 to July 1993. Since then, cross-sectional surveys have been carried out every 2-3 years. (FIGURE 1) In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Events are coded according to the International Classification for Primary Care (ICPC) [11] and the tenth edition of the International Classification of Diseases (ICD-10) [12]. Clinical information is obtained from the general practitioners working in the study district, hospital discharge records and the national morbidity registration (LMR) [13] that registers all hospital admissions. In the case of cancer validation, additional information is obtained from pathology reports (PATHAN) [14]. Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from the general practitioners involved in The Rotterdam Study.

1.3. Demographics of cancer

DISEASE BURDEN

Disability-adjusted-life-years (DALY's) are used to express disease burden in a population.[15] It is calculated on the basis of the number of lost life-years (due to early mortality) and the number of years alive with health problems (due to illnesss) weighted for the severity of the disease (disease-year equivalent). Three important aspects of public health are included in this measure: quantity (duration) of live, quality of live and the total number of persons involved. The National Institute of Public Health and the Environment (RIVM) of The Netherlands states that more than 50% of the total disease burden is caused by persons over 65 years, in which males contribute a greater portion than females.[16] Cardiovascular diseases and cancer are responsible for the majority of the disease burden in older persons, primarily by affecting mortality rates. Although the incidence of cardiovascular diseases is higher than that of malignancies, cancer specific mortality rates of females are surpassing those of cardiovascular diseases, as was previously mentioned. As a result of aging of the population and improved prevention and treatment strategies for cardiovascular diseases, the contribution of cancer to the disease burden grows on average with 1000 persons per year.[4] It was estimated that in the year 2005 4 out of 10 males and 3.5 out of 10 females (all ages) were faced with at least one cancer during lifetime. 66% of all female patients and 75% of all male patients are diagnosed with cancer at an age above 60 years. 10% of persons gets more than one cancer during lifetime, in which metastases are not included. Most of the time the type of such recurrent cancers (e.g. basal cell carcinoma or contralateral breast cancer) or cause (e.g. smoking induced lung- or laryngeal carcinoma) are identical. Additionally, multiple cancers can affect a person coincidentally. With aging of the population and life expanding therapies, this number will increase even more in the coming decades.

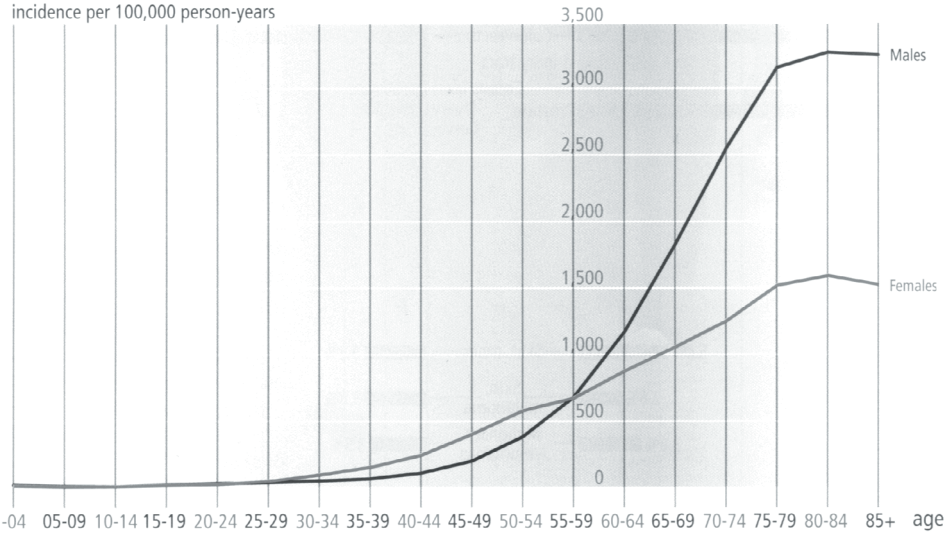

Figure 1. Age specific incidence rates of invasive tumours according to gender in 1998 (The Netherlands)

INCIDENCE, PREVALENCE AND MORTALITY

In the year 2005, a total of 74 500 incident cancers (all ages) were diagnosed in The Netherlands.[4] As much as 38 500 malignancy-related deaths were recorded in that same year. These numbers are more or less identical to other developed countries. Aging is the major risk factor for developing a malignancy. This can be explained by time needed to be exposed to harmful substances or influences, time to move through the different stages in the multistep carcinogenic process and the time it takes for a carcinoma to become clinically manifest. The extensive increase in cancer incidence at higher ages in men, as can be seen in FIGURE 1, is to an important extent caused by prostate- and lung cancer. The four most common malignancies (skin-related malignancies excluded) in developed countries are lung, colorectal, breast and prostate cancer (FIGURE 2 AND 3). Mainly these cancers affect cancer-related mortality.

In The Netherlands, the incidence rate of lung cancer in males has decreased with almost one third since 1989. For females, the incident rate has dramatically increased by 80% since the early 90s. These differences are explained by changes in smoking behaviour some decades ago. As a result of anti-smoking campaigns, the total number of males that uses tobacco has decreased drastically. However, female emancipation has led to an increase of the use of tobacco in women. Nevertheless, the disease burden of lung cancer is still 2.3-fold higher in

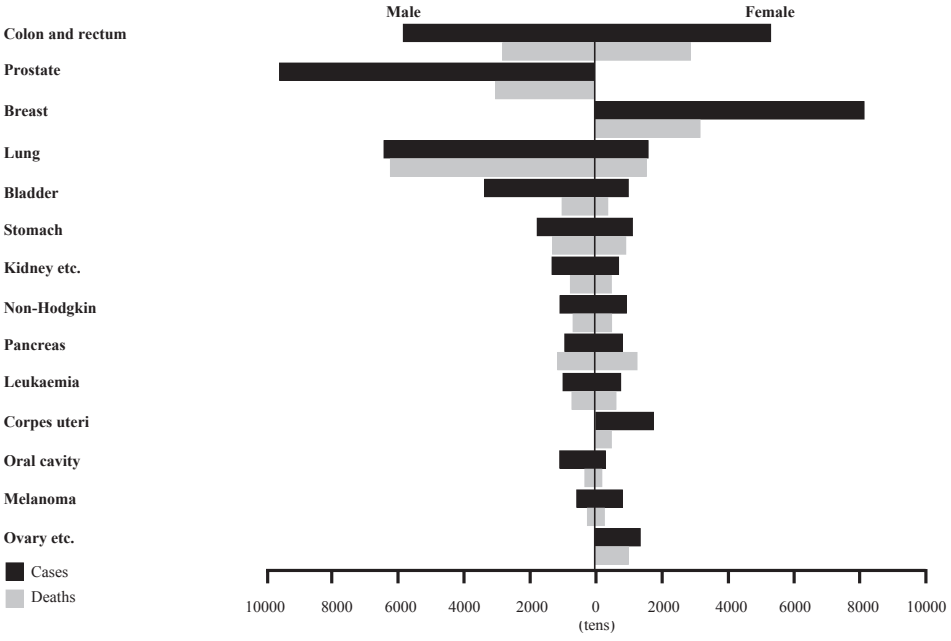

Figure 2. Incidence and mortality of cancer in persons aged 55 years and older in Western Europe (2002)

men than in women, but may converge in the near future. Since the decrease in males was somewhat higher than the increase in females, overall lung cancer incidence has declined over the last years.

Screening and prevention programs that started in the last one or two decades in The Netherlands as well, have influenced the incidence, prevalence and mortality rates of colorectal, breast and prostate cancer. Incidence rates increased while mortality rates decreased. For colorectal cancer (with high-risk patient screening programs) this resulted in slightly elevated incidence rates that were most clear in males. Mortality rates declined to a bigger extent over the last years. The population-based screening program for women aged 50-75 years that started in the early 90s has most definitely contributed to the major increase in breast cancer incidence as a consequence of early recognition. However, although cross-sectional mammograms are widely accepted among women in this age group for their proposed contribution to early detection and treatment, cost-benefit ratios are still matter of debate.

The striking increase in prostate cancer incidence rates since the early 90s is probably the result of the general introduction and use of measurements of prostate specific antigen (PSA) that is a marker of prostate cell growth. Although its specificity is relatively low since it cannot properly distinguish between prostate cancer and benign prostate hyperplasia, it more frequently resulted in early detection of prostate malignancies. This is probably the explanation for the decrease in incidence in men older than 75, who are detected with cancer at a younger age now. Nevertheless, lead-time bias (as a result of early detection) may be

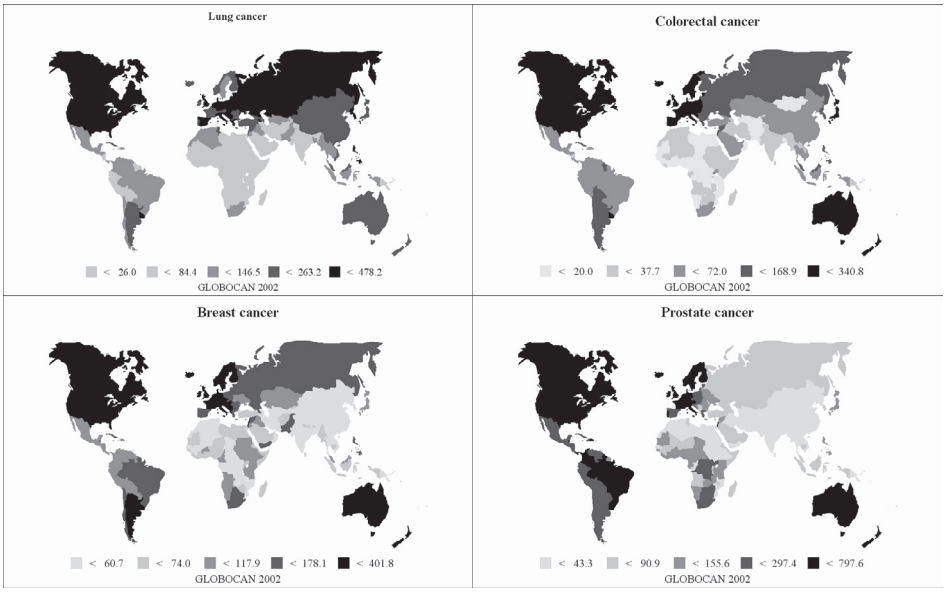

Figure 3. Geographic differences in cancer incidence of lung-, colorectal-, breast- and prostate cancer

of major importance in these cases and the question is whether it reflects a real increase of prostate cancer.

CANCER REGISTRATION IN THE NETHERLANDS

The aim of the in 1986 introduced National Cancer Registry (NCR) is to provide a deeper understanding of the character and quantity of cancer in The Netherlands.[4] For this purpose, it uses population-based information from automated registries of pathology laboratories (Patholologisch Anatomisch Landelijk Geautomatiseerd Archief (PALGA)) and hospital admission registries (Landelijke Medische Registratie (LMR)). (FIGURE 4) Nine regional cancer registries first register the information before it is submitted to the NCR. (FIGURE 5) Eventually, it covers more than 95% of the total Dutch population. The NCR collaborates in the International Association of Cancer Registries (IACR) that includes approximately 200 national cancer registries from all over the world.[17]

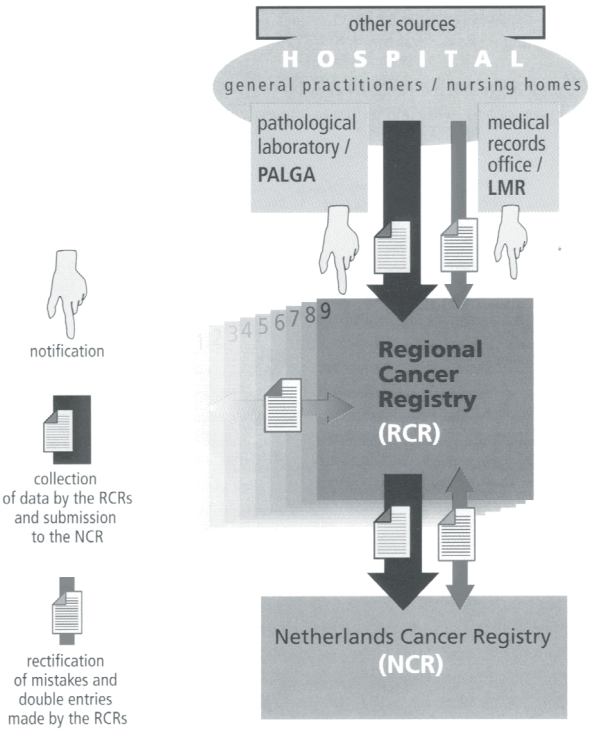

Figure 4. Data collection by Netherlands Cancer Registry

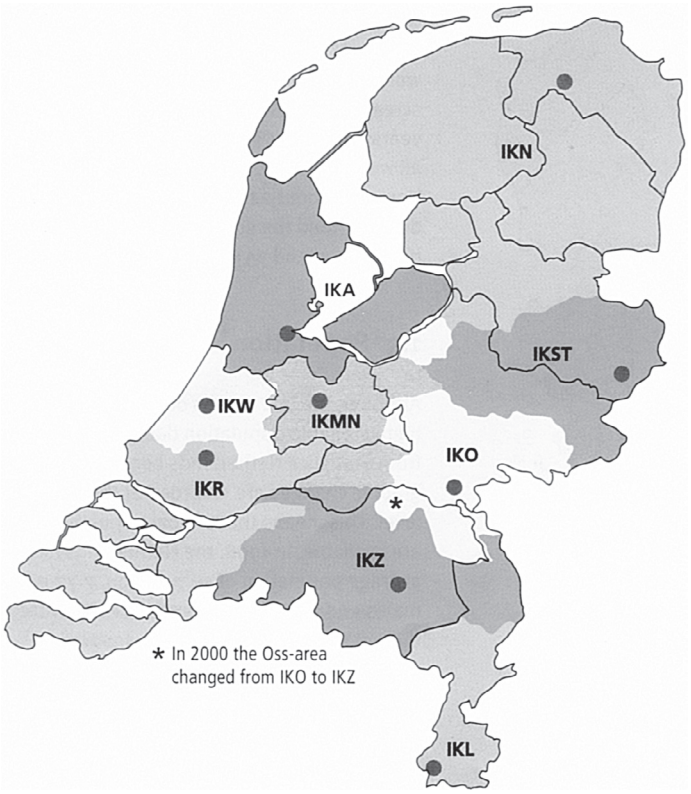

Figure 5. Regional cancer centers

CANCER REGISTRATION IN THE ROTTERDAM STUDY

The approach of case identification and validation of cancers in The Rotterdam Study was similar to the national registration. As a substitute of PALGA, a regional pathology registration (Pathologisch Anatomisch Laboratorium (PATHAN)) was used besides information from the LMR and general practioners working in the study district. For this thesis we focussed on the four most common non skin related cancers; i.e. lung, colorectal, breast and prostate cancer. Cases were subsequently validated by a physician (CS) on the basis of the International Classification of Primary Care (ICPC) (colorectal (D75), lung (R84), breast (X76) and prostate (Y77) cancer). The 10th version of the International Classification of Diseases (ICD-10) was used to distinguish between the subtypes non-sigmoid colon cancer (C18), sigmoid cancer (C19), rectal cancer (C20), anal cancer (C21), mesothelioma (C45) and bronchus carcinoma (C34) and to give an alternative code for breast (C50) and prostate (C61) cancers. The index date was defined as the earliest date found in the pathology reports.

During the follow up period between July 1989 and October 2004, a total of 792 cancers occurred. (TABLE 1) For lung and colorectal cancer a division on subtype was made. Unfortu-

nately, histology was not available for all cancers. In this thesis, bronchus carcinoma will be called lung carcinoma or lung cancer. Two (0.8%) of the 265 breast cancers were diagnosed in males and these were both incident cases. This percentage is almost equal to the 0.9% that is reported by the NCR.

Of the four cancers studied, incidence rates per 100,000 person-years over the years 1991 until 2004 are presented in FIGURE 6. Due to small sample sizes, the rates somewhat fluctuate. Nevertheless, results are in line with what was previously mentioned about the incidence in the Dutch population. Lung cancer incidence drastically declined after the year 1991 in males and increased in females. Colorectal cancer incidence somewhat increased, particularly in

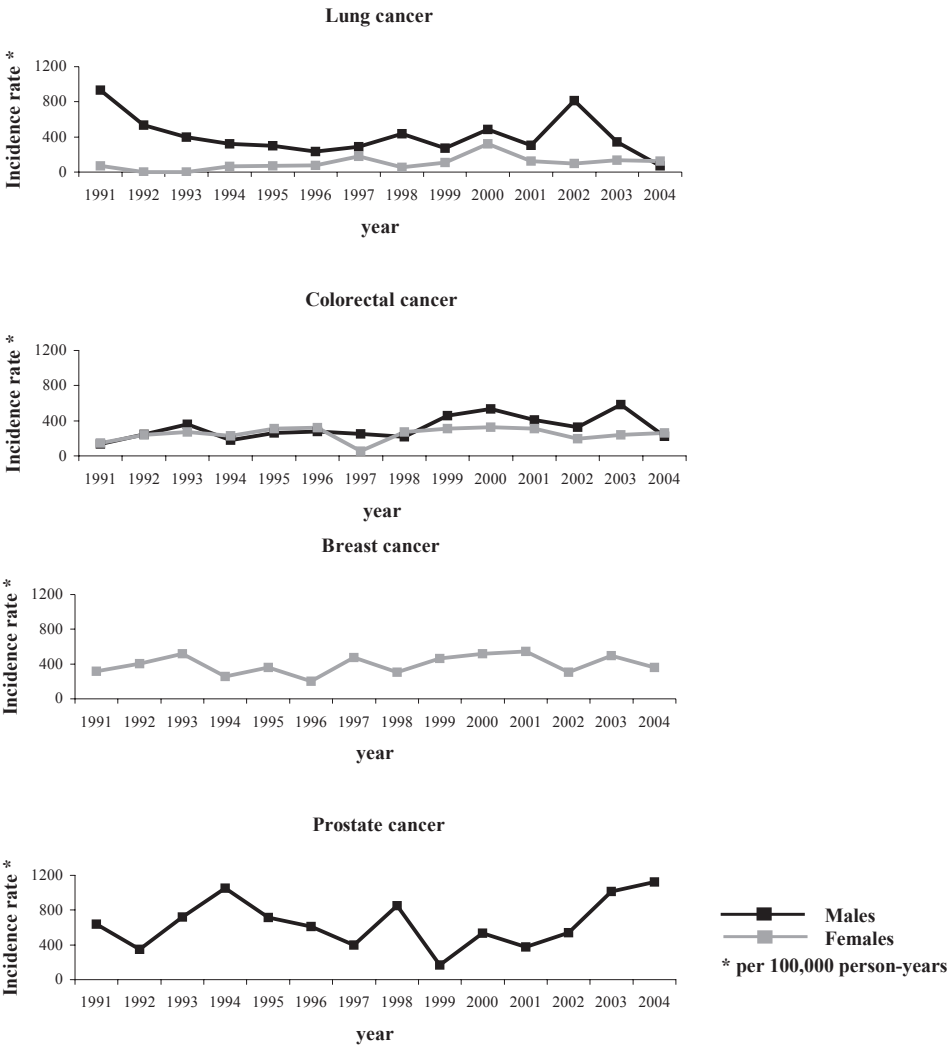

Figure 6. Incidence rates of four cancers in The Rotterdam Study

males. Incidence rates for breast cancer in females seemed stable over the years. The strongest fluctuations were found for prostate cancers. Although the incidence rates reported in FIGURE 6 do not include males who participated in special screening programs, the introduction of PSA measurements in the general medical practice might have caused these changes over the years.

FIGURE 7 shows the age-specific incidence rates per cancer in both The Rotterdam Study and the Dutch population.[4] Confidence bounds are displayed for the incidence rates found in The Rotterdam Study. As the majority of incidence rates of the Dutch population were

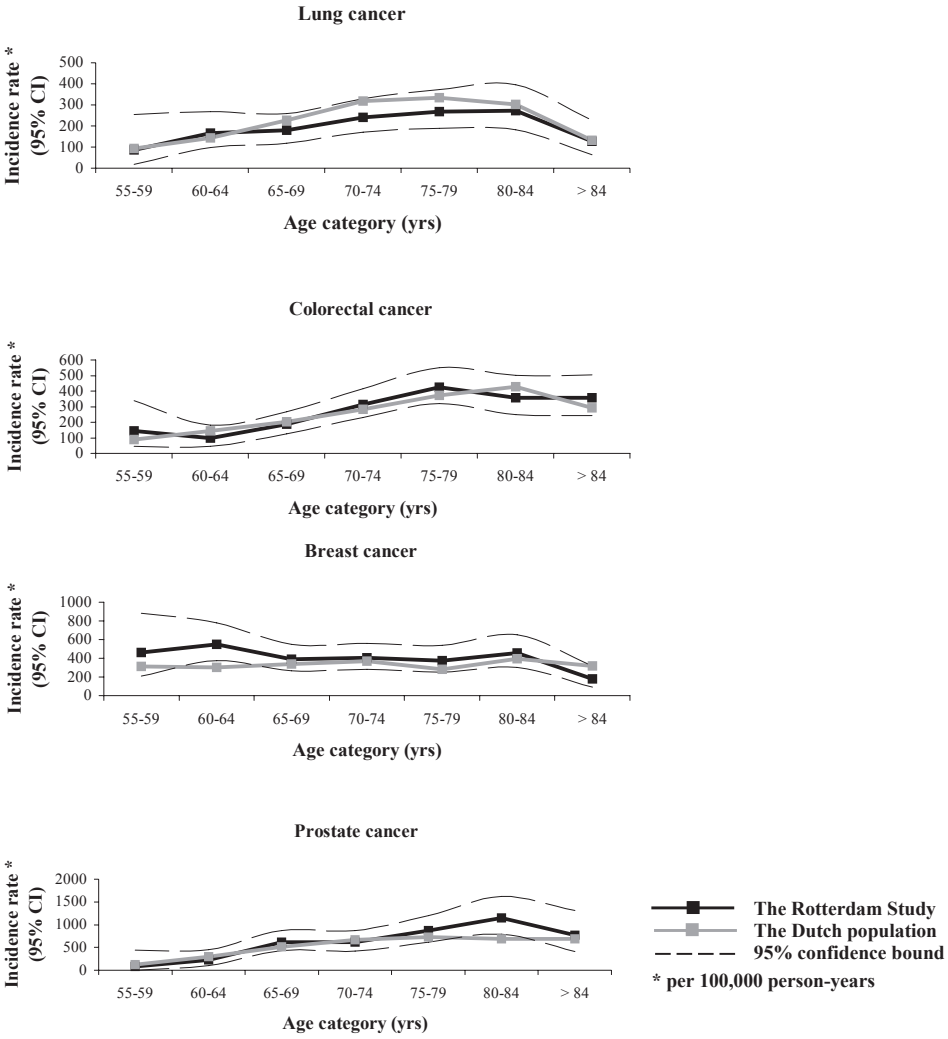

Figure 7. Age-specific incidence rates in The Rotterdam Study and the Dutch population

Table 1. Frequency of four common cancers in The Rotterdam Study

Cancer	Subtype	Histology [a]	All	Prevalent	Incident
Total cancers			959 (12%)	167	792
Lung cancer			187 (2.3%)	24	163
	Bronchus carcinoma		176 (2.2%)	24	152
		Adenoma	23	5	18
		SCC	50	13	37
	Mesothelioma		11 (0.1%)	0	11
Colorectal cancer			275 (3.4%)	60	215
	Non-sigmoid cancer		109 (1.4%)	20	89
	Sigmoid cancer		90 (1.1%)	16	74
	Rectal cancer		73 (0.9%)	24	49
	Anal cancer		3 (0.0%)	0	3
Breast cancer			265 (3.3%)	79	184 + 2
Prostate cancer			285 (3.6%)	57	228

SCC = squamous cell carcinoma. [a] Histology was not available for all bronchus carcinoma.

included in the 95% confidence intervals of the incidence rates of The Rotterdam Study, data were considered valid.

1.4. Genetic Variation

GENETIC VARIATION [18]

Sequence analyses suggest that 99.9% of the human genome is identical between individuals. Genome sequence variations exist at defined positions within genomes and are partly responsible for individual phenotypic characteristics. These characteristics include colour of the eyes or hair, but also a person's propensity towards complex disorders such as heart disease and cancer and the response to treatment. This makes genetic variation of great value for biomedical research. The variation in genome sequence is mainly due to one form of variability that is called **single nucleotide polymorphisms (SNPs).** Other forms of variation are (**variable number of) tandem repeat (VNTR) polymorphisms** and **deletion / insertion polymorphisms (DIPs)**. The most common version of an allele is called the wild type; the other version(s) are called variant(s). Most of the polymorphisms do not have a direct impact on disease or treatment susceptibility, but may rather serve as a marker to identify the real disease-associated gene. When these polymorphisms are found to be inherited with a particular trait, they may provide evidence of where the trait's gene is located on the genome.

Variable Number of Tandem Repeat Polymorphisms

Frequencies of VNTRs are estimated in a computer-based (in silico) setting at 100.000 in the total human genome. [18] They consist of variable length of identical or closely related sequence motifs that are repeated in tandem in a variable copy number. Based on the size of the tandem repeat units, VNTRs can be subdivided into *microsatellites* or short tandem repeats (STR) with approximately 1-6 repeat units and *minisatellites* with up to 100 repeat units.

Insertion / Deletion polymorphisms

The appearance of insertion / deletions is difficult to quantify, but is supposed to fall in between the frequencies of VNTRs and SNPs. It is hypothesized that VNTRs or inverted VNTRs predispose DNA to localized rearrangements. The rearrangements mainly occur between homologous repeats and are probably one of the causes of an insertion/deletion polymorphism.

Single nucleotide polymorphism

The frequency of SNPs is the highest. In the human genome, which is $3.3 \cdot 10^9$ base pairs, almost 13 million SNPs have been identified.[19] It is expected that in total there are 15-20 million SNPs, which means that they appear ones every 150-200 base pairs. It received its name due to single base substitutions of one nucleotide for another that occur in a population at a frequency of more than 1%. This latter distinguishes it from mutations that occur in less than 1% (better known as Mendelian genes). Survival of mutations leads to fixation of the allele in populations (see FIGURE 1). The substitutions can be subdivided in *transitions* and *transversions*. A nucleotide change of one purine (A,G) for another purine, or a pyramidine (C, T) for another pyramidine is entitled a transition. Transversions are nucleotide substitutions that

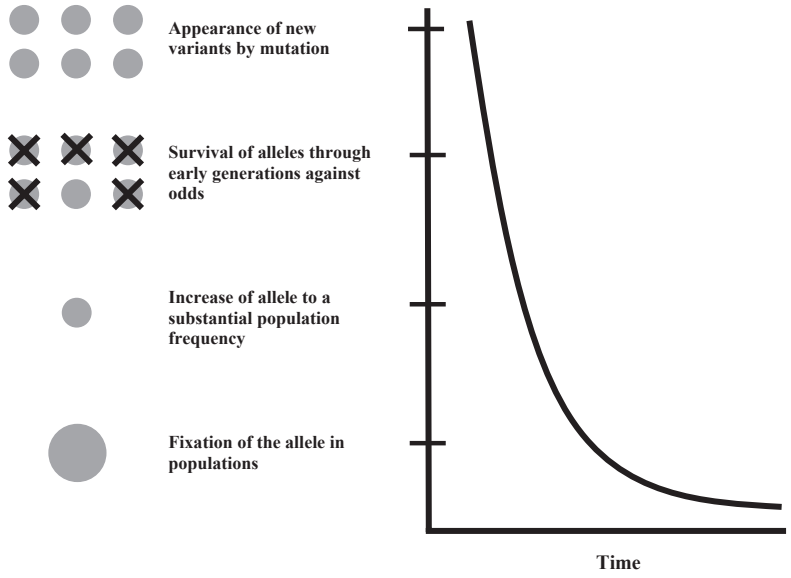

Figure 1. Life cycle of mutations into SNPs (Source: Bioinformatics for Geneticists page 45)

result in a change of a purine for a pyramidine or vice versa. Twenty-five percent of all SNPs in the human genome is accounted by G > A and C > T transitions.

CLASSIFICATION OF SNPS

SNPs may arise at any position in the genome (FIGURE 2). When SNPs appear in exons, that are the coding regions of the genome, the SNP is called a coding SNP (cSNP). These cSNPs can be subdivided into those that create a new codon that still codes for the same amino acid (synonymous coding SNP) and those that code for a new amino acid (nonsynonymous coding SNP). Synonymous coding SNPs can also be entitled as silent mutations. Additionally, SNPs

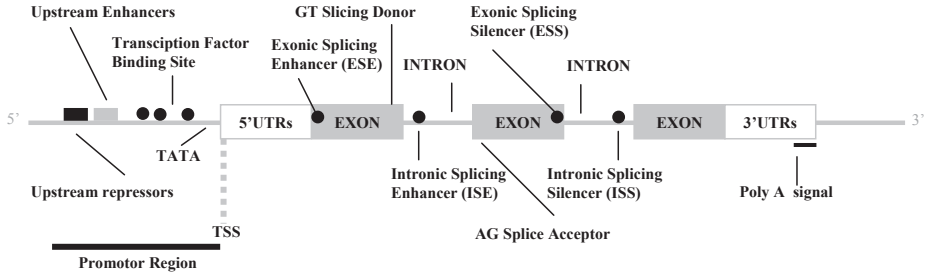

Figure 2. Possible genomic positions of polymorphisms

can appear in intronic regions or in the promoter region. The latter frequently results in an altered activity of the transcription of the gene and consequently a change in expression level of the corresponding protein.

HARDY WEINBERG EQUILIBRIUM

Until the beginning of the twentiest century, it was thought that over time dominant alleles would overwhelm recessive ones. This phenomenon was called 'genophagy' or 'gene eating'. In 1908 the English mathematician Godfrey Harold Hardy and the German physician Wilhelm Weinberg proved this theory to be wrong. (FIGURE 3) They showed that dominant alleles could just as easily decrease in frequency as recessive ones. Hardy and Weinberg stated that evolution is simply a change in frequency of alleles in the gene pool of a population (all genes of all persons considered to be part of a specific geographic group). No evolution would appear when seven conditions are met: no mutations, no natural selection, an infinitely large population, all members of the population breed, all mating is totally random, everyone produces the same number of offspring and nobody migrates in or out of the population. However, it is highly unlikely that these conditions are met and therefore evolution is the inevitable result. The equation that was called after Hardy and Weinberg can be used to determine the probable genotype frequencies in a population and to see how they change from one generation to another, since without outside forces the proportion of dominant and recessive alleles will remain constant from generation to generation. The equation is $p^2 + 2pq + q^2$, in which p and q are the allele frequencies of the dominant and recessive alleles. By using the equation it is possible to determine the genotype frequencies if phenotype frequencies are known and vice versa. In genetic associations studies, such as presented in this thesis, the Hardy Weinberg Equilibrium (HWE) is used to validate the result found in the study. Since deviations from HWE (indicated with p-values < 0.05) might not only indicate inbreeding or population stratification, but also genotyping problems and selection bias. For calculations of the HWE a simple χ^2-goodness-

Figure 3. Godfrey Harold Hardy (1877 – 1947) and Wilhelm Weinberg (1862 – 1937)

of-fit test is commonly performed that analyses differences between expected and observed allele frequencies. The genotype frequencies that must be used for these calculations should include those of the total study population (cohort study) or of controls (case-control study) as the initial hypothesis when performing an association study is that some specific alleles might be associated with diseases. So one would expect, for instance in a case-control study, that in cases there are deviations from HWE while controls are in HWE.

GENOTYPING

To study genotypes in association studies, high-throughput assays are available which are able to determine the requested genetic variation of a gene of interest. There are several

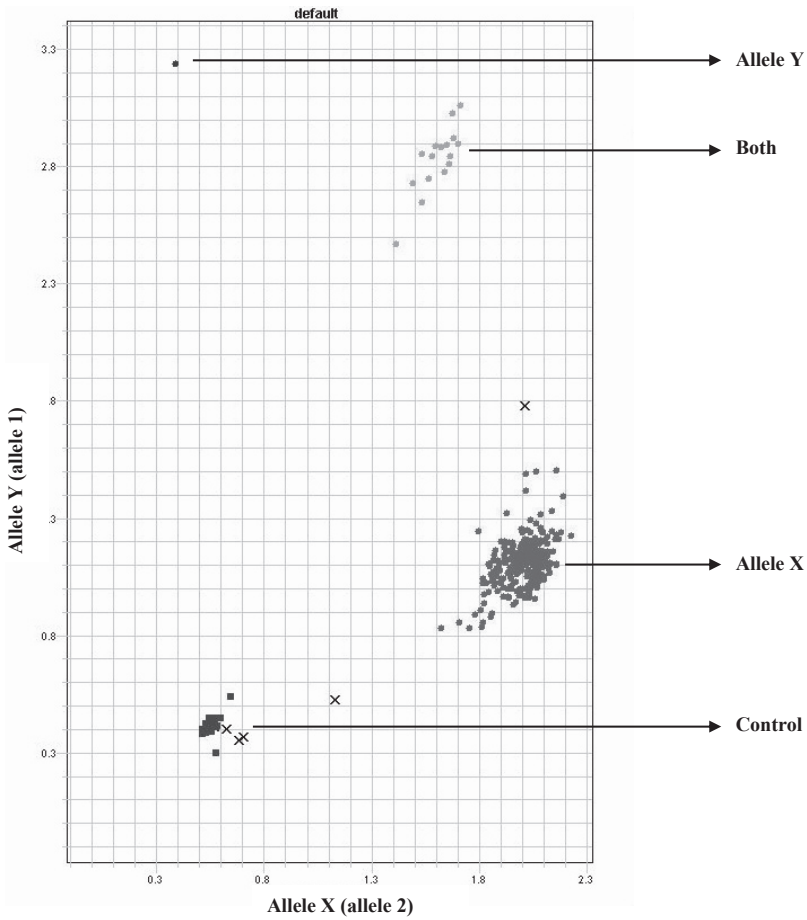

Figure 4. Output of a Taqman assay on the A-392G SNP of *CYP3A4*

methods for genotyping. For the analyses in our study, a real-time quantitative PCR method (Taqman) was used.[20] Genotyping results for an assay on *CYP3A4*1B* (A –392 G) are presented in FIGURE 4. The place of the individual dots represents the individual genotype. The left upper corner represents individuals homozygous for the Y-allele (in this example the G-allele). The X-allele area (in this example the A-allele) represents individuals homozygous for the X-allele. Person that both carry an X- and a Y-allele are placed in between. Dots in the left lower corner represent control samples. For individuals indicated with an X, genotyping data were undetermined. In this example, the groups for hetero- and homozygous genotypes can be easily distinguished. However, genotyping errors can exist when there is overlap between the groups, which makes it difficult to place individuals in the correct group.

DATABASES

Several databases provide information about SNPs. Two major classes of SNPs appear in these databases. Submitted SNPs (ss#) emerge from data of original observations of sequence variations. When these SNPs are checked for their redundancy and turn out to be unique it is given a reference SNP cluster number (rs#) that comprises all non-redundant annotated markers. Most commonly used is the dbSNP database [19] that is part of the database of the National Center for Biotechnology Information (NCBI). Another frequently used part of the NCBI database is the Online Mendelian Inheritance in Man (OMIM) that is available online since 1995.[21] This part of the NCBI gives a summary of the current state of knowledge of particular genes and disorders. It includes a variety of topics that range from biochemical and cytogenetic features, to diagnosis and clinical management and to allelic variants. Another useful database is the Human Genome Variation Database (HGVbase) that is the product of a collaboration between the Karolinska Institute (Sweden) and the European Bioinformatics Institute (UK).[22] Its objec-

Tabel 1. Databases for genetic association studies

Database	Website
Single Nucleotide Polymorphism database (dbSNP)	http://www.ncbi.nlm.nih.gov/SNP/index.html
Online Mendelian Inheritance in Man (OMIM)	http://www.ncbi.nih.gov/OMIM/index.html
Human Genome Variation Database (HGVbase)	http://hgvbase.cgb.ki.se
Single Nucleotide Polymorphism Consortium	http://snp.cshl.org
Perlegen Sciences	http://www.perlegen.com
Ensemble	http://www.ensembl.org/index.html
HUGO Genome Nomenclature Committee	http://www.gene.ucl.ac.uk/nomenclature
GeneCards	http://www.genecards.org/index.shtml
GeneSNPs	http://www.genome.utah.edu/genesnps
The Allele Frequency Database (ALFRED)	http://alfred.med.yale.edu/alfred
The pharmacogenetics and –genomics knowledge base	http://www.pharmgkb.org/index.jsp

tive is to provide a comprehensive catalog of normal human gene and genome variation that can be used as a research tool to help define the genetic component of human phenotypic variation. To provide public genomic data, the SNP consortium was organized as a non-profit foundation with the intention to develop up to 300,000 SNPs throughout the human genome that would be available to the public without intellectual property restrictions.[23] It was started in 1999, and to date, they discovered up to 1.8 million SNPs. Some of The SNP Consortium members are now in a process with the goal to determine the frequency of 60,000 SNPs in three major world populations (African American, Asian, Caucasian). This collaboration is called The SNP Consortium allele frequency/genotype project. Other useful websites are displayed in TABLE 1.

HAPLOTYPES

As was said before, the whole genome will probably include up to 10 million SNPs. To identify them all would be time consuming and expensive. Therefore, researchers are trying to downsize the number of genotypes by studying haplotypes. These refer to a set of single alleles or closely linked genes that tend to be inherited as a complete unit. It is thought that 65-85% of the human genome is organized in haplotype blocks. Each block comes in three or four common versions that capture the majority of genetic diversity throughout the entire human population. One block may contain a large number of SNPs, but only a few are enough to uniquely identify the common versions of the haplotype. These specific SNPs are called tagging SNPs. The common patterns of human DNA sequence variations are described by the HapMap Project.[24] For this purpose, DNA samples of 270 persons (90 Nigerians, 45 Japanese, 45 Han Chinese and 90 northern and western Europeans) are genotyped for at least 1 million SNPs across the genome. Results of these genotyping efforts will be a tool that will allow researchers to find genes and genetic variations that affect health and diseases. The blocks can be visualized by using the software package Haploview.[25] Inputted data can be of own material or dumped from the HapMap website. In this way, the percentage of linkage disequilibrium (allelic association), structure of haplotypes blocks, and tagging SNPs can be computed. FIGURE 5 shows an example of haplotype blocks in the *CYP3A* gene (a group of genes that was studied in this thesis) and the four most commonly studied SNPs. The darker the parts, the more linkage between SNPs was found, which eventually leads to the formation of haplotype blocks. The white parts indicate areas for which linkage could not be calculated as a result of insufficient or absent data. Only the *CYP3A5*3* SNP can be used as a tagging SNP for the first block, as it has in contrast to the other three a considerable allele frequency (80% in Caucasians). Studies on the association between haplotypes and disease outcome require specific software that reconstructs haplotypes of the individual genotypes in the studied population. As SNPs are detected individually, frequencies of combinations based on

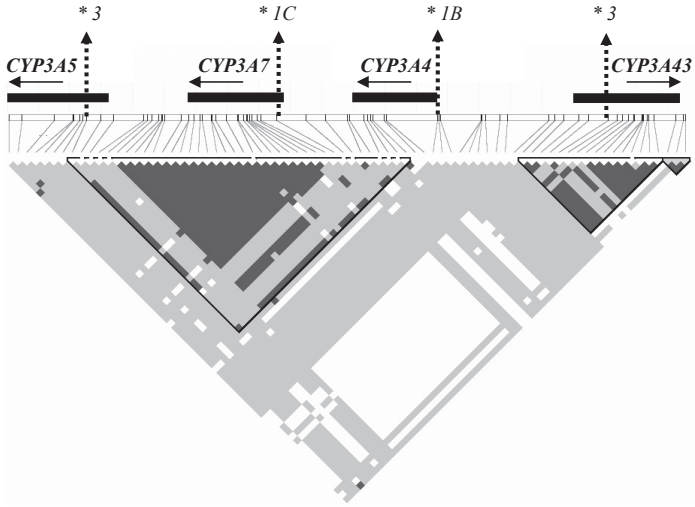

Figure 5. *CYP3A* structure, linkage disequilibrium blocks and SNPs

the haplotype blocks have to be calculated. The program PHASE estimates the most probable haplotype per individual, which can be merged into a statistical program, such as SPSS, for association analyses.[26] The advantage of another program called HaploStat is the implementation of weighted probabilities of haplotypes per individual, which results in more precise estimates.[27] However, the reconstructed haplotypes can only be studied in the statistical program R+ with special written syntaxes.[28] Since haplotype association studies are still in its infancy, only most straightforward analyses are possible to date with this software, but this will most probably be extended in the near future.

2. Cytochrome P450 and cancer

2.1. A critical appraisal of association studies of cytochrome P450 gene polymorphisms and susceptibility to cancer

ABSTRACT

Background: Polymorphisms of cytochrome P450 (CYP) enzyme-encoding genes involved in metabolic pathways of (pro)carcinogenic substances may be important for the susceptibility of an individual to environmental risk factors in relation to the induction or promotion of cancer. Human genome sequence mapping has brought new opportunities to identify etiological factors in non-mendelian complex diseases such as cancer.

Methods: We summarized the present knowledge on the association of CYP450 polymorphisms and cancer by conducting a search on *CYPs* and cancers and their associated terms in PubMed (including MEDLINE) until October 2005. The study was restricted to genetic polymorphisms as a susceptibility factor, therefore excluding literature on the expression of CYPs in tumor tissue.

Results: Most striking results were dissimilarity of studies in respect to ethnicity, composition of control groups, phenotype of interest, variant type and reference type. Moreover, some included specific substrates on which cytochrome P450 enzymes may act as effect modifiers, whereas others only studied genotypes. Most studies discussed cytochrome P450 enzymes with low family numbers that are primarily involved in the metabolism of tobacco and alcohol (*CYP1A1*, *CYP1B1*, *CYP2E1*) or those that are involved in hormonal pathways (*CYP17*, *CYP19*). The largest number of different CYPs were studied for prostate cancer. Half of the studies reported in this review were performed among Asian individuals, which indicates a lack of evidence in other populations since genetic association studies are proposed to be dependent on differences in ethnic allele frequencies.

Discussion: Apparently, further investigation of associations between CYP polymorphisms and carcinogenesis is needed, not only by studying single nucleotide polymorphisms but preferably with genome-wide screening and haplotype analyses. Furthermore, there is a need for standardization of study designs and methods to enable meta-analyses. More knowledge might lead to individualized treatments and provide risk profiles by combining genetic variants of different candidate genes.

BACKGROUND

Cancer is one of the most important causes of death throughout the world. From experimental and clinical research, it is commonly assumed that environmental- as well as genetic risk factors can influence carcinogenesis.[29,30] There is an increasing interest for interactions between genetic and environmental factors. This is caused by the fact that cancer is one of the diseases for which a similar degree of exposure to a certain substance may result in an individual difference in the incidence, which indicates a genetic predisposition in the etiology. For instance, not all persons who smoke develop a malignancy of the lung and even non-smoking individuals can get lung carcinoma.[31] Candidate genes for the interaction would encode for specific detoxification enzymes. Variation in such genes might be associated with an altered risk of carcinogenesis due to changes in the structure or quantity of particular encoded enzymes involved in either induction or promotion of cancer. Detoxification of drugs or endogenous- and environmental substances is supposed to take place in two phases. Phase I reactions mainly concern degradation of substances, whereas in phase II, intermediates of phase I reactions are conjugated to become more water-soluble to facilitate excretion. Several phase I and II enzyme-encoding genes have been described in the literature, such as *mEH* (microsomal epoxide hydrolases), *GSTM1* (gluthatione S-transferase, a phase II enzyme), and *COMT* (catechol-O-methyltransferase). Considering the large number of substrates they metabolize, phase I cytochrome P450 (CYP) enzymes were hypothesized to play an important role in the metabolism of potential (pro) carcinogens as well. As a consequence, a large number of studies on the association between variation in the cytochrome P450 gene and cancer development has been performed so far.

Previous review articles discussed the association between CYP enzyme *expression* in normal and tumor tissues and some functional aspects of CYP variation.[32-34] We review studies of the association between cytochrome P450 gene variation and cancer development. First, we discuss some general issues concerning cytochrome P450 enzymes and the consequences of genetic variation on its structure and physiological function. Thereafter, we will summarize and discuss the literature on the association between cytochrome P450 variants and cancer.

GENES

The biological function of cytochrome P450 enzymes

Cytochrome P450 designates for a group of enzymes that is most abundant in smooth endoplasmatic reticulum of hepatocytes and in epithelial cells of the small intestines. They are regulated by many transcriptional control factors, including pregnane X-receptor (PXR), constitutive androgen receptor (CAR) and the aryl hydrocarbon receptor (AHR). The main function of CYP enzymes is to catalyze oxidative reactions concerned in metabolic processes

of exogenous and endogenous substances. The most frequently studied biological activity of CYP enzymes is as part of a phase I reaction in which water insoluble drugs are metabolized into more water-soluble substances. Furthermore, some cytochrome P450 enzymes play a dominant role in the metabolism of xenobiotics, such as tobacco smoke and other exogenous toxins and in the conversion of hormones into their biologically (in)active substrate. The efficacy of these processes is not equal between persons. This results in interindividual variance of activation or inactivation of such compounds eventually leading to variations in character and quantity of metabolites. Genetic polymorphisms (defined as changes in the DNA structure with a frequency of at least 1% in the general population) are suggested to explain an important part of this variation.

Variation in cytochrome P450 enzymes

For a long time, the interpretation of mutations at a single gene (or locus) proposed by Mendel, was the only way of explaining the genetic basis of diseases. This explanation applied only to monogenetic diseases, mostly with a low incidence. With the discovery of numerous genes causing more common familial diseases, this Mendelian way of thinking was not appropriate to all genetically based diseases. These so-called complex diseases could not be explained by a simple monogenetic theory.[5] Cancer is such a complex disease in which multiple genetic and environmental factors are considered in the developmental process and for which genetic variation in candidate genes can be of substantial importance.

The activity of CYP enzymes can genetically be influenced in several ways. First, while variation in a **single nucleotide** in coding DNA can be silent, it may also lead to a change in transcription of RNA and subsequent translation into an amino acid. This may have consequences for the CYP protein structure or folding and can lead to an altered activity. Similarly, such a functional polymorphism may be the result of the **insertion or deletion** of one or several nucleotides, or of a variable number of tandem repeat units (**VNTRs**). Second, a genetic polymorphism might lead to completely abolished CYP activity if it alters the **transcriptional initiation site**, results in **splicing** or introduces a **stop codon**. Third, CYP activity may be increased when there is **gene duplication**. Additionally, zygosity is important for the risk profile, since an altered biological activity of a particular enzyme can result from either one (heterozygous) or two (homozygous) affected alleles.

It is likely that variation in one of the human CYP genes leading to a decreased or abolished activity of a particular CYP enzyme is associated with an increased tumor risk after exposure to a carcinogen. On the other hand, if cancer results from a carcinogenic metabolite, a decreased CYP activity may reduce the risk of cancer. The wild-type allele (variant with the highest frequency in the population where it was first detected) is then considered to be the risk allele.[35-40] Besides a direct risk on cancer development, polymorphisms can also be associated with an increased clinical or pathological severity at presentation.[41-44] Eventually,

combinations of genetic polymorphisms across different genes might result in a risk profile for clinically relevant individual susceptibility.[45-47]

Nomenclature cytochrome P450

The collective group of cytochrome P450 enzymes consists of approximately 1000 different known isoenzymes or subforms, each of which has a specific metabolizing effect. Approximately 57 of these subforms appeared to be important in humans.[48] The nomenclature of CYPs is based on the corresponding amino acid sequence. A *number* and *letter* indicate a CYP-family and –subfamily respectively. A *second number* is given to specific isoenzymes.[49] (FIGURE 1) Sometimes the family name originates from its enzymatic function (e.g., *CYP17*; steroid 17-alpha-hydroxylase). When variations in these genes are validated by dbSNP (a single nucleotide polymorphism database of the National Center for Biotechnology Information (NCBI)) and mapped to a specific location on the genome, they are indicated with a * and a number.[19] Genetic polymorphisms of CYPs may have important phenotypic consequences in relation to enzyme activities. Widely used phenotypic categories are: "poor- (PM)", "intermediate– (IM)", "extensive- (EM)" and "ultra rapid metabolizers (UM)". Unfortunately, inconsistency in this nomenclature is not unusual, partly because of changes in classification over the time.

Family	Subfamily	Isoenzyme	Polymorphism [a]
CYP1			
CYP2	A		
CYP3	B	1	
CYP4	C	2	
CYP5	D	…	* 2
CYP6	…	…	
CYP7			
CYP8			
…….			

[a] Only dbSNP validated variants are indicated with a * (see text)

Figure 1. Nomenclature CYPs

METHODS

For our review of the literature, we conducted a search using the English terms 'cytochrome P450', 'cyp', 'cancer', 'tumor', 'tumor', 'malignancies', and 'malignancy' in PubMed (including MEDLINE) during the period January 1966 – October 2005. Subsequently, title and abstract of papers, were screened for terms such as 'susceptibility' and 'risk' or 'risk factor', and relevant papers were copied and reviewed. Finally, the references of all relevant papers were checked for missing papers to complete our search. Each polymorphism found in this way was checked for the type of cancer and the exogenous or endogenous substance for which the relation was described, the number of risk alleles (homo- or heterozygous genotypes), the risk estimates that were found, study sample size, the ethnicity of study subjects, and Hardy Weinberg equi-

librium in control individuals. We focused on observational studies regarding the association between polymorphisms of cytochrome P450 genes and cancer, and excluded papers on cytochrome P450 expression in (tumor) tissue. First, we discuss the location, possible functionality and racial allele frequencies of those polymorphisms that were found with this literature search (TABLE 1) Subsequently, reported findings are described according to tumor type and cytochrome P450 isoenzyme. (TABLE 2) Finally, a summary is made per CYP isoenzyme for all associations with cancer types. (TABLE 3)

GENE VARIANTS AND FREQUENCY

Location, functionality and allele frequency of CYP polymorphisms

The importance of variation in cytochrome P450 enzyme-encoding genes for predisposition to diseases such as cancers differs per CYP-family. In general, CYPs with low family numbers (1-3) have less affinity for their substrates, are less conserved during evolution and are therefore more polymorphic than CYPs with higher family numbers (5-51) that are well conserved and have a high affinity for predominantly endogenous substrates.[50] Subsequently, substrate specificity in the first group of low family numbers can be divided into two classes; those that do not have important functional variability and metabolise both (pre) carcinogens and drugs, and those that seem to have functional variability but are only known for their drug metabolising properties.[34] This division can be observed in the number of studies per CYP family as well. Our literature search revealed that most studies investigated *CYP1A1*, *CYP1B1,* and *CYP2E1* enzymes, which are all polymorphic and active in the metabolism of (pre) carcinogens. However, variation in the steroid metabolising enzymes (*CYP17* and *CYP19*) was repeatedly mentioned as well.

CYP1A1 (chr. 15q22-24) plays a role in the activation of major classes of tobacco carcinogens such as polycyclic aromatic hydrocarbons (PAHs), aromatic amines (Aas) and benzo[a]pyrenes (BPs).[46,51-55] Another role is the catalization of the C-2 hydroxylation of estradiol.[56] Two polymorphisms, both located on exon 7 which are frequently studied are the T3801C (**2A*) and the A2455G (**2C*) substitution. The **2A* polymorphism is denoted with its polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) alias *MspI* (a specific restriction enzyme). The 2455G-polymorphism results in an amino acid change on codon 462 of isoleucine into valine, indicated as Val^{462}. These two variants can occur independently but a combined variant (**2B*) has been detected as well. All three variants result in an increased enzyme activity. The allele frequency of 3801C varies from 40 to 80 percent in Asians but seems to be less frequent in South Americans (4-28%). Although some report a frequency of 2-3% [57], the prevalence in Caucasians needs to be further investigated [58]. More consistent are the findings for 2455G that seems to be absent in Africans and African-Americans while in other ethnic groups it has an allele frequency of 10-20%.[59] The **4* (C2453A) polymorphism

results in a threonine to asparagine amino acid change on codon 461 and was found to be associated with a higher inducibility [60] and increased bioactivity [61]. This polymorphism seems only present in Caucasians (7%).[59] Other variants in the *CYP1A1* gene have been detected and studied as well, but these were reported only once and their functionality and allele frequency has to be further explored (see TABLE 1 for details).

Highly related to *CYP1A1* is its subfamily member ***CYP1A2*** (chr. 15q22-24). The genes that encode for these enzymes are located on chromosome 15 and share their 5' flanking region.[62] However, only for *CYP1A2* a polymorphism in this region (**1F*; C-163A) was identified that has been associated with a higher induction due to smoking [63], omeprazole [64], and by male gender [65] although this was not consistently found in all literature. Allele frequencies across populations occur in approximately 30-40 %, with highest rates in African individuals.

Besides their role in smoking-related metabolism, ***CYP1B1*** (chr. 2q22-p21) is involved in hormonal pathways. It is predominantly expressed extra-hepatically and overexpressed in tumor tissue.[34] Although a significantly decreased capacity to metabolize both estradiol [66] and benzo[a]pyrene [67] was found for **7* (a combination of C142G, Arg48Gly; G355T, Ala119Ser; C4326G, Leu432Val; and C4360G, Ala443Gly), our literature search only revealed associations of **2* (C142G and G355T) and **3* (C4326G) with cancer development. These individual genotypes might be responsible for a similarly decreased metabolism, although this needs further studying. Population diversity is not yet clear either.

Metabolic activation of tobacco carcinogens is also modified by *CYP2A6* [68], *CYP2A13* [69], *CYP2C19* [70], *CYP2E1* [68,71], *CYP3A4* [72] and *CYP3A5* [73]. They are involved in the oxidation of nicotine and the activation of *N*-nitrosamines such as 4-methylnitroso-1-(3-pyridyl)-1-butanone (NNK), *N*-dimethylnitrosamine (NDMA), *N*-diethylnitrosamine (NDEA), *N*-nitrosophenylmethyl-amine (NMPhA) and *N*-nitrosonornicotine (NNN). ***CYP2A6*** (chr. 19q13.2) mediates 90% of conversions of nicotine to the inactive cotinine [58] and possesses relatively long alkyl chains [68]. Corresponding with this finding was that a deletion of *CYP2A6* (**4*; G1846A, splicing defect) was associated with a lower tendency to smoke.[58,74-76] This particular polymorphism was proposed to increase smoking-related cancers since the concentration of nicotine might be longer active in the blood. However, the allele frequency was found to be low in Caucasians, with an allele frequency of 0.5-1% in contrast with a frequency of 7-22% in Asians.[77] Therefore, significant findings are not expected to be found in individuals of Caucasian descent due to this low allele frequency.

CYP2A13 (chr. 19q13.2) is particularly involved in the metabolism of NNKs. Three polymorphisms of importance have been found up till now. The first was a polymorphism leading to an Arg257Cys substitution in exon 5. The Cys[257] variant showed 37-56% reduced activity towards several substrates in comparison with the wild-type protein.[78] Ethnic variation in population frequency data has only been observed in Chinese populations.[79] A haplotype, in which carriage of two polymorphisms in a same genetic block is combined, of this variant and an Arg25Gln variant in exon 1, is denoted as **2*. The third variant is the **7* (Arg101stop) leading

Table 1. Cytochrome P450 gene variants

						Minor allele frequencies		
Gene	**Variant**	**RS number**	**Nucleotide change**	**Position**	**Effect on amino acid level**	**Caucasian**	**Asian**	**African**
CYP1A1				15q22-24				
	**2A*	rs 4646903	T 3801 C (MspI) [a]	exon 7 [b]	None	2 - 3 %	40 - 80 %	-
	**2B*	combination	T 3801 C (MspI) [a] + A 2455 G	exon 7 [b]	Ile 462 Val	NA	NA	NA
	**2C*	rs 1048943	A 2455 G	exon 7 [b]	Ile 462 Val	10 – 20 %	10 - 20 %	0 %
	-	rs 2606345	A –1738 C	5'UTR, intron 1	None	-	-	-
	-	-	T 6235 C	exon 7 [b]	None	-	-	-
	**4*	rs 1799814	C 2453 A (M4) [a]	exon 7 [b]	Thr 461 Asn	7 %	0 %	0 %
CYP1A2				15q22-24				
	**1F*	rs 762551	C –163 A	5'UTR	None	30 %	30 %	40 %
CYP1B1				2p22-p21				
	**2*	rs 1056827	C 142 G (m1) [a]	exon 1	Arg 48 Gly	-	-	-
		rs 10012	G 355 T (m2) [a]	exon 1	Ala 119 Ser	-	-	-
	**3*	rs 1056836	C 4326 G	exon 3	Leu 432 Val	-	-	-
CYP2A6				19q13.2				
	**4*	rs 8192789	NA	NA	Deletion CYP2A6	0.5 – 1 %	7 – 22 %	-
CYP2A13				19q13.2				
	**2*	rs 8192784	G 74 A	exon 1	Arg 25 Gln	-	-	-
		rs 8192789	C 3375 T	exon 5	Arg 257 Cys	-	-	-
	**7*	-	C 578 T	exon 2	Arg 101 stop	5 %	3 %	-
CYP2C9				10q24				
	**2*	rs 1799853	C 3608 T	exon 3	Arg 144 Cys	10 – 15 %	0 %	0 %
	**3*	rs 1057910	A 42614 C	exon 7	Ile 359 Leu	4 – 10 %	4 – 7 %	1 – 2 %
CYP2C19				10q24.1-q24.3				
	**2 (PM)*	rs 4244285	G 19154 A (M1) [a]	exon 5	Pro 227 Pro; Splicing defect	15 %	35 %	17 %
	**3 (PM)*	rs 4986893	G 17948 A (M2) [a]	exon 4	Trp 212 Ter	15 %	35 %	17 %
CYP2D6				22q13.1				
	**3 (PM)*	-	A 2549 Del	-	Frameshift	1-2 %	-	-
	**4 (PM)*	rs 1800716	G 1846 A	exon 4	Splicing defect	20 %	-	-
	**6 (PM)*	rs 5030655	T 1707 Del	exon 3	Trp 152 Frameshift	3 %	0 %	0 %
CYP2E1				10q24.3-qter				
	**1C/*1D*	-	8 tandem repeats [–2173 - –1946 bp]	5'UTR	None	-	-	-
	**5B*	rs 3813867	G –1293 C (PstI) [a]	5'UTR	None	-	-	-
		rs 2031920	C –1053 T (RsaI) [a]	5'UTR	None	-	-	-
	**6*	-	T 7632 / 7766 A (DraI) [a]	intron 6	None	-	-	-
	**7B*	rs 6413420 / rs 2070673	G –71 T / T –333 A	exon 1	None	-	-	-
CYP3A4				7q22.1				
	**1B*	rs 2740574	A –392 G	5'UTR	None	2 – 9 %	0 %	35 – 67 %
CYP3A5				7q22.1				
	**3*	rs 776746	A 6986 G	intron 3	Splicing defect	20 %	20 – 70 %	70 %
CYP3A43				7q22.1				
	**3*	rs 680055	C 36867 G	exon 10	Pro 340 Ala	96 %	100 %	61 %
CYP11A1				15q23-q24				
	-	-	$[TTTTA]_4$ deletion	-	None	-	-	-
CYP17				10q24.3				
	-	-	T 1931 C (MspI) [a]	promoter region	None	-	-	-
CYP19				15q21.1				
	-	-	$[TTTA]_{7-12}$ repeats	exon 6	None	-	-	-
	-	rs 28757190	C > T	exon 1	Arg 264 Cys	-	-	-
	-	rs 2236722	T > C	exon 10	Trp39Arg	-	-	-
	-	-	C > T	-	None	-	-	-

- = undefined, 5'UTR = 5' untranslated region, NA = not applicable. [a] The PCR-RFLP alias can be found between brackets. [b] Exon 6 in NCBI dbSNP.

to a nonsense mutation. Individuals homozygous for this nonsense mutation would not have a functional *CYP2A13* protein and were therefore expected to have a reduced sensitivity to xenobiotic toxicity resulting from *CYP2A13*-mediated metabolic activation.[80] This polymorphism was first proposed to be unique in Chinese individuals [80], but was later identified in French Caucasians as well [69].

CYP2C9 (chr. 10q24) and ***CYP2C19*** (chr. 10q24.1-q24.3) are both major drug-metabolizing enzymes. Additionally, *CYP2C19* was found to be involved in tobacco metabolism, as stated above. Nevertheless, their endogenous substrate has not yet been identified. The most relevant *CYP2C9* genetic variants are those that result in decreased enzymatic activity (3808T = **2* = Cys^{144} and 42614C = **3* = Leu^{359}). Both variants are mainly present in Caucasians with allele frequencies of 10-15% (**2*) and 4-10% (**3*). The **2* variant seems absent in persons from Asian or African descent, while **3* can be found in some Asian (4-7%) and African (1-2%) individuals.[81] Of the *CYP2C19* gene, several variants were found that resulted in absence of the enzyme effect, but only the common alleles **2* (19154A; splicing defect) and **3* (17948A; stop codon) were studied in respect to cancer.[82] These variants are most common in Asians (35%) and less in Africans (17%) or Caucasians (15%).[83]

CYP2D6 (chr. 22q13.1) takes a dominant place in drug-metabolism. This enzyme has special impact on drugs for the treatment of psychiatric and cardiovascular diseases and plays a role in the metabolism of tamoxifen and its adverse effects.[34] The *CYP2D6* gene has been intensively studied, resulting in a large number (> 40) of identified variant alleles. This polymorphic background provides a profound basis for phenotypic differences. Those with poor (absent or reduced) metabolism were studied most frequently. These are **3* (A2549del), **4* (G1846A) and **6* (T1707del). No studies explicitly report effects of ultra rapid metabolizers but some found results for those without a phenotypic poor metabolism. Overall, non-functional alleles in which group the poor metabolizers are embedded are present in 26, 50 and 65 percent of persons of Caucasian, Asian and African/African-American descent respectively.[65] The allele frequency in Caucasians for the specific **3*, **4* and **6* variants are 1-2 %, 20 % and 3 % respectively.[79] For Caucasians, the **4* variant is carried by 75% of all poor metabolizers.[79]

Besides their role in the tobacco metabolism in which they form primarily relatively short alkyl chains [68], ***CYP2E1*** (chr. 10q24.3-qtr) is responsible for the metabolism and activation of several other low-molecular-weight substances.[34] Physiological substrates are fatty acids [84] and gluconeogenetic precursors such as acetone and acetol.[85] *CYP2E1* is a major component of the microsomal ethanol-oxidizing system as well and is inducible by chronic alcohol consumption. It oxidizes ethanol to acetaldehyde and is involved in the metabolic activation of various carcinogens. Acetaldehyde has been suspected to play a pivotal role in the development of alcohol-related cancers because of its established carcinogenicity in laboratory animals.[41] Primarily three polymorphisms have been studied. Two were found in the 5'-regulatory region (**5*; G1293C, *PstI* and C1053T, *RsaI*). The third one was detected with *DraI* (**6*; T7632A). In Asian persons the variant alleles have been associated with enhanced enzyme activity, while in

Caucasian subjects the activity was lowered.[86] Interindividual variation in susceptibility seems to be affected by induction, for example by alcohol (**1D*; 8 repeats of 42 bp in the 5′ flanking region).[87] Allele frequencies are present for some populations and variants, but not all studies mention dbSNP reference numbers (rs), which makes comparisons difficult.[59]

Of particular importance is the ***CYP3A4*** (chr. 7q22.1) gene that encodes for the predominant expressed cytochrome P450 in the human liver. It is related to the metabolism of numerous compounds including hormones (catalyzes 6-beta-hydroxylation) and approximately 60% of drugs.[62] Several variants have been identified in this gene, but only one was frequently studied in respect to carcinogenesis. The -392G variant of this **1B* polymorphism, originally named *CYP3A4-V*, results in a 1.5 to 2-fold higher promoter activity in vitro.[88] Originally it was proposed that this variant decreases the oxidative deactivation of testosterone in vivo [89], but the increased activity in vitro does not support this theory. This variant seems to be rare in Asian individuals and is most frequent in African Americans (35-67%). 2-9% of Caucasians were found to be carrier of the variant allele.[65]

A family member that is genetically located 76983 base pairs near *CYP3A4* on chromosome 7 is ***CYP3A5*** (chr. 7q22.1). Expression is influenced by a gene polymorphism (**3*; A6986G) in intron 3. The 6986G variant results in a splicing defect through which there is absence of the expression of the enzyme. This consequently leads to a severely decreased enzyme activity.[62] The expression of the polymorphism varies widely between ethnicities, with 70% in the African American population and only 20% in Caucasians.[65] Allele frequencies among Asian individuals are in-between.[59] It was shown that this polymorphism is in close linkage with the *CYP3A4*1B* and results found for this latter variant might therefore be explained by the altered functionality in *CYP3A5*.[90]

To the other side of *CYP3A4* on chromosome 7, another family member is located within 43828 base pairs. The ***CYP3A43*** gene consist of 13 exons, with in exon 10 the most important variation until now. The amino acid sequence is identical for 75 % with *CYP3A4* and *CYP3A5*. Its expression is low (0.1 – 0.2%) and the contribution to *CYP3A* mediated metabolism is therefore thought to be minimal. Only one variant has been reported to be of potential interest. The 36867G variant is present in the majority of individuals with allele frequencies of 96 to 100% in Caucasians and Asians but only 61 % of African-American individuals are carrier. No functionalities are known.

Finally, cytochrome P450 families active in hormone synthesis and metabolism have been studied in cancer development. ***CYP11A*** (chr. 15q23-q24) catalyzes cholesterol-side chain cleavage to form pregnenolone out of cholesterol. ***CYP17*** (steroid 17-alpha-hydroxylase) (chr. 10q24.3) mediates both 17-alpha-hydroxylase and 17,20-lyase activity resulting in glucocorticoids and sex steroid formation respectively. A T to C transition in the 5` promotor region of the gene leads to an additional Sp-1-type (CCACC box) promotor site. This is proposed to result in an increased rate of transcription and consequently a higher level of circulating testosterone. Androgens are converted into estrogens by one of the most essential enzymes in

estrogen biosynthesis, aromatase (***CYP19***) (chr. 15q21.1). This enzyme catalyzes the reactions of androstenedione into estrone and testosterone into estradiol and is considered as the most important source of estrogens in post-menopausal women.[91] It is also referred as estrogen synthetase, and plays a role in the rate limiting transformation step of C19 androgens to C18 estrogens.[62] Polymorphisms in these genes found to be related to cancer development are scarce considering their potential impact. Functionality is only known for $[TTTA]_n$ repeats and a C to T substitution in exon 10, both in the *CYP19* gene and resulting in a higher activity of the enzyme. The exon 10 polymorphism and two others on *CYP19* (Trp39Arg and Arg264Cys) are believed to be in close linkage with the $[TTTA]_n$ repeats. Additionally, variants of hormone metabolizing genes are not properly defined or validated which makes it difficult to gather information about functionality and allele frequencies.

RESULTS

Lung cancer

Although the risk of lung cancer is highly related to tobacco use (90% of all lung cancer patients have a history of smoking), less than 20% of smokers develop the disease. This discrepancy has also been found for individuals with an equal number of pack-years history of smoking, indicating an important genetic factor in the etiology.[45] A number of candidate CYP gene variants have been associated with lung cancer. They all play a central role in the activation of tobacco-related carcinogens. NNK is the *N*-nitrosamine with the strongest association with lung cancer.

Although not all studies found a significant association between lung cancer and *CYP1A1*2* variants [46,51-55], clear evidence exists for an association with at least one of them. The Val[462] (**2C*) is frequently associated with an increased risk of lung cancer, also in individuals without a history of tobacco use [45,46,92-97] and across several populations (see TABLE 2). The association was demonstrated for squamous cell carcinoma as well as for adenocarcinoma of the lung. The **2A* 3801C was first considered to be responsible for the association, but is now believed to be in linkage disequilibrium with the Val[462] allele in at least some populations.[54] Most of the studies found a 2- to 3-fold risk increase for the separate **2* variants, although higher risks were found for the combined variant (**2B*).[45,93] In one study among 217 Chinese cases and 404 controls, an extremely increased risk was observed of nearly 30-fold for the synergistic effect of **2B* with tobacco use on squamous cell carcinoma incidence.[45] Although the study size seems to be sufficient, these risk ratios raise questions, since only a part of tobacco metabolism is mediated by *CYP1A1*. Moreover, a Japanese study [93] found an odds ratio of seven and no synergism was found among women of Chinese descent [95]. It remains unclear how these results should be interpreted. Additionally, the effect of *CYP1A1* genotype on lung cancer was frequently studied along with the genetic effects of GSTM1 (a phase II enzyme).[45,47,93]

Table 2. Association of cytochrome P450 gene polymorphisms and cancer

CYP	Variant	Substrate	OR [95% CI]	Cases / Controls	Population	HWE	Reference
Lung cancer							
Cyp1A1	**2A*	none	2.1 [1.2-3.7]	207 / 283	US Caucasians	-	94
		tobacco [a]	3.3 [1.4-7.6]	80 / 160	Japanese	0.85	92
	**2B*	tobacco	7.3 [2.1-25.1]	85 / 170	Japanese	0.83 / 0.84	93
		tobacco	29.9 [7.0-127.3]	217 / 404	Chinese	0.27 / 0.0004	45
	**2C*	none	OR > 1.0; p = 0.02	55 / 60	Turkish	-	97
		tobacco	2.9	108 / 99	Brazilians	0.16	96
		tobacco [a]	3.2 [1.0-10.3]	80 / 160	Japanese	1.00	92
		no tobacco	3.7 [1.9-7.3]	200 / 144	Chinese ♀♀	0.24	95
Cyp2A6	**4*	none	X^2 = 0.2; p=0.03	257 / 154	Japanese	-	68
		none [a]	0.3 [0.1-0.8]	492 / 402	Japanese	0.92	99
		tobacco	8.1 [3.2- 20.0]	151 / 326	Chinese	0.07	101
		tobacco [a]	0.2 [0.1-0.7]	1094 / 611	Japanese	0.26	74
Cyp2A13	**2*	tobacco	0.2 [0.1-0.7]	724 / 791	Chinese	0.38	100
	**7*	none	9.9 [1.9-52.2]	204 / 201	French	0.89	69
Cyp2C19	**2/*3 (PM)*	none [a]	3.2 [1.5-6.9]	212 / 372	Chinese	-	70
Cyp2E1	**5*	none	OR < 1.0; p < 0.05	195 / 206	Swedish	0.36	108
		none	0.5 [0.3-0.8]	164 / 181	Chinese	0.25	111
		none [a]	0.1 [0.0-1.0]	119 / 231	Taiwanese	<0.0001	110
		tobacco	OR < 1.0; p < 0.05	137 / 206	US Africans / Mexicans	-	109
	**6*	none	OR > 1.0; p < 0.05	74 / 73	Japanese	-	71
		tobacco [a]	OR = 22.7, p < 0.05	126 / 193	US Africans / Mexicans	-	107
Cyp3A4	**1B*	tobacco [a]	8.0 [2.1-30.3]	801 / 432	German ♀♀	0.27	72
Cyp3A5	**1*	none	1.5 [1.0-2.1]	133 / 270	Taiwanese	-	73
Esophageal cancer							
Cyp1A1	**1*	none	2.5 p < 0.05	62 / 38	North Chinese	0.19	35
	**2B*	none [a]	2.5 [1.2-5.3]	146 / 324	Taiwanese	0.46	112
		tobacco [a]	4.8 [1.6-14.8]	127 / 101	Chinese	0.67	113
		tobacco [a]	6.2 [2.1-20.1]	127 / 101	Chinese	0.67	114
Cyp2A6	**4*	tobacco	2.1 [1.0-4.5]	149 / 326	Chinese	0.07	101

Table 2. Continued

CYP	Variant	Substrate	OR [95% CI]	Cases / Controls	Population	HWE	Reference
Hepatocellular cancer							
Cyp1A1	**2*	tobacco/HBV	3.2 [1.2-8.3]	81 / 409	Taiwanese	0.83	137
Cyp2A6	**4*	none [a]	3.4 ($p < 0.05$)	44 / total population	Japanese	-	139
Cyp2D6	*EM*	none	6.4 [2.4-17.5]	75 / 200	Spanish	-	140
	**3 / *4 / *6 (PM)*	none [a]	0.1 [0.0-0.5]	91 / 90	Italian	0.21	138
Cyp2E1	**1*	tobacco [a]	OR > 1.0 $p = 0.001$	30 / 150	Taiwanese	-	141
	**5B*	none	5.8 [1.2-27.4]	78 / 138	Japanese	-	142
Breast cancer							
Cyp1A1	**2A*	none [a]	9.7 [2.0-47.9]	51 / 269	African – Americans ♀	-	147
		estrogens [a]	2.0 [1.0-4.0]	150 / 150	Chinese♀	0.06	148
		none	0.3 [0.1-0.8]	128 / 256	Brazilian♀	-	150
		none	OR < 1.0 $p = 0.004$	396 / 1936	Austrian ♀	0.18	146
		tobacco [a]	5.7 [1.5-21.3]	466 / 466	Caucasian♀	-	149
	-	none	0.6 [0.4-0.9]	195 / 272	Japanese♀	-	152
	**2C*	none	0.7 [0.5-1.0]	195 / 272	Japanese♀	-	152
		tobacco	5.2 [1.2-23.6]	216 / 282	Caucasian♀	-	151
		none	OR > 1.0 $p = 0.03$	396 / 1936	Austrian♀	-	146
		tobacco [a]	3.6 [1.1-11.7]	466 / 466	Caucasian♀	-	149
Cyp1B1	**3*	estrogens [a]	3.1 [1.0-9.1]	186 / 200	Shanghais♀	0.00009	153
		none	2.3 [1.3-4.3]	84 / 103	Turkish♀	0.03	154
Cyp17	-	estrogens [a]	2.0 [1.1-3.5]	109 / 117	Swedish♀	0.30	155
		none [a]	1.6 [1.0-2.6]	369 / 284	Australian♀	0.41	156
		none	2.5 [1.1-5.9]	174 / 285	Mixed♀	-	157
		none [a]	1.8 [1.1-3.1]	239 / 195	Japanese♀	0.21	158
Cyp19	*repeat*	none	2.4 [1.0-5.8]	182 (185) [b] / 252	Norwegian /Swedish ♀	-	159
		none	OR > 1.0 $p = 0.005$	462 / 618	Caucasian ♀	-	160
		none	OR > 1.0 $p = 0.028$	327 / 253	Caucasian ♀	-	161
		none	OR > 1.0 $p = 0.012$	327 / 253	Caucasian ♀	-	161
	-	none	0.4 [0.2-0.9]	204 / 200	Japanese♀	-	42
	-	alcohol	1.5 [1.1-2.2]	389 / 346	Korean ♀	-	162
	-	none [a]	1.5 [1.0-2.2]	481 / 236	Norwegian /Swedish ♀	0.03	163
	-	none	1.3 [1.1-1.5]	1355 / 2580	Mixed ♀	0.65	145

Table 2. Continued

CYP	Variant	Substrate	OR [95% CI]	Cases / Controls	Population	HWE	Reference
Prostate cancer							
Cyp1A1	**2B / *2C*	none	2.4 [1.7-3.3]	81 / 105	Japanese ♂	0.009 / 0.33	170
	**2C*	none [a]	2.4 [1.0-5.6]	115 / 200	Japanese ♂	0.19	171
Cyp1B1	**1*	none	OR >1.0 p = 0.009	245 (159) [b] / 222	Caucasian ♂	-	165
	**2*	none [a]	4.0 [1.7-9.4]	117 / 200	Japanese ♂	0.01 / 0.0003	164
Cyp2D6	**3 / *4 / *6 (PM)*	tobacco [a]	3.1 [1.1-8.9]	153 / 359	Danish ♂	-	38
Cyp3A4	**1B*	none	9.5 [2.5-35.2]	230 / 0 (only cases)	Caucasian ♂	-	89
	**1B*	none [a]	2.4 [1.1-5.4]	101 / (only cases)	African-American ♂	0.009	175
	**1B*	none	6.3 [2.7-17.3]		African-American ♂	0.35	174
	**1B*	none	4.1 [1.3-12.2]	84 / 136	African-American ♂	0.002	182
	**1B*	none	2.3 [1.1-4.5]	215 / 94	Caucasian ♂	< 0.0001	182
	**1B*	none	1.9 [1.0-3.6]	440 / 480	Caucasian ♂	0.77	90
	**1B*	none	0.2 [0.1-0.7]	440 / 480	Mixed ♂	-	176
	**1B*	none	0.5 [0.3-0.9]	622 / 396	Caucasian ♂	0.000014	173
Cyp3A5	**3*	none	0.4 [0.2-0.8]	440 / 480	Mixed ♂	0.00005	90
	**1*	none	0.2 [0.1-0.8]	260 / 212	Japanese ♂	0.84	177
CYP3A43	**3*	none	5.9 [1.1-31.2]	622 / 396	Caucasian ♂	0.61	173
	**3*	tobacco [a]	8.6 [1.1-87.0]	341 / 306	Mixed ♂	0.90	178
Cyp11A1	*[TTTTA] deletion*	none [a]	1.8 [1.1-3.0]	278 / 299	Japanese ♂	-	43
Cyp17	*A1*	none [a]	1.6 [1.0-2.5]	178 / 160	Swedish ♂	-	167
		none [a]	2.6 [1.4-4.8]	252 / 131	Japanese ♂	0.54	183
		none	2.5 [1.6-4.0]	384 / 360	Italian ♂	0.06	364
	A2	none [a]	2.8 [1.0-7.8]	63 / 126	Austrian ♂	0.09	180
		none	1.7 [1.0-3.0]	96 / 159	US Caucasian ♂	0.81	179
		none	2.4 [1.0-5.5]	101 / 200	Japanese ♂	0.002	181
		none	2.8 [1.0-7.4]	71 / 111	African American ♂	0.93	182
Cyp19	*4 [TTTA] repeat*	none [a]	1.8 [1.0-3.1]	99 [b] / 116	Japanese ♂	-	44
		none	1.6 [1.0-2.5]	226 / 156	French Caucasian ♂	-	186
		none	1.4 [1.0-2.0]	226 / 156	French Caucasian ♂	-	186
	UD	none	1.8 [1.0-3.1]	101 / 114	Japanese ♂	0.15	169
Head and neck cancer							
Cyp1B1	**2*	tobacco [a]	4.5 [2.6-8.0]	312 / 300	German	0.42	187
Cyp2E1	**7B*	tobacco	2.7 [1.5-4.9]	312 / 299	German	0.51	188
Cyp1B1	**2*	none [a]	3.0 [1.6-5.8]	211 / 200	Japanese	0.01 / 0.0003	189
	**3*	none [a]	2.2 [1.2-4.1]	211 / 200	Japanese	0.002	189

Table 2. Continued

CYP	Variant	Substrate	OR [95% CI]	Cases / Controls	Population	HWE	Reference
Renal cancer							
Cyp2C19	**2 / *3 (PM)*	none [a]	0.3 [0.1-0.7]	112 / 372	Chinese	-	70
Bladder cancer							
Cyp3A4	**1B*	epipodophylo.	0.1 [0.0-0.9]	129 / 0 (only cases)	Caucasian	-	39
Hematological cancer							
Cyp1B1	**2*	none [a]	3.3 [1.4-8.0]	113 / 202	Japanese	0.01 / 0.02	172
	**3*	none [a]	2.5 [1.1-5.7]	113 / 202	Japanese	0.02	172
Gynecological cancer							
Cyp17	*A2*	none	1.9 [1.3-2.8]	563 / 523	US Mixed	0.08	190
Cyp1A1	**1*	none [a]	2.9 p=0.0386	75 / 345	Caucasians	-	191
Dermatological cancer							
Cyp2D6	*EM*	none	15.5 [1.3-178.5]	926 / 0 (only cases)	English	-	40
	**4*	none [a]	2.2 [1.2-3.9]	-	English	-	192

OR = odds ratio, CI = confidence interval, HWE = Hardy Weinberg Equilibrium, EM = extensive metabolizer, PM = poor metabolizer. Only significant association studies are included in the table to avoid disarrangement. When synergy with substrates was studied, ORs of this interaction are presented. When authors studied polymorphisms as possible independent risk factors, no substrates are mentioned (none). [a] Results only found in homozygous carriers. [b] Hereditary cases.

Table 3. Association of cytochrome P450 gene and type of cancer

Cytochrome P450 gene	Lung	Esophagus	Stomach	Colorectal	Liver	Breast	Prostate	Head & neck	Kidney	Bladder	Blood	Female productive organs	Skin	No. of studies [a]
Cyp1A1	+	+	+	+	+	+	+	+	-	-	-	-	+	30
Cyp1A2	-	-	-	+	-	-	-	-	-	-	-	-	-	1
Cyp1B1	-	-	-	+	-	+	+	+	+	-	-	+	-	9
Cyp2A6	+	+	+	-	+	-	-	-	-	-	-	-	-	7
Cyp2A13	+	-	-	-	-	-	-	-	-	-	-	-	-	2
Cyp2C9	-	-	-	+	-	-	-	-	-	-	-	-	-	3
Cyp2C19	+	+	+	+	-	-	-	-	-	+	-	-	-	5
Cyp2D6	-	-	-	-	+	-	+	-	-	-	-	-	+	5
Cyp2E1	+	+	+	+	+	-	-	+	-	-	-	-	-	18
Cyp3A4	+	-	-	-	-	-	+	-	-	-	+	-	-	10
Cyp3A5	+	+	-	-	-	-	+	-	-	-	-	-	-	4
Cyp3A43	-	-	-	-	-	-	+	-	-	-	-	-	-	2
Cyp11A1	-	-	-	-	-	-	+	-	-	-	-	-	-	1
Cyp17	-	-	-	-	-	+	+	-	-	-	-	+	-	12
Cyp19	-	-	-	-	-	+	+	-	-	-	-	-	-	12

[a] Studies with significant associations

Despite its important role in tobacco metabolism, deletion of the *CYP2A6* gene (**4*) has a controversial association with lung cancer. Most studies, except one [98], were performed among Asian persons and both individual gene as well as gene-smoking associations were investigated. Decreased [68,74,99], equal [98,100] and increased risks [101] of lung cancer were found. An increased risk was found among Chinese and decreased risks among Japanese. As previously mentioned, allele frequencies differ between populations, even within Asia. This could explain the discrepancies found. Nevertheless, in populations in which the deletion genotype is common, lung cancer risk might be influenced by an altered metabolism of both tobacco and probably other compounds.

To our knowledge, there are only few studies on the association between *CYP2A13* polymorphisms and lung cancer, although it is one of the most important cytochrome P450 enzymes for the metabolic activation of the tobacco-specific carcinogen NNK. Both a 10-fold increased risk of small cell lung cancer [69] and a decreased risk of adenocarcinoma of the lung, mainly in those with moderate cigarette use [100], were found. Although these associations were detected for different variants (**7* and **2*, respectively), the functionality of these polymorphisms was assumed to be similar. These studies were not only dissimilar in exposure variant, however, but also differences in ethnicity and power may have played a part in the inconsistent findings.

A single study reported that homozygous *CYP2C19* 'poor metabolizer' carriers had a more than 3-fold increased risk of lung cancer. [70] Whether this association is genuine needs to be investigated.

Half of the studies in which relations between the **5* and **6* alleles of *CYP2E1* and lung cancer were investigated did not find any association at all.[41,86,96,102-106] In the remaining studies the **6* polymorphism was associated with an increased risk of lung carcinoma.[71,107] On the other hand, the **5B* polymorphism seems to be associated with a decreased risk of lung cancer.[108-111] *CYP2E1* seems to modify the effects of smoking on lung cancer, even though only a minority of studies considered the relation.[109]

Among German individuals (801 cases and 432 controls) a 2.25-fold increased risk of small cell lung cancer in homozygous carriers of *CYP3A4*1B* was found. Women homozygous for *CYP3A4*1B* showed a 3-fold [95% CI: 0.94 – 9.90; p=0.06] but non-significantly increased risk. This risk increase was not seen in men unless they had smoked for more than 20 pack years (OR 3.42; 95% CI: 1.65 – 7.14). In women with a smoking history, the risk was even higher (OR 8.00; 95% CI: 2.12 – 30.30) (TABLE 2).[72] *CYP3A5*, which is genetically situated near *CYP3A4* on chromosome 7, has also been associated with lung carcinoma. An Asian study of 133 cases and 270 controls discovered a higher allelic frequency in Taiwanese lung cancer patients of *CYP3A5*1* than in controls (OR 1.49; 95% CI: 1.04- 2.13).[73] Both *CYP3A4*1B* and *CYP3A5*1* are active forms. While evidence is still limited to these two studies, the fact that the associations were observed in different ethnic groups (Caucasians and Taiwanese) supports an involvement of *CYP3A* in lung cancer.

In conclusion, many different CYP isoenzymes have been studied in association with lung cancer, all of which are involved in the metabolism of tobacco pro-carcinogenic compounds. The effects of *CYP1A1* polymorphisms were most intensively studied and the risk increases with the **2* variants are more or less established. Even though *CYP2A6* is highly relevant for the metabolism of nicotine, polymorphisms in this gene are not consistently related with a specific risk pattern of malignancies of the lung until now. Although genes coding for enzymes concerned in the metabolism of tobacco were considered, not all studies investigated a synergistic effect of the gene with tobacco use. In studies in which this effect was considered, the particular CYP enzyme regularly modified smoking.

Esophageal cancer

Esophageal cancer is mostly associated with excessive intake of alcohol and other causes of mucosal damage. Like lung carcinoma, esophageal cancer is often related to smoking as well although the risk increase seems to be smaller than for lung cancer.

While all studies were performed in Asian individuals, contradictory results were found for the association between *CYP1A1* polymorphisms and esophageal cancer risk. Although several studies found a significant 2 to 3-fold risk increase among homozygous *CYP1A1*2B* carriers [112-114], others could not demonstrate an association [115,116] or even found a decreased risk in persons with variant alleles [35]. However, the latter study was small and studied a specific subgroup of north Chinese individuals, which could have led to spurious results.

The same research group that identified a higher risk of lung cancer in carriers of *CYP2A6*4,* found an increased risk of esophageal cancer in carriers of this variant who did not smoke. (OR 2.1; 95% CI: 1.0 – 4.5).[101] Although previous in vitro studies support the hypothesis that deletion of the gene leads to higher concentrations of active nicotine in the blood, this is the only association study.

Next to lung cancer risk, poor metabolizers of *CYP2C19* have been associated with a more than 3-fold increased risk of esophageal cancer in the same group.[70] This was, however, the only study found for this association.

Although one study among Chinese Kazakhs observed an 11-fold risk increase of esophageal cancer [117], the wild-type of *CYP2E1* (**1/*1*) has been associated with a 3- to 5-fold risk increase of esophageal cancer in other Chinese individuals.[118,119] In Japanese subjects, tandem repeats in the 5'flanking region of *CYP2E1* gene were associated with an increased risk of esophageal cancer.[41] The *CYP2E1* variant (**6*) was also associated with an increased risk of this type of cancer in South-African subjects [120] but the last two results have not yet been reproduced. All results established with the wild-type genotype were found without consideration of alcohol use or other potential substrates that would have made the pathophysiological mechanisms underlying this association more understandable.

*CYP3A5*3* was associated with an almost 2-fold lower risk of esophageal cancer among South-African subjects of mixed ancestry. When gene-environmental interaction with to-

bacco and alcohol use were considered, persons with the wild-type genotype were found to be twice as much susceptible for esophageal cancer both with tobacco and with alcohol use than those with this genotype who did not use one of these substrates.[121] Even though these results support the hypothesis, study size is rather small and results could not be reproduced in black South-Africans.

In conclusion, although not as consistent as with lung carcinoma, the *CYP1A1* Val^{462} seems associated with an increased risk of esophageal cancer. Also, people with the wild-type genotype of *CYP2E1* seem to have an increased risk. Most associations were found for patients with squamous cell carcinoma, possibly related to the strong etiologic association between alcohol and this type of cancer.

Gastric cancer

Important risk indicators for gastric cancer are alcohol use, smoking and dietary compounds such as salt and fruit.[122] Several CYP enzymes are involved in the metabolism of these substrates. Therefore, CYP gene polymorphisms are considered important in the interindividual variance in effect modifiers for the associations with gastric cancer.

Besides the fact that *CYP1A1* gene polymorphisms were not as intensively studied in gastric cancer as in lung- and esophageal cancer, results for **2* point suggest a reduced cancer risk. One Chinese cohort study that identified 90 gastric cancers out of 275 high risk participants found a risk reduction of 50% in Chinese individuals carrying the **2A* variant.[123] This cohort consisted of gastric mucosal dysplasia persons who are at increased risk for gastric cancer. It is possible, however, that genetic trait effects were overwhelmed by other stronger risk factors, diluting the results found for genotype differences.

A deletion of the *CYP2A6* gene has been associated with a 4-fold increased risk of tobacco-mediated gastric carcinoma, an association that seemed even stronger than with esophageal cancer.[124]

CYP2C19 genetic variants that result in a decreased activity (**2* and **3*) have not only been associated with an increased risk of lung cancers and esophageal cancers, but also with a 3-fold increased risk of gastric cancer.[70]

Both the wild-type [37,125], as well as the *CYP2E1*5* variant [126] were mentioned as a risk factor for gastric carcinoma in homozygous individuals, but these associations were not reproduced in another study.[127] However, one study focused on specific histological subtypes of gastric cancer [37], while others studied anatomical subtypes [125] or total gastric cancer incidence [126]. Additionally, exposures differed between study groups from alcohol [37], to smoking [125] as well as the independent role of *CYP2E1* polymorphisms [126].

In conclusion, there is yet little conclusive evidence for an association between cytochrome P450 polymorphisms and gastric cancer. Also associations in which specific substrates were considered and studies in non-Asian populations are sparse.

Colorectal cancer

Risk factors for sporadic (non-familiar) colorectal cancer (CRC) are multiple. Eating low amounts of red meat, three servings of vegetables a day and using multivitamins with folic acid have been associated with a lowered risk of CRC. Other determinants that seem to reduce the risk are oral contraceptives or post-menopausal hormones in women and the long-term use of aspirin.[122]

Although there might be several candidate genes in association with the above-mentioned preventive factors, only a few candidate genes were actually reported in the literature. *CYP1A1*2* was associated with an increased risk of sporadic CRC, both in smoking and non-smoking Caucasian individuals.[128,129] Although not significant, this variant was overrepresented among CRC cases in a Hungarian population as well.[130] Interestingly, the opposite result was found for the **4* polymorphism of the same gene, that is associated with an increased bioactivity too.[131] Moreover, a polymorphism in the promotor region of the gene for which functionality needs to be further studied was associated with a 50% reduced CRC incidence.[131] This Spanish group additionally reports increased risks of CRC for those with an intron 1 polymorphism of the *CYP1A2* gene (**1F*) and an exon 1 polymorphism of *CYP1B1*.[131] The **1F* polymorphism was also found associated with an increased risk of colorectal adenomas that are considered to be precancerous lesions.[132]

CYP2C9 gene variants corresponding with reduced enzyme activity were associated with CRC, although not consistent. Carriers of the **2* or **3* variant of the gene were found to be at a decreased risk of CRC, primarily of the proximal colon [133], although in one study, this could only be demonstrated for men.[133] In line with this, the wild-type genotype was associated with an increased risk of CRC.[134] However, a study among Spanish Caucasians found increased risks for those with the variant type.[131] Strange enough, they did not report this finding as statistically significant, as a result of which either one of these observations should be questioned. Two others were not able to find any association at all.[129,135] Although *CYP2C9* is important in the metabolism of NSAIDs and NSAID use is associated with a lower incidence of colorectal cancers, indications for a synergistic role of *CYP2C9* polymorphisms could not be found.[135] Further research should shed light on the effect of the slow metabolizing variants on colorectal cancer risk in non-Caucasians, the effect of interaction with non-steroidal anti-inflammatory drugs and whether this effect differs for proximal and distal located tumors.

Finally, two other CYP gene polymorphisms are associated with CRC risk. A 30% reduced risk was found in Caucasian carriers of an insufficient *CYP2C19* gene (**2*).[129] The combined variant of two polymorphisms in the untranslated region of *CYP2E1* on chromosome 10 (**2B*), related to increased risk of other cancers as well, was associated with an increased CRC risk among Hungarians.[130] However, this result could not be reproduced among Dutch Caucasians.[136] As functionality is not completely unravelled, these conflicting results are difficult to explain.[86]

Since colorectal cancer is considered as the final stage of the sequence from adenoma to carcinoma, studies on the association between cytochrome P450 gene polymorphisms and

adenoma gave similar results as for carcinoma. For colorectal cancer risk, it can be concluded that research on associations with CYP polymorphisms remained limited to some variation in the *CYP1A1* and *CYP2C9* gene that affect especially the proximal colon. Literature on gene-environmental interactions with substrates like NSAIDs or dietary features is still limited.

Hepatocellular cancer

In addition to excessive alcohol consumption, chronic infections with hepatitis B (HBV) and C viruses (HCV) are a major cause of hepatocellular carcinoma (HCC). Therefore, studies on the risk of cytochrome P450 polymorphisms and hepatocellular carcinoma are usually performed in hepatitis-infected patients, which may complicate the extrapolation of findings to the general population.

Once more, the *CYP1A1*2A* polymorphism was hypothesized to increase cancer risk. A more than 3-fold increased HCC risk was indeed found in Taiwanese hepatitis B virus carriers.[137] Nonetheless, a similar study in Italian hepatitis C virus patients could not reproduce this finding. [138] This could have been caused by the variety in allele frequencies. The **2A* variant allele was found in 40-80% of Asian individuals whereas it is only present in 2-3% of Caucasians.[57] Associations within low allele frequency areas can only be found when the association is strong enough, which is probably not the case in relations with single polymorphisms.

A Japanese study observed an increase in HCC risk for HCV positive homozygous carriers without a sufficient *CYP2A6* gene (**4*).[139] These results support the hypothesis of increased carcinogenesis but the study must be discussed for its design. It was a case-only study in which allele frequencies found in 44 HCV positive HCC patients were correlated with the frequency of the general healthy Japanese population. This could have led to spurious results. As the authors state themselves, further evaluation is required.

Spanish Caucasians who are extensive metabolizers of *CYP2D6* enzyme substrates have a 6 times increased risk of HCC.[140] In line with this finding, the poor metabolizers of this gene (**3*, **4* and **6*) had a 90% risk decrease of HCC.[138] Although this last study was performed in a population infected with hepatitis C virus, results suggest a genuine association.

Less consistent were results found for *CYP2E1*. Both the wild-type genotype [141] as the **5* variant that results in a higher inducibility [142], were associated with an increase risk of HCC. Both variants were also associated with an increased risk of gastric cancer. The wild-type was found to be the risk allele in Taiwanese and **5* in Japanese. This might indicate a variety in allele frequencies in these Asian countries, but other explanations could hold as well. Both risk increases found in Taiwanese were found in synergy with alcohol or tobacco. The risks found in Japanese were the individual results for the genotype. As we have seen earlier, the effects of the genetic traits could be overwhelmed by environmental factors at higher exposures. Another study that investigated the association of *CYP2E1* variants and HCC risk was unable to confirm this.[143]

In summary, poor metabolizers of *CYP2D6* seem to confer a reduced risk of hepatocellular carcinoma, especially in hepatitis C infected individuals but the overall evidence for an association between genetic polymorphisms and HCC risk remains restricted to subgroups of hepatitis infected patients.[144]

Breast cancer

Probably, many of the risk factors for breast cancer are directly or indirectly related to estrogen levels. Therefore, CYPs involved in the estrogen pathway are considered as important candidate genes for the susceptibility to breast carcinoma. Although there are several gene mutations that are associated with an increased risk of breast cancer, such as *BRCA-1*, *-2*, *CHK2*, *ATM*, and *STK11*, until now only a few CYP polymorphisms have been detected that seem to alter breast cancer risk.[42,145,146] Most of these studies did not study the interaction with estrogens. Only a few studies observed effects in subgroups of postmenopausal women, suggesting that the variant type acts as an effect modifier for the association between hormones and breast cancer in those with low levels of estrogens.

Contradictory results were found for *CYP1A1* variants that were candidate genes for the association with breast cancer due to their involvement in estradiol hydroxylation. Of the five studies on **2A*, all in different ethnical groups, three reported a positive association [147-149] and two a negative one [146,150]. The negative associations were found regardless of substrates such as estrogen or tobacco, two of the three positive ones considered the substrates, but the remainder found the highest risk increase. Dissimilarities in ethnicity, substrate consideration and study size could have resulted in these findings. When these observations are extended with the ones found for **2C*, resulting in an increased activity as well, more similarities are found with all positive associations [146,149,151] except for one [152]. This latter study was not only performed in Japanese, which makes it conceivable risk factors might also differ, but they additionally found another 3'noncoding region polymorphism (T6235C) that was associated with a reduced risk of breast cancer. This 3'noncoding variant was found to be associated, however, with a higher rate of lymph node metastasis. The study among Austrian individuals [146,149,151] found opposite results for **2A* and **2C*, which is not in line with the presumed hypothesis.

*CYP1B1*3* was associated with a 2-fold increased risk of breast cancer in Turkish and Shanghaian women.[153,154] Effect modification of the gene on estrogen levels was only studied specifically in the Shanghaian group and resulted in significant observations for homozygotes of the variant gene.[153] The Turkish study group postulated that women with high body mass indices have increased estrogen levels since estrogen is stored in fat tissue.[154] There may be a synergistic association of estrogen levels and the *CYP1B1* polymorphism on breast cancer risk.

A promoter variant of *CYP17* (T1931C), without known functionality, was repeatedly found to be associated with a 2-fold risk increase of fibroadenoma and carcinoma of the breasts.[155-158]

Results seem to be stable over ethnic groups. [TTTA] repeats of 8 or more in *CYP19* were consistently associated with an increased breast cancer risk as well, although these studies were all performed among Caucasians.[159-161] Several other polymorphisms in the aromatase gene (see TABLE 2) are associated with an altered breast cancer risk.[42,145,162,163] Most variant carriers are found to have an increased risk of cancer of the breast, but contradictory results have been found for a T to C substitution in exon 1.[42,162] The real effect of these polymorphisms is probably due to linkage disequilibrium with the previously mentioned [TTTA] repeats.[163]

Specific estrogen metabolizing genes are mostly associated with an increased breast cancer risk. Future studies should focus on specific substrate levels to understand more of estrogen metabolism by CYPs in the etiology of breast carcinoma.

Prostate cancer

Prostate cancer is considered to be the most important hormone-related cancer in men. Therefore, several CYP candidate genes were studied individually as well as in combination. They seem to be associated with the risk of prostate cancer under the influence of androgens.[43,44,89,90,164-169] Similar to studies on breast cancer, the synergistic effect of CYP polymorphisms and hormones requires more studies in subgroups of exposure for this particular cancer. CYPs are also frequently mentioned as tobacco compound metabolizers. As a consequence of the finding of an association between tobacco-induced CYPs and an increased risk of prostate carcinoma [38,170,171], a new tobacco-related tumor was found [172].

*CYP1A1*2C* conferred a more than 2-fold increased risk of prostate cancer in Japanese [170,171], although it remains unclear how many alleles must be affected for a significant effect [171]. Additive effects with the phase 2 *GSTM1* null variant (an allele that makes no gene product or whose product has no activity of any kind) were frequently found and this synergy may even increase prostate cancer risk in heterozygous *CYP1A1*2C* carriers.

Contradictory results were found for *CYP1B1* polymorphisms. Variant (**2*) [164] as well as wild-type genotype carriers [165] have been associated with an increased risk. The Caucasian study on **2* used the wild-type genotype as reference, however, while the other study was performed among Japanese and used a haplotype analysis. Hence, this contradiction needs further exploration.

In one study, poor metabolizing status of *CYP2D6* was associated with an increased risk of prostate cancer in 153 smoking Danish individuals (OR 3.10 [95% CI: 1.07 – 8.93]), but not in smoking Swedish men possibly due to low sample size or unknown coincidental factors.[38]

*CYP3A4*1B* was originally suggested to decrease the oxidative deactivation of testosterone, resulting in more dihydrotestosterone (DHT). DHT is the metabolite that induces prostate cancer. The variant was most common in African-Americans, who have the highest prostate cancer incidence in the world. Therefore, men with this variant were expected to have an increased prostate cancer risk. However, one study among 622 Caucasian prostate cancer patients found the opposite.[173] Others demonstrated that carriage of this variant type was

indeed associated with increased prostate cancer risk [174], or mainly in high stages (T3/T4) or Gleason scores ≥ 7.[89,90,175] In line with the finding that *CYP3A4*1B* increases high-grade prostate cancer was the decreased risk found for low-grade prostate cancer by another study.[176] This latter study used the same study population as Plummer et al.[90] This was mainly observed among relatively old men without any family history of prostate cancer. Studies were performed in Caucasian and African American men. It was argued that the *CYP3A4*, for which *in vivo* functionality lacks, is probably in linkage disequilibrium with a more functional variant. A candidate gene for this hypothesis was *CYP3A5*.[90] Persons with the **3* variant did not express the gene and therefore the protein, and seemed to have a decreased activity, possibly leading to increased testosterone levels.[62] In line with this hypothesis is an observation that *CYP3A5*3* decreases low-grade prostate cancer [90], although others found this association for *CYP3A5*1* [177]. The exact consequences to several substrate levels needs to be further evaluated to draw conclusions about prostate cancer risk. Subsequently, haplotype analyses were performed with combinations of the *CYP3A4* and *CYP3A5* variants. The combination of *CYP3A4*1B/CYP3A5*3* seemed to increase prostate cancer risk.[90] *CYP3A4*1B/CYP3A5*1* was both found to lower high [173] as well as low grade prostate cancer [90].

Close to the previous *CYP3A* families lies *CYP3A43* on chromosome 7, with the variant **3* of which the functionality needs to be proven yet. This polymorphism was more frequently expressed in prostate cancer patients with a familial background [173] or smokers [178].

A $[TTTTA]_4$ deletion of *CYP11A1* has been associated with an increased risk of metastatic prostate cancer (OR 1.79 [95% CI: 1.07 – 2.90]) and with a higher grade (Gleason score ≥ 8) or a more poorly differentiated tumor (OR 1.79 [95% CI: 1.08 – 2.97]). However, an increased risk of prostate carcinoma was not detected.[43]

Studies on *CYP17* gene variants, encoding for an enzyme of high interest in testosterone metabolism, are somewhat inconclusive. The majority of studies observed increased risks of prostate cancer for those who carried the *A2* variant of the gene, which is in line with the a priori hypothesis.[179-182] Others, however, observed increased risks for those with *A1* (wild-type) [167,183] and a large study of 590 cases and 782 controls could only report a borderline significant risk increase (OR 1.23 [95% CI: 0.99 – 1.54] in *A2* carriers [184]. A meta-analysis only revealed an increased risk of prostate cancer in *A2* carriers of African descent which was based on 3 studies.[185] Besides this, inconsistencies exist with respect to age at diagnosis. While two studies found the highest risk in youngest individuals [179,181], others observed this in older men [180] or no effect of age at all [183]. The study among African Americans additionally reported increased risks of high grade and stage prostate cancers in homozygous *A2* (C-allele) carriers (OR 7.1 [95% CI: 1.4 – 36.1].[182] One of the reasons for inconsistencies among studies is in the selection of controls. Besides healthy controls, also men with benign prostatic hyperplasia (BPH) were included.

Finally, prostate cancer risk has been associated with polymorphisms of *CYP19* (aromatase). Either short (7 or 8) repeats of 4[TTTTA] [44,186] as well as a C to T SNP in exon 7 [169], have been

proposed as risk factors for prostate cancer. Carriage of these variants was associated with an almost doubled risk of prostate cancer and the C to T mutation was also associated with a higher tumor grade (OR 2.59 [95% CI: 1.47 – 4.46]. In combination with either a polymorphism of the estrogen receptor (*ERα*) or a polymorphism of catechol-O-methyltransferase (*COMT*) an odds ratio of 3.00 [95% CI: 1.72 – 5.23] was found. When all three phase I and phase II polymorphisms were present in one individual they synergistically increased the risk to 6.30 [95% CI: 3.61 – 10.99].[169]

In conclusion, a variety of cytochrome P450 polymorphisms have been associated with an increased risk of prostate cancer risk of which polymorphisms of *CYP3A*, *CYP11A*, *CYP17* and *CYP19* seem to be most prominent. Future studies should focus on interactions between these subfamilies or a multilocus approach.

Other malignancies

Since only a limited number of studies on the association between cytochrome P450 polymorphisms and other malignancies have been performed, they are reviewed together.

Ninety percent of head and neck malignancies are squamous cell carcinoma that are generally considered tobacco-related. Two CYP enzymes that participate in tobacco-metabolism have been associated with an increased risk of head and neck malignancies, both in German individuals. First, persons with the *CYP1B1 *2/*2* polymorphism had an almost 5-times higher risk of such malignancies than those with the wild-type (**1/*1*), an observation which was accompanied by a 20 times higher likelihood of carrying mutations in the p53 suppressor gene.[187] Second, a study was performed demonstrating a 3-fold risk increase of head- and neck malignancies, in carriers of *CYP2E1*7B*. [188] The functionality of this last variant needs to be proven.

Until now, *CYP1B1* is the only cytochrome associated with renal cell carcinoma. One study in Japanese individuals found it to be increased in carriers of polymorphisms in the *CYP1B1* gene with decreased activity (**2* and **3*).[189] The same group also found evidence that these polymorphisms increased estrogen induced endometrium cancer.[172] People with one of the two variants and endometrium cancer also had an increased frequency of estrogen receptor and prostagen receptor positive status. As expected, an association between polymorphisms of the major estrogen metabolizing enzyme *CYP17* and ovarian cancer was found as well. Both homo- and heterozygous carriers of the *CYP17 MspI* variant (*A2*) had an increased risk, especially in women of 50 years and older and women with invasive serous cancer.[190]

Chinese individuals with *CYP2C19 *2* and **3*, exhibited a decreased risk of bladder cancer.[70] This study, however, requires confirmation in view of the low number of cases and the single ethnical group.

Since several years, it is known that anticancer drugs can induce leukaemia but only recently it was suggested that cancer patients with *CYP3A4*1B* who used anticancer drugs containing epipodophyllotoxins had a lower risk of developing a tumor-related leukaemia

than carriers of the wild-type version of the gene. Individuals with the wild-type genotype may have an increased production of potentially DNA-damaging reactive intermediates while the variant may be associated with decreased production of the epipodophyllotoxin catechol metabolite, the precursor of the potentially DNA-damaging quinone.[39]

For skin cancer, the wild-type of *CYP1A1* [191] and extensive metabolizers of *CYP2D6* [40] were both associated with basal cell carcinoma (BCC) incidence although the wild-type genotype of *CYP1A1* was associated with truncal BCCs most strongly. The *CYP2D6*4* variant was also associated with a doubled risk of malignant melanoma in homozygous individuals compared to non-carriers of this variant.[192]

In conclusion, for the malignancies mentioned in this paragraph, reliable association studies for cytochrome P450 are still limited. Since the few results that are found are consistent with the ones found for other tumors, future research seems to be valuable.

HARDY WEINBERG EQUILIBRIUM

It has been proposed to consistently include information about the Hardy Weinberg equilibrium (HWE) of study populations in all genetic association studies. Deviations from HWE might indicate inbreeding, population stratification, genotyping problems or can cause evidence for false associations. Therefore, to evaluate the results found in association studies, information about HWE must be provided. This test is commonly performed by a simple X^2-goodness-of-fit test that analyses differences between expected and observed allele frequencies, but also more profound methods are available. Authors should give *p*-values or allele frequencies per genotype in the study population or control subjects in order to validate their results. Unfortunately, not all studies point at this equilibrium. In our search 60 % (74 / 124) of the studies discussed HWE in some way (TABLE 2). For most of them, only allele frequencies were given. Out of this 60 %, one third was not in HWE when calculated on the basis of the numbers mentioned in the papers. Only a minority of studies mentioned this, but most of the others reported that their data were in HWE although this could not be reproduced with numbers provided in the paper.

DISCUSSION

Since the discovery of the cytochrome P450 enzyme family some 50 years ago, its important role in detoxification is increasingly recognized. Therefore, the hypothesis that variations in metabolic capacity might not only have consequences in the field of pharmacology but also aspects of carcinogenesis arose early. With the rapidly growing knowledge of genetic determinants resulting in a different protein structure or expression level of the numerous

cytochrome P450 enzymes, one might expect that the discovery of genetic polymorphisms associated with an increased or decreased risk of cancer will grow almost exponentially. However, current knowledge is still modest and contradictory. With the exception of tobacco-related cancers and polymorphisms in *CYP1A1*, *CYP2E1* genes, and hormone-related cancers and polymorphisms in *CYP17* and *CYP19* genes, there is yet little consistent evidence for an association with cancer. Furthermore, the proposed functionality is regularly in discordance with the results found. In view, however, of the hypothesized role of environmental toxins in carcinogenesis and the undisputed importance of the cytochrome P450 enzyme family, substantial findings are likely to lie ahead.

The development of cancer is a multifactorial process that probably starts with an initiation phase in which one or more cells develop permanent DNA damage, depending on the carcinogenicity and quantitative level of the causative exposure. If this damage is not repaired before a new division takes place, cell proliferation and DNA replication may be the first step towards a malignancy. It is likely that this does often not happen due to a large number of biological defence mechanisms. As soon, however, an uncontrolled growth and expansion occurs with complete loss of the original cellular structure and functions a tumor is considered as malignant, especially when metastasis occurs. In line with the complexity of cancer initiation and promotion, the activation and stimulation of the different phases is probably coordinated by a large number of substrates. According to this process, there are several important considerations regarding the association between cytochrome P450 gene polymorphisms and cancer development. **First**, it is unlikely that cytochrome P450 enzymes are a direct cause of cancer. It is more plausible to consider them as effect modifiers. Consequently, the likelihood of an association of a cytochrome P450 gene variant and cancer depends not only on the prevalence of a particular CYP gene polymorphism, but also on the exposure prevalence of the carcinogenic substance in the population. For endogenous substances such as estrogens and testosterone, the exposure level relates to gender. For exogenous substances of known toxicity, however, the exposure prevalence of the population will be low, because such substances will be avoided and interaction between a polymorphism and such a toxin on cancer in a population can hardly be demonstrated. Of course, psychotropic substances with a liability to abuse such as nicotine and alcohol will be frequently used despite known toxicity, and in such circumstances an adverse interaction between a CYP gene polymorphism and cancer is much more easily demonstrated. **Second**, studies done so far are not consistent with respect to the interaction of CYP polymorphisms and substrate levels. Whereas some authors focused on specific substrates and the interaction with CYP polymorphisms (CYPs as effect modifiers), others investigated the association between polymorphisms as independent risk factors for cancer. Results from these latter studies may underestimate the true risk because of misclassification if they are not stratified according to substrate exposure. **Third**, cytochrome P450 enzymes catalyze reactions that lead to an increase or decrease of substrate levels and their metabolites. A change in expression or activation of cytochrome P450 enzymes caused

by gene variants results in differences in substrate levels and might therefore alter disease risk. When the variant form works as an ultra-rapid metabolizer, more substrate will be metabolized per time unit. If this substance was a potential carcinogen before metabolisation, a decreased risk of cancer might be expected. If the substance, however, was a pro-carcinogen that is activated, an increased risk of cancer might be predicted. For poor metabolizers this works the other way around. Polymorphic expression of CYPs *in tissues* could also affect cancer development. This association was reviewed before and was therefore not considered here.[32,33] **Fourth**, one has to consider that CYPs originally play a role in the metabolism of endogenous and some exogenous substances and that their interaction with other exogenous compounds like tobacco and alcohol is rather a result of evolution than a primary function. And **fifth**, genetic trait effects might be overshadowed by other more important risk factors. Occasionally, results for the effect of CYPs were only significant in those that used low amounts of tobacco or alcohol or who were postmenopausal.

A potential explanation for the often contradictory results in this review is the large variation in methodology and population sizes. For the analyses, a case-control study design is regularly used, but if possible a prospective follow-up study seems more appropriate.[193] However, disadvantages of using a case-control approach are the use of poorly matched controls and the chance of bias. For some time it was thought that siblings would be the best controls, but this turned out to induce overmatching.[5] Overmatching occurs when genetic and environmental factors are more identical in cases and controls than between cases and the general population. Within families these factors are highly aggregated and overmatching then results in a loss of power and could consequently lead to an underestimation of the true risk estimate. Therefore, unrelated controls, matched for at least age, gender and ethnicity, are nowadays considered appropriate for genetic association studies. Most population sizes are relatively small, just like studies on other single nucleotide polymorphisms (SNPs).[194] Odds ratios often represent modest effects in a range of 1.1 – 1.5 and one single nucleotide polymorphism usually explains only 1-8% of the total disease risk in a population. Although this risk seems small, one has to consider the additive effect of multiple SNPs that might also include SNPs of other metabolizing genes. Eventually, a risk of 20-70% can be explained by combinations of SNPs. Ioannidis argued the need of one thousand or more subjects to obtain an adequately powered study.[194] The majority of association studies on SNPs, however, only consist of several hundreds (100-300). Since the design of a study with many thousands of subjects is practically not feasible due to high costs and other reasons of efficiency, one can use a meta-analysis to study if the association found in a group of smaller studies is in fact true. Since an adequate meta-analysis should contain thousands of persons, a total of ten or more smaller studies must be available to start a meta-analysis. At this time, however, there are two main limitations for a meta-analysis of the association of cytochrome P450 polymorphisms and cancer. First, studies in Asian study populations are overrepresented in this field (up to 50%). Therefore, allelic frequencies presented are primarily those found in

Asian people. Consequently, a meta-analysis at this moment cannot contribute to a better understanding of all polymorphic CYPs as risk factors in each ethnic group. Although there could be a large heterogeneity (58%) of allele frequencies between races, the genetic effects are less heterogeneous (14%).[195] The biological impact on the risk of diseases is then usually more or less consistent across different populations, although these relations might not be found due to insufficient power in low frequency areas or absence of specific substrates. Second, the number of studies that can be used in a specific meta-analysis is still too limited to be useful in adequate association analyses of SNPs. For example, there were approximately 600 positive associations between common genetic variants (SNPs) and specific diseases reported of which 166 had been studied 3 times or more.[196] Of these 166 studies, only 6 had a result that was similar in replicated studies. The number of studies done for SNPs of cytochrome P450 genes is even more limited and diverse.

So far, we discussed the restrictions of poorly (or over-) matched control groups and small sample sizes. Besides this some other statistical limitations should be considered. When the allelic frequencies of the study population are not checked for the Hardy-Weinberg equilibrium (HWE), there is a chance of genotyping error and/or selection bias. This happens, for example, with selective survival. Because the majority of cancer develops at a higher age, polymorphisms that are related to lethal diseases earlier in life are less frequent at higher ages. An underestimation as well as an overestimation of the result can be the consequence. Bias due to non-random population stratification can occur in any study in which cases and controls are not matched for their genealogical history.[197] As a consequence, a disease with a high prevalence in a subpopulation will be associated with any allele that has a high frequency in that specific population, regardless of the question whether the association is causal. Another potential limitation is that most association studies with a negative outcome are not published. If a meta-analysis of published data is performed, the effect of CYP polymorphisms tends to be overestimated. This phenomenon most frequently happens for those polymorphisms with a minority of positive outcome studies. Nevertheless, confounding bias is uncommon because cytochrome P450 gene polymorphisms are most probably effect modifiers rather than confounders. Yet, some studies focused on the confounding effect of exogenous substances. They considered the possibility that a CYP polymorphism would make a person more prone to nicotine or alcohol abuse. In this way, a genetic polymorphism might be seen as confounding factor. Nevertheless, definite conclusions from these studies cannot be drawn so far.

In conclusion, despite a compelling theoretical basis for the hypothesis that genetic polymorphisms encoding for cytochrome P450 enzymes may play a role in the development of cancer, the number of studies on this topic are scanty and give contradictory results. However, this topic warrants more detailed investigation, preferably on potential interaction with highly prevalent exposures like endogenous substances such as hormones, and exogenous substances such as toxins and food components. Association studies must be performed

in a standardized way in order to be more similar. They must use the same genetic variant per gene and construct haplotypes. Control groups must contain healthy individuals with the same allele frequencies. All assumptions must be explained thoroughly to later facilitate meta-analyses. An even better approach seems the use of large follow-up databases in order to get rid of the problems of a case-control design. A solution for the lack of common standards in identifying genetic determinants for complex diseases was proposed at a workshop at the end of 2004. The members recommend the creation of a network of all consisting human genome epidemiologists and invite persons around the world to join. Such a network permits broad, consistent and transparent assessment and replication of novel findings from individual studies. Functionality will be particularly been proven in those areas that include genetic effects with impact on multiple biological pathways that in turn affect many different diseases.[198] Most preferably, one should scrutinize each CYP gene for all polymorphisms in all ethnic groups and determine their functionality. This genome-wide screening would however imply tremendous work. Alternatively, for this moment, we can construct linkage disequilibrium blocks (LD-blocks) and use haplotype-tagging SNPs to analyze diversity across a gene without knowing which polymorphisms are the causal ones. Most of the time three or four common versions would capture the majority of genetic diversity throughout the entire human population. Large consortia are necessary to screen for functional causal SNPs in which standardization starts in the early phase of research. This standardization would allow to perform a prospective meta-analysis. Also the combination of several SNPs in several CYP and other genes must be considered to look at interaction and synergy. These combinations can lead to a genetic risk profile that might eventually be applied into clinical practice for prevention and treatment strategies.

2.2. *CYP3A* gene variation, steroid hormone levels and prostate cancer

ABSTRACT

Background: Polymorphisms in genes encoding for CYP3A enzymes, that play a role in steroid hormone metabolism, might affect steroid hormone serum levels and prostate cancer incidence or mortality.

Methods: Over 2400 male participants of The Rotterdam Study with available genotype data on *CYP3A* were included. General linear models were performed on different random subsets of hormone levels to study associations with genotype. Cox' proportional hazard models were used to study prostate cancer incidence and mortality among genotypes.

Results: *CYP3A4* G-allele carriage was associated with lower levels of estrone sulphate ($p = 0.005$) and higher levels of estradiol ($p = 0.04$) compared to non-carriers. *CYP3A5* A-allele carriage was associated with increased levels of estrone sulphate ($p = 0.02$), but higher mean levels of androstenedione and lower levels of testosterone did not reach significance anymore after adjustments for other *CYP3A* genotypes. *CYP3A7* G-allele carriage was associated with the highest number of significant differences in steroid hormone levels. Carriers of the allele resulting in continued enzyme expression during adulthood had decreased levels of DHEA sulphate ($p = 0.05$), androstenedione ($p = 0.006$), estrone ($p = 0.0001$) and estrone sulphate ($p = 0.003$) compared to mean levels of these hormones in homozygous wild type carriers. No associations of *CYP3A* genotype and prostate cancer incidence or mortality were observed.

Conclusion: Primarily the amount of CYP3A7 expression seems to affect steroid hormone levels. However, the variety of inducers and inhibitors of CYP3A and the circadian rhythm of some precursor hormones make analyses difficult. None of the *CYP3A* genotypes materially influenced testosterone level that is a precursor of the prostate growth and differentiation stimulating dihydrotestosterone. In line with this is our non-significant observation for prostate cancer incidence and mortality.

INTRODUCTION

An enumeration of evidence supports the role of androgens in prostate cancer risk [173,176]: Circulating levels of androgens have been reported as being highest in African-American men who experience the highest incidence of prostate cancer and lowest in Chinese men with the lowest prostate cancer incidence; castration at an early age prevents prostate cancer; aging is a major risk factor, suggesting that duration of exposure to androgens might be important; serum levels of testosterone metabolites such as dihydrotestosterone (DHT) correlate with prostate cancer and its growth and differentiation; androgen ablation in hormone positive prostate cancer patients reduces tumor size and disease burden. Therefore, candidate genes for the association with prostate cancer are those that are implicated in androgen production and metabolism. Numerous molecular genetic studies have been performed on polymorphic genes such as the androgen receptor (*AR*), prostate specific antigen (*PSA*), 5α-reductase type II (*SRD5A2*), cytochrome P450c17α (*CYP17*) and a putative hereditary prostate cancer susceptibility gene (*ELAC2*).[168]

Another group of candidate genes are those, which encode for the cytochrome P450 3A (CYP3A) enzyme subfamily that seem to determine bioavailability of testosterone as well.[21] CYP3A is the major subfamily of cytochrome P450 in the human liver and metabolizes many drugs, xenobiotics and endogenous compounds. The genetic coding consists of a cluster of four *CYP3A* genes on chromosome 7q21.1 (*CYP3A4*, *CYP3A5*, *CYP3A7* and *CYP3A43*) each of which contains 13 exons. They share 70-90% of metabolic properties, but are differentially regulated.[199,200] CYP3A4, CYP3A5 and CYP3A43 enzymes catalyze 2β-, 6β- and 15β-hydroxylase of mainly testosterone (T) and are involved in the formation of less biologically active metabolites.[201,202] CYP3A7 has a catalytic activity for 16α-hydroxylation of estrone (E1), dehydroepiandrosterone (DHEA) and DHEA sulphate (DHEAS).[203,204] (FIGURE 1)

Polymorphic expression of the four genes has been studied for alterations of biological functions and prostate cancer risk. For a promoter variant (A -293 G) of *CYP3A4*, designated as **1B*, functionality remains controversial. (TABLE 1) Although epidemiologic studies provide evidence for a decreased activity, with reduced conversion of testosterone to its less bioactive metabolites [89,90,174-176,205], molecular genetic *in vitro* studies concluded for a long time that no biological meaningful effects could be found [88,206-211]. Nevertheless, a re-calculation of the data of Westlind et al. by Rebbeck et al. showed a 2.9-fold increased 6β-hydroxylase activity in *CYP3A4*1B* carriers compared to non-carriers [212], that was reproduced by others as well [213], and almost all studies observed elevations in expression levels of the gene in variant carriers. Among Caucasians, 80% of the population carries the *CYP3A5*3* variant that results in aberrantly spliced mRNA with a premature stop codon. The decreased activity might result in a higher bioavailability of testosterone. It was originally been thought that CYP3A7 was exclusively expressed during fetal life, but a replacement of a part of the promoter (**1C*) with a sequence identical with the same region in the *CYP3A4* promoter, leads to a continued expression in adulthood. This higher activity might shunt DHEA to testosterone. A missense

Table 1. *CYP3A* gene variation and allele frequencies of males in The Rotterdam Study

Gene	SNP	ID number	Region	Nucleotide change	Effect	MAF	HWE	
							χ^2	*p*
CYP3A4	**1B*	rs2740574	promotor	A –392 G	Controversial	G: 4.2%	1.42	0.23
CYP3A5	**3*	rs776746	intron	A 6986 G	Splicing	A: 7.6%	0.42	0.52
CYP3A7	**1C*	rs11568825	promotor	G – 291 T	↑ expression	G: 6.1%	2.17	0.14

SNP = Single nucleotide polymorphism, ID = NCBI database identification, MAF = minor allele frequency, HWE = Hardy Weinberg Equilibrium.

mutation in exon 10 in the *CYP3A43* gene (**3*) was hypothesized to decrease the conversion of testosterone to 6β-hydroxytestosterone, leaving more testosterone available for the conversion to the active form, dihydrotestosterone.[178]

Aim of this study was to clarify the role of polymorphic CYP3A expression in the regulation of steroid hormone serum levels in males and to study the association between *CYP3A* gene variation and prostate cancer incidence and mortality.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population-based prospective cohort study that started with a baseline interview between July 1989 and July 1993. All inhabitants of the Rotterdam suburb Ommoord, aged 55 years and older, were invited. Of the 10,275 eligible subjects, 7983 (78%) agreed to participate.[6] The Medical Ethics Committee of the Erasmus Medical Center approved the study.

Participants were visited at home at the start of the study for a standardised interview on health state and socio-economic characteristics. Subsequently, an extensive physical examination and blood assessment followed at the research center. Since its start, cross-sectional surveys have been carried out periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from the general practitioners working in the study district.

Cohort definition

To explore associations between *CYP3A* genotypes and steroid hormone serum levels or prostate cancer, we included males who provided a blood sample and for whom genotypes were available (n_{cyp3a4} = 2535, n_{cyp3a5} = 2519, n_{cyp3a7} = 2451). Steroid hormone levels were measured in different random subsets of these males (see TABLE 2 for the individual numbers). For the

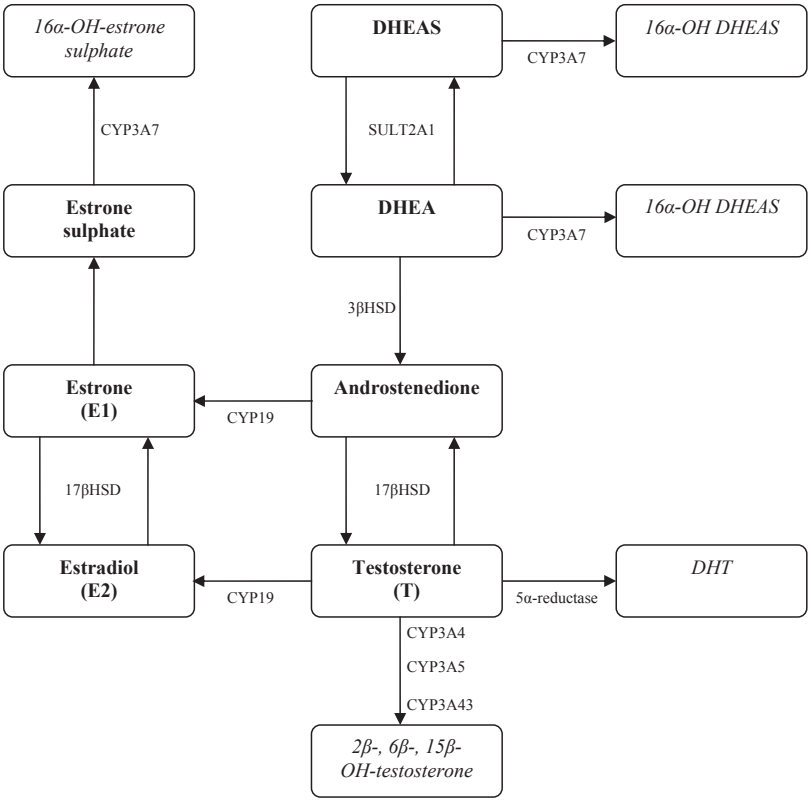

Figure 1. Catalytic activities of CYP3A enzymes in steroid metabolism

Hormone levels measured in this study are indicated **bold**, others are *italicized*. Abbreviations: **CYP3A4**; cytochrome P450 3A4, **CYP3A5**; cytochrome P450 3A5, **CYP3A7**; cytochrome P450 3A7, **CYP3A43**; cytochrome P450 3A43, **CYP19**; cytochrome P450 19 (aromatase), **SULT2A1**; dehydroepiandrosterone sulfotransferase, **3βHSD**; 3β-hydroxy-steroid-dehydrogenase, **17βHSD**; 17β-hydroxy-steroid-dehydrogenase, **-OH-**; hydroxy-, **DHT**; dehydrotestosterone

association with prostate cancer, males with a prevalent prostate cancer at baseline (n = 57) were excluded.

Exposure definition

Genotyping

Genotyping of the *CYP3A4*1B* (rs2740574), *CYP3A5*3* (rs776746), and *CYP3A7*1C* (rs11568825) was determined using TaqMan allelic discrimination assays on an ABI Prism 9700 HT sequence detection system.[20] Each assay consisted of two allele-specific minor groove binding (MGB) probes, labeled with the fluorescent dyes VIC and FAM. The primer and probe sequences, designed by Applied Biosystems by their Assay-by-Design service, are listed in TABLE 1. Polymerase chain reactions (PCR) were performed in a reaction volume of 2.0 µl, containing assay-specific primers, allele-specific TaqMan MGB probes, TaqMan Universal PCR Master Mix No AmpErase

Table 2. Sex-hormone levels per *CYP3A* genotype in males

N model 1 / 2 / 3 / 4	*CYP3A4*			*CYP3A5*			*CYP3A7*		
	AA	**AG + GG**	**p**	**AA+AG**	**GG**	**p**	**TT**	**TG + GG**	**p**
DHEA (µmol/L) N	253 / 237 / 238 / 223	22 / 21 / 22 / 21		40 / 37 / 39 / 37	232 / 221 / 218 / 207		228 / 218 / 216 / 206	41 / 40 / 39 / 38	
Model 1	10.9 (0.4)	11.0 (1.4)	0.93	11.2 (1.1)	10.8 (0.4)	0.68	10.8 (0.4)	11.5 (1.1)	0.56
Model 2	11.0 (0.4)	9.8 (1.8)	0.56	11.7 (1.4)	10.7 (0.5)	0.51	10.7 (0.5)	11.8 (1.1)	0.34
Model 3	10.8 (0.4)	11.2 (1.3)	0.78	11.4 (1.0)	10.7 (0.4)	0.49	10.8 (0.4)	11.8 (1.0)	0.35
Model 4	11.0 (0.4)	9.8 (1.7)	0.54	11.8 (1.2)	10.7 (0.4)	0.45	10.7 (0.4)	11.8 (1.0)	0.29
DHEA sulphate (µmol/L) N	703 / 642 / 238 / 223	62 / 58 / 22 / 21		115 / 108 / 39 / 37	642 / 592 / 218 / 207		622 / 588 / 216 / 206	116 / 112 / 39 / 38	
Model 1	4.4 (0.1)	4.3 (0.4)	0.86	4.2 (0.3)	4.4 (0.1)	0.36	4.5 (0.1)	3.8 (0.3)	$6 \cdot 10^{-3}$
Model 2	4.4 (0.1)	4.5 (0.5)	0.82	4.3 (0.3)	4.4 (0.1)	0.66	4.5 (0.1)	3.8 (0.3)	$1 \cdot 10^{-2}$
Model 3	4.7 (0.2)	4.5 (0.6)	0.75	4.2 (0.5)	4.7 (0.2)	0.30	4.8 (0.2)	3.8 (0.5)	$3 \cdot 10^{-2}$
Model 4	4.6 (0.2)	4.6 (0.8)	0.92	4.3 (0.6)	4.7 (0.2)	0.53	4.8 (0.2)	3.8 (0.5)	$5 \cdot 10^{-2}$
Androstenedione (nmol/L) N	694 / 636 / 250 / 234	57 / 53 / 20 / 19		109 / 102 / 39 / 36	632 / 587 / 228 / 217		614 / 583 / 224 / 214	111 / 106 / 40 / 39	
Model 1	4.4 (0.1)	4.7 (0.2)	0.23	4.6 (0.2)	4.3 (0.1)	0.17	4.5 (0.1)	4.1 (0.2)	$4 \cdot 10^{-2}$
Model 2	4.4 (0.1)	4.7 (0.3)	0.31	4.5 (0.2)	4.4 (0.1)	0.59	4.4 (0.1)	4.1 (0.2)	$4 \cdot 10^{-2}$
Model 3	4.8 (0.1)	5.3 (0.4)	0.21	5.3 (0.3)	4.7 (0.1)	$4 \cdot 10^{-2}$	4.9 (0.1)	4.2 (0.3)	$7 \cdot 10^{-3}$
Model 4	4.8 (0.1)	4.9 (0.5)	0.80	5.2 (0.3)	4.7 (0.1)	0.20	4.9 (0.1)	4.1 (0.3)	$6 \cdot 10^{-3}$
Testosterone (nmol/L) N	677 / 618 / 608 / 555	59 / 55 / 53 / 49		109 / 101 / 98 / 91	620 / 572 / 554 / 513		602 / 571 / 542 / 514	108 / 102 / 95 / 90	
Model 1	11.5 (0.1)	11.0 (0.5)	0.31	10.8 (0.4)	11.6 (0.2)	$5 \cdot 10^{-2}$	11.5 (0.2)	11.2 (0.4)	0.53
Model 2	11.4 (0.2)	11.5 (0.6)	0.93	10.8 (0.5)	11.6 (0.2)	0.13	11.5 (0.2)	11.2 (0.4)	0.52
Model 3	11.4 (0.1)	10.6 (0.5)	0.16	10.5 (0.4)	11.4 (0.2)	$2 \cdot 10^{-2}$	11.3 (0.2)	11.3 (0.4)	1.00
Model 4	11.3 (0.2)	11.0 (0.6)	0.68	10.6 (0.4)	11.4 (0.2)	0.10	11.3 (0.2)	11.3 (0.4)	0.97
Free testosterone (nmol/L) N	508 / 461 / 501 / 454	47 / 43 / 45 / 49		84 / 78 / 83 / 77	463 / 426 / 454 / 418		454 / 429 / 446 / 421	78 / 75 / 77 / 74	
Model 1	0.3 (0.0)	0.2 (0.0)	0.10	0.3 (0.0)	0.3 (0.0)	0.16	0.3 (0.0)	0.3 (0.0)	0.61
Model 2	0.3 (0.0)	0.2 (0.0)	0.29	0.3 (0.0)	0.3 (0.0)	0.58	0.3 (0.0)	0.3 (0.0)	0.72
Model 3	0.3 (0.0)	0.2 (0.2)	$6 \cdot 10^{-2}$	0.2 (0.0)	0.3 (0.0)	0.12	0.3 (0.0)	0.3 (0.0)	0.85
Model 4	0.3 (0.0)	0.2 (0.0)	0.14	0.3 (0.0)	0.3 (0.0)	0.66	0.3 (0.0)	0.3 (0.0)	0.86
Estrone (pmol/L) N	673 / 615 / 606 / 554	59 / 55 / 53 / 49		112 / 104 / 100 / 93	613 / 566 / 551 / 510		594 / 563 / 537 / 509	111 / 107 / 97 / 94	
Model 1	91.5 (2.0)	81.1 (6.9)	0.14	82.6 (5.0)	92.3 (2.1)	$7 \cdot 10^{-2}$	95.0 (2.5)	70.7 (5.0)	$< 1 \cdot 10^{-5}$
Model 2	91.4 (2.2)	91.2 (8.9)	0.98	84.2 (6.4)	92.7 (2.3)	0.23	95.1 (2.2)	71.8 (5.1)	$< 1 \cdot 10^{-4}$
Model 3	90.0 (2.0)	76.6 (6.9)	$6 \cdot 10^{-2}$	80.3 (5.1)	90.6 (2.1)	$6 \cdot 10^{-2}$	92.7 (2.5)	70.6 (5.1)	$< 1 \cdot 10^{-4}$
Model 4	90.3 (2.2)	82.2 (8.9)	0.39	84.0 (6.4)	90.7 (2.3)	0.35	93.1 (2.2)	71.1 (5.3)	$1 \cdot 10^{-4}$
Estrone sulphate (pmol/L) N	616 / 560 / 568 / 513	50 / 47 / 48 / 45		96 / 88 / 91 / 83	568 / 519 / 521 / 475		543 / 513 / 499 / 470	98 / 94 / 92 / 88	
Model 1	472.3 (14.2)	363.5 (50.2)	$4 \cdot 10^{-2}$	462.3 (36.3)	463.6 (14.9)	0.98	497.3 (15.2)	302.0 (35.8)	$< 1 \cdot 10^{-6}$
Model 2	479.1 (15.0)	295.7 (63.2)	$6 \cdot 10^{-3}$	544.7 (45.4)	451.4 (15.9)	$7 \cdot 10^{-2}$	494.7 (15.4)	302.5 (36.1)	$1 \cdot 10^{-6}$
Model 3	469.8 (12.2)	401.1 (42.3)	0.12	477.0 (30.6)	462.2 (12.7)	0.65	486.1 (13.2)	376.0 (31.2)	$1 \cdot 10^{-3}$
Model 4	478.5 (13.1)	317.8 (54.2)	$5 \cdot 10^{-3}$	549.8 (39.2)	450.8 (13.9)	$2 \cdot 10^{-2}$	481.6 (13.4)	379.4 (31.3)	$3 \cdot 10^{-3}$
Estradiol (pmol/L) N	534 / 485 / 422 / 382	50 / 47 / 43 / 40		93 / 88 / 78 / 71	479 / 446 / 379 / 351		470 / 445 / 376 / 356	92 / 87 / 69 / 66	
Model 1	47.4 (1.1)	50.5 (3.5)	0.40	45.3 (2.5)	47.7 (1.1)	0.39	48.6 (1.1)	42.1 (2.5)	$2 \cdot 10^{-2}$
Model 2	46.4 (1.1)	56.3 (4.4)	$4 \cdot 10^{-2}$	40.5 (3.2)	48.6 (1.2)	$3 \cdot 10^{-2}$	48.2 (1.2)	42.8 (2.6)	$6 \cdot 10^{-2}$
Model 3	47.6 (1.0)	54.7 (3.3)	$4 \cdot 10^{-2}$	50.6 (2.4)	47.3 (1.1)	0.22	48.4 (1.1)	46.4 (2.6)	0.49
Model 4	47.1 (1.1)	56.4 (4.2)	$4 \cdot 10^{-2}$	45.9 (3.1)	48.4 (1.2)	0.48	48.2 (1.1)	46.9 (2.7)	0.65
Free estradiol (pmol/L) N	357 / 318 / 330 / 296	36 / 33 / 36 / 33		65 / 59 / 63 / 57	317 / 292 / 293 / 272		317 / 297 / 294 / 277	56 / 54 / 54 / 52	
Model 1	1.4 (0.0)	1.4 (0.1)	0.69	1.3 (0.1)	1.4 (0.0)	0.80	1.4 (0.0)	1.3 (0.1)	0.46
Model 2	1.3 (0.0)	1.5 (0.1)	0.26	1.2 (0.1)	1.4 (0.0)	0.20	1.3 (0.0)	1.3 (0.1)	0.88
Model 3	1.3 (0.0)	1.5 (0.1)	0.16	1.4 (0.1)	1.3 (0.0)	0.31	1.3 (0.0)	1.4 (0.1)	0.52
Model 4	1.3 (0.0)	1.5 (0.1)	0.17	1.3 (0.1)	1.4 (0.0)	0.67	1.3 (0.0)	1.4 (0.1)	0.37
Cortisol (nmol/L) N	321 / 300	26 / 25		49 / 45	294 / 280		288 / 276	51 / 49	
Model 1	355.0 (8.3)	389.0 (29.4)	0.27	369.9 (21.5)	355.6 (8.7)	0.54	363.0 (8.7)	325.4 (20.8)	0.10
Model 2	355.9 (8.8)	394.8 (37.0)	0.32	355.2 (27.4)	359.5 (9.3)	0.89	364.6 (9.1)	326.9 (21.7)	0.11
Cortisol / DHEAS ratio N	298 / 279	25 / 24		47 / 44	271 / 259		270 / 258	46 / 45	
Model 1	144.8 (10.5)	159.7 (36.6)	0.70	139.8 (26.8)	149.2 (11.1)	0.75	141.2 (11.1)	182.6 (27.0)	0.16
Model 2	144.3 (11.4)	198.4 (47.8)	0.29	110.5 (34.9)	155.0 (12.1)	0.25	142.1 (11.6)	186.0 (27.9)	0.15

N = number of persons included in the model. DHEA = dehydroepiandrosterone. DHEAS = dehydroepiandrosterone sulphate. Values are expressed as means (standard error). Precursor hormone [DHEAS: DHEA], [DHEA: DHEAS], [androstenedione: DHEA], [(free) testosterone: androstenedione], [estrone: androstenedione], [estradiol: estrone, testosterone], [free estradiol: estrone, testosterone], [estrone sulphate: estrone]. Model 1: adjusted for age, body mass index, smoking status and number of pack years, alcohol intake, physical activity, corticosteroid use and time of blood assessment. Model 2: adjusted for age, body mass index, smoking status and number of pack years, alcohol intake, physical activity, corticosteroid use, time of blood assessment and other *CYP3A* genotypes. Model 3: adjusted for age, body mass index, smoking status and number of pack years, alcohol intake, physical activity, corticosteroid use, time of blood assessment, and precursor hormone. Model 4: adjusted for age, body mass index, smoking status and number of pack years, alcohol intake, physical activity, corticosteroid use, time of blood assessment, precursor hormones and other *CYP3A* genotypes.

UNG (2X) and genomic DNA(1 ng). The thermal profile consists of an initial denaturation step at 95°C for 15 minutes, followed by 40 cycles of denaturation at 92°C for 15 seconds, and annealing and extension at 60°C for 1 minute. Genotypes were scored by measuring allelic-specific fluorescence using the SDS 2.2.2 software for allelic discrimination.[20]

Outcome definition

Hormone measurements

Blood samples were drawn by venapuncture from non-fasting subjects and collected in 5-ml tubes containing 0.5 ml sodium citrate solution. All tubes were stored on ice before and after blood sampling. Levels of steroid hormones were measured in plasma. Platelet-free plasma was obtained by two-stage centrifugation, first for 10 min at 1600 x *g* at 4°C and then for 30 min at 7000 x *g* at 4°C. Platelet-free samples were immediately frozen in liquid nitrogen and transferred to the laboratory where they were stored at -80°C until hormone estimations were performed. Plasma levels of dehydroepiandrosterone (DHEA), DHEA sulphate (DHEAS), androstenedione, testosterone (T), estrone (E1), estrone sulphate (E1S), estradiol (E2), cortisol and sex steroid hormone binding globulin (SHBG) were estimated in 12 separate batches of samples using coated tube or double antibody RIAs, purchased from Diagnostic Systems Laboratories.[214] Because of the relatively small volumes of plasma available, all values reported are single sample estimations. Intra-assay coefficients of variation, determined on the basis of duplicate results of internal quality control pools with three different levels of each analyte, were below 15% for all assays. Because inter-assay variations were relatively large (33% DHEA, 30% DHEAS, 33% androstenedione, 19% testosterone, 18% estrone, 29% estrone sulphate, 25% estradiol, 24% cortisol and 14% SHBG), we multiplied all concentrations within a batch with a factor, which made results for the internal quality control pools comparable. This reduced inter-assay variations to 3%, 10%, 8%, 7%, 18%, 7%, 14%, 7%, and 3%, respectively, and was considered justified because the relative differences in the results for the high and middle internal quality control pools per batch were comparable, as evidenced by correlations between these results ($r = 0.99$ for DHEA, $r = 0.58$ for DHEAS, $r = 0.91$ for androstenedione, $r = 0.75$ for testosterone, $r = 0.60$ for estrone, $r = 0.93$ for estrone sulphate, $r = 0.21$ for estradiol, $r = 0.93$ for cortisol and $r = 0.91$ for SHBG; all $P < 0.05$, except for E2). Assays were performed

blind with respect to information on the subject. As a measure of bioavailable testosterone and estradiol, non-SHBG-bound hormone was calculated on the basis of hormone, SHBG, and albumin levels, and respective affinity constants according to the method described by Södergård *et al.* [215] and van den Beld *et al.* [216]. We additionally investigated the ratio of free cortisol over DHEAS, because DHEAS may antagonize the actions of cortisol.[217-219]

Prostate cancer case identification and validation

Three different databases were used for prostate cancer case identification. First, cases diagnosed by general practitioners in the research area were collected (International Classification of Primary Care (Y77)). Second, the national registry of all hospital admissions was consulted to detect all malignancy related hospital admissions for study participants. Third, regional pathology databases were linked to The Rotterdam Study to identify cases (International Classification of Diseases-10 (C61)). A physician (CS) subsequently validated prostate cancer cases on the basis of medical records of the general practitioner, discharge letters, and pathology reports. Only pathologically confirmed cases were considered in the analyses. The index date was defined as the earliest date found in the pathology reports. Cases for whom information about differentiation grade was available were additionally divided into low- and high-differentiated cases. Early onset was defined as a diagnosis before the age of 75.

Covariates

For analyses on hormone levels, the following potential confounders were considered: age, body mass index (BMI, kg/m^2), smoking status (never, current, past), alcohol intake (grams/day), ability to be physically active, corticosteroid use, precursor hormones, *CYP3A* status, and time of blood assessment (in two hour intervals).

For analyses on prostate cancer age, prostate cancer screening (PSA and rectal examination), BMI and other *CYP3A* genotypes were considered to be potential confounders. Age at diagnosis was considered a confounder in the prostate cancer specific mortality analysis.

Statistical analyses

First, we compared characteristics of participants with and without blood samples using a Student's t-test for continuous variables and a χ^2-test for dichotomous variables. Hardy Weinberg Equilibria (HWE) were calculated with a χ^2-goodness-of-fit statistic. Genotypes were dichotomised. Heterozygous persons were combined with homozygous carriage of the minor allele (*CYP3A4*: AA *vs* AG + GG, *CYP3A5*: AA + AG *vs* GG, *CYP3A7*: TT *vs* TG + GG) to optimise power. Missing indicators were used to study the effect of missing values. Missing data on covariates were imputed by single imputation using the Expectation Maximization (EM) algorithm.[220] We anticipated that the level of precursor hormone influences hormone levels. Therefore, the following adjustments for precursor hormones (when available) were made: DHEAS adjusted for DHEA, DHEA adjusted for DHEAS, androstenedione adjusted for DHEA, (free) testosterone

adjusted for androstenedione, estrone (E1) adjusted for androstenedione, estrone sulphate (E1SO4) adjusted for estrone, (free) estradiol (E2) adjusted for estrone and testosterone. To avoid confounding by other *CYP3A* genotypes, the remainder of the genotypes was included in the statistical model to explore the *CYP3A* specific genotype effects. Since none of the covariates changed the estimates by more than 10%, all potential confounders were included in the following models. Age-, BMI-, smoking-, alcohol-, physical activity-, corticosteroid use-, time of blood assessment (model 1), model 1 covariates with other *CYP3A* genotype- (model 2), model 1 covariates with precursor hormone level- (model 3) and model 1 covariates with both precursor hormone level- and other *CYP3A* genotype- (model 4) adjusted general linear models were used to analyse differences in mean (± standard error of the mean (SEM)) hormone serum levels between genotypes. For cortisol levels no adjustments were made for precursor hormones. Cox' proportional hazard models were used to analyse age-, screening-, BMI- (model 1) and other *CYP3A* genotype- (model 2) adjusted prostate cancer risks per genotype. Finally, age at diagnosis- (model 1) and other *CYP3A* genotype-adjusted (model 2) 'overall', 'prostate cancer specific' and 'subgroups of prostate cancer specific' ('low differentiated', 'high differentiated', 'early onset', 'late onset') mortality was studied per genotype. P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant. All analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA)

RESULTS

Individuals of whom the genotype was unknown were on average older, relatively more frequently smoker, and had a shorter follow-up time than those for whom genotype data were available. All SNPs were in HWE (TABLE 1).

CYP3A4 variant carriage (AG or GG) was associated with a significant increase in mean estradiol level compared to the wild type (56.4 ± 4.2 vs 47.1 ± 1.1 *p*mol/L, respectively) (TABLE 2). However, no significant difference was observed for free estradiol levels. *CYP3A4* G-allele carriers showed 160 *p*mol/L lower mean estrone sulphate levels than persons with the wild type ($p = 0.005$). A significant decrease in mean estrone sulphate level was also found among *CYP3A5* G-allele homozygotes compared to persons with at least one copy of the A-allele (549.8 ± 39.2 *vs* 450.8 ± 13.9 *p*mol/L, respectively). Although the first two models resulted in significantly increased mean serum testosterone levels among *CYP3A5* G-allele homozygotes, this effect disappeared after introducing the status of other *CYP3A* genotypes into the model. No association of *CYP3A5* genotype and free (or biologically active) testosterone was observed. In the current study, *CYP3A7* genotype had the strongest impact on steroid hormone levels. Carriers of the G-allele resulting in continued enzyme expression during adulthood had decreased levels of DHEA sulphate ($p = 0.05$), androstenedione ($p = 0.006$), estrone ($p = 0.0001$) and

estrone sulphate (p = 0.003) compared to mean levels of these hormones in homozygous wild type carriers. The influence of other *CYP3A* genotypes on *CYP3A7* effects seems negligible.

No associations between any of the three *CYP3A* genotypes and prostate cancer incidence were found (TABLE 3). Combinations between *CYP3A* pairs and triplets showed no significant differences between carriers and non-carriers as well (data not shown). Primarily not for the combination of *CYP3A4* AG or GG (**1B*) and *CYP3A5* AA or AG (**1*) among all males (HR 0.88; 95%CI 0.50-1.54), males with early onset (< 75 years) prostate cancer (HR 0.57; 95% CI 0.23-1.40), all low differentiated cancers (HR 1.08; 95%CI 0.47-2.50), early onset low differentiated cancers (HR 0.37; 95%CI 0.05-2.60), all high differentiated cancers (HR 1.06; 95%CI 0.46-2.45) and early onset high differentiated cancers (HR 0.97; 95%CI 0.35-2.69). The combination of *CYP3A4* AG or GG (**1B*) and *CYP3A5* AG or GG (**3*) also resulted in non-statistically significant hazard ratios (all males: HR 0.50; 95%CI 0.07-3.55 and early onset cases: HR 1.09; 95%CI 0.15-7.80). Further subgroup analyses were not possible as a consequence of insufficient power. Significant differences in prostate cancer risk were also not observed when we compared low- with high-differentiated prostate cancers (data not shown). None of the genotypes were associated with overall or prostate cancer specific mortality. (TABLE 4) Although a borderline-significant 3-fold increase in mortality rate was observed in *CYP3A5* A-allele carriers with late onset prostate cancer, this association disappeared after considering other *CYP3A* status.

Table 3. Association between *CYP3A* genotypes and prostate cancer risk

		Model 1 HR (95% CI)			Model 2 HR (95% CI)		
Genotype	Group	All cases	Low differentiation	High differentiation	All cases	Low differentiation	High differentiation
Any *CYP3A4*1B*	Total sample	0.89 (0.53-1.51)	0.94 (0.41-2.17)	0.95 (0.41-2.19)	0.82 (0.09-7.35)	0.97 (0.34-2.74)	1.19 (0.40-3.50)
	Early onset < 75 years	0.60 (0.27-1.37)	0.31 (0.04-2.25)	0.87 (0.32-2.41)	0.73 (0.27-2.00)	0.25 (0.03-2.13)	0.87 (0.24-3.11)
Any *CYP3A5*1*	Total sample	1.25 (0.82-1.90)	1.04 (0.55-1.97)	0.99 (0.54-1.83)	1.12 (0.15-8.21)	0.98 (0.44-2.19)	1.20 (0.51-2.85)
	Early onset < 75 years	1.23 (0.72-2.11)	1.12 (0.44-2.88)	0.95 (0.47-1.93)	1.13 (0.56-2.29)	0.71 (0.26-1.98)	0.91 (0.36-2.35)
Any *CYP3A7*1C*	Total sample	0.89 (0.58-1.36)	0.61 (0.26-1.40)	1.16 (0.63-2.15)	3.37 (0.91-12.46)	0.62 (0.27-1.44)	1.26 (0.68-2.35)
	Early onset < 75 years	1.11 (0.67-1.83)	0.59 (0.18-1.94)	1.35 (0.68-2.68)	1.12 (0.67-1.88)	0.61 (0.19-2.00)	1.48 (0.74-2.95)

HR = hazard ratio, CI = confidence interval. Early onset is calculated out of date of diagnosis. Homozygous wild type (*1/*1) was reference for the analyses of *CYP3A4* and *CYP3A7*. Homozygous mutant (*3/*3) was reference for the analyses of *CYP3A5*. Model 1: Adjusted for age at baseline, body mass index and screening. Model 2: Adjusted for age at baseline, body mass index, screening and other *CYP3A* genotypes.

Table 4. Association between *CYP3A* genotype and mortality

		Mortality					
		HR (95% CI)					
		All cause	**All prostate cancer**	**Low differentiated prostate cancer**	**High differentiated prostate cancer**	**Early onset (< 75 years) prostate cancer**	**Late onset (≥ 75 years) prostate cancer**
*CYP3A4*1B*	Model 1	0.61 (0.22-1.70)	0.51 (0.18-1.41)	0.20 (0.03-1.43)	1.17 (0.30-4.49)	2.07 (0.47-9.23)	0.29 (0.07-1.21)
	Model 2	0.54 (0.13-2.32)	0.42 (0.10-1.76)	0.08 (0.01-1.14)	1.79 (0.17-19.15)	2.13 (0.26-17.23)	0.35 (0.03-3.91)
*CYP3A5*1*	Model 1	1.41(0.68-2.91)	1.50 (0.74-3.05)	1.66 (0.58-4.71)	1.20 (0.41-3.52)	0.65 (0.25-1.68)	3.21 (0.98-10.51)
	Model 2	0.80 (0.28-2.25)	0.77 (0.28-2.15)	0.51 (0.15-1.73)	1.29 (0.17-10.05)	0.70 (0.19-2.58)	1.16 (0.16-8.56)
*CYP3A7*1C*	Model 1	0.63 (0.15-2.61)	0.66 (0.16-2.71)	0.45 (0.06-3.40)	1.10 (0.14-8.38)	0.93 (0.12-7.26)	0.61 (0.08-4.61)
	Model 2	0.70 (0.16-3.05)	0.78 (0.18-3.32)	1.54 (0.15-16.07)	1.00 (0.13-7.75)	0.48 (0.04-5.49)	0.84 (0.11-6.50)

HR = hazard ratio, CI = confidence interval. Model 1: adjusted for age. Model 2: adjusted for age and other *CYP3A* genotypes.

DISCUSSION

This population-based study was subdivided into two parts. In the first part, cytochrome P450 *3A* genotypes were associated with baseline steroid hormone levels. Mean estrone sulphate levels were found to be significantly lower in variant carriers of *CYP3A4* (AG + GG) and *CYP3A7* (TG + GG) and persons homozygous for the *CYP3A5* G-allele compared to the remainder of the genotypes. Furthermore, the *CYP3A4* G-allele was associated with higher levels of estradiol. Persons with continued expression of CYP3A7 enzymes (G) showed in addition to decreased levels of estrone sulphate, lower levels of DHEA sulphate, androstenedione and estrone in comparison with persons that do not express CYP3A7 in adulthood. In the second part, no significant associations were found between individual and combined *CYP3A* genotypes and prostate cancer incidence or mortality.

The role of CYP3A enzymes in steroid hormone metabolism is widely accepted. However, very limited data are available on *CYP3A* gene variations and steroid metabolism and these mainly focus on testosterone. Furthermore, most studies comprise an *in vitro* setting. The present study not only included multiple *CYP3A* genes, it also explored the effect of gene variation on different steroid hormone levels *in vivo.* In two previous studies, it was observed that *CYP3A4* G-allele carriers showed increased 6β-hydroxylase activity.[212,213] Therefore, the authors proposed a stimulation of the downstream testosterone metabolism pathway resulting in lower testosterone levels. Although our results show slightly decreased levels of both serum testosterone and its biological active component free testosterone in persons with *CYP3A4* AG or GG, no statistically significant differences could be found. That *in vivo* data do not confirm the *in vitro* results is most probably explained by the role of the counterbalancing pituitary-adrenal (cortisol) and pituitary-gonadal (testosterone) pathways. Furthermore, our results are to a certain extent similar to those found in a study that used two different subgroups.[221] The first subgroup is a random sample of both genders from The Rotterdam Study and the second

comprises males of 73 years or older randomly selected from a medium-sized town in The Netherlands. Both significantly decreased serum levels of DHEA sulphate and estrone were consistently found in persons with the variant allele of the *CYP3A7* gene, encoding for enzymes which are normally only expressed during fetal live. Nevertheless, in the present study, decreased serum levels of estrone sulphate and androstenedione were additionally observed in *CYP3A7* G-allele carriers. No data on estrone sulphate were available in the previous study. The contradiction in results found for androstenedione might be explained by the use of other covariates. For our first statistical model, that shows the highest similarity with the model used in the study of Smit *et al.*, we observed a significant decrease in androstenedione levels with a p-value of 0.04. After introducing the level of the precursor hormone DHEA ($p = 0.007$) and additionally other *CYP3A* genotypes ($p = 0.006$), the difference in serum androstenedione levels between *CYP3A7* TG + GG and TT became much more significant. Another reason might be the mixture of genders in the previous study. The results for DHEA sulphate, estrone and estrone sulphate follow directly from the proposed increased CYP3A7 activity in G-allele carriers.[203,204] The diminished effects in models 3 and 4 might be a consequence of the adjustment for precursor hormones with a circadian rhythm, as p-values remain similar in a model without the precursor hormone. As was previously put forward [221], a short half-life of DHEA together with a high metabolic clearance rate [222] can explain the fact that no differences were found for mean DHEA levels among genotypes. Since there are strong correlations between DHEA, DHEAS and androstenedione [223] these might explain the significant difference observed for *CYP3A7* genotypes. That *CYP3A7* variation has the highest impact on steroid hormone levels seems to be a result of the variety of targets of its encoding enzyme in the steroid hormone pathway, whereas the other *CYP3A* enzymes are related to the degradation of testosterone and estrogens for which levels are kept more or less constant by other counterbalancing pathways.

Our data do not confirm the association between *CYP3A* genotypes and prostate cancer or advanced stages of prostate cancer as was observed in some previous studies (see also TABLE 5).[89,90,174-177,205,224] When only participants of Caucasian descent are considered, results of published studies are contradictory, varying from associations with significantly increased high-grade or late onset prostate cancers [89,205] and with decreased low-grade prostate cancer [90] to non-significant findings [225] or even inverse relations with prostate cancer [173]. Our power-induced choice to define early onset as a prostate cancer diagnosis before the age of 75 might explain our non-significant findings. Nevertheless, studying the role of CYP enzymes in humans is complicated since a broad range of inducers and inhibitors determines its expression, which might overrule the genetically induced variability. This would argue for using larger cohorts and to attempt to control for as many inducers and inhibitors as possible. However, as this study shows, the effect of *CYP3A* genotype variation on serum testosterone and free testosterone level, as precursors of the prostate cancer growth- and differentiation determining metabolite dihydrostestosterone, is very limited. This supports the thoughts of polymor-

Table 5. Previous *CYP3A* variant association studies on prostate cancer risk

CYP	Variant	Design	No. cases/controls	Ethnicity	Result	OR (95%CI)	Reference
CYP3A4	**1B*	Case-only	230 / 0	CA	↑ high grade PrCa	2.1 (1.1-4.1)	Rebbeck (1998) [89]
					↑ PrCa in persons without positive family history	2.7 (1.2-5.6)	
					↑ PrCa in persons aged > 63 years	6.7 (2.5-17.7)	
					↑ PrCa in high grade, aged > 63, no family history	9.5 (2.5-35.2)	
		Case-control	174 / 116	AA	↑ PrCa	p < 0.01	Paris (1999) [175]
		Case-only	103	AA	↑ high grade PrCa in persons aged > 65 years	2.4 (1.1-5.4)	
		Case-control (N)			↑ PrCa in BPH patients	6.3 (2.7-17.3)	Tayeb (2002) [174]
		Case-control	84 / 136	AA	↑ PrCa	4.1 (1.3-12.2)	Kittles (2002) [205]
			215 / 94	CA	↑ PrCa	2.3 (1.1-4.5)	
			77 / 82	NG	NS	-	
			376 / 312	ALL	↑ PrCa	2.4 (1.4-3.9)	
		Case-control (N)	21 / 379	CA	NS	-	Tayeb (2003) [225]
		Case-sibling	440 / 480	CA + AA	↑ high grade PrCa	1.9 (1.0-3.6)	Plummer (2003) [90]
					↓ low grade PrCa	0.1 (0.0-0.5)	
					↓ PrCa in persons < 63 years	0.5 (0.3-1.0)	
		Case-control	530 / 230	CA	↓ PrCa	0.5 (0.3-0.9)	Zeigler-Johnson (2004) [173]
			77 / 54	AA	NS	-	
		Case-sibling	440 / 480	CA + AA	↓ low grade PrCa	0.2 (0.1-0.7)	Loukola (2004) [a 176]
		Case-control	145 / 103	AA	↑ high grade PrCa	11.9 (1.62-533.0)	Bangsi (2006) [224]
			177 / 145	CA	NS	-	
CYP3A5	**3*	Case-sibling	440 / 180	CA + AA	↓ low grade PrCa	0.4 (0.2-0.8)	Plummer (2003) [90]
		Case-control	55 / 177	CA	NS	-	Zeigler-Johnson (2004) [173]
			81 / 65	AA	NS	-	
	**1*	Case-control	260 / 212	JP	↓ low grade PrCa	0.2 (0.1-0.8)	Zhenhua (2005) [177]
					↓ low stage PrCa	0.3 (0.1-1.0)	
Both	**1B + *3*	Case-sibling	433 / 469	CA + AA	↑ PrCa	8.6 (1.1-87.0)	Plummer (2003) [90]
				AA	↑ PrCa	2.9 (1.4-6.2)	
	**1B + *1*			CA	↓ low grade PrCa	0.1 (0.0-0.5)	
	**1B + *1*	Case-control	132 / 241	CA + AA	↓ high grade PrCa	0.4 (0.2-0.9)	Zeigler-Johnson (2004) [173]
					↓ high grade PrCa in persons aged < 60 years	0.2 (0.1-0.8)	

CA = Caucasian, AA = African-American, NG = Nigerian, JP = Japanese, PrCa = prostate cancer, BPH = benign prostate hyperplasia, N = nested (case-control), NS = not significant, ↑ increased risk, ↓ decreased risk. [a] Table presents *CYP3A* associations of the most commonly studied SNPs and prostate cancer. Loukola et al. additionally studied other SNPs, see reference for more detail

phisms in *CYP3A* genes being in linkage with a real, not yet detected, disease-associated gene. Polymorphisms in *CYP3A* genes, as in this study, were mainly used for their proposed *in vitro* functionality. Since this study does not support important *in vivo* influences, it might be interesting to use a haplotype or genome-wide approach to explore whether *CYP3A* genetic variability is essential in induction or promotion of prostate cancer or in prostate cancer specific mortality (see HapMap for the definition of haplotype blocks [24]).

Observational studies are prone to selection, information and confounding bias. Although in this study persons without blood samples, who appeared to be older, were excluded, this most probably did not result in selection bias. Aging is associated with at least some of the hormone levels [226] and with prostate cancer, but seems not to be associated with *CYP3A* genotype since these were all in Hardy-Weinberg equilibrium. Information bias is unlikely as both exposure and outcome measurements were performed in advance of the research question. However, the use of pathologically confirmed cases might have slightly diminished the risk estimates since males with an undetected prostate cancer were included as controls. As far as we know, we adjusted for most known confounders. Finally, our choice to combine heterozygous with homozygous of the minor allele, might have diluted the effect.

In conclusion, this study shows associations between *CYP3A* genotypes and some steroid hormone levels, with the most distinct role for a common variant allele of the *CYP3A7* (**1C*) gene that was associated with decreased levels of estrone, estrone sulphate, DHEA sulphate, and androstenedione. However, no indications for an association with prostate cancer incidence or mortality were observed. Additional studies should focus on the meaning of differences in hormone levels in both males and females on disease risks. Furthermore, the role of *CYP3A* genes on prostate cancer risk needs a haplotype or genome wide approach and might be independent of testosterone level.

2.3. *Cyp3A* gene variation, steroid hormone levels and breast cancer

ABSTRACT

Background: Polymorphisms in genes encoding for CYP3A enzymes, that play a role in steroid hormone metabolism, might affect steroid hormone serum levels, age at menarche and menopause and breast cancer incidence and mortality.

Methods: Over 3600 postmenopausal female participants of The Rotterdam Study who provided a blood sample were included. General linear models were performed on different random subsets of hormone levels to study associations with genotype. Student's t-tests were used to explore differences in age at onset of menarche or menopause and total estrogen-exposure period. Cox' proportional hazard models were used to study associations between genotype and breast cancer incidence and mortality.

Results: Statistically significant associations were found for *CYP3A5* GG, which results in absence of CYP3A5 enzyme expression, and decreased levels of DHEA ($p = 0.03$) and *CYP3A7* TG/GG, which results in continued expression of CYP3A7 enzymes during adulthood, and decreased levels of DHEA sulphate ($p = 0.03$), estrone ($p = 0.05$) and estrone sulphate ($p = 0.05$) compared to the remainder of the genotypes, although results were highly dependent on the model used. The relationship between estrogen levels and *CYP3A* genotypes was of borderline significance. Women with the *CYP3A4* GG-genotype, for which functionality remains to be proven, had a significantly increased estrogen-exposure period compared to those with AA or AG genotypes ($p = 0.01$) and *CYP3A5* AG- or GG-genotype resulted in an older age at menopause compared to women with an AA genotype ($p = 0.05$). No associations were found with breast cancer incidence or mortality.

Conclusion: Although some associations between *CYP3A* genotypes and steroid hormone levels were observed, results are less pronounced than we found in a previous study among males. This might be a consequence of gender-specific expression levels of CYP3A enzymes. Furthermore, age at onset of menarche and menopause and the calculated estrogen-exposure period, all commonly accepted risk factors for breast cancer, were related with some of the *CYP3A* genotypes. However, no associations were observed for the genotypes and breast cancer incidence or mortality, which is in line with a previous study. Future studies might benefit from stratification on estrogen- or progesterone-receptor status in tumor tissue.

INTRODUCTION

Estrogens have been associated with breast cancer induction and promotion.[227] They stimulate the division of breast epithelial cells, which leads to a higher chance of mutations and increased growth rates. Several established risk factors for breast cancer are related to increased levels or exposure times of estrogens: early menarche, late menopause, nulliparity, late age at first birth, marital status, oral contraceptive use, hormone replacement therapy, low physical exercise levels, and obesity. Candidate genes for breast cancer are those that are implicated in the metabolism of estrogens. Extensive research has been performed on the role of the estrogen receptor and aromatase (cytochrome P450 19) genes and their polymorphic appearances.[42,145,159-163,228-234] Other well-characterized enzymes that are involved in steroid metabolism are those of the cytochrome P450 3A family.[201-204,235,236] *CYP3A* is the major subfamily of cytochrome P450 in the human liver and metabolizes many drugs, xenobiotics and endogenous compounds. The genetic coding consists of a cluster of four *CYP3A* genes on chromosome 7q21.1 (*CYP3A4*, *CYP3A5*, *CYP3A7* and *CYP3A43*) each of which contains 13 exons. Although the four *CYP3A* genes are differently regulated, they share 70-90% of metabolic properties.[199,200] Nevertheless, few studies have investigated the role of *CYP3A* in breast cancer and only two point at the polymorphic expression of these enzymes by exploring the association between *CYP3A4*1B* and breast cancer incidence [209] or age at onset of menarche [237].

The objective of this study was to investigate the role of polymorphic CYP3A expression in the regulation of multiple steroid hormone serum levels in postmenopausal females and to study the association between *CYP3A* gene variation and proxies of reproductive life and the incidence and mortality of breast cancer.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population-based prospective cohort study that started with a baseline interview between July 1989 and July 1993. All inhabitants of the Rotterdam suburb Ommoord, aged 55 years and older, were invited. Of the 10,275 eligible subjects, 7983 (78%) agreed to participate.[6] The Medical Ethics Committee of the Erasmus Medical Center approved the study.

Participants were visited at home at the start of the study for a standardised interview on health state and socio-economic characteristics. Subsequently, an extensive physical examination and blood assessment followed at the research center. Since its start, cross-sectional surveys have been carried out periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Information on vital status is obtained regularly

from municipal health authorities in Rotterdam and from the general practitioners working in the study district.

Cohort definition

Postmenopausal females who provided a blood sample and for whom genotypes were available (n_{cyp3a4} = 3707, n_{cyp3a5} = 3695, n_{cyp3a7} = 3606) were included. For analyses on steroid hormone level, females who used any form of hormone replacement therapy at the time of blood sampling (n = 97) were excluded. Information on all three *CYP3A* genotypes was available for 3385 of females. For the analyses on breast cancer, females with a prevalent breast cancer (n = 79) were excluded. Information on all three *CYP3A* genotypes was available for 3414 of females. FIGURE 1 presents the numbers of individuals in each cohort. Steroid hormone levels were measured in different random subsets of these females (see TABLE 2 for the individual numbers).

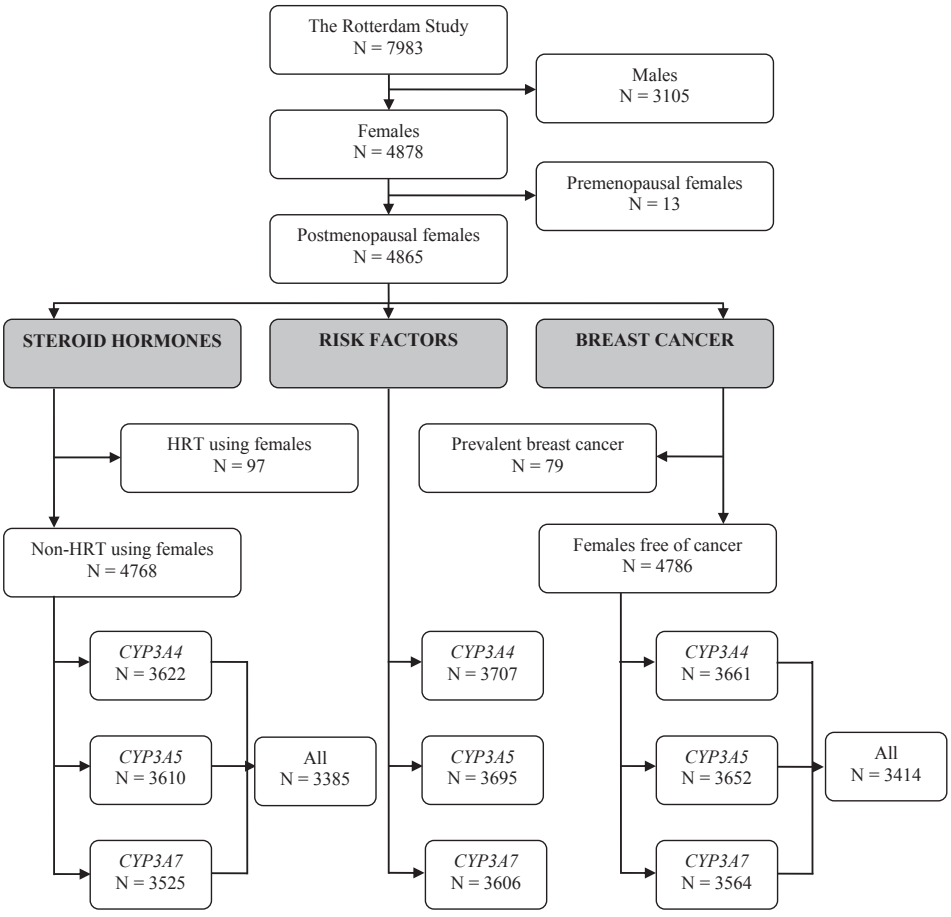

Figure 1. Study cohorts

Table 1. Baseline characteristics

Characteristic	Postmenopausal female N = 4865
Mean age (SD), yrs	71.7 (10.3)
Mean body mass index (SD), kg/m²	26.7 (4.1)
Mean pack years (SD), py	9.0 (16.6)
Smoking status, N (%)	
never	2542 (52.3%)
past	1242 (25.5%)
current	804 (16.5%)
missing	277 (5.7%)
Physical activity, N (%)	
no difficulty	2193 (45.1%)
with some difficulty	1100 (22.6%)
with much difficulty	408 (8.4%)
unable to do	929 (19.1%)
missing	235 (4.8%)
Mean total energy intake (SD), kcal/day	1786.5 (403.8)
Mean alcohol intake (SD), gram/day	6.1 (10.0)
Mean vitamin A intake (SD), mg/day	0.8 (0.4)
Mean vegetable intake (SD), gram/day	329.2 (139.1)
Hormone replacement therapy, N (%)	97 (2.0%)
Corticosteroid use, N (%)	95 (2.0%)
Mean age at onset menarche (SD), yrs	13.7 (1.8)
Mean age at onset menopause (SD), yrs	48.9 (5.2)
Mean years since menopause (SD), yrs	21.6 (10.9)
Mean number of children (SD), N	2.1 (1.7)
Time of blood assessment (range), hrs:min	11:41 (07:40 – 18:57)
Genotypes	
CYP3A4 [a] AA	3448
AG	254
GG	5
CYP3A5 [b] AA	23
AG	490
GG	3182
CYP3A7 [c] TT	3263
TG	336
GG	7

N = number of persons, SD = standard deviation. Missings were imputed by an expectation maximization algorithm. None of the participants used pituitary or hypothalamic hormones. HWE: [a] *CYP3A4*: $\chi^2 = 0.02$, p = 0.9; [b] *CYP3A5*: $\chi^2 = 0.7$, p = 0.38; [c] *CYP3A7*: $\chi 2 = 0.29$, p = 0.6

Table 2. Sex-hormone levels per *CYP3A* genotype in non-HRT using postmenopausal females

	CYP3A4			*CYP3A5*			*CYP3A7*		
N model 1 / 2 / 3 / 4	**AA**	**AG + GG**	***p***	**AA+AG**	**GG**	***p***	**TT**	**TG + GG**	***p***
DHEA (µmol/L) N	296 / 277 / 285 / 266	19 / 18 / 18 / 17		30 / 30 / 30 / 30	288 / 265 / 276 / 253		274 / 260 / 264 / 250	37 / 35 / 35 / 33	
Model 1	11.0 (0.4)	12.0 (1.5)	0.50	12.8 (1.1)	10.7 (0.4)	0.08	11.1 (0.4)	10.5 (1.1)	0.73
Model 2	11.1 (0.4)	11.1 (1.7)	0.97	12.9 (1.3)	10.9 (0.4)	0.15	11.0 (0.4)	11.7 (1.1)	0.54
Model 3	10.9 (0.3)	12.9 (1.3)	0.13	12.9 (1.0)	10.6 (0.3)	0.03	11.0 (0.3)	11.9 (0.9)	0.32
Model 4	11.0 (0.3)	12.1 (1.5)	0.46	11.8 (1.4)	10.3 (0.4)	0.46	10.9 (0.3)	12.1 (0.9)	0.46
DHEA sulfate (µmol/L) N	838 / 787 / 285 / 266	67 / 61 / 18 / 17		114 / 109 / 30 / 30	789 / 739 / 276 / 253		785 / 751 / 264 / 250	103 / 97 / 35 / 33	
Model 1	2.7 (0.1)	2.6 (0.2)	0.65	2.6 (0.2)	2.7 (0.1)	0.63	2.7 (0.1)	2.3 (0.2)	0.03
Model 2	2.7 (0.1)	2.7 (0.3)	0.91	2.6 (0.2)	2.7 (0.1)	0.71	2.7 (0.1)	2.3 (0.2)	0.05
Model 3	3.1(0.1)	2.3(0.4)	0.09	2.6 (0.3)	3.0 (0.1)	0.19	3.1 (0.1)	2.8 (0.3)	0.32
Model 4	3.1(0.1)	2.4 (0.5)	0.17	2.9 (0.4)	3.0 (0.1)	0.17	3.1 (0.1)	2.7 (0.3)	0.17
Androstenedione (nmol/L) N	810 / 760 / 286 / 268	66 / 60 / 19 /18		106 / 100 / 29 / 29	767 / 720 / 279 / 257		755 / 722 / 265 / 252	104 / 98 / 36 / 34	
Model 1	3.6 (0.1)	3.3 (0.2)	0.21	3.4 (0.2)	3.6 (0.1)	0.21	3.6 (0.1)	3.7 (0.2)	0.66
Model 2	3.6 (0.1)	3.5 (0.3)	0.76	3.4 (0.2)	3.6 (0.1)	0.47	3.6 (0.1)	3.7 (0.2)	0.54
Model 3	4.4 (0.1)	4.2(0.4)	0.69	4.0 (0.3)	4.4 (0.1)	0.25	4.3 (0.1)	4.6 (0.3)	0.48
Model 4	4.4 (0.1)	4.6 (0.5)	0.71	4.0 (0.4)	4.4 (0.1)	0.71	4.4 (0.1)	4.6 (0.3)	0.71
Testosterone (nmol/L) N	839 / 788 / 753 / 705	65 / 59 / 60 / 54		113 / 107 / 100 / 94	789 / 740 / 710 / 665		778 / 745 / 698 / 667	108 / 102 / 98 / 92	
Model 1	1.4 (0.0)	1.3 (0.1)	0.59	1.4 (0.1)	1.4 (0.0)	0.88	1.4 (0.0)	1.3 (0.1)	0.37
Model 2	1.4 (0.0)	1.4 (0.1)	0.91	1.4 (0.1)	1.4 (0.0)	0.92	1.4 (0.0)	1.3 (0.1)	0.41
Model 3	1.3 (0.0)	1.3 (0.1)	0.51	1.3 (0.1)	1.3 (0.0)	1.00	1.3 (0.0)	1.2 (0.1)	0.47
Model 4	1.3 (0.0)	1.4 (0.1)	0.27	1.2 (0.1)	1.3 (0.0)	0.27	1.3 (0.0)	1.2 (0.1)	0.27
Free testosterone (nmol/L) N	671 / 630 / 658 / 619	49 / 43 / 48 / 42		85 / 80 / 83 / 78	638 / 593 / 626 / 583		622 / 593 / 610 / 582	85 / 80 / 84 / 79	
Model 1	2.6 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.89	2.6 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.80	2.6 • 10-2 (0.0)	2.5 • 10-2 (0.0)	0.88
Model 2	2.6 • 10-2 (0.0)	2.7 • 10-2 (0.0)	0.56	2.5 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.74	2.6 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.92
Model 3	2.5 • 10-2 (0.0)	2.7 • 10-2 (0.0)	0.47	2.6 • 10-2 (0.0)	2.5 • 10-2 (0.0)	0.75	2.5 • 10-2 (0.0)	2.5 • 10-2 (0.0)	0.97
Model 4	2.5 • 10-2 (0.0)	2.8 • 10-2 (0.0)	0.31	2.4 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.31	2.5 • 10-2 (0.0)	2.6 • 10-2 (0.0)	0.31
Estrone (pmol/L) N	773 / 732 / 685 / 645	64 / 59 / 58 / 53		110 / 104 / 95 / 89	727 / 687 / 646 / 609		740 / 710 / 655 / 626	85 / 81 / 78 / 72	
Model 1	48.2 (1.3)	50.0 (4.6)	0.71	47.5 (3.5)	48.9 (1.4)	0.71	49.5 (1.4)	41.1 (4.0)	0.05
Model 2	48.3 (1.4)	54.5 (5.8)	0.31	45.1 (4.3)	49.3 (1.5)	0.38	49.6 (1.4)	41.1 (4.1)	0.05
Model 3	46.0 (1.3)	51.6 (4.5)	0.23	46.3 (3.6)	46.8 (1.4)	0.90	47.4 (1.4)	40.0 (4.0)	0.09
Model 4	45.8 (1.4)	58.8 (5.8)	0.07	41.8 (4.5)	47.4 (1.5)	0.07	47.4 (1.4)	40.1 (4.1)	0.07
Estrone sulfate (pmol/L) N	692 / 655 / 615 / 583	60 / 55 / 55 / 51		100 / 95 / 96 / 91	652 / 615 / 575 / 543		665 / 636 / 597 / 572	77 / 74 / 65 / 62	
Model 1	176.0 (7.8)	170.6 (26.6)	0.84	180.9 (20.6)	175.3 (8.0)	0.80	183.0 (8.0)	134.4 (23.8)	0.05
Model 2	178.7 (8.3)	171.5 (34.5)	0.84	182.9 (26.0)	177.4 (8.7)	0.85	183.2 (8.3)	134.6 (24.6)	0.06
Model 3	185.3 (6.5)	166.1 (21.8)	0.40	184.6 (16.6)	184.6 (6.8)	1.00	189.4 (6.6)	160.3 (20.3)	0.18
Model 4	189.4 (6.9)	161.0 (27.5)	0.33	195.7 (20.4)	185.7 (7.3)	0.33	190.0 (6.9)	160.9 (21.0)	0.33
Estradiol (pmol/L) N	575 / 554 / 532 / 511	44 / 40 / 42 / 38		77 / 73 / 73 / 69	545 / 521 / 504 / 480		552 / 532 / 512 / 492	64 / 62 / 59 / 57	
Model 1	18.8 (0.9)	21.1 (3.2)	0.48	18.5 (2.4)	19.3 (0.9)	0.75	19.5 (0.9)	15.3 (2.6)	0.14
Model 2	18.6 (0.9)	24.5 (4.1)	0.17	16.2 (3.0)	19.4 (0.9)	0.33	19.5 (0.9)	14.8 (2.7)	0.10
Model 3	19.0 (0.9)	21.7 (3.3)	0.41	18.8 (2.5)	19.5 (0.9)	0.79	19.6 (0.9)	16.1 (2.8)	0.23
Model 4	18.8 (1.0)	25.1 (4.3)	0.16	16.3 (3.2)	19.6 (1.0)	0.16	19.6 (0.9)	15.5 (2.8)	0.16
Free estradiol (pmol/L) N	428 / 411 / 427 / 410	33 / 29 / 33 / 29		55 / 51 / 55 / 51	409 / 389 / 408 / 388		411 / 394 / 410 / 393	48 / 46 / 48 / 46	
Model 1	0.5 (0.0)	0.6 (0.1)	0.32	0.5 (0.1)	0.5 (0.0)	0.72	0.5 (0.0)	0.4 (0.1)	0.21
Model 2	0.5 (0.0)	0.7 (0.1)	0.08	0.4 (0.1)	0.5 (0.0)	0.22	0.5 (0.0)	0.4 (0.1)	0.16
Model 3	0.5 (0.0)	0.6 (0.1)	0.33	0.5 (0.1)	0.5 (0.0)	0.62	0.5 (0.0)	0.4 (0.1)	0.21
Model 4	0.5 (0.0)	0.7 (0.1)	0.08	0.4 (0.1)	0.5 (0.0)	0.08	0.5 (0.0)	0.4 (0.1)	0.08
Cortisol (nmol/L) N	373 / 343	26 / 23		44 / 42	355 / 324		337 / 319	51 / 47	
Model 1	356.8 (6.9)	349.4 (26.3)	0.79	358.1 (20.2)	356.1 (7.0)	0.93	358.7 (7.5)	345.4 (19.8)	0.41
Model 2	356.7 (7.3)	360.4 (32.1)	0.91	363.8 (23.8)	356.1 (7.5)	0.91	359.5 (8.8)	362.0 (23.3)	0.91
Cortisol / DHEAS ratio N	353 / 324	24 / 21		43 / 41	335 / 304		321 / 303	46 / 42	
Model 1	250.8 (13.7)	250.3 (52.9)	0.99	261.0 (40.0)	251.9 (14.1)	0.84	244.0 (14.4)	294.5 (38.4)	0.22
Model 2	248.2 (14.6)	236.2 (65.7)	0.86	262.8 (47.1)	245.4 (15.3)	0.86	245.2 (15.0)	263.6 (40.7)	0.86

N = number of persons included in the model, DHEA = dehydroepiandrosterone, DHEAS = dehydroepiandrosterone sulphate, HRT = hormone replacement therapy. Values are expressed as means (standard error). Precursor hormone [DHEAS: DHEA], [DHEA: DHEAS], [androstenedione: DHEA], [(free) testosterone: androstenedione], [estrone: androstenedione], [estradiol: estrone, testosterone], [free estradiol: estrone, testosterone], [estrone sulfate: estrone]. Model 1: adjusted for age, body mass index, smoking, alcohol intake, physical activity, corticosteroid use, years since menopause, and time of blood assessment. Model 2: adjusted for age, body mass index, smoking, alcohol intake, physical activity, corticosteroid use, years since menopause, time of blood assessment and other *CYP3A* genotypes. Model 3: adjusted for age, body mass index, smoking, alcohol intake, physical activity, corticosteroid use, years since menopause, time of blood assessment and precursor hormone. Model 4: adjusted for age, body mass index, smoking, alcohol intake, physical activity, corticosteroid use, years since menopause, time of blood assessment, precursor hormone, and other *CYP3A* genotypes.

Exposure definition

Genotyping

Genotyping of the *CYP3A4*1B* (rs2740574), *CYP3A5*3* (rs776746), and *CYP3A7*1C* (rs11568825) was determined using TaqMan allelic discrimination assays on an ABI Prism 9700 HT sequence detection system.[20] Each assay consisted of two allele-specific minor groove binding (MGB) probes, labeled with the fluorescent dyes VIC and FAM. The primer and probe sequences, designed by Applied Biosystems by their Assay-by-Design service, are listed in TABLE 1. Polymerase chain reactions (PCR) were performed in a reaction volume of 2.0 µl, containing assay-specific primers, allele-specific TaqMan MGB probes, TaqMan Universal PCR Master Mix No AmpErase UNG (2X) and genomic DNA(1 ng). The thermal profile consists of an initial denaturation step at 95°C for 15 minutes, followed by 40 cycles of denaturation at 92°C for 15 seconds, and annealing and extension at 60°C for 1 minute. Genotypes were scored by measuring allelic-specific fluorescence using the SDS 2.2.2 software for allelic discrimination.[20]

Outcome definition

This study focussed on three different types of outcome: (1) steroid hormone levels, (2) measures of reproductive life such as ages at menarche or menopause and estimated total estrogen-exposure period, and (3) breast cancer incidence and mortality.

Hormone measurements

Blood samples were drawn by venapuncture from non-fasting subjects and collected in 5-ml tubes containing 0.5 ml sodium citrate solution. All tubes were stored on ice before and after blood sampling. Levels of steroid hormones were measured in plasma. Platelet-free plasma was obtained by two-stage centrifugation, first for 10 min at 1600 x *g* at 4°C and then for 30 min at 7000 x *g* at 4°C. Platelet-free samples were immediately frozen in liquid nitrogen and transferred to the laboratory where they were stored at -80°C until hormone estimations were performed. Plasma levels of dehydroepiandrosterone (DHEA), DHEA sulphate (DHEAS), androstenedione, testosterone (T), estrone (E1), estrone sulphate (E1S), estradiol (E2), cortisol and sex steroid hormone binding globulin (SHBG) were estimated in 12 separate batches of samples using coated tube or double antibody RIAs, purchased from Diagnostic Systems

Laboratories.[214] Because of the relatively small volumes of plasma available, all values reported are single sample estimations. Intra-assay coefficients of variation, determined on the basis of duplicate results of internal quality control pools with three different levels of each analyte, were below 15% for all assays. Because inter-assay variations were relatively large (33% DHEA, 30% DHEAS, 33% androstenedione, 19% testosterone, 18% estrone, 29% estrone sulphate, 25% estradiol, 24% cortisol and 14% SHBG), we multiplied all concentrations within a batch with a factor, which made results for the internal quality control pools comparable. This reduced inter-assay variations to 3%, 10%, 8%, 7%, 18%, 7%, 14%, 7%, and 3%, respectively, and was considered justified because the relative differences in the results for the high and middle internal quality control pools per batch were comparable, as evidenced by correlations between these results (r = 0.99 for DHEA, r = 0.58 for DHEAS, r = 0.91 for androstenedione, r = 0.75 for testosterone, r = 0.60 for estrone, r = 0.93 for estrone sulphate, r = 0.21 for estradiol, r = 0.93 for cortisol and r = 0.91 for SHBG; all $P < 0.05$, except for E2). Assays were performed blind with respect to information on the subject. As a measure of bioavailable testosterone and estradiol, non-SHBG-bound hormone was calculated on the basis of hormone, SHBG, and albumin levels, and respective affinity constants according to the method described by Södergård *et al.* [215] and van den Beld *et al.* [216]. We additionally investigated the ratio of free cortisol over DHEAS, because DHEAS may antagonize the actions of cortisol.[217-219]

Age at menarche and menopause

Female participants were asked to report the ages of menarche and menopause. We additionally calculated the total estrogen exposure period by subtracting the age at onset of menarche from the age at onset of menopause per female individual.

Breast cancer case identification and validation

Three different databases were used for breast cancer case identification. First, cancer cases diagnosed by general practitioners in the research area were collected (International Classification of Primary Care (X76)). Second, the national registry of all hospital admissions was consulted to detect all malignancy related hospital admissions for study participants. Third, regional pathology databases were linked to The Rotterdam Study to identify cases (International Classification of Diseases-10 (C50)). A physician (CS) subsequently validated breast cancer cases on the basis of medical records of the general practitioner, discharge letters, and pathology reports. Only pathologically confirmed cases were considered in the analyses. The index date was defined as the earliest date found in the pathology reports.

Covariates

For analyses on hormone levels, the following potential confounders were considered: age, body mass index (BMI, kg/m^2), smoking status (never, current, past), alcohol intake (grams/

day), ability to be physically active, corticosteroid use, years since menopause, precursor hormones, *CYP3A* status, and time of blood collection (in two hour intervals).

For analyses on breast cancer, age, BMI, smoking status and number of pack years, alcohol, vitamin A, vegetables and total energy intake, hormone replacement therapy, ability to be physically active, age at menarche and menopause, number of children and other *CYP3A* genotypes were considered to be potential confounders. Age at diagnosis was considered to be a confounder in the breast cancer specific mortality analyses.

Statistical analyses

First, we compared characteristics of participants with and without blood samples using a Student's t-test for continuous variables and a χ^2-test for dichotomous variables. Hardy Weinberg Equilibria (HWE) were calculated with a χ^2-goodness-of-fit statistic. Genotypes were dichotomised. Heterozygous persons were combined with homozygous carriage of the minor allele (*CYP3A4*: AA *vs* AG + GG, *CYP3A5*: AA + AG *vs* GG, *CYP3A7*: TT *vs* TG + GG) to optimise power. Missing indicators were used to study the effect of missing values. Missing data on covariates were imputed by single imputation using the Expectation Maximization (EM) algorithm.[220] We anticipated that the level of precursor hormone influences hormone levels. Therefore, the following adjustments for precursor hormones (when available) were considered: DHEA sulphate (DHEAS) adjusted for DHEA, DHEA adjusted for DHEA sulphate, androstenedione adjusted for DHEA, (free) testosterone adjusted for androstenedione, estrone (E1) adjusted for androstenedione, estrone sulfate (E1SO4) adjusted for estrone, (free) estradiol (E2) adjusted for estrone and testosterone. To avoid confounding by other *CYP3A* genotypes, the remainder of the determined genotypes was included in the statistical model to explore the *CYP3A* specific genotype effects. No potential confounders were included in the analyses on age at onset of menarche and menopause. Since none of the covariates changed the estimates by more than 10%, all potential confounders were included in the following models. Age-, BMI-, smoking-, alcohol-, physical activity-, corticosteroid use-, years since menopause-, time of blood assessment (model 1), model 1 covariates with other *CYP3A* genotype- (model 2), model 1 covariates with precursor hormone level- (model 3) and model 1 covariates with both precursor hormone level- and other *CYP3A* genotype- (model 4) adjusted general linear models were used to analyse differences in mean (± standard error of the mean (SEM)) hormone serum levels between genotypes. For cortisol levels no adjustments were made for precursor hormones. Student's t-tests were used to calculate differences in mean age at menarche and menopause and mean estrogen exposure periods (age at menopause – age at menarche) between genotypes. Cox' proportional hazard models were used to analyse age-, BMI-, smoking-, alcohol-, vitamin A-, vegetable-, and total energy intake-, physical activity-, age at menarche-, age at menopause, number of children-, hormone replacement therapy-, corticosteroid use-, (model 1) and other *CYP3A* genotype- (model 2) adjusted breast cancer risks per genotype. Finally, age at diagnosis- (model 1) and other *CYP3A* genotype- (model 2)

adjusted 'overall', 'breast cancer specific', 'early onset (diagnosis breast cancer < 65 years)' and 'late onset (diagnosis breast cancer ≥ 65 years)' mortality was studied per genotype. P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant. All analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA)

RESULTS

Individuals of whom the genotype was unknown were on average older, relatively more frequently smoker, and had a shorter follow-up time than those for whom genotype data were available. TABLE 1 shows the baseline characteristics of all postmenopausal female participants. The three *CYP3A* SNPs were in HWE.

Borderline but non-significantly increased levels of estrone (58.8 ± 5.8 *vs* 45.8 ± 1.4 pmol/L) and free estradiol (0.7 ± 0.1 *vs* 0.5 ± 0.0 nmol/L) and decreased levels of DHEA sulphate (2.3 ± 0.4 *vs* 3.1 ± 0.1 µmol/L) were found in *CYP3A4* variant carriers (AG and GG) compared to non-variant carriers (AA) (TABLE 2). For *CYP3A5* similar results were found for estrone (47.4 ± 1.5 *vs* 41.8 ± 4.5 pmol/L) and free estradiol (0.5 ± 0.0 *vs* 0.4 ± 0.1 pmol/L). Levels were higher in women homozygous for GG than in those with no or only one G-allele (AA and AG). Homozygous G-allele carriers had statistically significantly lower levels of DHEA compared to those with AA or AG (10.6 ± 0.3 *vs* 12.9 ± 1.0 µmol/L, respectively). G-allele carriage of the *CYP3A7* gene was significantly related to decreased levels of DHEA sulphate (2.3 ± 0.2 *vs* 2.7 ± 0.1 µmol/L, respectively), estrone (41.1 ± 4.1 *vs* 49.6 ± 1.4 pmol/L, respectively), estrone sulphate (134.4 ± 23.8 *vs* 183.0 ± 8.0 pmol/L, respectively) and borderline non-significantly

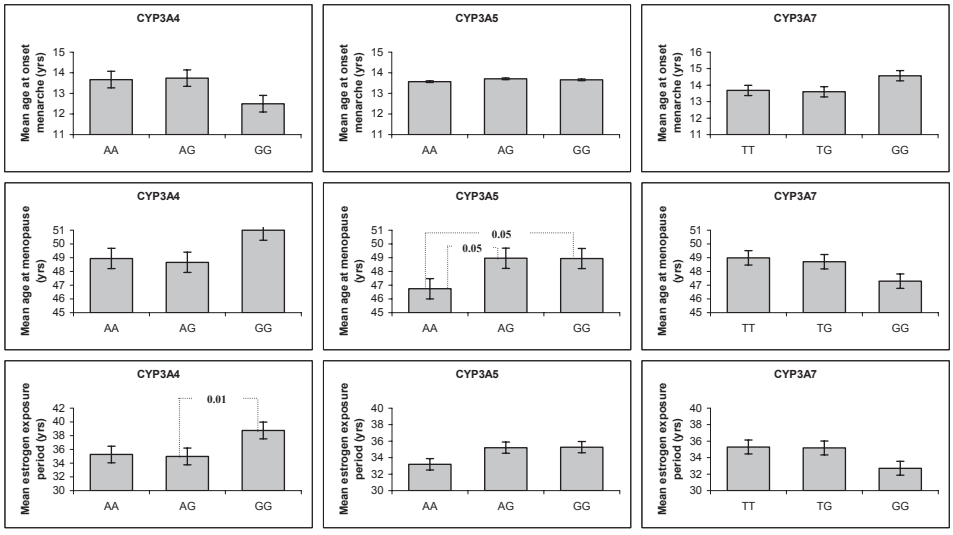

Figure 2. Mean 'age at onset of menarche', 'age at onset of menopause' and 'estrogen exposure period' per *CYP3A* genotype

to free estradiol (0.4 ± 0.1 *vs* 0.5 ± 0.0 pmol/L, respectively) compared to women who were homozygous for the T-allele. However, results were strongly dependent on the model used. Other steroid hormone levels seemed not to be related with *CYP3A* genotypes.

CYP3A gene variation was associated with risk factors for breast cancer such as age at onset of menarche and menopause and the total estrogen-exposure period (FIGURE 2). Females who carried two copies of the *CYP3A4* G-allele seemed to be younger at onset of menarche (GG; 12.5 ± 0.6 *vs* AG; 13.7 ± 0.1 and AA; 13.7 ± 0.0 years) and older at onset of menopause (GG; 51.0 ± 1.0 *vs* AG; 48.7 ± 0.3 and AA; 48.9 ± 0.1 years) compared to persons with at least one wild type allele. Although these single differences did not reach statistical significance, the calculated total period of estrogen exposure showed a significant increase in cumulative exposure time ($p = 0.01$) of almost 4 years for homozygous *CYP3A4* G-allele carriers in contrast to heterozygotes. *CYP3A5* genotype was not associated with onset of menarche, but menopause was found to appear 2.5 years earlier in AA-genotype individuals compared to both heterozygous ($p = 0.05$) and homozygous ($p = 0.05$) G-allele carriers. Exposure periods were not statistically significantly different. None of the periods was associated with *CYP3A7* genotype. Nevertheless, homozygosity of the G-allele resulted in a non-significant older age at menarche, a younger age at menopause and consequently a shorter period of estrogen exposure. Breast cancer incidence (TABLE 3) and mortality (TABLE 4) seemed not to be affected by *CYP3A* genotype.

Table 3. Hazard ratios of breast cancer incidence by *CYP3A* genotype

				HR (95% CI)		
		Genotype	**N cases/ cohort**	**Model 1**	**Model 2**	**Model 3**
CYP3A4	Total sample	AA	133 / 3183	1.00 (reference)	1.00 (reference)	1.00 (reference)
		AG + GG	9 / 231	0.92 (0.47-1.81)	0.89 (0.45-1.74)	0.83 (0.37-1.84)
	Early onset	AA	27 / 3079	1.00 (reference)	1.00 (reference)	1.00 (reference)
	(< 65 years)	AG + GG	1 / 223	0.46 (0.06-3.37)	0.44 (0.06-3.25)	0.82 (0.06-11.52)
CYP3A5	Total sample	GG	121 / 2941	1.00 (reference)	1.00 (reference)	1.00 (reference)
		AA + AG	21 / 473	0.95 (0.60-1.51)	0.99 (0.62-1.58)	0.93 (0.54-1.61)
	Early onset	GG	26 / 2847	1.00 (reference)	1.00 (reference)	1.00 (reference)
	(< 65 years)	AA + AG	2 / 455	2.28 (0.54-9.38)	2.21 (0.52-9.38)	2.01 (0.30-13.48)
CYP3A7	Total sample	TT	131 / 3090	1.00 (reference)	1.00 (reference)	1.00 (reference)
		TG + GG	11 / 317	0.77 (0.41-1.42)	0.78 (0.42-1.44)	0.77 (0.42-1.43)
	Early onset	TT	25 / 2986	1.00 (reference)	1.00 (reference)	1.00 (reference)
	(< 65 years)	TG + GG	3 / 316	1.00 (0.30-3.31)	1.02 (0.30-3.44)	0.97 (0.29-3.28)

HR = hazard ratio, CI = confidence interval. Model 1: adjusted for age. Model 2: adjusted for age, BMI, smoking status, pack years of smoking, alcohol intake, age at menarche, age at menopause, number of children, physical activity, HRT use, corticosteroid use, vitamin A intake, vegetable intake, and total energy intake. Model 3: adjusted for age, BMI, smoking status, pack years of smoking, alcohol intake, age at menarche, age at menopause, number of children, physical activity, HRT use, corticosteroid use, vitamin A intake, vegetable intake, total energy intake, and other *CYP3A* genotype.

Table 4. Association between *CYP3A* genotype and mortality

		Mortality			
		HR (95% CI)			
		All cause	All breast cancer	Early onset (< 65 years) breast cancer	Late onset (≥ 65 years) breast cancer
*CYP3A4*1B*	Model 1	0.66 (0.28-1.55)	1.00 (0.36-2.77)	NA	1.10 (0.39-3.06)
	Model 2	1.16 (0.34-3.99)	1.32 (0.38-4.53)	NA	1.41 (0.41-4.83)
*CYP3A5*1*	Model 1	1.88 (0.93-3.81)	1.24 (0.56-2.77)	NA	1.13 (0.50-2.53)
	Model 2	1.97 (0.75-5.19)	1.36 (0.52-3.58)	NA	1.28 (0.49-3.38)
*CYP3A7*1C*	Model 1	1.15 (0.57-2.34)	1.10 (0.44-2.78)	1.98 (0.17-23.22)	0.94 (0.34-2.62)
	Model 2	1.35 (0.63-2.91)	1.18 (0.47-2.97)	1.50 (0.12-18.35)	1.00 (0.36-2.80)

HR = hazard ratio, CI = confidence interval, NA = not applicable. Model 1: adjusted for age at diagnosis. Model 2: adjusted for age at diagnosis and other *CYP3A* genotypes.

DISCUSSION

In this prospective population-based cohort study on the association between *CYP3A* genotypes and three types of outcome, we encountered several interesting findings. First, significant associations were found between *CYP3A5* homozygous G-allele carriage and decreased DHEA serum levels, and between *CYP3A7* G-allele carriage and decreased levels of DHEA sulphate, estrone, and estrone sulphate. Although not significant, levels of other hormones seemed to be influenced by *CYP3A* genotype as well. Second, the *CYP3A4* GG genotype was associated with a longer estrogen-exposure period compared to any of the other *CYP3A4* genotypes and the *CYP3A5* AA genotype was associated with a younger age at menopause compared to the *CYP3A5* AG and GG genotypes. Third, despite these hypothesis-strengthening findings that some *CYP3A* genotypes are associated with risk factors for breast cancer, no associations of any of the *CYP3A* genotypes and breast cancer incidence or mortality were found. Since no data were available for the *CYP3A43* gene, polymorphic expression of CYP3A43 was not included.

Although the role of CYP3A enzymes in steroid hormone metabolism is commonly accepted, studies on the effect of genetic variation in genes encoding for these enzymes are scanty. Studies mainly focus on testosterone levels and comprise an *in vitro* setting. The current study not only included multiple *CYP3A* genes and its most commonly studied polymorphisms but also investigated several steroid hormone levels, all within an *in vivo* setting. The biological function of polymorphic expression of CYP3A4 induced by the **1B* gene polymorphism (G-allele) is controversial and preferably studied for its impact on testosterone metabolism. While epidemiologic studies hypothesized that the polymorphic form would result in lower metabolic rates [89,90,174-176,205], some *in vitro* studies found an almost 3-fold increased 6β-hydroxylase activity converting testosterone to its less active metabolites [212]. In the present study, no

effects of *CYP3A* genotype on testosterone or free testosterone levels were found. That *in vivo* data do not confirm the *in vitro* results is most probably explained by the role of the counterbalancing pituitary-adrenal (cortisol) pathways. In spite of a lack of evidence for genetic variation, consistency exists for the role of CYP3A in estrogen metabolism. The enzyme subfamily catalyzes the 2-, 4-, 16α- and 17β-hydroxylations of estrone and estradiol.[203,235,236] Results of the current study show, although not statistically significantly different, increased levels of estrone, estradiol and free estradiol and decreased levels of estrone sulphate in *CYP3A4* variant carriers. These findings endorse the role of CYP3A4 enzymes in estrogen metabolism and indicate a decreased activity in variant carriers. A previous study observed moderate effects of *CYP3A5* polymorphisms on estrogens, which might explain our less pronounced findings for the gene encoding for this enzyme.[235] Finally, the results found for *CYP3A7* were mainly similar to those of *CYP3A4* and *CYP3A5*, with exclusion of the results for estradiol that leads to lower levels in variant carriers. As shown by Lee *et al.*, CYP3A7 shares the catalytic activity of the CYP3A subfamily for 16α-hydroxylase of estrone which might be an explanation for the finding in our study.[203] Overall, studies on CYPs are highly influenced by a large number of inducers and inhibitors, which might have overruled or antagonized the influences of genetic variation. Moreover, in this study we used several models to explore differences between *CYP3A* genotypes. This revealed more insight into the effects of confounders such as precursor hormone level and the remainder of *CYP3A* gene variation. As demonstrated by these results, adjustments of hormone levels without a circadian rhythm for precursor hormones with a circadian rhythm resulted in dilution or vanishing of the differences observed in a previous model. For these particular hormones (DHEA sulphate and estrone sulphate), models without adjustments for precursor hormones represent probably more adequate results than models that include these levels. Although adjustment for other *CYP3A* genotypes reveals more insight into the *CYP3A* specific genotype effects, one should recognize that only a minority of genetic variation within each gene was considered and that potential influences on both hormone levels and disease outcomes are probably explained by a combination of *CYP3A* gene variations and/or other gene variations. This would argue for a haplotype or genome-wide association approach. Next to this, combining heterozygous persons with those homozygous for the minor allele, might have led to a dilution of the effect as measures of reproductive life (**FIGURE 2**) show some significant differences for homozygous minor allele carriers. However, for hormone levels low numbers forced us to combine groups. In a previous study on *CYP3A* genotype and steroid hormone levels in males, we found highly significant associations of primarily *CYP3A7* genotype and serum levels of DHEA sulphate, androstenedione, estrone and estrone sulphate (unpublished data). The current study shows that for females these results are only to a certain part similar and less pronounced. Gender-related differences in the metabolism of xenobiotics are commonly encountered in traditional laboratory rodents [238], but are debatable in humans [239-242]. However, recent data indicate regulation of expression of CYP3As by sex hormones.[243,244] For cytochrome P450 enzymes, the expression level

is normally regulated by constitutive androstane receptors (CAR) and pregnane X-receptors (PXR). Sex hormones bind as ligands to these receptors and stimulate (de)activation of the promotor.[245,246] As a consequence of this mechanism, several studies observed higher activities of mainly CYP3A4 and CYP3A5 in females, since expression levels were not inhibited by androgens.[239-242] As males have lower expression levels, consequences of genotype differences might have more impact on these levels than in females with high expression levels.

A previous study on 137 healthy nine-year old American girls (of whom 41 of Caucasian descent) reported an allele-dose dependent association between the *CYP3A4* G-allele (**1B*) and onset of puberty, defined as a Tanner breast development stage of ≥ T2B. 90% of girls homozygous for the G-allele had reached puberty in contrast to 56% or 40% in heterozygous or homozygous wild type carriers, respectively.[237] This observation is underlined by results from the current study. Females with the *CYP3A4* GG genotype reported a (although not statistically significant) younger age at menarche. Additionally, they reported higher ages at menopause, which resulted in a statistically significantly longer estrogen-exposure period for these females, compared to AA or AG genotype carriers. Furthermore, our study shows statistically significantly higher ages at menopause in G-allele carriers of *CYP3A5*. Although the observations can directly be explained by our observation of increased estrone and estradiol levels in *CYP3A4* G-allele carriers and *CYP3A5* homozygous G-allele carriers, it remains uncertain what the impact of the genotypes is on steroid hormone levels in younger females.

Despite the moderate evidence found for effects of some *CYP3A* genotypes on steroid hormone levels, including estrogens, and established risk factors of breast cancer such as ages at menarche and menopause, no association with breast cancer incidence or mortality was observed in the current study. Although this is in line with results from a previous study on a large group of almost 1000 breast cancer cases [209], future research might stratify on estrogen- (ER) and progesterone-receptor (PR) status in tumor tissue. The effect of estrogens on ER or PR positive breast cancers might be different and could as yet reveal hormone-*CYP3A* gene-disease associations.

Observational studies are prone to selection, information and confounding bias. For this study persons without a blood sample, who were on average older, were excluded. Although aging is associated with at least some of the hormone levels [226] and with breast cancer, it was not associated with genotype as all genes were in Hardy Weinberg equilibrium. Consequently, selection bias is unlikely. Information bias seems also unlikely, as both exposure and outcome measurements were performed much earlier and prior to formulating the research question. However, the inter-individual variation of steroid hormone levels and the self-reported ages at menarche and menopause could have introduced some misclassification. Nevertheless, this misclassification is most probably non-differential and irrespective of genotype. The use of pathologically confirmed cases might have slightly diminished a genuine association if females with an undetected breast malignancy served as controls. Confounding for most accepted risk factors was adjusted for and hormone levels were explored by using different models.

In conclusion, some associations between *CYP3A* genotypes and steroid hormone levels were observed. However, results are less pronounced than we observed in a previous study among males, possibly as a consequence of gender-specific expression levels of CYP3A enzymes. Although age at onset of menarche and menopause and the calculated estrogen-exposure period, all commonly accepted risk factors for breast cancer, were related with some of the *CYP3A* genotypes, no association was observed for the genotypes and breast cancer incidence or mortality. Future studies might benefit from stratification on estrogen- or progesterone-receptor status in tumor tissue.

2.4. Cytochrome P450 (CYP) *2C9*2* and *3 variant alleles do not increase the beneficial effects of NSAIDs on colorectal cancer

ABSTRACT

Background: Because CYP2C9 enzymes are involved in NSAID metabolism, we investigated whether *CYP2C9*2* and **3* variant alleles, encoding for enzymes with lower activity, increased the protective effects of NSAIDs on colorectal cancer.

Methods: Individual and combined associations of NSAIDs and *CYP2C9*2* and **3* variant alleles with colorectal cancer were studied in 7757 Caucasian individuals of The Rotterdam Study, a population-based prospective cohort since 1990. Additive and multiplicative effect modification models were used to examine drug-gene interactions.

Results and conclusion: There were 212 incident cases of colorectal cancer during follow-up. A reduced risk of colorectal cancer was observed in individuals who used NSAIDs for more than a year (HR 0.45; 95%CI 0.28-0.71), and in carriers of a *CYP2C9* variant allele associated with lower enzymatic activity (HR 0.67; 95%CI 0.47-0.96). The combination of both determinants was associated with a further risk reduction but without synergy. This might indicate independent pathophysiological mechanisms.

INTRODUCTION

Colorectal cancer is the second leading cause of cancer related death in the western world [247] and is considered as the final stage of the sequence from adenoma to carcinoma by accumulation of genetic mutations in epithelial cells [248]. This may result from exposure to carcinogens and mutagens, which can be activated by xenobiotic-metabolizing enzymes.[249] Furthermore, tumor development is dependent on vascularisation, cell proliferation and apoptosis.

Cytochrome P450 (CYP) enzymes, that metabolise most endogenous and exogenous substrates, are mainly expressed in the human liver, but also in normal intestinal epithelium and in colon adenocarcinoma.[250-253] The CYP2C subfamily accounts for 20% of all CYP in the liver, CYP2C9 being the main isoform.[254] The CYP2C9 isoform is capable of activating specific carcinogens and is related to the formation of DNA adducts.[255] Besides their role in (de)toxification, there is a physiological role for CYP2C9 in the metabolism of arachidonic acid, forming epoxyeicosatrienoic acids through epoxygenase.[256,257] An increase of arachidonic acid metabolism that results in decreased arachidonic acid levels, induces neo-vascularisation and cell growth and inhibits apoptosis [258-261], conditions that might lead to tumor development. Additionally, CYP2C9 metabolizes non-steroidal anti-inflammatory drugs (NSAIDs), which have been associated with a decreased risk of colorectal cancer.[262,263] In individuals of Caucasian descent, two allelic variants (**2* and **3*) of the *CYP2C9* gene are relatively common. These **2* and **3* variants have been shown to result in a lower enzyme activity for several substrates, both in vitro and in vivo, compared to the activity of the wild-type allele.[81,251,264,265]

The objective of this cohort study was to investigate whether *CYP2C9*2* and **3* variant alleles are associated with an increase of the protective effect of NSAIDs on colorectal cancer risk in a Caucasian population.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population–based prospective cohort study among inhabitants of the Rotterdam suburb Ommoord. Between July 1989 and July 1993, all persons aged 55 years and older were invited (n=10,275). In total 7,983 subjects (78%) participated (4,878 women and 3,105 men). The design, ethical approval and rationale behind this study have been described earlier.[6]

At baseline, a home interview was performed followed by two visits at the research center for clinical examinations. Blood samples were collected and DNA isolated. Baseline data collection was performed from October 1990 to July 1993. Since then, participants have been re-examined periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with

files from general practitioners. Information on medication use is available for all participants since January 1991. The seven computerized pharmacies that cover the research area are linked to one network. In this way, the date of prescription, the total amount of drug units per prescription and the prescribed defined daily dosage (DDD) are available per drug defined by an Anatomical Therapeutical Chemical (ATC) code [10]. Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from general practitioners in the study district.

Cohort definition

Pharmacy data were available for 7857 (98%) subjects. Persons with a diagnosed colorectal cancer before January 1st 1991 (n = 48) or who died or were lost to follow up before this date (n = 52) were excluded from the analyses. This resulted in a study cohort of 7757 (97%) individuals. Genotypes were determined in 6378 persons for whom genetic material was isolated and the assay was successfully performed. Follow-up time was defined as the period between January 1st 1991 and a diagnosis of colorectal cancer, death, or the end of the study period on October 1st 2004 whichever came first.

Exposure definition

NSAID use

The exposure of interest included both non-aspirin and aspirin NSAIDs. The following drugs were used by study subjects: acetylsalicylic acid, carbasalate calcium, diflusinal, sulindac, nabumeton, naproxen, ibuprofen, diclofenac, diclofenac/misoprostol, tolmetin, indometacin, piroxicam, ketoprofen, dexketoprofen, flurbiprofen, azapropazon, meloxicam, celecoxib, etirocoxib, and rofecoxib. Time-dependent exposure variables were defined by reference to the date of diagnosis of a colorectal cancer (index date) and to calculate cumulative dose and duration for each case and the remainder of participants in the cohort until that date. Given this approach, subjects were eligible as controls as long as they were not a case or censored. Consequently, participants were used in several case-sets. The methodology of time-varying exposure has been described by Clayton and Hill.[266]

Genotyping

DNA was extracted using standard procedures and stored at -20°C until used for DNA amplification. *CYP2C9*2* (rs1799853) and *CYP2C9*3* (rs1057910) were determined using 2-ng genomic DNA with the Taqman Prism 7900HT 384 wells format allelic discrimination assay.[267] Primer and probe sequences were optimized by using the SNP assay-by-design service. These allelic variants occur at appreciable frequency in the Caucasian population. The ALlele FREquency Database (ALFRED) reports allele frequencies of 13 % for *CYP2C9*2* and 7 % for *CYP2C9*3*.[79] The variants are the result of amino acid substitutions at position 144 (*CYP2C9*2*; $Arg_{(\mathbf{C}GC)} \rightarrow Cys_{(\mathbf{T}GC)}$) and position 359 (*CYP2C9*3*; $Ile_{(\mathbf{A}TT)} \rightarrow Leu_{(\mathbf{C}TT)}$). Since other variants are extremely

rare among Caucasians, persons without **2* or **3* were considered as having the wild-type genotype (**1*).

Case identification and validation

Three different databases were used for case identification. First, cases diagnosed by general practitioners in the research area were collected (International Classification of Primary Care (D75)). Second, the national registry of all hospital admissions was consulted to detect all malignancy related hospital admissions for study participants. Third, regional pathology databases were linked to The Rotterdam Study to identify cases. Subsequently, colorectal cancer cases were validated by a physician (CS) based on medical records of the general practitioner, discharge letters, and pathology reports. The 10th edition of the International Classification of Diseases (ICD-10) was used to distinguish between the anatomical locations non-sigmoid colon (C18), sigmoid colon (C19) and rectal (C20) cancer. Because of their low incidence and proposed different pathophysiology, anal cancers (n = 3) were not included as cases. Only pathologically confirmed cases were considered in the analyses. The index date was defined as the earliest date found in the pathology reports.

Covariates

On the basis of medical literature the following co-variables, assessed at baseline, were considered as potential confounders: age, gender, body mass index (BMI) (kg/m^2), total energy intake (kcal/day), alcohol (grams/day), vegetable (grams/day), fruit (grams/day), meat (grams/day), fiber (grams/day), and selenium intake (grams/day), hypercholesterolemia (total cholesterol > 6,5 mmol/L), physical activity (without difficulty, with some difficulty, with much difficulty, unable to do) and smoking (total packyears). The method and validation of dietary assessment in The Rotterdam Study has been described elsewhere.[7]

Statistical analyses

Genotype proportions and allele frequencies were tested for deviations from the Hardy-Weinberg equilibrium using a χ^2-test. Cox proportional hazard models were used to study associations between NSAID use or *CYP2C9* variant alleles and colorectal cancer risk. A first model adjusted for age and gender. A second model was made with those co-variables that changed the point estimate by more than 10% or which were independent risk factors for the outcome, according to the literature.

NSAID use was both studied as a dichotomized variable (never / ever use) and in categories of cumulative duration (never use / 1-365 days use / > 365 days use). The analyses were performed for total NSAID use and non-aspirin NSAIDs and aspirin NSAIDs separately. We defined five categories of drug exposure: never use of NSAIDs, 1-365 days of non-aspirin NSAID, > 365 days of non-aspirin NSAID, 1-365 days of aspirin use and > 365 days of aspirin use in which the "never use" category served as a reference while the other categories could partly overlap

when an individual used both aspirin- and non-aspirin NSAIDs during the study period. Trend analyses were performed to quantify a duration-effect response.

The association between genotype and colorectal cancer was studied in the total cohort, and in a subgroup of non-NSAID users to investigate the drug-independent effect of *CYP2C9* variant alleles. Carriers were defined as having at least one variant allele. The homozygous wild-type genotype (**1/*1*) served as the reference category. In addition to the association with the total group of colorectal cancers, the effect on anatomical subtypes (non-sigmoid colon, sigmoid colon, rectum) was investigated.

The combined effect of NSAIDs and *CYP2C9* genotype was studied by using the following groups: non-carriers without NSAID use (reference), variant carriers without use, non-carriers with use, variant carriers with use. Analyses were performed for the total colorectal cancer group as well as for anatomical subtypes. Drug-gene interactions were studied for the separate exposure subgroups (non-aspirin and aspirin NSAIDs) as well. When the interaction with non-aspirin NSAIDs was studied, the reference group was composed of only those without non-aspirin use and the analyses were adjusted for aspirin use. In a similar way, interaction of aspirin NSAIDs and genotype was studied. Trend analyses were performed on all these groups. The sequence for which the trends hold is based on the results of the separate analyses of NSAID use and *CYP2C9* variant allele carriage on colorectal cancer risk. Effect modification was studied with both additive (biological) and multiplicative interaction models. The relative excess risk reduction due to interaction (RERI) was used to evaluate departures from an additive scale. SAS software (Statistical Analysis Software version 8.2, Cary, NC) was used to derive regression coefficients (3) and covariance matrices (9). The obtained numbers were used to calculate RERIs ($RR_{combination} - RR_{exposure\ A} - RR_{exposure\ B} + 1$) and their corresponding 95% confidence limits. [268,269] If there is no biological interaction RERI is equal to 0. Interaction terms were added to the model to identify multiplicative effect modification. All analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA). P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant.

RESULTS

Individuals of whom the genotype was unknown (n = 1379) were on average older, relatively more frequently female and smoker, and had a shorter follow-up time than those for whom genotype data were available. Baseline characteristics of the study group are presented in TABLE 1. During a mean follow-up time of 9.8 years, 212 colorectal cancers (3 anal cancers not included) occurred. This was three percent of our cohort and corresponds with the incidence of colorectal cancer in the general Dutch population in persons aged 55 years and over.[1] The mean age was 68.1 years and 38.2 % were males. Ninety-six percent of the population used an NSAID at any time during the study period. Non-aspirin NSAIDs were taken by 60 percent

Table 1. Baseline characteristics

Characteristic	Total participants N = 7757
Genotyped	6378
Colorectal cancer cases (% of total participants) (genotyped)	212 (3%) (184)
Age, mean (SD)	68.1 (8.47)
Male gender, N (%)	2963 (38.2 %)
Body mass index (BMI), mean (SD), N (%)	26.4 (3.7)
Underweight (< 18.5)	54 (0.7 %)
Normal weight (18.5 – 24.9)	2645 (34.1 %)
Overweight (25.0 – 29.9)	3367 (43.4 %)
Obesity (30 – 39.9)	1047 (13.5 %)
Extreme obesity (≥ 40)	23 (0.3 %)
Smoking status, total pack years (SD), N (%)	26.7 (23.1)
Never	2723 (35.1 %)
Former	3149 (40.6 %)
Current	1699 (21.9 %)
Physical activity, N (%)	
Without difficulty	4608 (59.4 %)
With some difficulty	1598 (20.6 %)
With much difficulty	496 (6.4 %)
Unable to do	931 (12.0 %)
Hypercholesterolaemia (> 6,5 mmol/L), N (%)	3731 (48.1 %)
Total energy intake (kcal/day), mean (SD)	1967 (501)
Vegetable intake (grams/day), mean (SD)	350 (137)
Fruit intake (grams/day), mean (SD)	230 (132)
Meat intake (grams/day), mean (SD)	108 (47)
Fiber intake (grams/day), mean (SD)	17 (5)
Selenium intake (grams/day), mean (SD)	33 (10)
Fat intake (grams/day), mean (SD)	40 (19)
Alcohol consumption (grams/day), mean (SD)	10 (15)
Genotypes, N (%) [a]	
*CYP2C9 *1/*1*	4229 (66.3%)
*CYP2C9 *1/*2*	1339 (21.0 %)
*CYP2C9 *1/*3*	593 (9.3%)
*CYP2C9 *2/*2*	102 (1.6 %)
*CYP2C9 *2/*3*	89 (1.4 %)
*CYP2C9 *3/*3*	26 (0.4%)

N = number of persons, SD = standard deviation. [a] Hardy Weinberg $\chi^2 = 1.55$ ($p = 0.34$). Percentages do not sum up to 100 % due to missing values.

and aspirin NSAIDs by 30 percent of the population while 21 percent used both types during the study period. Mean duration of non-aspirin and aspirin NSAIDs use was 90 and 280 days, respectively. Genotype data were in Hardy Weinberg equilibrium ($\chi^2 = 1.55$; $p = 0.35$). 33.7% of

Table 2. Association between cumulative NSAID use and colorectal cancer

Cumulative days of NSAID use	N	Model 1 HR (95% CI)	N	Model 2 [b] HR (95% CI)
		Any NSAID use [c f]		
No use [a]	82	1.00 (reference)	63	1.00 (reference)
1-365 days use	90	0.72 (0.53-0.98)	69	0.65 (0.45-0.92)
> 365 days use	40	0.48 (0.32-0.72)	32	0.45 (0.28-0.71)
		Non-aspirin NSAID use [d]		
No use [a]	82	1.00 (reference)	63	1.00 (reference)
1-365 days use	73	0.77 (0.55-1.06)	59	0.72 (0.50-1.05)
> 365 days use	4	0.33 (0.12-0.91)	3	0.29 (0.09-0.95)
		Aspirin NSAID use [e]		
No use [a]	82	1.00 (reference)	63	1.00 (reference)
1-365 days use	17	0.48 (0.28-0.82)	10	0.34 (0.17-0.67)
> 365 days use	36	0.55 (0.36-0.83)	29	0.51 (0.32-0.81)

N = number of cases, HR = hazard ratio, CI = confidence interval. [a] No use is defined as no use of non-aspirin or aspirin NSAIDs during the study period. [b] Complete case analyses. [c] Model 1: adjusted for age and gender. Model 2: adjusted for age, gender, smoking, energy intake, physical activity, body mass index and hypercholesterolaemia. [d] Model 1: adjusted for age, gender and aspirin use. Model 2: adjusted for age, gender, smoking, energy intake, physical activity, body mass index, hypercholesterolaemia and aspirin use. [e] Model 1: adjusted for age, gender and non-aspirin use. Model 2: adjusted for age, gender, smoking, energy intake, physical activity, body mass index, hypercholesterolaemia and non-aspirin use. [f] Trend significant at 0.001 level.

the study population carried at least one variant allele. Allele based frequencies of *CYP2C9*2* and *CYP2C9*3* were 12.8% and 5.8 %, respectively.

Ever use of NSAIDs was associated with a 37% risk reduction of colorectal cancer (HR 0.63; 95% CI 0.47-0.85). Duration of use was inversely related to colorectal cancer incidence (p-trend = 0.001). (TABLE 2) Both aspirin- and non-aspirin NSAIDs were associated with a significant risk reduction for colorectal cancer, especially after more than one year of cumulative use. Total energy intake was the only potential confounder that changed the point estimate of NSAID use on the age and gender adjusted colorectal cancer risk by more than 10 %. The specified dietary factors did not change the risk when adjusted for the total intake. Other potential confounders were put into the model because they were considered as potential risk factors in the medical literature. Dietary data were not available for 23.1% of the population. Missing status of these and other factors had no effect on the association between NSAID use and colorectal cancer risk (p-values between 0.11 – 0.51). Therefore, complete case analyses with a second model that consisted of age, gender, total energy intake, physical activity, body mass index, hypercholesterolemia and the specified exposure were performed.

Carriage of a *CYP2C9* variant allele was also associated with a risk reduction (60%), primarily in the proximal parts of the colorectal tract. (TABLE 3) This risk reduction subsists for colon carcinoma in the analyses among non-NSAID users, although not significant anymore, probably due to insufficient power. The number of cases were too small to investigate the individual effects of *CYP2C9*2* and **3* on cancer risk. Overall, the reduced risk of colorectal

Table 3. Association between *CYP2C9* genotype and colorectal cancer

	All participants				Non-NSAID users			
	N	Model 1 HR (95%CI)	N	Model 2 HR (95%CI)	N	Model 1 HR (95%CI)	N	Model 2 HR (95%CI)
Colorectal cancer	184		155		71		60	
CYP2C9 wt	132	1.00 (reference)	115	1.00 (reference)	50	1.00 (reference)	44	1.00 (reference)
CYP2C9 variant	52	0.79 (0.57-1.09)	40	0.67 (0.47-0.96)	21	0.94(0.56-1.56)	16	0.78 (0.44-1.38)
Non-sigmoid colon	78		66		34		30	
CYP2C9 wt	61	1.00 (reference)	55	1.00 (reference)	27	1.00 (reference)	25	1.00 (reference)
CYP2C9 variant	17	0.56 (0.33-0.96)	11	0.38 (0.20-0.72)	7	0.59 (0.26-1.35)	5	0.43 (0.16-1.13)
Sigmoid colon	65		52		21		16	
CYP2C9 wt	46	1.00 (reference)	37	1.00 (reference)	14	1.00 (reference)	11	1.00 (reference)
CYP2C9 variant	19	0.82 (0.48-1.40)	15	0.79 (0.44-1.45)	7	1.09 (0.44-2.69)	5	1.00 (0.35-2.88)
Rectum	41		37		16		14	
CYP2C9 wt	25	1.00 (reference)	23	1.00 (reference)	9	1.00 (reference)	8	1.00 (reference)
CYP2C9 variant	16	1.28 (0.69-2.41)	14	1.21 (0.62-2.36)	7	1.73 (0.64-4.66)	6	1.57 (0.53-4.62)

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAID = non-steroid anti-inflammatory drugs, wt = wild type. Model 1: Adjusted for age and gender. Model 2: Adjusted for age, gender, smoking, energy intake, physical activity, body mass index and hypercholesterolaemia.

cancer associated with NSAID use seemed to be stronger than that associated with variant allele carriage.

Combinations of both variant allele carriage and NSAID use resulted in more protection than either of the factors alone. (TABLE 4) This effect was primarily seen in proximal parts with a significant trend for non-sigmoid colon cancer. However, significant effect modification on a multiplicative or additive scale did not occur (p-values all > 0.05). The results were similar for aspirin- and non-aspirin NSAIDs.

DISCUSSION

This prospective population-based cohort study demonstrates associations between NSAID use, *CYP2C9* variant allele carriage and colorectal cancer incidence. Duration of NSAID use was inversely related to the incidence of colorectal cancer. Since both non-aspirin and aspirin NSAIDs have been associated with a decreased risk of colorectal cancer in former studies, we combined both types of NSAIDs and additionally performed separate analyses. No obvious differences were observed in the protective effect of aspirin and non-aspirin NSAIDs. Carriage of a *CYP2C9* variant allele was associated with a lower risk of non-sigmoid colon cancers, even in non-NSAID users, although for this last group no significance was reached, probably because of power problems. A combination of variant allele carriage and use of NSAIDs resulted in a larger reduction of colorectal cancer risk than one of the determinants independently.

Table 4. Combined effect of NSAID use and *CYP2C9* genotype on colorectal cancer risk

	Total NSAID use [a]		Non-aspirin NSAID use [b]		Aspirin NSAID use [c]	
	N	HR (95%CI)	N	HR (95%CI)	N	HR (95%CI)
Colorectal cancer						
CYP2C9 wt, no use	50	1.00 (reference)	57	1.00 (reference)	100	1.00 (reference)
CYP2C9 variant, no use	21	0.92 (0.56-1.54)	25	0.94 (0.59-1.51)	39	0.79 (0.55-1.15)
CYP2C9 wt, use	82	0.64 (0.45-0.93)	75	0.87 (0.61-1.24)	32	0.62 (0.41-0.93)
CYP2C9 variant, use	31	0.47 (0.30-0.74)	27	0.60 (0.38-0.96)	13	0.50 (0.28-0.89)
Trend [d]	184	p = 0.001	184	p = 0.06	184	p = 0.003
Non-sigmoid colon						
CYP2C9 wt, no use	27	1.00 (reference)	28	1.00 (reference)	51	1.00 (reference)
CYP2C9 variant, no use	7	0.57 (0.25-1.31)	10	0.77 (0.37-1.58)	11	0.44 (0.23-0.85)
CYP2C9 wt, use	34	0.50 (0.30-0.85)	33	0.81 (0.48-1.37)	10	0.38 (0.19-0.77)
CYP2C9 variant, use	10	0.29 (0.14-0.60)	7	0.33 (0.14-0.77)	6	0.45 (0.19-1.07)
Trend [d]	78	p < 0.001	78	p = 0.02	78	p = 0.002
Sigmoid colon						
CYP2C9 wt, no use	14	1.00 (reference)	18	1.00 (reference)	32	1.00 (reference)
CYP2C9 variant, no use	7	1.10 (0.44-2.62)	8	0.95 (0.42-2.20)	15	0.94 (0.51-1.74)
CYP2C9 wt, use	32	0.93 (0.49-1.77)	28	1.02 (0.55-1.87)	14	0.88 (0.46-1.68)
CYP2C9 variant, use	12	0.67 (0.30-1.47)	11	0.76 (0.35-1.64)	4	0.50 (0.17-1.41)
Trend [d]	65	p = 0.32	65	p = 0.61	65	p = 0.25
Rectum						
CYP2C9 wt, no use	9	1.00 (reference)	11	1.00 (reference)	17	1.00 (reference)
CYP2C9 variant, no use	7	1.73 (0.65-4.66)	7	1.38 (0.53-3.56)	13	1.57 (0.76-3.23)
CYP2C9 wt, use	16	0.63 (0.27-1.46)	14	0.77 (0.34-1.75)	8	0.85 (0.36-2.01)
CYP2C9 variant, use	9	0.68 (0.27-1.76)	9	0.96 (0.39-2.37)	3	0.63 (0.18-2.17)
Trend [d]	41	p = 0.20	41	p = 0.67	41	p = 0.50

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAIDs = non-steroid anti-inflammatory drugs, wt = wild type. No use is defined as no use of the NSAID type of interest. For the total NSAIDs group this means no use of any non-aspirin or aspirin NSAID, for the non-aspirin NSAIDs group this means no non-aspirin NSAID use and for the group of aspirin NSAIDs this means no aspirin NSAID use. [a] Adjusted for age and gender. [b] Adjusted for age, gender and aspirin use. [c] Adjusted for age, gender and non-aspirin use. [d] Sequence for which the trend holds is based on the results of the separate analyses that NSAID use seems to be more protective than carriage of a variant allele.

This seems to be due to independent pathophysiological mechanisms, since both additive and multiplicative interaction terms were not significantly different from the combined risk reduction. The influence on different regulatory pathways in the arachidonic acid metabolism and the influence CYP2C9 seems to have on the formation of DNA adducts might explain these independent mechanisms.[255]

Regular and long-term use of NSAIDs [263,270-276] and the effect of *CYP2C9* genotype [129,131,133-135] on colorectal cancer risk have both been studied before. Currently, there is increasing evidence that regular or long-term use of NSAIDs protects against malignancies in the gastro-intestinal tract but final proof from randomized clinical trials is hardly available. In most observational

studies and in line with our results, cumulative NSAID use is associated with a reduced cancer risk already after 1 to 2 years of cumulative use.[270-273] Trials however, report a minimum required period of use of more than ten years.[274,275]

More conflicting are the results of studies that investigated the CYP2C9-cancer relation. Two previously published studies were in line with our finding that *CYP2C9* gene variant allele carriers had a decreased colon cancer risk.[133,134] Nevertheless, some others found inconsistent results.[129,131,135] The different risk per anatomical subtype might be the result of a different pathophysiological process that depends on a variety of environmental and genetic factors for proximal and distal cancers [277], which would argue for performing studies by anatomical site.

The objective of our study was to investigate whether *CYP2C9*2* and **3* variant alleles are associated with an increase of the protective effect of NSAIDs on colorectal cancer risk. The question whether there is synergism between *CYP2C9* variant alleles and NSAID use have been studied before in association with colon adenoma [278,279] and colorectal cancer [135,280]. Besides differences in outcome definition, exposure definitions and methodology in these studies differed. One of the previously published studies only presented results on the association with aspirin use [279], while others reported separate results for aspirin and non-aspirin NSAIDs [135,278] or aspirins and ibuprofen [280]. Similar to the first three studies, we did not find effect modification on a multiplicative scale. Additive modification of the drug effects was found in the first published study in approximately 500 cases and a similar number of controls.[278] The protective effect of aspirin on colorectal cancer risk appeared to be absent in those who carried a variant allele. Our results indicate the opposite with the lowest risk in both aspirin and non-aspirin NSAID users who are variant carriers. Unlike earlier studies, we studied potential effect modification on an additive and multiplicative scale in both non-aspirin and aspirin users in one cohort to investigate whether there was synergism between *2C9* variant alleles and NSAIDs use on colorectal cancer. For additive interaction we provided information about significance by using RERIs. None of the interaction terms was significant and consequently, based on our population-based study, there is no strong evidence that there is synergism and that *CYP2C9* variant carriage enhances the potential protective effects of NSAIDs on colorectal cancer risk. This seems in contrast with results from previous studies that *CYP2C9* variant allele carriers may accumulate NSAIDs and enhance their activity by a decreased metabolism.[281-283] However, it is more or less established that duration of use is more important than dose in the prevention of colorectal cancer.[272] It is therefore possible that accumulation of an NSAID, resulting in higher serum levels of the drug, does not add so much extra protection against colorectal cancer. The observed reduced risks in persons with both factors present might therefore be considered as the sum of two independent risk factors. Hence, these factors most probably act primarily through different pathophysiological pathways. Nevertheless, enhancement of the effects of NSAIDs on colorectal cancer by *CYP2C9* variants can still be present for specific NSAIDs, as was observed in a recently published American case-control study [280] Due to a

lack of power, we were not able to study the interaction between *CYP2C9* variant alleles and NSAIDs for individual products.

Observational studies might have some limitations. Selection bias due to including only those for whom blood samples were available seems unlikely since genotype data were in Hardy Weinberg equilibrium. Even though this group was slightly younger and possibly healthier, selection bias would not explain the risk reduction, which we found in NSAID users. Because data on exposure and outcome were prospectively gathered without knowledge of the research hypothesis, information bias is unlikely. One of the strengths of this study was the use of pharmacy data on a day-to-day basis with little room for misclassification of the drug exposure due to recall bias. Misclassification of the anatomical subtypes would, if present, be random and lead to underestimation of the true estimates as well. Confounding was adjusted for in the analyses.

In conclusion, NSAID use and *CYP2C9* variant alleles are both associated with a reduced risk of colorectal cancer, primarily of the non-sigmoid colon. Both aspirin and non-aspirin NSAIDs account for this effect. Variant allele carriers who used NSAIDs experienced the strongest reduction in risk. This seems to be due to independent mechanisms, and not as a consequence of interaction. Nevertheless, interaction might be present for specific NSAIDs.

3. Inflammation and cancer

3.1. A population-based study on NSAID use and colorectal cancer risk

ABSTRACT

Background: This study explored the association between cumulative NSAID use and colorectal cancer in a prospective population-based study.

Methods: Hereto, we included 7621 participants of 55 years and older from The Rotterdam Study. During a mean follow-up time of 11.5 years, 195 colorectal (colon, sigmoid and rectal) cancers occurred. Cases were validated based on medical records, discharge letters and pathology reports. Details on NSAID use since January 1st 1991 were available on a day-to-day basis. Cumulative duration, dose and time since last intake were used as exposure measures.

Results and conclusion: More than one year of cumulative use appeared to be associated with a reduced colorectal cancer risk [Any NSAID: HR 0.49, 95%CI 0.33-0.75; Non-aspirin NSAIDs: HR 0.25, 95%CI 0.10-0.63; Aspirin NSAIDs: HR: 0.58, 95%CI 0.37-0.90]. Although current use was associated with a stronger risk reduction than use in the past, a protective effect of NSAIDs on colorectal cancer remains demonstrable up to 5 years after discontinuation. Daily dose of NSAIDs seems to be of less importance.

INTRODUCTION

Already in the early 1980s, Waddell and Loughry found evidence for a protective effect of the newer non-steroid anti-inflammatory drug (NSAID) sulindac on colorectal cancer risk in patients with familial adenomatous polyposis coli.[284] Up till now, the association between use of NSAIDs and cancer has been examined in respect to several sites [122,285-287], but most thoroughly for colorectal cancer. NSAIDs inhibit the conversion of arachidonic acid into prostaglandins and harmful by-products by acting on cyclo-oxygenase (COX) enzymes. The resulting increased levels of arachidonic acid and the decreased levels of prostaglandins beneficially influence apoptosis and other cellular homeostatic mechanisms. Consequently, prolonged use of NSAIDs might inhibit the early stages of carcinogenesis.[288-291] Inhibition of the activity of COX-2 expression mainly in aberrant crypt foci or early adenomatous lesions is the proposed mechanism for this.[292] Since decades, the adenoma to carcinoma sequence in colorectal cancer is well established.[293] Therefore, the association between NSAID use and the risk of adenomas has been studied many times as well. For both adenomas and carcinomas, risk appeared to be reduced by approximately 30 to 50 percent, especially in the proximal parts of the colon.[294] Observational studies, animal models and clinical trials indicate that the effect is most pronounced during early stages of carcinogenesis. However, it remains unclear what the minimally required dose and duration of use is to reduce its risk. Whereas in some observational studies cumulative NSAID use is already reported to significantly reduce cancer risk after 1 to 2 years [270-273], trials on colorectal polyps report a minimally required period of use of more than ten years [274,275]. Although some proposed a dose-relationship [270,274,290,295], dosage was believed to be of less importance by others [272].

In most of the above-mentioned studies, NSAID use was assessed by means of standardized questionnaires. This can be effective to obtain information about medication use for drugs that are easily available over-the-counter. However, recall bias is a great problem of such an assessment. Therefore, some studies made use of numbers of prescriptions available from a pharmacy database.[272,276] However, limited information was available on potential confounding factors and these latter studies did either not take into account duration [276] or used a short follow up period [272].

The objective of this study was to explore the association between cumulative NSAID use and colorectal cancer in a prospective population-based cohort of individuals 55 years and older and to study whether the hazard changed after discontinuation of NSAIDs.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population-based prospective cohort study that started with a baseline interview between July 1989 and July 1993. All inhabitants of the Rotterdam suburb Ommoord aged 55 years and older were invited. Of the 10,275 eligible subjects, 7983 (78%) agreed to participate. The Medical Ethics Committee of the Erasmus Medical Center approved the study. The rationale and design of The Rotterdam Study have been described elsewhere.[6]

All participants were visited at home at the start of the study for a standardised questionnaire in which their health state and socio-economic status were recorded. Subsequently, a physical examination followed at the research center. Since the start of The Rotterdam Study, cross-sectional surveys have been carried out periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Information on medication use is available for all participants since January 1st 1991. The seven computerized pharmacies that cover the research area are linked to one network. In this way, the date of prescription, the total amount of drug units per prescription and the prescribed daily dosage are available for every drug. Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from the general practitioners working in the study district.

Cohort definition

To ensure a minimum of 1 year medication history, participants diagnosed with lung, colorectal, breast or prostate cancer (n = 220) before January 1st 1992 were excluded from the analyses. 123 persons were excluded since they died or were lost to follow up before this date. Additionally, participants diagnosed with cancer after January 1st 1992 but before their date of entrance in The Rotterdam Study (n = 19) were excluded as well. This resulted in a study cohort of 7621 (96%) participants. Follow-up time was defined as the period between January 1st 1992 and a diagnosis of colorectal cancer, death, loss to follow up or the end of the study period (October 2004), whichever came first. Persons with one of the other previously mentioned cancers were censored at date of diagnosis.

Exposure definition

The exposure of interest included both non-aspirin and aspirin NSAIDs. To study the effect of cumulative use, three categories were defined a priori: no use, 1-365 days of use and > 365 days of use during the follow-up period. Cut-off points were chosen on the basis of current knowledge that a significant risk reduction was seen after one year of cumulative use in some of the previously published studies [272,273] and as a restriction of sample size. Furthermore, cumulative duration was distinguished into current and past use. Current use was defined as

a final NSAID prescription within 90 days prior to the index date. Past use was defined as a last prescription more than 90 days before the index date. The index date was set on the date of diagnosis in cases and the corresponding date in controls. We divided exposure into 5 mutually exclusive categories: no use, cumulative duration of 1-365 days in past users, cumulative duration of 1-365 days in current users, cumulative duration of > 365 days in past users and cumulative duration of > 365 days in current users. Additionally, the effect of discontinuation of use was examined in six exposure categories: no use, current use for 1-365 days, current use for more than 365 days, discontinuation of use since less than 1 year, discontinuation of use since 1-5 years, and discontinuation of use since more than 5 years. Finally, cumulative duration of use was studied for those with an average defined daily dose (DDD) per day less than 1 and equal or more than 1. A defined daily dose is the recommended dose for an adult given for the most important indication.[9] The average defined daily dose per individual per day was calculated by dividing the cumulative amount of DDD by the cumulative duration of use.

Case identification and validation

Three different databases were used for case identification. First, cases diagnosed by general practitioners in the research area were collected (International Classification of Primary Care (D75)). Second, the national registry of all hospital admissions was consulted to detect malignancy related hospital admissions for study participants. Finally, regional pathology databases were linked to The Rotterdam Study to identify cases. Subsequently, colorectal cancer cases were validated by a physician (CS) based on medical records of the general practitioner, discharge letters, and pathology reports. The tenth edition of the International Classification of Diseases (ICD-10) was used to distinguish between the anatomical locations non-sigmoid colon (C18), sigmoid colon (C19) and rectal (C20) cancer. Because of their low incidence and probably different pathophysiology, anal cancers (n = 3) were not included. Only cases confirmed by pathology reports were considered in the analyses. The index date was defined as the earliest date found in the pathology reports.

Covariates

The following covariates were considered as potential confounders. Age, gender, body mass index (BMI) in kg/m^2, physical activity (level of difficulty with physical activeness), total pack years of smoking, rheumatoid arthritis (RA), osteoarthritis (OA), C-reactive protein level (mg/L), intake of fat (grams/day), alcohol (grams/day), vegetables (grams/day), meat (grams/day), fruit (grams/day), fibres (grams/day), selenium (grams/day), total energy intake (kcal/day), hypercholesterolemia (total cholesterol above 6.5 mmol/L).

Statistical analyses

Since exposure to NSAIDs may vary over time, the risk of colorectal cancer was calculated with a Cox' proportional hazard model with NSAIDs exposure as a time-dependent covari-

ate. In this way, the date of diagnosis of a colorectal cancer is set as the index date. Strata are defined with one incident case and all non-cases from the remainder of the cohort as a reference group. Given this approach, subjects were eligible as controls as long as they were not a case.[266] On the index date, the current or past cumulative duration and dose during the study period is calculated for each participant until that date and categorized as stated above. For all analyses, 'no use' was defined as the absence of a prescription for any non-aspirin or aspirin NSAID during the study period. A statistical model was built that adjusted for those factors that changed the point estimate by more than 10% or which were independent risk factors for the outcome according to the literature. Missing indicators were used to study the effect of missing values. When missing status was found to be no confounder, missing data on covariates were imputed by single imputation using the Expectation Maximization (EM) algorithm.[220]

We anticipated that persons with a prescription within the first months of 1992 were probably prevalent users. Therefore, to reduce misclassification of the duration of use, an additional analysis was performed in those who had no prescription in the year before (1991). We further hypothesized that persons with an undiagnosed colorectal cancer might have increased their use of NSAIDs because of cancer-related pain, shortly before diagnosis. Therefore, to avoid potential protopathic bias, an extra analysis was performed in which current users with a cumulative duration of use of less than 90 days were excluded. P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant. All statistical analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA).

RESULTS

During a mean follow up time of 11.5 years, 195 colorectal cancer cases (anal cancers not included) occurred. Baseline characteristics of the participants in the total study cohort and of the participants with a colorectal cancer are presented in TABLE 1.

Analyses for potential confounders did not change the point estimate of NSAID exposure on colorectal cancer risk substantially. Therefore, a model was built of those covariates that were proposed in the literature as potential risk factors for colorectal cancer. We excluded all dietary factors, since these did not individually affect the risk when additionally adjusted for total energy intake. Ever use of NSAIDs was associated with a decreased risk of incident colorectal cancer [HR 0.61; 95%CI 0.45-0.83] and was similar for non-aspirin and aspirin use. (TABLE 2) When cumulative duration was considered, time of use was inversely related with colorectal cancer incidence (p_{trend} for any NSAID = 0.008). However, this was most pronounced for non-aspirin NSAIDs that were used for more than a year [HR 0.25; 95%CI 0.10-0.63]. For the anatomical locations, the hazard ratios for any NSAID use were 0.38 [95%CI 0.12-1.25], 0.19 [95%CI 0.03-1.43] and 0.23 [95%CI 0.03-1.73] for non-sigmoid, sigmoid and rectum cancer,

Table 1. Baseline characteristics

Characteristic	Participants with colorectal cancer N = 195	Total participants N = 7621
Gender, N (%)		
Female	109 (55.9 %)	4709 (61.8 %)
Male	86 (44.1 %)	2912 (38.2 %)
Age (yrs), mean (SD), N (%)	70.4 (7.8)	68.1 (8.5)
55-64 yrs	52 (26.7 %)	3186 (41.8 %)
65-74 yrs	87 (44.6 %)	2812 (36.9 %)
75-84 yrs	48 (24.6 %)	1326 (17.4 %)
≥ 85 yrs	8 (4.1 %)	297 (3.9%)
Body mass index (kg/m^2), mean (SD)	25.9 (3.4)	26.4 (3.7)
Total pack years of smoking (yrs), mean (SD)	30.4 (28.6)	26.7 (23.1)
Total energy intake (grams/day), mean (SD)	1979 (499)	1967 (501)
Alcohol intake (grams/day), mean (SD)	10 (16)	10 (15)
Fat intake (grams/day), mean (SD)	43 (20)	40 (19)
Vegetable intake (grams/day), mean (SD)	367 (177)	350 (137)
Fruit intake (grams/day), mean (SD)	204 (120)	230 (132)
Fiber intake (grams/day), mean (SD)	16 (6)	17 (5)
Selenium intake (grams/day), mean (SD)	32 (10)	33 (10)
Meat intake (grams/day), mean (SD)	115 (50)	108 (47)
Hypercholesterolaemia, N (%)	112 (57.4 %)	3666 (48.1 %)
Rheumatoid arthritis, N (%)	3 (1.5 %)	99 (1.3 %)
Osteoarthritis, N (%)	23 (11.8 %)	754 (9.9%)
Physical activity, N (%)		
Without difficulty	111 (56.9 %)	4527 (59.4 %)
With some difficulty	37 (19.0 %)	1570 (20.6 %)
With much difficulty	11 (5.7 %)	488 (6.4 %)
Unable to do	33 (16.9 %)	915 (12.0 %)
Missing	3 (1.5 %)	121 (1.6 %)

N = number of persons, SD = standard deviation. Missings were imputed by a single method (Expectation Maximization) for analyses.

Table 2. Hazard ratios of cumulative duration of NSAID use on colorectal cancer risk[a]

	Any NSAIDs		Non-aspirin NSAIDs[b]		Aspirin NSAIDs[c]	
	N	HR [95% CI]	N	HR [95% CI]	N	HR [95% CI]
No use	76	1.00 (reference)	76	1.00 (reference)	76	1.00 (reference)
Any use	119	0.61 (0.45-0.83)	108	0.62 (0.45-0.87)	48	0.51 (0.34-0.77)
1-365 days	80	0.68 (0.49-0.95)	105	0.67 (0.48-0.93)	13	0.41 (0.22-0.75)
> 365 days	39	0.49 (0.33-0.75)	3	0.25 (0.10-0.63)	35	0.58 (0.37-0.90)

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAIDs = non-steroidal anti-inflammatory drugs. No use is defined as no non-aspirin or aspirin use. Since exposure groups can overlap, numbers of cases per NSAID class can be more than for the total groups of NSAIDs. [a] Adjusted for age, gender, body mass index, C-reactive protein level, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, and hormone replacement therapy. Additionally the NSAID subclasses are adjusted for the other class: [b] Non-aspirin NSAIDs for aspirins and [c] Aspirin NSAIDs for non-aspirin NSAIDs.

Table 3. Hazard ratios of cumulative duration of NSAID use in current and past users on colorectal cancer risk

	Any NSAIDs		Non-aspirin NSAIDs		Aspirin NSAIDs	
	N	HR [95% CI]	N	HR [95% CI]	N	HR [95% CI]
No use	76	1.00 (reference)	76	1.00 (reference)	76	1.00 (reference)
≤ 365 days use in past	64	0.74 (0.53-1.04)	89	0.72 (0.51-1.00)	10	0.53 (0.28-1.02)
≤ 365 days current use	16	0.98 (0.60-1.61)	16	0.93 (0.54-1.59)	3	0.47 (0.17-1.30)
> 365 days use in past	13	0.67 (0.38-1.15)	2	0.29 (0.09-0.92)	11	0.64 (0.32-1.25)
> 365 days current use	26	0.60 (0.38-0.94)	1	0.23 (0.06-0.93)	24	0.68 (0.42-1.12)
	195	$P_{trend} = 0.04$	184	$P_{trend} = 0.006$	124	$P_{trend} = 0.11$

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAIDs = non-steroidal anti-inflammatory drugs. No use is defined as no non-aspirin or aspirin use. Since exposure groups can overlap, numbers of cases per NSAID class can be more than for the total groups of NSAIDs. [a] Adjusted for age, gender, body mass index, C-reactive protein level, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, and hormone replacement therapy. Additionally the NSAID subclasses are adjusted for the other class: [b] Non-aspirin NSAIDs for aspirins and [c] Aspirin NSAIDs for non-aspirin NSAIDs.

respectively. The additional analysis in persons without a prescription in the year before 1992 revealed similar results as well. (Data not shown) When cumulative use was considered separately in current and past users, there was hardly a decline of the risk reduction in those who were no longer using NSAID on the index date but who had been using for more than a year cumulatively in the past (TABLE 3). Additionally, persons who discontinued use of NSAIDs remained on average on the same level of risk reduction (HR 0.60-0.70). (FIGURE 1) Five years after discontinuation, risk was still decreased by 30-40%, but none of the associations for the prior classified groups reached statistical significance, probably due to insufficient numbers. We additionally assessed whether the effect of cumulative duration was different for those with low (DDD per day < 1) and high (DDD per day ≥ 1) doses of NSAID use. In the category 1-365 days of use, the hazard ratios were identical for low (0.83, 95%CI 0.54-1.26) and high dose (0.82, 95%CI 0.60-1.12). For the category > 365 days of use, the hazard ratios were 0.15 (95%CI 0.02-1.10) for low dose and 0.37 (95%CI 0.13-1.01) for high dose but the numbers in this category were too low for a meaningful comparison.

DISCUSSION

In this population-based prospective cohort study, cumulative use of NSAIDs for more than one year was associated with a reduced risk for colorectal cancer. Duration of use seemed to be more important for non-aspirin NSAID use than aspirin use. Results for anatomical subgroups were less reliable due to low numbers. The association was largely independent of NSAID dose. After discontinuation of NSAID use, the protective effect remained visible for at least 5 years. The reduced risk in our study is in line with results from most [263,270,272-274,276,288,296-298] but not all [275,299] earlier studies. The large majority of these studies, however, used questionnaires to assess NSAID use in the past and defined 2-3 tablets per week as regular use. To avoid NSAID

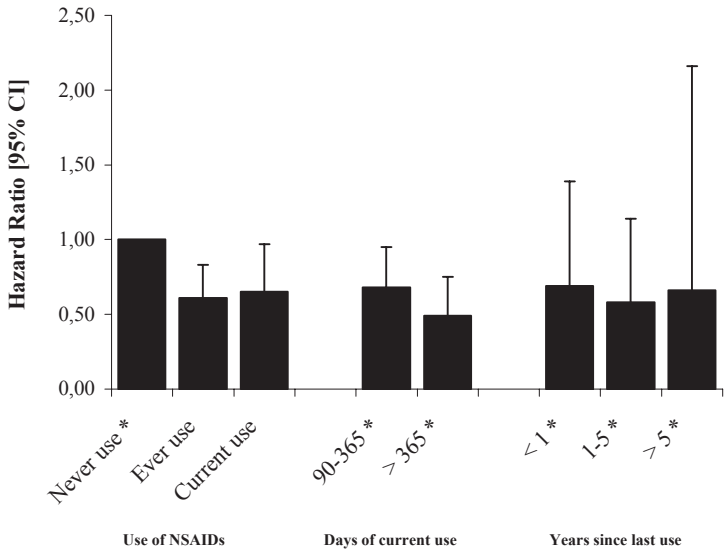

Figure 1. Effect of NSAID use and discontinuation of use on colorectal cancer risk [a]

[a] Adjusted for age, gender, body mass index, C-reactive protein level, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, and hormone replacement therapy. * Included in mutually exclusive categories

exposure misclassification by recall bias, some studies were published that used numbers of prescriptions derived from pharmacy databases. Two of these studies [272,276] examined the association of NSAID use and colorectal cancer. Only one of these studies explored the effect of duration but this study had limited access to potential covariates.[272] In our study, we combined several strengths of The Rotterdam Study, in that we had precise data on drug use over almost 15 years, and extensive information on potential confounders. Moreover, we were able to study the association with colorectal cancer after discontinuation of NSAIDs.

Considering the role NSAIDs play in inhibiting primarily early stages of carcinogenesis, it seems plausible that longer duration of NSAID use postpones the time a malignancy becomes clinically manifest. Similarly, our finding that the protective effect remains visible up to 5 years after discontinuation can be explained. When NSAIDs reduce growth in early preclinical stages, it may take some time before the process is fully resumed after stopping NSAIDs as our results suggest that there is no acceleration of growth after cessation. Former studies suggested that chemoprevention by NSAIDs was predominantly caused by COX-2 inhibitors (coxibs), since COX-2 expression was up-regulated in colon adenoma and colorectal carcinoma.[300] Because in our study only a minority of the drugs was a coxib and nevertheless a reduced risk was found, this explanation has to be reconsidered. First, the effect of non-selective COX inhibitors – the majority of NSAIDs in both our study (83%) and in daily life – might be as effective on COX-2 as selective COX-2 inhibitors. Second, there may be other mechanisms via which NSAIDs might reduce cancer risk.[301]

Observational studies may have limitations in the form of selection bias, information bias, and confounding. Selection bias in our study is unlikely since we used a prospective population-based cohort encompassing eighty percent of the source population whose participation was probably independent of NSAID use. As both disease status and drug use were prospectively gathered without knowledge of the research hypothesis, information bias seems unlikely. Since the database only contained information about prescriptions from 1992 onwards, however, and misclassification of duration could have biased our results, we performed an additional analysis in those without prevalent use of NSAIDs. This analysis revealed no substantial differences with the first analysis. NSAIDs are available over-the-counter (OTC) in The Netherlands since 1996 but to a limited extent and in a relatively low daily-recommended dose. Moreover, it is unlikely that persons with a prescribed NSAID also use them over-the-counter, since all NSAIDs on prescription – including long-term use - are fully reimbursed. Nevertheless, we performed an additional analysis in which we included self-reported information on NSAID use at baseline, and observed similar results. Although we performed an additional analysis in which we excluded current users with less than 90 days cumulative use, there was still an increase in the hazard ratio of current users with less than 1-year cumulative use. Nevertheless, the increase was not as pronounced as without the 90 days exclusion. The three months exclusion period might have been too short to fully eliminate protopatic bias. The total follow-up period in which we performed our research might have been too short to draw definite conclusions about discontinuation of NSAID use on colorectal cancer risk. The proposed effect of the drug on early stages of carcinogenesis makes it necessary to include longer follow up periods.

In conclusion, our results demonstrate that both non-aspirin and aspirin NSAID use are inversely related with colorectal cancer risk. However, duration of use seems to be more important for non-aspirin NSAID use. The inverse association remains demonstrable for at least five years after discontinuation.

3.2. The protective effect of NSAIDs on cancer and influence of COX-2 expression

ABSTRACT

Background: Inhibition of COX-2 enzymes has been the most frequently proposed mechanism for the beneficial effects of NSAIDs on carcinogenesis. The aim of this study was to explore the role of cumulative NSAID use on four common non-skin related cancers and the effect of diminished COX-2 expression as a result of a less affective *COX-2* gene.

Methods*:* 7621 participants of The Rotterdam Study were included. In a mean follow up time of 10 years, 720 colorectal, lung, breast or prostate cancers occurred. Cumulative NSAID use was calculated per NSAID class. Individual associations of NSAID use and *COX-2* genotype on cancer risk were explored with Cox' proportional hazard models. Next, the association of NSAIDs and cancer stratified by *COX-2* genotype was studied. Finally, combinations of NSAID use and *COX-2* genotype on survival times were investigated.

Results: All NSAID classes were associated with a reduced risk of colorectal cancer but not with other cancers. No associations between *COX-2* genotype and incident cancer, overall or cancer specific mortality were observed. COX-selective NSAIDs showed modest further risk reduction in persons with normal expression levels of the enzyme. Survival times were more than twice as long for colorectal patients that both carried a variant of the *COX-2* gene and used NSAIDs in the five years prior to diagnosis in comparison to those with normal expression levels of COX-2 enzyme who did not use NSAIDs ($p = 0.007$)

Conclusion: Our results confirm the protective effect of NSAID use on colorectal cancer and show that this is similar for all subclasses. COX-2 expression seems to be important in predicting the survival of colorectal cancer. However, COX-2 independent effects must be considered, both for colorectal cancers as other types of cancer.

INTRODUCTION

The analgesic effect of plants, like willow trees and myrtle leaves, containing the aspirin precursor salicylic acid have been described since 400-1500 A.D.[302,303] Non-steroid anti-inflammatory drugs (NSAIDs), with aspirin analogous effects, are currently the most frequently used drugs in the world.[304] NSAIDs inhibit the conversion of arachidonic acid (AA) into prostaglandins (PGs) and thromboxanes (Txs) by acting on a cyclooxygenase enzyme (COX) which catalyses this conversion. Inhibition of the pathway results in increased levels of arachidonic acid and decreased levels of prostaglandins. PGs perform a central role in physiological processes such as pain, fever, homeostasis and immunity. Reduction of PGs diminishes these effects and COX-inhibitors are therefore believed to be analgesic, anti-inflammatory and antipyretic.

It has been postulated in the past that prostaglandins can also play a role in tumor growth.[303] When prostaglandins are synthesized, a number of potentially harmful substances are produced as well. These substances include organic free radicals, peroxides, and activated oxygen compounds.[303] They activate environmental carcinogens, cause direct DNA damage, stimulate proliferation, inhibit and suppress apoptosis and anti-tumor immunity, and stimulate metastasis.[62] Inhibition of prostaglandin synthesis by NSAIDs was therefore hypothesized to be preventative for the development of cancer. In 1983 Waddell and Loughry found the first evidence for a protective role of NSAIDs on tumor growth. Familial adenomatous polyposis coli (FAP) patients who were treated with the newer NSAID sulindac experienced a dramatic regression of their rectal polyps.[284] Since then, the association between NSAIDs and cancer has frequently been studied. A protective effect was found for colon cancer [270-274,276,288,295,296], but risks of other gastrointestinal cancers [122,305], pancreatic cancer [306], breast cancer [287,307] and genitourinary malignancies [285,286] also seem to be diminished by regular NSAID use. Usefulness of long-term NSAID use as chemoprevention for tumor growth, however, seems to be limited because of serious adverse events such as gastrointestinal haemorrhage. Therefore, the distinction between COX-1 and COX-2 inhibiting NSAIDs might be relevant. COX-1 (PGHS-1) enzymes are expressed in a wide range of tissues and they act as a predominantly constitutive housekeeper in the regulation of homeostatic processes. COX-2 (PGHS-2) enzymes are only permanently expressed in some specific tissues, but its expression in other tissues depends largely on inflammatory conditions. Since cancer is often preceded or accompanied by inflammation, and selective COX-2 inhibiting NSAIDs result in less gastro intestinal side effects, studies increasingly focus on the association of specific COX-2 inhibitors (coxibs) and a reduction of cancer risk.

We hypothesized that the inhibiting effect of NSAIDs on cancer might depend on the degree of COX-2 expression. Therefore, the aim of this study was to explore the role of cumulative NSAID use on four common non-skin related cancers in carriers of a less effective *COX-2* gene.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population-based prospective cohort study that started with a baseline interview between July 1989 and July 1993. All inhabitants of the Rotterdam suburb Ommoord, aged 55 years and older, were invited. Of the 10,275 eligible subjects, 7983 (78%) agreed to participate. The Medical Ethics Committee of the Erasmus Medical Center approved the study. The rationale and design of The Rotterdam Study have been described elsewhere.[6]

All participants were visited at home at the start of the study for a standardised questionnaire in which they were asked for their health state. Dietary data were collected with a 170-item semi-quantitative food frequency questionnaire (SFFQ) adapted for use in the elderly. The two-step dietary assessment comprised a simple self-administered questionnaire (20 min) followed by a structured interview with trained dieticians (20 min) based on the complete questionnaire. This questionnaire has been validated for use in the elderly.[7] A physical examination followed subsequently at the research center. Since the start of The Rotterdam Study, cross-sectional surveys have been carried out periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Information on day-to-day drug use is available for all participants since January 1st 1991. The seven computerized pharmacies that cover the research area are linked to one network. In this way, the date of prescription, the total amounts of drug units per prescription and the prescribed daily dosage are available per drug defined by an Anatomical Therapeutical Chemical (ATC) code.[10] Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from the general practitioners working in the study district Ommoord.

Cohort definition

To ensure a minimum of 1 year medication history, participants diagnosed with lung, colorectal, breast or prostate cancer (n = 220) before January 1st 1992 were excluded from the analyses. 123 persons were excluded since they died or were lost to follow up before this date. Additionally, participants diagnosed with cancer after January 1st 1992 but before their date of entrance in The Rotterdam Study (n = 19), were excluded as well. This resulted in a study cohort of 7621 (96 %) participants. Genotypes were determined in 6147 individuals (77 %). Follow-up was defined as the period between January 1st 1992 and a diagnosis of one of the four cancers of interest, death, loss to follow-up or the end of the study period (October 2004), whichever came first.

Exposure definition

NSAIDs use

All prescriptions during the study period were filled in one of the seven fully computerized pharmacies in the research area. The exposure of interest included both non-aspirin and aspirin NSAIDs. The non-aspirin NSAIDs were divided into non-selective COX inhibitors, COX-1 inhibitors and COX-2 inhibitors. Agents used by individuals in this study were classified as follows: acetyl salicylic acid, carbasalate calcium and diflusinal (*aspirins*), sulindac, nabumeton, naproxen, ibuprofen, diclofenac and diclofenac/misoprostol (*non-selective COX inhibitors*), tolmetin, indometacin, piroxicam, ketoprofen, dexketoprofen, flurbiprofen and azapropazon (*COX-1 inhibitors*), meloxicam, celecoxib, etirocoxib, and rofecoxib (*COX-2 inhibitors*).

Time-dependent exposure to NSAIDs was used by defining the date of diagnosis of a cancer as the index date and to calculate cumulative duration for each participant until that date. Only cumulative duration was considered since this was found to be more important in prevention than dose.[272] Given this approach, subjects were eligible as controls as long as they were not a case. Consequently, participants were used in several case-sets. Clayton and Hill previously described the methodology of time-varying exposure.[266]

Cyclooxygenase-2 gene polymorphism

Genomic DNA was extracted from samples of peripheral venous blood according to standard procedures. 1-2 ng genomic DNA was dispensed into 384-wells plates using a Caliper Sciclone ALH3000 pipetting robot.[308] Genotypes were determined using the Taqman allelic discrimination assay. The Assay-by-Design service was used to set up a Taqman allelic discrimination assay for the *COX-2* (*PGHS2*) G-765C polymorphism (rs20417) (Primers Fw: TGCTTAGGACCAGTATTA TGAGGAGAA Rv: CCCCCTCCTTGTTTCTTGGAA. Probes FAM: CCTTTCCCCCCTCTCT, VIC: CTTTCCCGCCTCTCT).[267] This upstream polymorphism (chromosome 1q25.2-q25.3) showed to result in approximately 30% lower expression levels, probably by affecting the Sp1 binding site that is responsible for activating prescription.[309] The PCR reaction mixture included 2 ng of genomic DNA in a 2 µl volume and the following reagents: FAM and VIC probes (200 nM), primers (0.9 uM), 2x Taqman PCR master mix.[310] Reagents were dispensed in a 384-well plate using the Deerac Equator NS808.[311] PCR cycling reaction were performed in 384 wells PCR plates in an ABI 9700 PCR system and consisted of initial denaturation for 15 minutes at 95° C, and 40 cycles with denaturation of 15 seconds at 95° C and annealing and extension for 60 seconds at 60° C.[267] Results were analysed by the ABI Taqman 7900HT using the sequence detection system 2.22 software.[267] To confirm the accuracy of genotyping results, 332 (5%) randomly selected samples were re-genotyped with the same method. No inconsistencies were observed.

Case identification and validation

Three different databases were used for case identification. First, cases diagnosed by general practitioners in the research area were collected. Second, the national registry of all hospital admissions was consulted to detect malignancy related hospital admissions for study participants. Finally, regional pathology databases were linked to The Rotterdam Study to identify cases. A physician (CS) subsequently validated cases based on the International Classification of Primary Care (ICPC) (lung (R84), colorectal (D75), breast (X76) and prostate (Y77) cancer). The 10th version of the International Classification of Diseases (ICD-10) was used to distinguish between the subtypes non-sigmoid colon cancer (C18), sigmoid cancer (C19), rectal cancer (C20), anal cancer (C21), mesothelioma (C45) and bronchus carcinoma (C34) and to give an alternative code for breast (C50) and prostate (C61) cancers. Anal cancers (n = 3) and mesothelioma (n = 10) were censored at the date of diagnosis. Only cases confirmed by pathology were considered in the analyses. The index date was defined as the earliest date found in the pathology reports.

Covariates

The following covariates were considered as potential confounders: age, gender, body mass index (BMI), smoking (total pack-years), ability to be physically active (without difficulty / with some difficulty / with much difficulty / unable to do), rheumatoid arthritis, osteoarthritis, C-reactive protein level, cholesterol intake, vitamin intake and supplement use (total / vitamin D / vitamin D), mineral intake and supplement use (total / calcium / selenium), total energy intake (kcal/day), and intake of other dietary products (meat, fat, saturated fat, monounsaturated fat, polyunsaturated fat, fibres, vegetables, fruits, eggs, carbohydrates, carotenoids). Additional potential confounders in women were: age at onset of menarche and menopause, use of hormone replacement therapy (HRT), and number of children.

Statistical analyses

Genotype proportions and allele frequencies were tested for deviations from the Hardy Weinberg equilibrium using a χ^2-goodness-of-fit test. For all analyses on cancer incidence, Cox' proportional hazard models were used. First, the individual association of NSAID use and cancer risk was explored. NSAIDs were divided into the following classes: All NSAIDs, non-aspirin NSAIDs, aspirin NSAIDs, non-selective COX inhibitors, COX-1 inhibitors, and COX-2 inhibitors. Cumulative duration was categorized in three exposure groups: no use, 1-365 days cumulative use and > 365 days cumulative use during the study period, according to some previous studies that report a protective effect for colorectal cancer after 1 year cumulative use.[272,273,295] For all analyses, 'no use' was defined as the absence of a prescription for any non-aspirin or aspirin NSAID during the study period. For adjustments, two models were built. The first model adjusted for age and gender (or only age in the analyses on breast and prostate cancer). The second model adjusted for age, gender, cumulative use of other classes of NSAIDs and co-

variates that changed the point estimate by more than 10% or which were independent risk factors of the outcome according to the literature. Missing indicators were used to study the effect of missing values. When missing status was found to be no confounder, missing data on covariates were imputed by single imputation using the Expectation Maximization (EM) algorithm.[220] Second, the individual age, gender, smoking and C-reactive protein-adjusted associations of *COX-2* genotype and cancer incidence or mortality were studied. Associations were explored for carriage (GC and CC) as well as heterozygous (GC) or homozygous (CC) carriers. Homozygosity of the G-allele was set as the reference. Finally, the association of all classes of NSAID use (yes / no) and the different cancer types was studied stratified by *COX-2* variant carriage (non-carriers / carriers). Effect modification was studied with both additive (biological) and multiplicative interaction models. The relative excess risk reduction due to interaction (RERI) was used to evaluate departures from an additive scale. SAS software (Statistical Analysis Software version 8.2, Cary, NC) was used to derive regression coefficients (3) and covariance matrices (9). The obtained numbers were used to calculate RERIs ($RR_{combination} - RR_{exposure\ A} - RR_{exposure\ B} + 1$) and their corresponding 95% confidence limits. [268,269] If there is no biological interaction RERI is equal to 0. Interaction terms were added to the model to identify multiplicative effect modification. We anticipated that the effect of COX-2 enzyme expression on the benefit of NSAIDs on carcinogenesis was only relevant when expression was up-regulated. Therefore, an additional analysis was performed on the association of NSAID use in the five years before the index date, at a time an undiagnosed tumor must already be present, and the occurrence of cancer. Age at diagnosis and gender-adjusted general linear models were used to explore the associated between combinations of *COX-2* genotype and NSAID use five years prior to diagnosis, and survival times. All analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA). P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant.

RESULTS

Characteristics of the study population are presented in TABLE 1. Individuals of whom the genotype was unknown (n = 1474) were at baseline on average older, relatively more frequently female and smoker, and had a shorter follow-up time compared to the ones with genotype data. During a mean follow-up time of 9.7 (SD 4.1) years, 720 cancers occurred. Mean age of the study population was 70 years (range: 55 – 106) and 39% were males. During the study period, 69% of the population used an NSAID. Aspirin NSAIDs were taken by 30% and non-aspirin NSAIDs by 60%. 21% used both types during the study period. The majority of non-aspirin NSAID users had a prescription of a non-selective COX inhibitor. Only 2.3% of the study population used a specific COX-2 inhibitor. The frequency of the C-allele of the *COX-2* gene was 14.3%. Genotype data were in Hardy Weinberg equilibrium.

Table 1. Characteristics

Characteristic	Total participants N = 7621
Genotyped	6147
Total cases, N	
Any cancer	720
Lung cancer	134
Colorectal cancer	195
Non-sigmoid colon	81
Sigmoid colon	71
Rectum	43
Breast cancer	175
Prostate cancer	216
Age, mean yrs (SD)	70.3 (9.7)
Gender, N (%)	
Male	2949 (38.7 %)
Female	4672 (61.3 %)
NSAID use, N (%) [a]	5251 (68.9 %)
Aspirin NSAIDs	2279 (29.9 %)
Non-aspirin NSAIDs	4534 (59.5 %)
Non-selective COX	4268 (56.0%)
COX-1 selective	1036 (13.6%)
COX-2 selective	175 (2.3%)
Genotypes, N (%) [b]	
GG	4535 (73.8 %)
GC	1472 (23.9 %)
CC	140 (2.3%)

N = number of persons, SD = standard deviation. [a] % of total population. [b] Hardy Weinberg $\chi^2 = 2.51$ (2p = 0.11).

None of the covariates changed the point estimate of the association of NSAIDs and cancer by more than 10%. Therefore, additional to an age and gender adjusted model, a second model was built of only those covariates that were mentioned in literature as independent risk factors for the outcome. Total cancer risk seemed not to be associated with use or duration of use of any of the NSAID classes. (TABLE 2) Since differences in estimates were marginal between the models, only the second model was used for further subanalyses. Stratification on cancer site revealed, as was expected from previous results in literature, a protective effect for colorectal cancer. (TABLE 3) Both aspirin and non-aspirin NSAID use was inversely related to this risk. Duration of use seemed to be more important for non-aspirin NSAIDs. Although small numbers hampered the subdivision of non-aspirin NSAIDs in the COX-classes, both COX non-selective and COX selective inhibitors seemed to contribute to chemoprevention of colorectal cancer. When colorectal cancers were subdivided into their anatomical location, significant results were primarily detected for proximal cancers. (TABLE 4) Nevertheless, point estimates were below 1 for the other locations as well. *COX-2* genotype seemed not to be

Table 2. Association between different NSAID classes and overall cancer risk

	Any cancer		
		Model 1	Model 2
	N	HR (95%CI)	HR (95%CI)
Any NSAIDs	720		
No use	230	1.00 (reference)	1.00 (reference)
Any use	490	0.93 (0.78-1.08)	0.91 (0.77-1.08)
1-365 days	312	0.94 (0.79-1.12)	0.93 (0.78-1.11)
> 365 days	178	0.87 (0.70-1.07)	0.86 (0.70-1.06)
Non-aspirin NSAIDs [a]	667		
No use	232	1.00 (reference)	1.00 (reference)
Any use	435	0.93 (0.79-1.11)	0.94 (0.79-1.12)
1-365 days	398	0.96 (0.81-1.14)	0.95 (0.79-1.13)
> 365 days	37	0.80 (0.56-1.14)	0.79 (0.55-1.13)
Aspirin NSAIDs [b]	446		
No use	232	1.00 (reference)	1.00 (reference)
Any use	214	0.84 (0.68-1.03)	0.82 (0.67-1.01)
1-365 days	71	0.79 (0.60-1.03)	0.79 (0.60-1.05)
> 365 days	143	0.83 (0.66-1.05)	0.84 (0.67-1.06)
Non-selective COX [c]	649		
No use	232	1.00 (reference)	1.00 (reference)
Any use	417	0.95 (0.79-1.03)	0.95 (0.80-1.14)
1-365 days	387	0.97 (0.82-1.15)	0.95 (0.80-1.14)
> 365 days	30	0.86 (0.58-1.28)	0.86 (0.57-1.28)
COX-1 selective [d]	330		
No use	232	1.00 (reference)	1.00 (reference)
Any use	98	0.97 (0.73-1.27)	0.95 (0.72-1.25)
1-365 days	93	0.95 (0.74-1.22)	1.00 (1.00-1.00)
> 365 days	5	0.51 (0.21-1.25)	1.00 (1.00-1.00)
COX-2 selective [e]	249		
No use	232	1.00 (reference)	1.00 (reference)
Any use	17	0.62 (0.30-1.27)	0.62 (0.30-1.26)
1-365 days	16	0.61 (0.35-1.08)	0.61 (0.29-1.26)
> 365 days	1	0.60 (0.08-4.38)	0.74 (0.09-6.34)

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAID = non-steroidal anti-inflammatory drug, COX = cyclo-oxygenase. No use is defined as no non-aspirin or aspirin use. Since exposure groups can overlap, numbers of cases per NSAID class can be more than for the total group of NSAIDs. Model 1: Adjusted for age and gender. Model 2: Adjusted for age, gender, body mass index, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, C-reactive protein level and additionally: [a] aspirins, [b] non-aspirin NSAIDs, [c] aspirins and COX 1 and COX 2 selective NSAIDs, [d] aspirins and COX aselective and COX-2 selective NSAIDs, [e] aspirins and COX aselective and COX-1 selective NSAIDs.

Table 3. Association between different NSAID classes and risk of separate cancer types

	Lung cancer		Colorectal cancer		Breast cancer		Prostate cancer	
	N	HR (95%CI)	N	HR (95%CI)	N	HR (95%CI)	N	HR (95%CI)
Any NSAIDs	134		195		175		216	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	93	0.94 (0.64-1.39)	119	0.61 (0.45-0.83)	131	1.21 (0.84-1.73)	143	1.02 (0.76-1.37)
1-365 days	60	1.01 (0.67-1.52)	80	0.68 (0.49-0.95)	88	1.19 (0.81-1.73)	83	0.99 (0.72-1.37)
> 365 days	33	0.83 (0.51-1.35)	39	0.49 (0.33-0.75)	43	1.27 (0.80-2.00)	60	1.07 (0.75-1.54)
Non-aspirin NSAIDs [a]	127		184		167		189	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	86	1.07 (0.72-1.61)	108	0.62 (0.45-0.87)	123	1.20 (0.83-1.74)	116	1.01 (0.74-1.40)
1-365 days	83	1.12 (0.75-1.69)	105	0.67 (0.48-0.93)	105	1.18 (0.81-1.72)	106	1.00 (0.72-1.37)
> 365 days	3	0.41 (0.12-1.36)	3	0.25 (0.10-0.63)	18	1.40 (0.79-2.50)	10	1.41 (0.71-2.83)
Aspirin NSAIDs [b]	84		124		89		150	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	43	0.90 (0.56-1.45)	48	0.51 (0.34-0.77)	45	1.05 (0.65-1.71)	77	1.04 (0.73-1.47)
1-365 days	14	0.88 (0.47-1.66)	13	0.41 (0.22-0.75)	17	0.94 (0.52-1.70)	27	1.18 (0.75-1.87)
> 365 days	29	0.91 (0.54-1.55)	35	0.58 (0.37-0.90)	28	1.16 (0.67-2.02)	50	0.97 (0.65-1.43)
Non-selective COX [c]	126		183		162		179	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	85	1.10 (0.74-1.73)	107	0.68 (0.49-0.95)	118	1.19 (0.82-1.74)	106	0.99 (0.71-1.37)
1-365 days	83	1.20 (0.79-1.80)	102	0.71 (0.51-0.98)	104	1.18 (0.80-1.72)	98	0.97 (0.70-1.35)
> 365 days	2	0.39 (0.09-1.66)	5	0.37 (0.15-0.95)	14	1.37 (0.72-2.62)	8	1.55 (0.72-3.34)
COX-1 selective [d]	56		96		77		102	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	15	1.06 (0.54-2.08)	20	0.56 (0.31-1.03)	33	1.23 (0.72-2.10)	29	1.30 (0.80-2.13)
1-365 days	15	1.14 (0.59-2.24)	20	0.62 (0.34-1.12)	30	1.25 (0.73-2.15)	27	1.29 (0.78-2.12)
> 365 days	0	NA	0	NA	3	0.99 (0.28-3.49)	2	1.57 (0.37-6.65)
COX-2 selective [e]	44		79		54		74	
No use	41	1.00 (reference)	76	1.00 (reference)	44	1.00 (reference)	73	1.00 (reference)
Any use	3	2.42 (0.53-11.02)	3	0.29 (0.06-1.45)	10	0.92 (0.30-2.89)	1	1.17 (0.12-11.69)
1-365 days	3	2.45 (0.54-11.08)	3	0.29 (0.06-1.48)	9	0.90 (0.29-2.83)	1	1.18 (0.12-11.80)
> 365 days	0	NA	0	NA	1	1.93 (0.17-21.61)	0	NA

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAID = non-steroidal anti-inflammatory drug, COX = cyclo-oxygenase. No use is defined as no non-aspirin or aspirin use. Since exposure groups can overlap, numbers of cases per NSAID class can be more than for the total group of NSAIDs. Lung cancer: adjusted for age, gender, C-reactive protein level and pack years of smoking. Colorectal cancer: Adjusted for age, gender, body mass index, C-reactive protein level, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, and hormone replacement therapy. Breast cancer: adjusted for age, body mass index, C-reactive protein level, pack years of smoking, hormone replacement therapy, age at onset of menarche and menopause, and number of children. Prostate cancer: adjusted for age, body mass index, C-reactive protein level and pack years of smoking. And additionally: [a] aspirins, [b] non-aspirin NSAIDs, [c] aspirins and COX 1 and COX 2 selective, [d] aspirins and COX aselective and COX-2 selective NSAIDs, [e] aspirins and COX aselective and COX-1 selective NSAIDs.

Table 4. Association between different NSAID classes and anatomical location of colorectal cancer

	Colon cancer		Sigmoid cancer		Rectum cancer	
	N	HR (95%CI)	N	HR (95%CI)	N	HR (95%CI)
Any NSAIDs	81		71		43	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	48	0.56 (0.35-0.90)	45	0.74 (0.44-1.23)	26	0.54 (0.28-1.03)
1-365 days	33	0.65 (0.39-1.07)	30	0.78 (0.45-1.35)	17	0.60 (0.30-1.20)
> 365 days	15	0.42 (0.22-0.80)	15	0.65 (0.33-1.27)	9	0.44 (0.19-1.04)
Non-aspirin NSAIDs [a]	77		66		41	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	44	0.58 (0.35-0.95)	40	0.69 (0.39-1.20)	24	0.62 (0.31-1.24)
1-365 days	41	0.62 (0.37-1.02)	39	0.74 (0.43-1.29)	23	0.67 (0.34-1.33)
> 365 days	3	0.33 (0.10-1.12)	1	0.16 (0.02-1.21)	1	0.21 (0.03-1.62)
Aspirin NSAIDs [b]	52		44		28	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	19	0.49 (0.26-0.92)	18	0.57 (0.29-1.11)	11	0.48 (0.21-1.13)
1-365 days	7	0.52 (0.22-1.22)	3	0.28 (0.08-0.95)	3	0.39 (0.11-1.39)
> 365 days	12	0.47 (0.22-0.97)	15	0.76 (0.37-1.54)	8	0.53 (0.21-1.36)
Non-selective COX [c]	77		66		40	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	44	0.64 (0.39-1.07)	40	0.76 (0.43-1.33)	23	0.76 (0.38-1.50)
1-365 days	41	0.66 (0.40-1.09)	39	0.79 (0.45-1.39)	22	0.79 (0.39-1.56)
> 365 days	3	0.48 (0.14-1.63)	1	0.24 (0.03-1.84)	1	0.41 (0.05-3.16)
COX-1 selective [d]	44		34		18	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	11	0.65 (0.29-1.50)	8	0.77 (0.30-2.01)	1	0.80 (0.10-6.21)
1-365 days	11	0.73 (0.32-1.65)	8	0.84 (0.32-2.17)	1	0.93 (0.12-7.06)
> 365 days	0	NA	0	NA	0	NA
COX-2 selective [e]	35		27		17	
No use	33	1.00 (reference)	26	1.00 (reference)	17	1.00 (reference)
Any use	2	0.57 (0.08-4.14)	1	0.26 (0.02-4.15)	0	NA
1-365 days	2	0.62 (0.09-4.51)	1	0.27 (0.02-4.47)	0	NA
> 365 days	0	NA	0	NA	0	NA

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAID = non-steroidal anti-inflammatory drug, COX = cyclo-oxygenase. No use is defined as no non-aspirin or aspirin use. Since exposure groups can overlap, numbers of cases per NSAID class can be more than for the total group of NSAIDs. Adjusted for age, gender, body mass index, C-reactive protein level, pack years of smoking, cholesterol, physical activity, total energy intake, rheumatoid arthritis, osteoarthritis, and hormone replacement therapy. And additionally: [a] aspirins, [b] non-aspirin NSAIDs, [c]aspirins and COX 1 and COX 2 selective, [d] aspirins and COX aselective and COX-2 selective NSAIDs, [e] aspirins and COX aselective and COX-1 selective NSAIDs.

Table 5. Association between COX-2 genotype and cancer

Cancer type	Genotype	All participants N = 6147		Non-users of NSAIDs N = 1905	
		N	HR (95%CI) [a]	N	HR (95%CI) [a]
Any cancer	GG	452	1.00 (reference)	152	1.00 (reference)
	GC	148	1.02 (0.85-1.23)	39	0.89 (0.62-1.26)
	CC	12	0.83 (0.47-1.47)	5	1.08 (0.44-2.64)
	GC + CC	160	1.01(0.84-1.21)	44	0.91 (0.65-1.27)
Lung cancer	GG	81	1.00 (reference)	24	1.00 (reference)
	GC	19	0.73 (0.44-1.21)	5	0.72 (0.27-1.89)
	CC	3	1.15 (0.36-3.34)	0	NA
	GC + CC	22	0.77 (0.48-1.23)	5	0.64 (0.24-1.68)
Colorectal cancer	GG	123	1.00 (reference)	50	1.00 (reference)
	GC	36	0.90 (0.62-1.30)	12	0.81 (0.43-1.53)
	CC	5	1.26 (0.51-3.08)	2	1.39 (0.34-5.72)
	GC + CC	41	0.93 (0.65-1.32)	14	0.86 (0.48-1.56)
Breast cancer	GG	106	1.00 (reference)	31	1.00 (reference)
	GC	41	1.24 (0.87-1.78)	8	0.90 (0.41-1.95)
	CC	1	0.32 (0.05-2.30)	0	NA
	GC + CC	42	1.16 (0.81-1.66)	8	0.85 (0.39-1.85)
Prostate cancer	GG	141	1.00 (reference)	48	1.00 (reference)
	GC	52	1.15 (0.84-1.58)	14	0.99 (0.54-1.80)
	CC	3	0.66 (0.21-2.07)	3	1.68 (0.52-5.42)
	GC + CC	55	1.11 (0.81-1.51)	17	1.07 (0.61-1.86)

N = number of cases, HR = hazard ratio, CI = confidence interval, NSAIDs = non-steroidal anti-inflammatory drugs. [a] Adjusted for age, gender, smoking and C-reactive protein.

associated with cancer incidence in the total cohort or in a subgroup of non-NSAID users. (TABLE 5) In the analyses stratified on *COX-2* gene carriership, significant results were found for colorectal cancer in the (major) group of non-variant carriers (TABLE 6). However, estimates of the hazard ratio seem quite similar for both carriers and non-carriers in the groups of total NSAIDs, non-aspirin NSAIDs and COX-aselective NSAIDs. Nevertheless, the estimates for COX-1 and COX-2 selective NSAIDs were lower in the non-carriers compared to the carriers (TABLE 6) or the total group of participants (TABLE 4). No additive or multiplicative effect modification was observed. The more specific analyses on cumulative NSAID use in the 5 years prior to the index date revealed similar results as for the total use of NSAIDs during the study period in regard to cancer incidence with and without stratification for *COX-2* genotype. (data not shown) However, when survival times were investigated, colorectal cancer patients without *COX-2* variant carriage that did not use any NSAIDs in the 5-year pre-diagnostic period had

Table 6. Protective effect NSAIDs stratified for COX-2 polymorphism

		No variant carrier N = 4535		Variant carrier N = 1612		Additive interaction	Multi-plicative interaction
		N	HR (95% CI)	N	HR (95% CI)	RERI (95% CI)	p-value
All NSAIDs	Any cancer	300	0.90 (0.73-1.11)	116	0.90 (0.62-1.29)	0.16 (-0.19 – 0.51)	0.43
	Lung cancer	57	1.11 (0.67-1.82)	17	0.68 (0.24-1.92)	0.17 (-0.63 – 0.97)	0.81
	Colorectal cancer	73	0.56 (0.38-0.82)	27	0.63 (0.32-1.24)	0.15 (-0.40 – 0.70)	0.67
	Breast cancer	75	1.07 (0.69-1.82)	34	1.30 (0.57-2.93)	0.46 (-0.30 – 1.22)	0.30
	Prostate cancer	93	1.00 (0.69-1.43)	38	1.08 (0.59-1.96)	0.05 (-0.65 – 0.75)	0.83
Non-aspirins	Any cancer	268	0.94 (0.75-1.16)	107	0.94 (0.64-1.38)	0.06 (-0.35 – 0.47)	0.39
	Lung cancer	51	1.32 (0.79-2.22)	17	0.68 (0.23-2.02)	0.11 (-0.87 – 1.10)	0.68
	Colorectal cancer	68	0.57 (0.38-0.86)	23	0.69 (0.34-1.42)	-0.16 (-0.88 – 0.56)	0.88
	Breast cancer	70	1.05 (0.66-1.66)	34	1.39 (0.61-3.18)	0.52 (-0.34 – 1.39)	0.22
	Prostate cancer	78	1.03 (0.69-1.52)	33	1.16 (0.62-2.20)	0.03 (-0.74 – 0.80)	0.77
Aspirins	Any cancer	134	0.82 (0.63-1.06)	48	0.69 (0.44-1.09)	0.04 (-0.36 – 0.44)	0.83
	Lung cancer	27	1.00 (0.55-1.84)	8	0.69 (0.20-2.37)	0.17 (-0.71 – 1.05)	0.91
	Colorectal cancer	31	0.50 (0.31-0.83)	9	0.40 (0.16-1.01)	-0.17 (-0.88 – 0.53)	0.78
	Breast cancer	27	0.96 (0.52-1.75)	12	1.14 (0.39-3.32)	0.50 (-0.39 – 1.39)	0.34
	Prostate cancer	49	1.02 (0.66-1.57)	19	0.86 (0.42-1.76)	-0.02 (-0.79 – 0.74)	0.93
COX	Any cancer	254	0.94 (0.76-1.18)	104	0.95 (0.65-1.40)	0.07 (-0.35 – 0.49)	0.33
aselective	Lung cancer	50	1.42 (0.84-2.40)	17	0.86 (0.29-2.56)	0.09 (-0.99 – 1.17)	0.65
	Colorectal cancer	68	0.63 (0.42-0.96)	23	0.77 (0.37-1.58)	-0.27 (-1.07 – 0.54)	0.87
	Breast cancer	66	1.02 (0.64-1.63)	33	1.40 (0.61-3.23)	0.55 (-0.30 – 1.41)	0.21
	Prostate cancer	70	1.00 (0.67-1.50)	31	1.12 (0.58-2.14)	0.11 (-0.64 – 0.86)	0.68
COX-1	Any cancer	57	0.89 (0.62-1.28)	24	1.06 (0.58-1.92)	0.13 (-0.45 – 0.72)	0.47
selective	Lung cancer	11	1.33 (0.59-2.98)	1	1.43 (0.10-19.57)	-0.50 (-1.58 – 0.57)	0.36
	Colorectal cancer	11	0.38 (0.17-0.85)	3	0.90 (0.21-3.78)	-0.34 (-1.27 – 0.60)	0.83
	Breast cancer	16	1.02 (0.51-2.06)	10	1.58 (0.49-5.13)	0.98 (-0.30 – 2.27)	0.14
	Prostate cancer	18	1.34 (0.70-2.57)	10	1.72 (0.69-4.28)	0.37 (-0.90 – 1.65)	0.56
COX-2	Any cancer	8	0.53 (0.20-1.41)	5	0.17 (0.02-1.41)	0.45 (-0.63 – 1.54)	0.30
selective	Lung cancer	2	2.44 (0.39-15.11)	0	NA	-0.65 (-2.24 – 0.95)	0.98
	Colorectal cancer	2	0.17 (0.02-1.33)	0	NA	-0.77 (-1.86 – 0.31)	0.97
	Breast cancer	3	1.37 (0.21-8.86)	5	0.56 (0.06-5.14)	2.92 (-0.38 – 6.22)	0.30
	Prostate cancer	1	1.02 (0.10-10.40)	0	NA	-0.43 (-1.35 – 0.48)	0.99

N = number of cases, HR = hazard ratio, CI = confidence interval, RERI = relative excess risk (reduction) due to interaction, NSAID = non-steroidal anti-inflammatory drug, COX = cyclo-oxygenase.

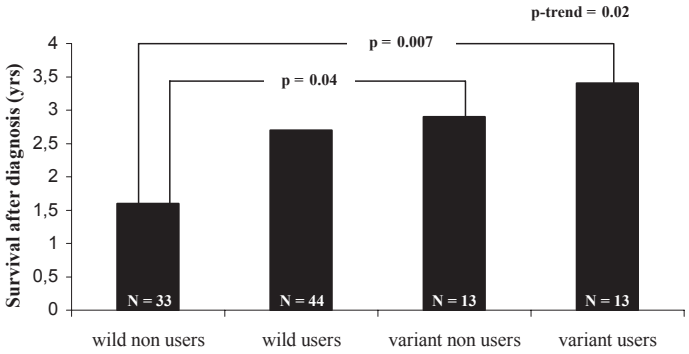

Figure 1. Years of survival after diagnosis stratified for NSAID use and *COX-2* genotype

significantly shorter survival times than non-users (p = 0.04) or users (0.007) that carried the variant. (FIGURE 1) Patients that both carried the variant genotype and used NSAIDs were found to live more than twice as long after diagnosis than those without the variant or NSAID use (3.4 years *vs* 1.6 years). These differences were not observed for the other cancers.

DISCUSSION

This prospective-population based cohort study adds to the large body of research that implicates NSAIDs and cyclooxygenases in carcinogenesis. The aim of this study was to clarify the role of different classes of NSAIDs on the risk of four common non-skin related cancers and the effect of genetic variation in the *COX-2* gene that is associated with decreased COX-2 expression levels. We found inverse relations between all NSAID classes and colorectal cancer, especially for proximally located tumors. There was no evident association with any of the other cancers. *COX-2* genotype by itself was not found to be associated with cancer incidence and overall or cancer specific mortality. When NSAID use was stratified by *COX-2* genotype, individuals without the variant type, with presumably normal expression levels of COX-2 enzyme, who used one of the COX specific NSAIDs, experienced a greater risk reduction than person with variant carriage or the total group of participant. However, no multiplicative effect modification could be observed. An interesting finding was that survival times were found to be twice as long for colorectal cancer patients with a variant allele who used NSAIDs in comparison to non-NSAID using patients with the wild type. These differences were not observed for the other cancers.

Previous studies most consistently reported reduced risks of colorectal [270-274,276,288,295,296] and breast cancer [287,307,312-314] in regular or long term users of aspirin or other NSAIDs. Associations with lung or prostate cancer are less conclusive. This study observed only reduced risks for colorectal cancer and results were in line with previous studies showing also the most pro-

nounced associations with proximally located tumors. It shows that reduction is not limited to one specific class of NSAIDs. This might be explained by the knowledge that each NSAID class has COX-1 and COX-2 inhibiting properties, with more or less affinity to one in specific. No associations of NSAID use and breast cancer incidence were found. However, in our population, a large part of non-aspirin NSAID users consumed ibuprofen, the drug that showed no risk reduction in a previous study as well.[314] This observation addresses the importance of considering individual drug effects since each drug has its own affinity and IC50 (half maximal inhibitory concentration) for inhibiting COX1/2 and effects differ per tissue site.[302] Additionally, regular aspirin use was only associated with a lower breast cancer risk in hormone receptor positive postmenopausal women in one previous study.[314] Although most of the women in our study population were postmenopausal, we had no information on receptor status and could therefore not stratify. Moreover, the required period of use remains controversial for all cancers and for breast cancer as well. A couple of months [314], years [287,313] or even decades [312] have been suggested. Our division in more or less than one-year cumulative use, based on previous knowledge on the inverse association with colorectal cancer risk, might have been too short to draw conclusions about breast cancer, which may have a very long induction and latent period.

The expression level of COX-2 enzymes differs between tissues and is up-regulated in specific situations such as inflammation or cancer. The expression is associated with cell proliferation, angiogenesis, inhibition of apoptosis and local immune-down modulation [302] and was found to be partly influenced by variation in its encoding gene [309]. According to The Cancer Genome Anatomy Project, tumor tissue in lung, colorectal tract, breast and prostate all express the gene that encodes for the COX-2 enzyme.[315] However, in contrast to the other three, *COX-2* gene expression was not found in normal gastrointestinal tract tissue. This is in line with previous reports on the incidence of lung [316] and colorectal cancer [317-319]. However, a case-control study among 500 colorectal adenoma patients and controls reports a reduced risk of adenoma in homozygous C-allele carriers that do not use any NSAIDs.[320] As COX-2 expression is the least up-regulated in the early stages of the adenoma-carcinoma sequence [321] and the conversion of arachidonic acid into prostaglandins is not immobilized by NSAIDs, the inhibiting effects on adenoma found for the CC-genotype could have been considered. The already expected extremely up-regulated expression in our colorectal cancer patients might be the reason that we could not reproduce this finding in our cohort. Overall, other variations in the *COX-2* gene seem to be more promising in prediction of cancer risk, but functionality of these SNPs remains to be proven.[316,317,322-324]

Both NSAID use and *COX-2* gene variant carriage (with less COX-2 expression) are believed to result in a decreased conversion of arachidonic acid into prostaglandins. As was previously explained, both increased arachidonic acid levels and decreased prostaglandin levels are found to be anti-carcinogenic. Next to the possible individual effects, there are generally two potential pathophysiological mechanisms in which NSAID use and *COX-2* genotype might

be related. First, variant carriage results in lower expression levels of the COX-2 enzyme. This might result in a decrease in targets for NSAIDs and therefore less benefit of its protective effects. Second, the diminished COX-2 expression might lead to a more rapid inhibition of all COX-2 enzymes. In this case, the protective effects of NSAIDs are anticipated to be clearer in variant carriers. In a comparable way, the effect of variant carriage might be most pronounced in non-NSAID users. Previous studies discussed issues that can be subsumed to either the first or the second mechanism. It was shown that not only expression levels of COX-2 enzyme are reduced in variant carriers of the gene, but also the C-reactive protein level response is diminished in these persons.[309] Subjects with the highest levels of CRP are most likely to benefit from NSAIDs when it comes to myocardial infarction and stroke.[325] Although this might be explained by factors that are associated with severity of the disease, it could in part be the direct result of decreased COX-2 expression and be analogous for carcinogenesis. This might support the first mechanism. On the contrary, the significant inverse association of homozygous carriers (CC) of the *COX-2* gene and colorectal cancer adenoma in non-NSAID users that was found in a previous study would be classified in the second mechanism postulated.[320] In the present study, our finding that estimates seem lower in persons with normal expression levels of the COX-2 enzyme argues for the first mechanism. The observation that this was primarily the case for cyclooxygenase specific inhibitors agrees to the proposed interaction of COX-2 enzyme and NSAIDs, although no significant interaction could be observed. We could not reproduce the inverse association found in non-NSAID using homozygous variant carriers, which would have endorsed the second mechanism. Since variant carriage and NSAID use, regardless of the other determinant, both prolonged survival times of colorectal cancer patients compared to patients without any of these determinants, these effects must be primarily regarded independently. This independent effect has been addressed before. Both *in vitro* [326] and *in vivo* [327,328] studies showed COX-2-independent effects of cyclooxygenase-2-inhibitors with proposed other targets in the 15-lipoxygenase-1 enzyme, PPARs and vitamin D receptor. These COX-2-independent effects might explain why we did not observe differences in incidence or survival-times in combinations for *COX-2* genotype and NSAID use in other cancers than colorectal cancer. Although other explanations, like insufficient follow-up time, small numbers, no information about stage and grade and other competing risk indicators could have played a role as well.

As with all observational studies, selection bias, information bias and confounding must be considered. Although persons without blood samples were on average older, more frequently female and smoker and had shorter follow-up times, genotype data were in Hardy Weinberg equilibrium and spurious associations are therefore unlikely to be found. As both disease status and drug use were prospectively gathered without knowledge of the research hypothesis, information bias is unlikely as well. However, NSAIDs are available over-the-counter (OTC) in The Netherlands since 1996 but to a limited extent and in a relatively low daily-recommended dose. Moreover, it is unlikely that persons with a prescribed NSAID also use them over-the-

counter, since all NSAIDs on prescription – including long-term use - are fully reimbursed. Confounding was adjusted for in the analyses. A great strength of the present study was its information on day-to-day NSAID use in different classes. However, results were somewhat hampered by small numbers. Therefore, future studies must included more cases to be able to identify differences in multiple subgroups, include longer follow-up times to elucidate effects on other cancers as well and include information on tumor differentiation.

In conclusion, this study shows inverse relations between all classes of NSAIDs and colorectal cancer. The inhibiting effects of COX-selective NSAIDs are slightly influenced by the level of COX-2 enzyme expression. Survival times of colorectal cancer patients were twice as long for those that either used NSAIDs in the 5-years prior to diagnosis or who carried the C^{765} variant of the *COX-2* gene compared to patients without any of these determinants. The fact that there was no evidence for modification, would argue for independent effects. Results were inconclusive for all other types of cancer.

3.3. C-reactive protein levels, variation in the *CRP* gene and cancer

ABSTRACT

Background: It remains unclear if inflammation itself may induce cancer, is a result of tumor growth or a combination of both. The aim of this study was to examine whether C-reactive protein (CRP) levels and gene haplotypes were associated with an altered risk of colorectal-, lung-, breast-, or prostate cancer.

Methods: 7017 participants of ≥ 55 years of The Rotterdam Study were eligible for analyses. Mean follow-up time was 10.2 years. High-sensitivity CRP (hs-CRP) measurements were performed to identify additional values of 0.2 – 1.0 mg/L compared to standard procedures. Genotypes of the CRP gene were determined with the Taqman allelic discrimination assay.

Results: High levels (> 3 mg/L) of CRP were associated with an increased risk of incident cancer (Hazard Ratio 1.4; 95% CI: 1.1-1.7) compared to persons with low levels (< 1 mg/L), also after introducing a potential latent period of five years. Although CRP seems to affect several cancer sites, the association was strongest for lung cancer (Hazard Ratio 2.8; 95% CI: 1.6-4.9). A *CRP* haplotype associated with decreased CRP levels was associated with an increased lung cancer risk of 2.6 (95% CI 1.6-4.4) in homozygous carriers.

Conclusion: Baseline CRP levels appear to be a biomarker of chronic inflammation preceding lung cancer, even after subtracting a 5-year latent period. Furthermore, a *CRP* gene haplotype associated with low CRP blood levels was relatively common in patients with lung cancer. Both chronic inflammation and impaired defence mechanisms resulting in chronic inflammation might explain these results.

INTRODUCTION

Since 1863, when Virchow hypothesized "that the origin of cancer was at sites of chronic inflammation", studies have been carried out on this topic. Although in general the hypothesis remains unproven, it is widely accepted that chronic inflammation increases the risk of cancer. For instance, inflammatory bowel diseases, human immunodeficiency virus, viral hepatitis B, human papilloma virus and rheumatoid arthritis are all associated with an increased risk of specific types of cancer.[329] Already in the nineteenth century, it was suggested that the combination of specific chronic inflammatory disease compounds, tissue injuries and ongoing inflammation result in enhanced cell proliferation that may eventually lead to tumor growth.[330] Although extensive research in the last decades showed that this explanation was far too easy and cell proliferation alone can not induce carcinogenesis, the finding of Virchow initiated a worldwide search for mechanisms underlying tumorigenesis, specifically those interacting with inflammation.

Nowadays, two hypotheses are studied concerning the association between inflammation and cancer. First, the induction hypothesis states that chronic inflammation results in excessive cell proliferation and activation of a cascade of cellular actions that can lead to ***induction*** of irreversible DNA damage. Persistent irritation and inflammation subsequently promote these initiated cells resulting in tumor growth, progression of metastatic disease and immunosuppression.[329] Second, the immune ***response*** of the host is studied as a consequence of tumor growth itself. In both hypotheses, products of inflammatory processes are believed to be biomarkers.[331-333]

The acute phase C-reactive protein is an inflammatory cell compound that is associated with a wide range of diseases including atherosclerosis and diabetes mellitus.[334,335] Since cancer is related to several forms of inflammation, CRP levels have also been implicated. The response hypothesis was frequently studied in the past and is currently more or less accepted.[331-333] However, the induction hypothesis remains unproven due to conflicting study results.[336-341]

The objective of this study was to investigate the association of CRP levels at baseline and the subsequent risk of developing one of the four most common non-skin related cancers (colorectal, lung, prostate and breast cancer). Additionally, we studied the association between variation in the *CRP* gene and these cancers.

METHODS

Setting

Data were obtained from The Rotterdam Study, a population-based prospective cohort study that started with a baseline interview between July 1989 and July 1993. All inhabitants of the Rotterdam suburb Ommoord, aged 55 years and older, were invited. Of the 10,275 eligible

subjects, 7983 (78%) agreed to participate.[6] The Medical Ethics Committee of the Erasmus Medical Center approved the study.

Participants were visited at home at the start of the study for a standardised interview on health state. Dietary data were collected with a 170-item semi-quantitative food frequency questionnaire (SFFQ) adapted for use in the elderly. The two-step dietary assessment comprised a simple self-administered questionnaire (20 min) followed by a structured interview with trained dieticians (20 min) based on the complete questionnaire. This questionnaire has been validated for use in the elderly.[7] Subsequently, a physical examination followed at the research center. Since the start of The Rotterdam Study, cross-sectional surveys have been carried out periodically. In addition, participants are continuously monitored for major events, including cancers, which occur during follow-up, through automated linkage with files from general practitioners. Information on vital status is obtained regularly from municipal health authorities in Rotterdam and from the general practitioners working in the study district.

Cohort definition

Only persons who provided blood samples at baseline were eligible for analyses. (FIGURE 1) Of the 7983 Caucasian persons in The Rotterdam Study, 7081 (89%) donated blood. Persons with a history of one of the cancers of interest at baseline (n = 64) were excluded from the analyses. This resulted in a study cohort of 7017 subjects. Follow-up time was defined as the period between baseline and a first diagnosis of one of the four cancers, death, loss to follow up or the end of the study period (October 2004), whichever came first.

Exposure definition

C-reactive protein level measurement

Non-fasting blood was collected at baseline and stored at -20°C until CRP measurement. Stability of CRP was not affected by this storing temperature.[342] High-sensitivity CRP (hs-CRP) measurements were performed using Rate Near Infrared Particle Immunoassay.[343] By using a high-sensitivity measurement, additional values of 0.2– 1.0 mg/L could be detected compared to standard CRP measurement procedures. Three categories of exposure were used according to advice of the American Heart Association (AHA). The categories of low (< 1.0 mg/L), average (1.0 to 3.0 mg/L) and high (>3.0 mg/L) correspond to approximate tertiles of hs-CRP among > 40000 adults in > 15 populations allowing adequate definitions of the distribution.[344] Besides categories, we studied the association between cancer and continuous CRP levels using a logarithmic transformation of CRP levels to obtain a normal distribution. Since acute infections during blood sampling could bias our results, only values below 10 mg/L were considered in the analyses.

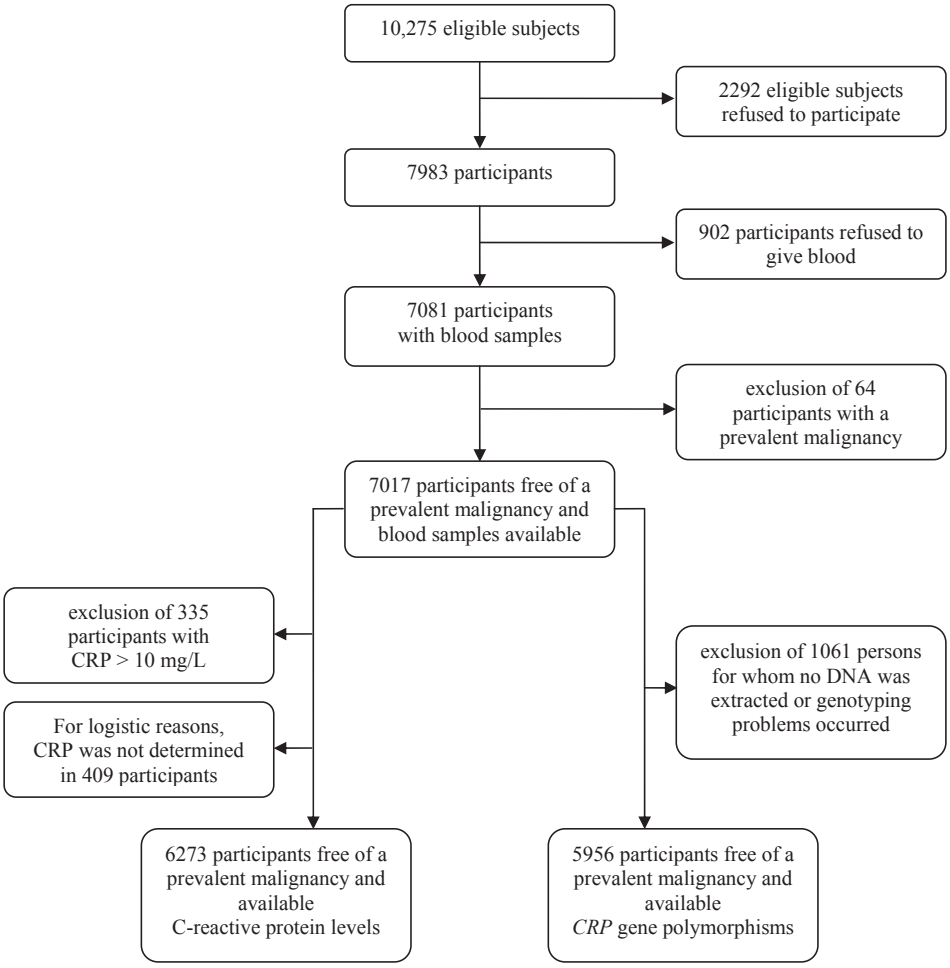

Figure 1. Cohort definition

Genotyping

The Seattle Program for Genomic Applications constructed four common *CRP* gene haplotypes in individuals of European descent.[345] We identified the three "haplotype tagging" SNPs (single nucleotide polymorphisms); rs1130864 (1184 C>T), rs1205 (2042C>T) and rs3093068 (2911 C>G) with a Taqman allelic discrimination assay.[267] Genotype data were used for each of the 3 polymorphisms to infer the haplotype alleles present in the population by using the program PHASE, which implements a Bayesian statistical method for reconstructing haplotypes from population genotype data.[346] This resulted in six haplotypes, but the fifth and sixth haplotype were present in <0.001 of the alleles and were therefore not used in the analyses. Haplotype alleles were coded 1 through 4 in order of decreasing frequency in the population. Haplotype 1 = C-T-C, haplotype 2 = T-C-C, haplotype 3 = C-C-C and haplotype 4 = C-C-G, following the

allele coding from 1184 C>T, 2042C>T and 2911 C>G. Haplotype 1 corresponds with the H1 and H2 haplotypes in the study of Carlson, both associated with a significant decreased CRP level in relation to the other haplotypes.[347,348]

Case identification and validation

Regional pathology databases were linked to The Rotterdam Study to identify cancer cases in more than 95% of persons within The Rotterdam Study. Cases were subsequently validated by a physician (CS) based on the International Classification of Primary Care (ICPC) (colorectal (D75), lung (R84), breast (X76) and prostate (Y77) cancer). The 10th version of the International Classification of Diseases (ICD-10) was used to distinguish between the subtypes non-sigmoid colon cancer (C18), sigmoid cancer (C19), rectal cancer (C20), anal cancer (C21), mesothelioma (C45) and bronchus carcinoma (C34) and to give an alternative code for breast (C50) and prostate (C61) cancers. The index date was defined as the earliest date found in the pathology reports.

Covariates

The following covariates were considered as potential confounders: age, gender, body mass index (BMI), smoking (total pack-years), alcohol, physical activity (assessed by a questionnaire; without difficulty, with some difficulty, with much difficulty and unable to do), use of NSAIDs, immunosuppressants and/or statins, age at onset of menarche and menopause, use of hormone replacement therapy, number of children in women, atherosclerosis, hypercholesterolemia (below or above 6.5 mmol/L), vitamin intake and supplement use (total / vitamin D / vitamin E), mineral intake and supplement use (total / calcium / selenium), total energy intake (kcal/day), and intake of other dietary products (meat, fat, saturated fat, monounsaturated fat, polyunsaturated fat, fibres, vegetables, fruits, eggs, carbohydrates, carotenoids).

Statistical analyses

First, we compared characteristics of participants with and without blood samples using a Student's t-test for continuous variables and a χ^2-test for dichotomous variables. Subsequently, a Cox proportional hazards model was built with age, gender and separate potential risk factors as covariates. Persons with colitis (n=19), anal cancers (n=3), mesothelioma (n=10) and men with breast cancer (n=2) were censored at the date of diagnosis. Participants who died or were lost to follow-up contributed person-time until the mortality date or date of last follow-up. In the analyses for the various cancers, we censored on date of diagnosis of other cancers. Missing indicators were used to study the effect of missing values. Missing data on covariates were imputed by single imputation using the Expectation Maximization (EM) algorithm.[220] Cox proportional hazard models were used to analyse the association between CRP levels and cancer. All analyses were done for tertiles of CRP, and for CRP levels following a natural logarithmic-transformation, because of the skewed absolute values of CRP in our population.

In a first model, we adjusted for age and gender. A second model adjusted for covariates that changed the point estimate by more than 10% or covariates, which were independent risk factors of the outcome according to the literature. To study whether there was evidence for the induction hypothesis, we additionally introduced a latent period of 5 years. Hereto, we restricted the analyses to those subjects who were free of cancer for at least 5 years following the baseline CRP assessment. Stratified analyses were performed by gender, smoking status, and subgroups of colorectal cancer (non-sigmoid colon, sigmoid colon, rectum) and lung cancer (squamous cell carcinoma and adenocarcinoma). Additionally, analyses were executed in non-smokers and female non-HRT users, to explore the smoking and HRT independent association of CRP and cancer.

Hardy Weinberg Equilibrium (HWE) was calculated with a chi-square statistic. Age- and gender-adjusted Cox models were used to study the association of haplotypes of the *CRP* gene and cancer. To study whether CRP levels operate as an intermediate for the association between haplotypes and cancer, we placed both covariates simultaneously in the model. Interaction terms and stratified analyses were introduced to explore effect modification.

Analyses were performed for all cancers together, and for the specific cancer types separately. Trends were calculated by including categorical variables as a continuous value in the Cox model. P-values below the conventional level of significance ($p < 0.05$) were considered statistically significant. All statistical analyses were performed with SPSS software (version 11.0.1; SPSS Inc., Chicago, USA).

RESULTS

Baseline characteristics

During a mean follow-up time of 10.2 years, a total of 780 incident cancers occurred. For the analyses of CRP level and genetic variation, this resulted in 706 and 677 incident cases, respectively. Baseline characteristics are presented in TABLE 1. Mean age in the total cohort was 69.6 years and 40% were men.

C-reactive protein levels

6273 participants without a prevalent malignancy and with CRP levels ≤ 10 mg/L, were available. The logarithm of serum levels of CRP was associated with a 19-20 percent risk increase of developing any of the four cancers during the follow-up period. (TABLE 2) After consideration of a five-year latent period the effect somewhat diminished to a significant 14-15 percent. Results were most pronounced in the highest tertile; with 37-38 and 24-26 percent risk increases, respectively. No effect modification by gender was observed ($p = 0.34$). Subsequently, analyses were performed according to tumor type.

Table 1. Baseline characteristics

Characteristic	Total participants N = 7017
Genotyped	5956
Male gender (%)	2808 (40.0 %)
Age (yrs, mean ± SD)	69.6 ± 9.2
55-64	2580 (36.8%)
65-74	2480 (35.3%)
75-84	1463 (20.8%)
≥ 85	494 (7.0%)
Body mass index (mean ± SD)	26.3 ± 3.7
underweight (< 18.5)	75 (1.1%)
normal weight (18.5-24.9)	2576 (36.7%)
overweight (25.0-29.9)	3166 (45.1%)
obesity (30.0-39.9)	960 (13.7%)
extreme obesity (≥ 40)	24 (0.3%)
missing	216 (3.1%)
Smoking status (mean total pack years)	27.5 ± 24.3
never	2410 (34.3%)
former	2842 (40.5%)
current	1565 (22.3%)
missing	200 (2.9%)
Physical activity	
without difficulty	3872 (55.2 %)
with some difficulty	1425 (20.3%)
with much difficulty	471 (6.7%)
unable to do	1106 (15.8%)
missing	143 (2.0%)
Hypercholesterolaemia (mean ± SD)	6.6 ± 1.2
absent	3491 (49.8%)
present	3483 (49.6%)
missing	43 (0.6%)
Genotypes	
CRP 1184 C > T	
CC	2782 (46.7 %)
CT	2591 (43.5 %)
TT	583 (9.8 %)
CRP 2042 C > T	
CC	2662 (44.7 %)
CT	2687 (45.1 %)
TT	607 (10.2 %)
CRP 2911 C > G	
CC	5274 (88.5 %)
CG	653 (11.0 %)
GG	29 (0.5 %)

N = number of persons, SD = standard deviation. Missings were imputed by a single method (Expectation Maximization) for analyses.

Table 2. The association between C-reactive protein levels at baseline and cancer

Cancer type	CRP level	Total follow-up			Follow-up > 5 years		
		No. cases/ cohort	Model 1 HR (95% CI)	Model 2 HR (95% CI)	No. cases/ cohort	Model 1 HR (95% CI)	Model 2 HR (95% CI)
Any cancer [a]	< 1 mg/L	202 / 1839	reference	reference	176 / 1671	reference	reference
	1-3 mg/L	293 / 2659	1.10 (0.91-1.31)	1.10 (0.92-0.32)	240 / 2351	1.06 (0.87-1.28)	1.06 (0.87-1.30)
	3-10 mg/L	211 / 1775	1.37 (1.12-1.66)	1.38 (1.13-1.69)	156 / 1409	1.24 (1.00-1.54)	1.26 (1.01-1.58)
	logarithm	706 / 6273	1.19 (1.10-1.30)	1.20 (1.10-1.31)	572 / 5431	1.14 (1.04-1.24)	1.15 (1.04-1.27)
Colorectal cancer [b]	< 1 mg/L	61 / 1839	reference	reference	48 / 1671	reference	reference
	1-3 mg/L	82 / 2659	0.96 (0.69-1.34)	0.94 (0.67-1.31)	64 / 2351	0.99 (0.68-1.44)	0.93 (0.64-1.36)
	3-10 mg/L	46 / 1775	0.93 (0.63-1.37)	0.93 (0.63-1.37)	26 / 1409	0.74 (0.46-1.19)	0.71 (0.44-1.15)
	logarithm	189 / 6273	1.04 (0.89-1.22)	1.04 (0.88-1.22)	138 / 5431	0.94 (0.78-1.13)	0.91 (0.75-1.11)
Lung cancer [c]	< 1 mg/L	17 / 1839	reference	reference	15 / 1671	reference	reference
	1-3 mg/L	50 / 2659	2.21 (1.28-3.84)	1.95 (1.12-3.38)	35 / 2351	1.80 (0.98-3.30)	1.61 (0.88-2.96)
	3-10 mg/L	50 / 1775	3.60 (2.07-6.26)	2.78 (1.59-4.85)	30 / 1409	2.69 (1.44-5.00)	2.11 (1.13-3.96)
	logarithm	117 / 6273	1.70 (1.37-2.11)	1.51 (1.21-1.88)	80 / 5431	1.62 (1.25-2.10)	1.45 (1.12-1.89)
Breast cancer [d]	< 1 mg/L	47 / 1092	reference	reference	42 / 999	reference	reference
	1-3 mg/L	80 / 1666	1.21 (0.84-1.74)	1.16 (0.80-1.67)	70 / 1486	1.20 (0.82-1.76)	1.14 (0.77-1.69)
	3-10 mg/L	57 / 1032	1.68 (1.14-2.47)	1.59 (1.05-2.41)	46 / 822	1.58 (1.04-2.40)	1.48 (0.94-2.33)
	logarithm	184 / 3790	1.30 (1.10-1.54)	1.28 (1.07-1.54)	158 / 3307	1.26 (1.05-1.50)	1.23 (1.02-1.50)
Prostate cancer [e]	< 1 mg/L	80 / 747	reference	reference	72 / 672	reference	reference
	1-3 mg/L	91 / 993	0.92 (0.68-1.25)	0.94 (0.70-1.28)	79 / 865	0.91 (0.66-1.26)	0.94 (0.68-1.30)
	3-10 mg/L	71 / 743	1.19 (0.86-1.64)	1.23 (0.89-1.71)	61 / 587	1.20 (0.85-1.69)	1.25 (0.88-1.79)
	logarithm	242 / 2483	1.10 (0.95-1.27)	1.12 (0.97-1.30)	212 / 2124	1.08 (0.93-1.25)	1.10 (0.94-1.29)

HR = hazard ratio, CI = confidence interval, CRP = C-reactive protein. Cox proportional hazard models were used for analyses. HR is Hazard Ratio. Model 1: Adjusted for age and gender. Model 2: Adjusted for main risk factors; [a] Adjusted for age, gender, body mass index, smoking, cholesterol, physical activity, NSAID use, fruit, selenium and total energy intake. [b] Adjusted for age, gender, smoking, NSAID use, cholesterol, physical activity, hormone use, fruit, selenium and total energy intake. [c] Adjusted for age, gender and smoking. [d] Adjusted for age, smoking, body mass index, age at menarche and menopause, hormone use and number of children. [e] Adjusted for age, smoking and body mass index.

The 189 incident colorectal cancers were analysed in total and in subgroups (78 non-sigmoid colon, 67 sigmoid and 44 rectal cancers). No significant results were found for the total group of colorectal cancers (TABLE 2) but high levels of CRP were significantly associated with a 2-fold risk increase of non-sigmoid colon cancer (TABLE 3). The effect diminished and became non-significant after excluding the predetermined latent period of five years.

There was an age-, gender- and smoking-adjusted hazard ratio for incident lung cancer of 2.8 when having CRP levels at baseline within the highest tertile. After the five-year latent period, this hazard ratio was reduced to a 2.1-fold risk increase in the same tertile. (TABLE 2) Additional adjustments for drugs used for chronic obstructive lung diseases or stratification by smoking status did not materially change the risk estimates (data not shown). When histological subtypes of lung cancer were considered, the increased risk could only be observed for squamous cell carcinoma of the lung. (TABLE 3) For breast cancer, increased risks were

Table 3. Stratified analyses of C-reactive protein level on cancer

Cancer type	CRP level	Total follow-up			Follow-up > 5 years		
		No. cases/ cohort	Model 1 HR (95% CI)	Model 2 HR (95% CI)	No. cases/ cohort	Model 1 HR (95% CI)	Model 2 HR (95% CI)
				Colorectal cancer			
Non-sigmoid colon	< 1 mg/L	16 / 2659	reference	reference	11 / 1671	reference	reference
	1-3 mg/L	35 / 2659	1.55 (0.86-2.80)	1.54 (0.85-2.80)	27 / 2351	1.77 (0.87-3.57)	1.72 (0.85-3.48)
	3-10 mg/L	27 / 1775	2.03 (1.09-3.77)	2.07 (1.10-3.90)	14 / 1409	1.71 (0.77-3.77)	1.68 (0.75-3.74)
	logarithm	78 / 6273	1.46 (1.13-1.90)	1.48 (1.14-1.94)	52 / 5431	1.29 (0.94-1.77)	1.28 (0.93-1.77)
Sigmoid colon	< 1 mg/L	26 / 2659	reference	reference	21 / 1671	reference	reference
	1-3 mg/L	29 / 2659	0.83 (0.49-1.42)	0.84 (0.49-1.44)	25 / 2351	0.91 (0.51-1.64)	0.88 (0.49-1.58)
	3-10 mg/L	12 / 1775	0.60 (0.30-1.20)	0.62 (0.31-1.24)	8 / 1409	0.53 (0.23-1.19)	0.51 (0.23-1.17)
	logarithm	67 / 6273	0.89 (0.68-1.15)	0.90 (0.69-1.17)	54 / 5431	0.85 (0.64-1.14)	0.84 (0.62-1.13)
Rectum	< 1 mg/L	19 / 2659	reference	reference	16 / 1671	reference	reference
	1-3 mg/L	18 / 2659	0.66 (0.34-1.26)	0.57 (0.30-1.10)	12 / 2351	0.55 (0.26-1.17)	0.47 (0.22-1.00)
	3-10 mg/L	7 / 1775	0.44 (0.18-1.04)	0.38 (0.16-0.92)	4 / 1409	0.34 (0.11-1.01)	0.29 (0.09-0.87)
	logarithm	44 / 6273	0.76 (0.55-1.05)	0.70 (0.50-0.99)	32 / 5431	0.68 (0.46-0.99)	0.60 (0.40-0.90)
				Lung cancer			
Adenocarcinoma	< 1 mg/L	3 / 2084	reference	reference	2 / 1897	reference	reference
	1-3 mg/L	10 / 2104	0.58 (0.08-4.34)	0.52 (0.06-4.44)	8 / 1865	0.68 (0.07-7.03)	1.27 (0.04-43.89)
	3-10 mg/L	3 / 2085	0.67 (0.12-3.69)	0.63 (0.11-3.55)	2 / 1669	0.34 (0.04-3.19)	0.07 (0.00-3.48)
	logarithm	16 / 6273	0.66 (0.29-1.47)	0.67 (0.30-1.48)	12 / 5431	0.24 (0.05-1.23)	0.09 (0.00-1.97)
Squamous cell carcinoma	< 1 mg/L	8 / 2084	reference	reference	6 / 1897	reference	reference
	1-3 mg/L	9 / 2104	1.62 (0.55-4.76)	1.56 (0.52-4.64)	7 / 1865	2.46 (0.63-9.64)	1.74 (0.42-7.28)
	3-10 mg/L	14 / 2085	1.93 (0.71-5.23)	2.35 (0.82-6.74)	5 / 1669	0.97 (0.22-4.35)	1.91 (0.39-9.41)
	logarithm	31 / 6273	1.20 (0.77-1.86)	1.30 (0.82-2.07)	18 / 5431	0.93 (0.48-1.79)	1.27 (0.61-2.61)

HR = hazard ratio, CI = confidence interval, CRP = C-reactive protein. Cox proportional hazard models are used for analyses. Colorectal cancer: Model 1: Adjusted for age and gender. Model 2: Adjusted for age, gender, smoking, NSAID use, cholesterol, physical activity, hormone use, fruit, selenium and total energy intake. Lung cancer: Model 1: Adjusted for age and gender. Model 2: Adjusted for age, gender and total pack years of smoking.

observed both in the total analyses as after exclusion of the 5-year latency period. For prostate cancer, no association was found.

CRP haplotypes

The 5956 participants without a history of any of the cancers of interest who were genotyped were used for analyses. The three tagging SNPs were in HWE (1184 C>T $\chi^2 = 0.32$ p = 0.57; 2042 C>T $\chi^2 = 3.49$p = 0.06, 2911 C>G $\chi^2 = 3.23$ p = 0.067). Homozygozity of haplotype 1 increased lung cancer by 2.6-fold. (TABLE 4) As a result of the design used, haplotype 3 (that can be 'tagged' by the common versions of the three tagging SNPs), showed a decreased risk. It refers to the other haplotypes of which haplotype 1, that showed an increased risk, is the most common. When serum levels of CRP and *CRP* haplotypes were simultaneously put in the age- and gender-adjusted model, the effect of *CRP* gene haplotype on cancer incidence did not

Table 4. The association between CRP haplotypes and cancer

		Haplotype 1 [CTC]		Haplotype 2 [TCC]		Haplotype 3 [CCC]		Haplotype 4 [CCG]	
		No. cases / cohort	HR (95% CI)	No. cases / cohort	HR (95% CI)	No. cases / cohort	HR (95% CI)	No. cases / cohort	HR (95% CI)
Any cancer	Absent	280 / 2664	1.00 (reference)	326 / 2788	1.00 (reference)	338 / 2917	1.00 (reference)	614 / 5280	1.00 (reference)
	Heterozygote	321 / 2685	1.16 (0.99-1.36)	293 / 2589	0.99 (0.85-1.16)	274 / 2525	0.91 (0.77-1.06)	59 / 649	0.82 (0.63-1.07)
	Homozygote	76 / 607	1.22 (0.95-1.58)	58 / 579	0.85 (0.64-1.12)	65 / 514	1.02 (0.78-1.33)	4 / 27	1.31 (0.49-3.50)
	trend	677 / 5956	p = 0.05	677 / 5956	p = 0.37	677 / 5956	p = 0.58	677 / 5956	p = 0.26
Colorectal cancer	Absent	78 / 2664	1.00 (reference)	92 / 2788	1.00 (reference)	94 / 2917	1.00 (reference)	167 / 5280	1.00 (reference)
	Heterozygote	92 / 2685	1.18 (0.87-1.60)	79 / 2589	0.93 (0.69-1.26)	79 / 2525	0.95 (0.70-1.28)	22 / 649	1.13 (0.72-1.76)
	Homozygote	19 / 607	1.12 (0.68-1.85)	18 / 579	0.94 (0.56-1.55)	16 / 514	0.91 (0.53-1.54)	0 / 27	-
	trend	189 / 5956	p = 0.39	189 / 5956	p = 0.67	189 / 5956	p = 0.65	189 / 5956	p = 0.92
Lung cancer	Absent	39 / 2664	1.00 (reference)	51 / 2788	1.00 (reference)	67 / 2917	1.00 (reference)	108 / 5280	1.00 (reference)
	Heterozygote	51 / 2685	1.34 (0.88-2.03)	53 / 2589	1.12 (0.77-1.65)	39 / 2525	0.66 (0.45-0.98)	5 / 649	0.38 (0.16-0.93)
	Homozygote	23 / 607	2.64 (1.57-4.41)	9 / 579	0.82 (0.40-1.66)	7 / 514	0.58 (0.27-1.27)	0 / 27	-
	trend	113 / 5956	p = 0.001	113 / 5956	p = 0.93	113 / 5956	p = 0.03	113 / 5956	p = 0.03
Breast cancer	Absent	72 / 1563	1.00 (reference)	80 / 1668	1.00 (reference)	83 / 1739	1.00 (reference)	155 / 3164	1.00 (reference)
	Heterozygote	89 / 1639	1.15 (0.84-1.57)	81 / 1544	1.15 (0.84-1.57)	68 / 1501	0.92 (0.67-1.27)	15 / 370	0.82 (0.48-1.39)
	Homozygote	11 / 350	0.72 (0.38-1.35)	11 / 340	0.71 (0.38-1.34)	21 / 312	1.29 (0.80-2.08)	2 / 18	2.28 (0.56-9.24)
	trend	172 / 3552	p = 0.83	172 / 3552	p = 0.81	172 / 3552	p = 0.63	172 / 5956	p = 0.86
Prostate cancer	Absent	102 / 1101	1.00 (reference)	111 / 1120	1.00 (reference)	115 / 1178	1.00 (reference)	211 / 2116	1.00 (reference)
	Heterozygote	99 / 1046	1.03 (0.78-1.36)	93 / 1045	0.91 (0.69-1.20)	93 / 1024	0.92 (0.70-1.21)	17 / 279	0.69 (0.42-1.13)
	Homozygote	29 / 257	1.25 (0.83-1.89)	26 / 239	1.01 (0.66-1.56)	22 / 202	1.06 (0.67-1.67)	2 / 9	3.03 (0.75-12.21)
	trend	230 / 2404	p = 0.37	230 / 2404	p = 0.78	230 / 2404	p = 0.86	230 / 2404	p = 0.37

HR = hazard ratio, CI = confidence interval. Cox proportional hazard models were used for analyses. Adjusted for age and gender (except for breast and prostate cancer).

materially change. There was no effect modification observed (p = 0.77) and no differences were found in stratified analyses. These results indicate that CRP levels and *CRP* haplotypes are probably independent risk indicators.

DISCUSSION

In this prospective cohort study on the association between CRP determinants and four common non-skin related malignancies, high levels of CRP were associated with a 38% increase in overall cancer risk and a 26% increase in cancer risk after subtracting a latent period of 5 years. Sub-analyses showed that the association was predominantly observed for squamous cell carcinoma of the lung, non-sigmoid colon and breast cancers and was still significant after 5 years for lung and breast cancer. The majority of results could be explained by the response hypothesis in which increased CRP levels at baseline are a biomarker for an underlying, probably not yet discovered, malignancy. However, for lung cancers, this might be different. Since survival times for these cancers are short, increased CRP levels five years before diagnosis probably preceded cancer and were not a marker for secondary inflammation around an undiagnosed cancer. This might be in line with the induction hypothesis. Our results resemble findings of two previous studies [337,339], but are at variance with ones found in a study of 28 345 women of 45 years and older from the USA [341]. Since our results seem to indicate that the risk increase is mainly observed for lung cancer, the power of the US study with 32 lung cancer cases might have been too low to draw conclusions. Other explanations for the differences might be the restriction to female gender, since point estimates in our analyses where somewhat lower than in males, a younger population, and the relatively short follow-up time. Three previously published studies reported contradictory outcomes regarding the effect of CRP baseline levels on non-sigmoid colon cancer risk.[336,338,340] We observed that CRP baseline measurements were only associated with an increased non-sigmoid colon cancer risk in the first years after the assessment at a time when the tumor might already have been present. Also in the Women's Health Study (WHS) and the CLUEII cohort significant results were only found during the first years of follow-up. However, when they stratified on smoking status, Erlinger et al. observed a significant risk increase after five years in non-smoking individuals with non-sigmoid colon cancer.[340] Our results slightly support this finding, since, although not significant, point estimates for non-sigmoid colon cancer risk in non-smokers after five years were higher than in the analyses regardless of smoking status.

One of the most important limitations of baseline levels is the single measurement in time. Although it has been postulated that CRP levels are more or less stable over time [349], it is well established that CRP levels are easily influenced by age, smoking and underlying inflammations. Therefore, we hypothesized that variation in the *CRP* gene might be a better determinant for cancer risk, because it represent lifetime exposure to a particular 'steady

state' CRP level or a less vigorous acute inflammatory response. Previous studies showed that the combination of SNPs C^{1184}, T^{2042}, and C^{2911} (the CTC haplotype or haplotype 1) is associated with decreased CRP serum levels. We found that in individuals with the CTC-haplotype, lung cancer risk was increased 2.6-fold for those who were homozygous with a significant allele-effect relationship. It seems possible that subjects carrying the haplotype associated with decreased CRP serum levels have a less efficient response to external noxious agents. An impaired defence mechanism, in the form of reduced CRP response might result in prolonged inflammation and increased tissue damage. Of course, this is only one potential explanation. Moreover, we cannot exclude that the *CRP* gene haplotype is in linkage disequilibrium with other polymorphisms on the genome that are responsible for the increased lung cancer risk and that our finding is not causal.

Some potential limitations must be considered. First, we included subjects for whom blood was available. Those who did not donate blood were on average older, more frequently women and had a shorter follow up time. However, baseline cancers were excluded and there is little reason to assume that selective inclusion of patients with relatively high CRP levels and later tumors caused a spurious association. If selection bias occurred, it is more likely that blood sampling from relatively healthy elderly led to an underestimation of the true risk increase. Misclassification bias is unlikely as data collection during follow-up was performed without knowledge of the research hypothesis. Misclassification of CRP levels seems unlikely. Confounding for age, gender, BMI, smoking, hypercholesterolemia, physical activity, NSAID use, fruit, selenium and total energy intake was adjusted for in the analyses. In the case of lung cancer, use of medication against chronic obstructive lung disease did not influence our results. That this made little difference might also be explained by the possibility that CRP level elevation is an intermediate in the causal chain between such risk factors and lung cancer although it is unlikely that CRP itself is a risk factor for cancer.

In conclusion, we found that high baseline values of CRP and a particular *CRP* gene haplotype were both consistently associated with an increased risk of lung cancer. Subgroup analyses of cancers at sites for which inflammation is believed to be one of the main risk factors (colon cancers and squamous cell carcinoma of the lung) showed doubled risk profiles in the highest CRP tertile. These findings seem to endorse the induction hypothesis as originally postulated by Virchow.

4. General Discussion

INTRODUCTION

Improved medical care is one of the factors, which results in aging of the population. Although life-expansion is one of the main goals of this improvement, it also increases the incidence of diseases that are strongly correlated with age, such as cancer. Continuation of the broad range of studies on cancer etiology, treatment and prognosis is therefore inevitable. As a consequence of the efforts made in cancer research over the past decades, individualized approaches are state of the art nowadays. However, many of these programs can still benefit from fine-tuning to optimize personalized medicine. An important way to achieve this is to consider genetic variation.

It has been calculated that genetic factors represent 5 to 10 percent of all risk factors for cancer.[4] Obviously, this cannot be more than a very crude overall estimation and the role of genetic variation may be substantially higher if, for example in cancer research, both induction and promotion are included. Besides single mutations such as those in the *BRCA-1* and *BRCA-2* gene, genetic variability in candidate genes became of increased interest since the 1970s when Harris proposed that common functional variation could explain some of the inherited variation in susceptibility to common diseases. [350] In the 1980s it became possible to clone genes of interest and genetic markers within it (usually restriction fragment length polymorphisms).[196] The development of genome-wide genetic maps in the 1990s allowed researchers to study genome-wide linkage to disease status and enhanced the impact of genetic association studies even more.[351,352] Genetic variability is responsible for the 0.1% of the genome that determines individual phenotypes and was proposed to affect not only a person's propensity towards diseases, but also the response to therapy and prognosis. Candidate genes are those that encode for enzymes with a plausible biological action. For cancer, a wide range of candidate genes exists. This thesis focussed on genes that encode for cytochrome P450 enzymes and genes that are involved in inflammation. Furthermore, analyses were restricted to the four most common non-skin related cancers.

In this chapter the main findings of the studies described in this thesis are reported and placed in a broader perspective. Methodological considerations, strengths and limitations for both epidemiological research in general as for research facilitated in The Rotterdam Study in particular are discussed. A reflection of the relevance and potential clinical implications of our observations is given and recommendations for future research are put forward.

MAIN FINDINGS AND PERSPECTIVE

As for cancer a large variety of candidate genes can be studied, this thesis was restricted to two topics with potentially high impact. First, cytochrome P450 related gene polymorphisms were considered, as genes that encode for this enzyme family are involved in a broad range

of metabolic pathways. These pathways comprise those of endogenous (i.e. hormones) or exogenous (i.e. drugs or other xenobiotics such as alcohol or tobacco) compounds that are all related to cancer. Detailed insight into the effects of polymorphic CYP expression might potentially lead to new ways of optimizing personalized medicine. Second, as of the end of the 19th century, when Virchow hypothesized "that the origin of cancer was at sites of chronic inflammation", the association between inflammatory conditions and cancer has been studied many times. Results are encouraging, but diverse at the same time. While some observed associations with microorganisms and anti-inflammatory medications, others focussed on inflammatory markers for which the cause-effect relationship with cancer remains to be explored. New insight into associations between inflammatory related determinants, genetic variation and cancer, could therefore add to the body of evidence of this intriguing relationship that is most probably involved in all aspects of cancer such as prevention, induction, promotion, treatment and prognosis.

Cytochrome P450 and cancer

Since the time that the human genome sequence mapping was nearly completed, the number of association studies of genetic polymorphisms and specific disease outcomes has grown exponentially. A PubMed search for terms such as 'polymorphism' and 'genotype' gives over 100,000 and 150,000 hits, respectively. This quantifies the impact genetic variability has on medical research nowadays. Nevertheless, for specific disease outcomes and individual genotypes, the number of studies is much less impressive. In 2002, an estimation was made that only a minority of genetic variations with a previously published positive association was studied for three times or more and that only 4% reported similar results in replicated studies.[196] This lack of consistency was also observed in our review on association studies between cytochrome P450 and cancer susceptibility, reported in chapter 2.1. Although the importance of the cytochrome P450 (CYP) family in (de)toxification was rapidly recognized after its discovery some 50 years ago, the number of association studies between CYP polymorphisms and cancer is still modest in perspective of its potential consequences. Next to this, comparability is highly affected by inconsistencies in study designs, control groups, phenotypes and genotypes of interest and population of choice, factors that were recognized by others as well.[196,353] Furthermore, our review aimed at providing an insight into cancer susceptibility affected by cytochrome P450 polymorphisms measured in serum leucocytes. Previous review articles focussed on the association between polymorphic CYP expression in (tumor) tissue and cancer.[32-34] Both strategies have their pros and cons and these will be discussed in the clinical implications section of this chapter. As became clear from our review, many studies indeed focussed on candidate genes but too little on a more complete causal biological pathway. As we reviewed, *CYP3A* gene polymorphisms are regularly associated with prostate cancer in Caucasians [89,90,176,205], but not with breast cancer [209]. Nevertheless, *CYP3A* genes are considered candidate genes for hormone related cancer association studies since the enzymes for which

they encode are involved in the metabolism of steroid hormones. Therefore, we performed studies on the association between genetic variation of CYP3A enzymes and hormone levels, which also included prostate- and breast cancer as endpoints (chapter 2.2. and 2.3.). Although both studies showed associations of some of the *CYP3A* genotypes with hormone levels, no associations were found for either of the cancers. For prostate cancer, results from our study were not in accordance with some other studies reported in literature [89,90,176,205], although studies with the most appropriate designs report insignificant findings as well [173,224,225]. Moreover, our negative results for the association with both cancers are in line with the results found for hormone levels. Testosterone levels that are considered to correlate with levels of the tumor growth- and differentiating hormone dihydrotestosterone were not associated with genotype. For breast cancer, some of the *CYP3A* genotypes that seemed to have a minor influence on estrogens were related to the age at onset of menarche and/or menopause as well. Although both hormone levels and reproductive life are associated with breast cancer, no associations were found between the genotypes and breast cancer. The studies presented in chapter 2.2. and 2.3. not only included multiple *CYP3A* gene variations and hormone levels in an *in vivo* setting, it additionally gave more insight into the potential mechanisms. As our results suggest, the association between *CYP3A* and prostate cancer remains questionable, not only as a result of inconsistent observations, but also since the association with testosterone cannot be clearly demonstrated. Together with the observations for breast cancer, it proves at least that *CYP3A* genotyping is not useful in predicting individual susceptibility to hormone related cancers, even though we cannot exclude that our negative findings for breast- and prostate cancer were caused by a limited sample size. Nevertheless, our results show some interesting associations of genotypes and hormone levels with gender-specificity that might initiate more detailed examinations of CYP3A involvement in steroid hormone metabolism.

An example of a study to explore the potential role of CYP polymorphisms as effect modifiers is reported in chapter 2.4. The study presented here aimed at increasing insight into the importance of *CYP2C9* slow metabolizing variants on the protective effect of NSAIDs on colorectal cancer incidence. Previous studies focussed on nonaspirin NSAIDs or aspirins and used different methods to explore effect modification. [135,278,280,289] Our study included both nonaspirin NSAIDs and aspirins and studied additive and multiplicative effect modification. Additionally, it included subanalyses on anatomical location, something that has been shown to be important as different colorectal locations are probably differentially affected by risk factors.[277] Both NSAID use and *CYP2C9* variant carriage were associated with a reduced risk of colorectal cancer, especially of proximal parts, which is in line with previous studies.[129,131,133-135,262,263,270-276] The observation that no effect modification was found could have been expected since duration of NSAID use was found to be more important than dose in the prevention of colorectal cancer. Nevertheless, the importance of dose was suggested in some studies that found effect modification of *CYP2C9* variants on NSAID use in the prevention of colorectal cancer. By studying different types of NSAIDs and the use of two different statistical methods

for effect modification, our study provided additional evidence that duration of NSAID use is indeed more important than dose in the prevention of colorectal cancers. The associations found between slow-metabolizing variants and some NSAID levels *in vitro* seemed therefore not to have any significant clinical impact.

Inflammation and cancer

Despite the fact that an association between inflammation and cancer is commonly accepted, the underlying mechanisms are still a matter of debate. Studies have shown associations between viruses (i.e. Human Immunodeficiency Virus or Human Papilloma Virus) or diseases with an enhanced immunological status (i.e. rheumatoid arthritis or inflammatory bowel diseases) and cancer development. Furthermore, it has consistently been shown that tumor growth results in cell damage that consecutively leads to activation of the immune system. Several post-operative studies report associations between levels of inflammatory markers, such as C-reactive protein (CRP), and tumor load.[331-333] As a consequence of the heterogenous association of inflammation and cancer, we introduced an induction (as in the first examples) and a response (as in the final examples) hypothesis to distinguish between several aspects of cancer growth.

The first study that was described in this chapter contributes to the large body of evidence that NSAIDs protect against colorectal cancer, especially those of the proximal colon.[262,263,270-276] Contributions of our study to the existing literature, were the analyses with cumulative medication use, the population based design, which resulted in the availability of a large number of potential confounders and the inclusion of an analysis on the effect of NSAIDs after cessation. Although limited by a small sample size, the analysis on cessation revealed that a protective effect for colorectal cancer could still be found five years after discontinuation of the drug. This is in line with other studies that report the most important effects of cancer prevention in early cancer stages.[288-291] Effects of cessation are therefore not to be expected to become visible already in the first years after the discontinuation but might require a longer follow-up. However, we simultaneously observed an effect of medication already after 1 year of cumulative use, which seems to be in contrast to the previous explanation. Therefore, it might be possible that our findings indicate effects of NSAIDs in both early and later stages. An explanation for the protective effect of NSAIDs on carcinogenesis that is most frequently put forward is inhibition of cyclooxygenases (COX) that convert arachidonic acid into prostaglandins and harmful by-products such as peroxides and free radicals. Arachidonic acid is believed to result in anti-tumor effects that include reduction of neovascularisation and enhancement of apoptosis, whereas the by-products induce cell damage. NSAID use leads to the preferred situation of high arachidonic acid levels and low prostaglandin- and by-product levels. Considering the importance of COX enzymes, we hypothesized that the protective effect of NSAIDs might depend on the degree of COX expression. Therefore, the second study of chapter 3 was conduced in which cumulative NSAID use was studied stratified for *COX-2* genotype. The variant type of this gene had been found to be associated with 30% lower

expression levels of the corresponding protein.[309] Surprisingly, *COX-2* genotype did not affect the inverse association between NSAID use and cancer. As a result, we hypothesized that the difference in COX-2 expression levels introduced by the *COX-2* genotype might only be revealed in active inflammatory stages. This had been brought forward in a previous study as well, in which it was reported that the inverse association of NSAIDs and cancer was most pronounced in persons with the highest C-reactive protein levels.[325] Therefore, we studied NSAID use in the five years before the index date. In this time, a cancer is most probably already present but in its preclinical stage and therefore not detected. Nevertheless, results were similar. These results and the insignificant differences between COX-specific NSAIDs on the inverse association underline the importance of other targets for NSAIDs than only COX, which has also been proposed by others.[326-328]

In the final study, we aimed at gaining more insight into the induction hypothesis. A number of statistical methodological procedures were used to study the association between baseline C-reactive protein levels and *CRP* haplotypes that were related to CRP levels and cancer induction (see paragraph on methodological considerations). Although CRP levels were proposed as a proxy of inflammation by others as well [337,339], the effect of *CRP* haplotype on cancer was never studied before. We studied associations with the four most common cancers and were only able to find an association with lung cancer. The insignificant findings for breast- and prostate cancer were not surprising, as these cancers are supposed to be hormone-dependent and not the result of chronic inflammation. We hypothesized that inflammation might be associated with all cancers exposed to exogenous stimuli, such as colorectal and lung cancers. Anatomical and histological subtypes of these cancers (squamous cell carcinoma of the lung and nonsigmoid colon cancers) showed the highest risk in persons with elevated CRP levels at baseline. As we included a lag-time of five years and lung cancers have a high potential to lead to early mortality, the results presented in this final study seem to point at an influence of chronic respiratory inflammation on lung cancer.

METHODOLOGICAL CONSIDERATIONS

Setting

Cancer is a common disease with a relatively high incidence rate, especially in older individuals. Therefore, The Rotterdam Study, that is a cohort-study of individuals of 55 years and older must be considered appropriate for research questions on cancer. For reasons of efficiency and impact on disease burden, the four most common non-skin related cancers (i.e. lung, colorectal, breast and prostate cancer) were studied, that all have the highest incidence in older individuals. Breast cancer at younger ages has an important impact on public health as well, but these are mainly related to single genes like *BRCA-1* or *BRCA-2* and were not included in this thesis.

Analyses and effect modification

For most of the analyses performed, Cox' proportional hazard models were used since cancer has a recognizable endpoint in time. Follow-up time was calculated as the time between the start of entering The Rotterdam Study until a pathologically confirmed diagnosis of cancer, death or the end of the study period. As cancer cases were validated until the end of October 2004, this date was set as study end date for all analyses. By calculating the follow-up time until a diagnosis of one of the four cancers of interest, it was possible to censor for other cancers in the analyses on individual malignancies.

For all analyses we proposed some potential confounders based on their known association with the outcome and therefore frequently reported in the literature. Statistical models were intended to be built with those covariates that changed the point estimate by ten or more percent. Nevertheless, none of the covariates had such an effect on the point estimate. Therefore, models were built that included the most important available risk factors for the outcome that were previously mentioned by others.

Genes and their polymorphic protein expression were considered as being potential effect modifiers, possibly leading to a diagnosis of cancer at a younger age. There are two types of effect modification; additive (or biological) effect modification and multiplicative effect modification. These require different statistical methods. Multiplicative effect modification, an explicit mathematical approach that estimates whether the effect of a combination of two factors is more than the *product* of the individual ones, is most frequently used. However, FIGURE 1 shows that this may lead to the rejection of the possibility of biological effect modification. In this thesis we described biomedical research, and it seemed therefore reasonable to study additive effect modification (estimate of the combination of two factors is more than the *sum* of the individual estimates), as well. Additive effect modification has three expression measures; synergy index (SI), relative excess risk due to interaction (RERI) and attributable proportion (AP). They all follow from the risk ratios of the four possible combinations (i.e. neither of the factors, factor A or factor B and both factors simultaneously).[268,269] (TABLE 1) For

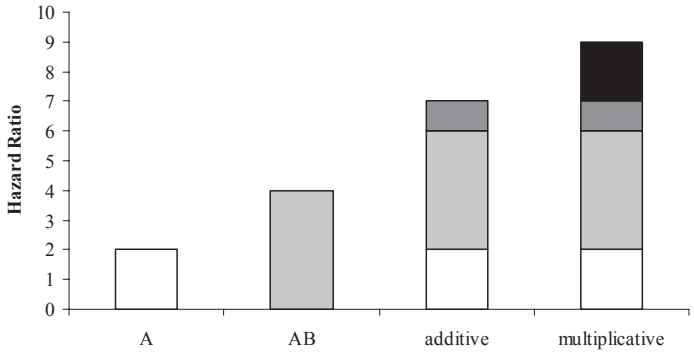

Figure 1. Additive and multiplicative effect modification

Table 1. Additive effect modification measures

		Factor B	
		no	yes
Factor A	no	RR_{00}	RR_{01}
	yes	RR_{10}	RR_{11}

SI = $[RR_{11} - 1] / [[RR_{10} - 1] + [RR_{01} - 1]]$

RERI = $RR_{11} - RR_{10} - RR_{01} + 1$

AP = $RERI / RR_{11}$

SI = synergy index, RERI = relative excess risk (reduction) due to interaction, AP = attributable proportion, RR = relative risk.

our analyses, we used the RERI as this estimate is most easy to interpret. As we additionally explored preventative estimates, the RERI was used as the relative excess risk *reduction* due to interaction, but the principal remains the same. Both types of effect modification should be presented in future studies.

Use of medication

Information about medication use until 1991 was reported by the participants during the interview at the start of their inclusion in The Rotterdam Study. Information about medication use since 1991 was available for all participants on a day-to-day basis by using computerized pharmacy data. This made it possible to calculate the cumulative dose and duration per drug for each participant, which results in a higher precision of information about medication use and a 'power boost' of participants. To study cumulative medication use, a complex dataset has to be built.[266] Case-sets are made that include one case and the remainder of the cohort as controls. The case in the first case-set is the first case, which occurs during follow up. In each subsequent case-set previous cases have been censored and deleted, and person-time experience comes from the then occurring case and the remainder of the cohort which is still alive and eligible at the date of diagnosis of the cancer case. Eventually, the number of case-sets is identical to the number of events under study. After building these case-sets, the cumulative dose and duration of use for the drug of interest until the index date, is calculated per case-set for all including participants. The index date is defined as the date of diagnosis of the case representing the case-set. This approach makes it possible to compare cumulative duration/dose at the time of a diagnosis of cancer with the cumulative duration/dose of all other persons in the cohort at exactly the same time. Note that the use of different case-sets implies that a control participant can attribute different cumulative durations/doses to each case-set and that it is even possible that a person serves as a control in a number of case-sets before he becomes a case himself and is then included in his 'own' case-set. (FIGURE 2) Next, it is possible to investigate whether there are differences in disease outcomes when medication use is stratified in current and past use. Finally, these two exposure measures (duration/dose and current/past use) can be studied simultaneously.

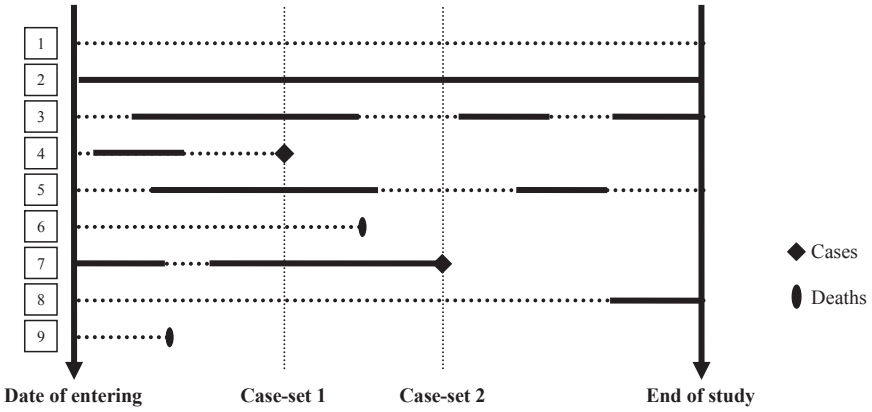

Figure 2. Methodology of time-varying medication exposure

Participant 1 in FIGURE 2 contributes person-time to the total follow-up period but is never a user. Participant 2 contributes person-time to the total follow-up period and uses the medication of interest during this whole period. Participant 3 contributes person-time to the total follow-up period and uses medication intermittently. He contributes another cumulative duration-time to case-set 1 than 2 and is current user in case-set 1 and past user in case-set 2. Participant 4 contributes person-time until he becomes a case. At the date of diagnosis, this participant is a past user. Participant 5 is a current user in case-set 1 and a past user in case-set 2. Participant 6 serves as control in case-set 1 and has no cumulative medication use. Since he dies between the two case-sets, this person is not included in case-set 2. Participant 7 serves as a control for case-set 1 and becomes a case in case-set 2. Participant 8 contributes person-time to the total follow-up period and only uses medication in the final period. Since no cases occur after he starts using, this medication use is not included in the analyses. Participant 9 dies before the first case(-set) and is therefore not included in the analyses.

Induction and promotion

The studies presented in chapter 3.2 and 3.3 include some methodological considerations to explore differences in effects on induction and/or promotion.

Chapter 3.2. studied the effect of NSAID use and *COX-2* genotype. In the initial analyses, cumulative duration of NSAID use was considered as presented in the previous paragraph. (FIGURE 3, number 1, example participant 7) These analyses aimed at studying the effect of NSAIDs and *COX-2* on the induction of cancer. To study promotion, we postulated that the effect of COX-2 expression as a result of *COX-2* gene variation might only be important when the system was biologically active, as in inflammatory stages. We therefore subtracted five years from the index date and calculated the cumulative duration of NSAID use for these five years to study promotion. (FIGURE 3, number 2) The data on medication use before these five years were used to minimize misclassification of the exposure, as we included only persons

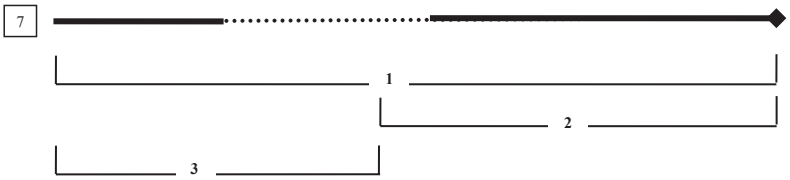

Figure 3. Methodology of analyzing induction and promotion

in the analyses who did not use NSAIDs in the years, which preceded the specified five-year period. (FIGURE 3, number 3)

Chapter 3.3. studied the effect of C-reactive protein levels and haplotypes on cancer. As was explained in the beginning of the current chapter, studies on inflammation and cancer are heterogeneous as a result of factors that contribute either to the induction or to the response hypothesis. Therefore, we included a minimum of five-year lag-time between the date of entering The Rotterdam Study, the time at which blood samples were drawn for C-reactive protein measurements, and the first cancer case. As we go back to FIGURE 2, this could have meant that the cancer case of case-set 1 occurred within 5-years after inclusion in The Rotterdam Study and was therefore excluded of this specific analysis. By following this approach, we were able to partly exclude the rise in CRP induced by tumor growth itself. Although for breast cancers with their long latent period, a period of 5 years may be too short, for lung cancers this seems long enough to distinguish between induction and promotion. Although the choice of the five-year subtraction is arbitrary, the size and follow-up period in The Rotterdam Study precluded the inclusion of longer latent periods because the number of remaining cases would be too low.

STRENGTHS AND LIMITATIONS

As in all observational cohorts, studies performed in The Rotterdam Study are potentially prone to selection, information and confounding bias. However, careful design and analysis might deal with most of these problems. In the following, these three types of bias will be discussed in the perspective of the studies presented in this thesis, although specific considerations of bias are discussed in the individual chapters.

All cohorts comprise a selection of the total population. Whether there is bias depends on the effect of this selection on the true association between the determinant and outcome. As only persons of 55-years and older were asked to participate to The Rotterdam Study, this might have led to the inclusion of relatively healthy elderly, since they were able to survive for at least 55-years. Self-selection is possible whenever the group of people being studied decides whether they will participate. Again, healthier persons are more likely to participate. On top of this, within the total number of participants, only persons who provided blood samples

at baseline were included in our analyses. Our study cohorts therefore included healthier persons at baseline. Nevertheless, this probably did not affect our results as we merely studied genotypes and their interaction with other determinants. Furthermore, genotypes were in Hardy Weinberg equilibrium in all studies presented, which indicates that participation was independent of genotype. Another important selection criteria, that was introduced by using data from The Rotterdam Study, is the inclusion of persons from one genetic background. 99.8% of all participants are from Caucasian descent. Although it is recommended to study populations with the same racial background to excluded population stratification [354] (a form of selection bias as a result of ethnical differences) as we did in our studies, one must consider the generalizability of the results. Interactions of genes with other factors are especially dependent on the allele frequencies of the genotypes and the occurrence of the factor of interest in a population.

Misclassification bias is unlikely as data collection was performed without knowledge of the research hypothesis. Nevertheless, some issues must be discussed. First, although all genotypes were in HWE, the methodology of the Taqman assay might induce some non-differential misclassification as genotypes are allocated subjectively according to the groups of dots as presented in chapter 1.4. Nevertheless, this random misclassification was limited by using control samples, performing re-do's and allowing three persons to evaluate the Taqman assay results. Second, although we had access to medication use on a day-to-day basis, no or limited information about use before January 1st 1991 and over-the-counter use was available. In our studies, we tried to minimize the potential misclassification induced by these factors. We performed some additional analyses, excluding uncertainties about medication use before 1991. Additionally, over-the-counter use of NSAIDs was considered less important as in The Netherlands analgesics are only to a limited extent freely available since 1996 and in relatively low recommended doses. Moreover, it is unlikely that persons with a prescribed NSAID also use them over-the-counter, since all NSAIDs on prescription – including long-term use - are fully reimbursed. Misclassification of malignancies is also unlikely, as all cancers were pathologically confirmed. Nevertheless, preclinical and therefore undiagnosed cancers might have diminished the effects.

Confounding was adjusted for in the analyses and is thought to be less important in genomic association studies. However, potential effects of a large range of inducers and inhibitors, such as several drugs, that result in altered biological functions of the enzymes, must not be ruled out. Another point of discussion is the censoring for other cancers. As we only had data on the four most common cancers, it is possible that cancers that were not included in our validation influenced results. Nevertheless, as the four malignancies studied in this thesis account for the majority of the disease burden by cancer, this is most probably of less importance. In studies on medication use, specific types of confounding are possible, e.g. confounding-by-indication and its specific form confounding-by-severity. In our studies, another type of bias might have been possible. Protopathic bias occurs when a pharmaceutical agent is prescribed for an

early manifestation of a disease that has not yet been recognized.[355] NSAIDs might have been prescribed to reduce pain that is the result of tumor growth. However, since we were mainly interested in long-term use and additionally included some specific subanalyses to reduce this type of bias, spurious associations are not likely to result from this issue.

Next to the three types of bias discussed above, other intriguing aspects of epidemiological studies must be mentioned. As a result of logistic problems or selective participation, most epidemiological studies encounter missing data for several covariates. Missing status can either be random (as with logistic problems) or non-random (as with selective participation). The impact of missing determinants as a result of the refusal of providing blood samples was discussed above. The studies presented in this thesis were most hampered by missing status (23%) of dietary factors. Although missing status was correlated with older and less healthier individuals, it was randomly distributed over the genotypes and therefore unlikely to have caused spurious associations. However, it might have led to an underestimation of the true risk. We dealt with this issue in several ways. First, we studied the impact of missing status on the association and observed that missing status was not a confounder. Therefore, we performed complete case analyses, in which we excluded all persons with missing values on covariates. As this resulted in a major reduction of the power, we imputed missings with a single imputation method, i.e. expectation maximization (EM).[220] EM alternates between performing an expectation (E) step, which computes an expectation of the likelihood by including the latent variables as if they were observed, and a maximization (M) step, which computes the maximum likelihood estimates of the parameters by maximizing the expected likelihood found on the E step. The parameters found on the M step are then used to begin another E step, and the process is repeated. Hazard ratios for the complete case analyses and for the analyses with imputed data were similar, although confidence bounds were narrower in the last analyses and were therefore reported in the results of the performed studies.

Limitations by a low sample size may occur especially in those studies on genetic associations that are thought to result in only minor differences in the risk per genotype.[194] However, if the association is strong, relatively small cohorts may still be able to detect genotype differences. Only a minority of significant associations were observed in the studies presented in this thesis. Our studies were definitely hampered by a limited study size. Rough power calculations show that for minor effects thousands of cases are needed. Nevertheless, with 150-200 cases we could have detected risk estimates around 4-6, considering the individual allele frequencies, the incidence of the diseases, a power of 80% ($\beta = 0.2$) and 95% confidence ($\alpha = 0.05$).

According to the issues put forward in the previous paragraph, most potential limitations were encountered by appropriate study designs and additional subanalyses. Furthermore, the strengths of the available information on day-to-day medication and the availability – besides genotypes – of intermediate outcomes such as hormone levels, made it possible to study much more then mere associations between genotype and cancer.

CLINICAL IMPLICATIONS AND FUTURE DIRECTIONS

Despite the exponential growth of genetic association studies, this domain is still in its infancy. Direct clinical implications need further studying. Nevertheless, the growing mass of results, to which the studies described in this thesis add, show not only some of the pitfalls of genetic association studies but also interesting areas that may guide future research. Moreover, the number of useful results from studies with potential clinical implications increases. An example of such a clinical application is the drug-gene interaction between anticoagulating drugs and slow-metabolizing variants of the *CYP2C9* gene.[356,357]

In view of results found in the studies presented in this thesis, we are able to stress the importance of some specific topics. First, although it would have been nice to find a strong association between polymorphic genes that could have been measured out of DNA from serum leucocytes and cancer as reviewed in chapter 2.1., polymorphic expression levels in specific tissues have probably much more potential to be found associated with specific disease outcomes as this approach minimizes the effect of other counterbalancing factors. As many inducers and inhibitors affect protein expression, consideration of polymorphic expression levels reduces the bias of other known and unknown factors. Second, our results of the two studies on *CYP3A* genotypes show some interesting associations with hormone levels, which should encourage future studies to investigate the *CYP3A* genes in more detail. Third, results of the *CYP2C9* genotype study stress the importance of a clinical hypothesis in the beginning of an association study. Although *in vitro* studies showed impact of *CYP2C9* on NSAID levels, a drug-gene interaction was not expected to be found *in vivo*, as duration of NSAID use in generally known to be much more important than dose in the prevention of colorectal cancer. [272] The studies on inflammation and cancer reveal that it is important to distinguish between the different stages of cancer development. Furthermore, chapter 3.2 shows that NSAIDs might have COX-independent properties, which could initiate new strategies for colorectal cancer prevention. Additionally, the results of the C-reactive protein levels and haplotype study have the most clinical impact, as it underlines the importance of reducing chronic inflammatory stages to prevent cancer development. In summary, the first part of this thesis, on cytochrome P450 gene variation and cancer, indicates that more detailed and standardised approaches are necessary in future studies on this topic. The second part of this thesis, on inflammation and cancer, indicates the importance of reducing inflammation to prevent at least some cancers and underlines the importance of studies that explore COX-independent effects of NSAIDs.

The conclusion that further research is necessary is the inevitable result of a relatively new research field, but it is a knockdown argument that can benefit from suggestions that are more explicit. Therefore, in the following, different strategies for genetic association studies will be discussed. The aim of genetic association studies is to identify disease-causing genes. Until recently, the ways to do this could roughly be distinguished in two approaches. The first approach uses linkage disequilibrium to identify areas of the genome that may carry a gene

associated with the phenotype of interest. After identifying these areas that mostly include hundreds of genes, fine-mapping of the regions through the evaluation of candidate genes is necessary after which specific polymorphisms in this gene are detected. As this method was for a long period time-consuming and often not feasible due to high costs, a second approach became the preferred method. It builds on the existing biological knowledge and chooses candidate genes that are involved in specific pathophysiological pathways of the disease or phenotype of interest. Polymorphisms within these genes are additionally chosen based on their potential functional significance to the encoded protein. The polymorphisms of the *CYP3A*, *CYP2C9*, and *COX-2* genes studied in this thesis were all identified through this second approach. Recently, an increasing number of studies uses a haplotype approach that includes blocks of linkage disequilibrium in which only a few common polymorphims are needed to "tag" more than 90% of the diversity of common haplotypes (with a frequency above 5%).[358,359] This resulted in specific databases on haplotype tagging SNPs as was mentioned in chapter 1.4. and that was used for the haplotype approach in the association between the C-reactive protein gene and cancer in this thesis. The most recent contribution is the introduction of chips to perform genome-wide screens. These chips make it easy to perform high-throughput genotyping of 500.000 SNPs at ones. Genome-wide assocation studies will become state of the art in the following years.

With these new genotyping availabilities within reach, some scientists, geneticists and clinicians argue about the relevance of studying genotype-phenotype associations.[360-362] They prefer to look at genome-phenome associations in which "phenome" stands for the complex network of several phenotypes. They introduce models of biological systems that are reductions of separate genotype-phenotype spaces. In this situation, one factor is strongly dependent on other factors that are by themselves pleiotropic as well. The models are intended to make it possible to represent and investigate biology in detail. FIGURE 4 gives an example of a biological system of which many express the complexity of a phenome. The phenotypic

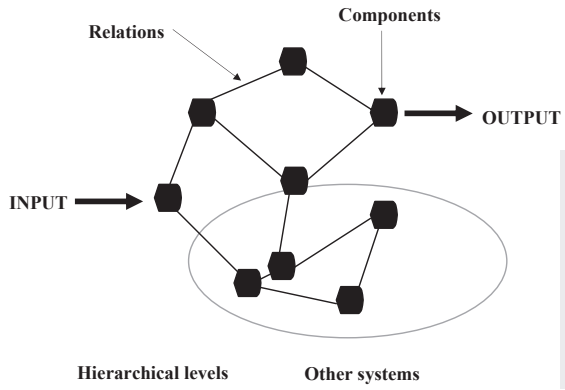

Figure 4. Schematic overview of a biological system

representation of a biological system is multidimensional and can be extended or reduced according to newly acquired knowledge.

As also seen in our studies, polymorphisms affecting "proximal" phenotypes that are located closely to the *input* factor in FIGURE 4, such as RNA or protein expression or circulating protein levels, have been shown associated far more consistently and often stronger than polymorphisms and "distal" phenotypes that are located on a great distance of the original *input* factor, such as disease end-points (*output*). This inconsistency can be explained by the complexity of the network in which the disease end-point is involved and in which the gene of interest exerts its effect. It was measured that the distance between a genotype and a phenotype is a function of the complexity of the network of components.[363] When you think of complex traits or diseases such as cancer as an expression of a complex biological system, genes will only be related with this complex disease through networks of relationships. This idea makes it unlikely that the current method of studying genotype-disease phenotype is very rewarding. But how can these complex networks be studied? It needs thousands of cases for which multiple cohorts have to be merged into one giant mega-study. Additionally, it must include hundreds of genes and covariates. One can understand that for practical reasons this is far from feasible at the moment. However, it is good to have this idea in mind when performing genetic association studies with a candidate gene approach.

The idea of complex biological systems and the fact that their components may belong to several other biological systems or that one biological system is part of a larger biological system seems to fit with the idea of the multistep process explained by Rothman and described in the first chapter of this thesis. It underlines not only the necessity of multiple component causes to lead to a sufficient cause, but it also stresses the importance of considering that one component cause may be involved in several sufficient causes that either result in the same disease of interest or other diseases. More practically, a sufficient cause will only lead to a disease when all component causes are met. Otherwise, the process will not proceed, and considering association studies, will not lead to significant findings.

Finally, in light of all these scientific high standing procedures it must be recognized that the most important factor in a human interface model is of course the human.

5. Literature

1. Visser OSS, van Dijck JAAM. *Incidence of cancer in the Netherlands 1999/2000*, (Vereniging van Integrale kankercentra, Utrecht, 2003).
2. Rothman KJ. Causes. *Am J Epidemiol* 1976; 104: 587-592.
3. Rothman KJ. Causes. 1976. *Am J Epidemiol* 1995; 141: 90-95; discussion 89.
4. Integrale Kanker Centra (IKC), available at *http://www.ikc.nl/vvik/kankerregistratie*. 2007.
5. Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2000; 405: 847-856.
6. Hofman A, Grobbee DE, de Jong PT and van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* 1991; 7: 403-422.
7. Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE and Witteman JC. Dietary assessment in the elderly: validation of a semi-quantitative food frequency questionnaire. *Eur J Clin Nutr* 1998; 52: 588-596.
8. Miller SA, Dykes DD and Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; 16: 1215.
9. World Health Organization (WHO), available at *http://www.who.int*. 2007.
10. Collaborating Centre for Drug Statistics Methodology, available at *http://www.whocc.no.* 2007.
11. International Code for Primary Care (ICPC), available at *www.who.int/entity/classifications/icd/adaptations/icpc2/en/index.* 2007.
12. International Classification of Diseases - tenth version. *http://www.who.int/classifications/icd/en*. 2007.
13. Landelijke Medische Registratie (LMR), available at *http://www.prismant.nl.* 2007.
14. Stichting Laboratorium Pathologie en Cytologie (PATHAN), available at *http://www.pathan.nl.* 2007.
15. Homedes N. The Disability-Adjusted Life Year (DALY) definition, measurement and potential use, available at *http://www.worldbank.org/html/extdr/hnp/hddflash/workp/wp_00068.html.* 2007.
16. Rijksinstituut voor Volksgezondheid en Milieu (RIVM), available at *http://www.rivm.nl.* 2007.
17. International Association of Cancer Registries (IACR), available at *http://www.iacr.com.fr.* 2007.
18. Chattopadhyay A. Genetic Variations Resources, available at *http://www.ncbi.nlm.nih.gov/Class/NAWBIS/Modules/Variation/var18.html.* 2007
19. NCBI SNP database, available at *http://www.ncbi.nlm.nih.gov/SNP/index.html.* 2007.
20. AppliedBiosystems, Nieuwerkerk aan den IJssel, The Netherlands, available at *http://www.appliedbiosystems.com/europe.* 2007.
21. Online Mendelian Inheritance in Man (OMIM), available at *http://www.ncbi.nih.gov/OMIM/index.html*. 2007.
22. Human Genome Variation Database (HGVbase), available at *http://hgvbase.cgb.ki.se*. 2007.
23. SNP Consortium, available at *http://snp.cshl.org*. 2007.
24. International HapMap Project, available at *http://www.hapmap.org*. 2007.
25. HaploView, available at *http://www.broad.mit.edu/mpg/haploview*. 2007.
26. PHASE Software, available at *http://www.stat.washington.edu/stephens/home.html.* 2007.
27. HaploStat software, available at *http://alla.cs.gsu.edu/~software/diseaseAssociation.html.* 2007.
28. R-Project for statistical computing software, available at *http://www.r-project.org.* 2007.
29. Ponder BA. Cancer genetics. *Nature* 2001; 411: 336-341.
30. Rothman N, Wacholder S, Caporaso NE, Garcia-Closas M, Buetow K and Fraumeni JF, Jr. The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. *Biochim Biophys Acta* 2001; 1471: C1-10.
31. Thun MJ, Henley SJ, Burns D, Jemal A, Shanks TG and Calle EE. Lung cancer death rates in lifelong nonsmokers. *J Natl Cancer Inst* 2006; 98: 691-699.
32. Agundez JA. Cytochrome p450 gene polymorphism and cancer. *Curr Drug Metab* 2004; 5: 211-224.

33. Oyama T, Kagawa N, Kunugita N, Kitagawa K, Ogawa M, Yamaguchi T, Suzuki R, Kinaga T, Yashima Y, Ozaki S, Isse T, Kim YD, Kim H and Kawamoto T. Expression of cytochrome P450 in tumor tissues and its association with cancer development. *Front Biosci* 2004; 9: 1967-1976.
34. Rodriguez-Antona C and Ingelman-Sundberg M. Cytochrome P450 pharmacogenetics and cancer. *Oncogene* 2006; 25: 1679-1691.
35. Wang LD, Zheng S, Liu B, Zhou JX, Li YJ and Li JX. CYP1A1, GSTs and mEH polymorphisms and susceptibility to esophageal carcinoma: study of population from a high- incidence area in north China. *World J Gastroenterol* 2003; 9: 1394-1397.
36. Chen SY, Liu TY, Shun CT, Wu MS, Lu TH, Lin JT, Sheu JC, Santella RM and Chen CJ. Modification effects of GSTM1, GSTT1 and CYP2E1 polymorphisms on associations between raw salted food and incomplete intestinal metaplasia in a high-risk area of stomach cancer. *Int J Cancer* 2004; 108: 606-612.
37. Suzuki S, Muroishi Y, Nakanishi I and Oda Y. Relationship between genetic polymorphisms of drug-metabolizing enzymes (CYP1A1, CYP2E1, GSTM1, and NAT2), drinking habits, histological subtypes, and p53 gene point mutations in Japanese patients with gastric cancer. *J Gastroenterol* 2004; 39: 220-230.
38. Wadelius M, Autrup JL, Stubbins MJ, Andersson SO, Johansson JE, Wadelius C, Wolf CR, Autrup H and Rane A. Polymorphisms in NAT2, CYP2D6, CYP2C19 and GSTP1 and their association with prostate cancer. *Pharmacogenetics* 1999; 9: 333-340.
39. Felix CA, Walker AH, Lange BJ, Williams TM, Winick NJ, Cheung NK, Lovett BD, Nowell PC, Blair IA and Rebbeck TR. Association of CYP3A4 genotype with treatment-related leukemia. *Proc Natl Acad Sci U S A* 1998; 95: 13176-13181.
40. Ramachandran S, Fryer AA, Smith AG, Lear JT, Bowers B, Hartland AJ, Whiteside JR, Jones PW and Strange RC. Basal cell carcinomas: association of allelic variants with a high-risk subgroup of patients with the multiple presentation phenotype. *Pharmacogenetics* 2001; 11: 247-254.
41. Itoga S, Nomura F, Makino Y, Tomonaga T, Shimada H, Ochiai T, Iizasa T, Baba M, Fujisawa T and Harada S. Tandem repeat polymorphism of the CYP2E1 gene: an association study with esophageal cancer and lung cancer. *Alcohol Clin Exp Res* 2002; 26: 15S-19S.
42. Miyoshi Y, Iwao K, Ikeda N, Egawa C and Noguchi S. Breast cancer risk associated with polymorphism in CYP19 in Japanese women. *Int J Cancer* 2000; 89: 325-328.
43. Kumazawa T, Tsuchiya N, Wang L, Sato K, Kamoto T, Ogawa O, Nakamura A, Kato T and Habuchi T. Microsatellite polymorphism of steroid hormone synthesis gene CYP11A1 is associated with advanced prostate cancer. *Int J Cancer* 2004; 110: 140-144.
44. Suzuki K, Nakazato H, Matsui H, Koike H, Okugi H, Ohtake N, Takei T, Nakata S, Hasumi M and Yamanaka H. Association of the genetic polymorphism of the CYP19 intron 4[TTTA]n repeat with familial prostate cancer risk in a Japanese population. *Anticancer Res* 2003; 23: 4941-4946.
45. Song N, Tan W, Xing D and Lin D. CYP 1A1 polymorphism and risk of lung cancer in relation to tobacco smoking: a case-control study in China. *Carcinogenesis* 2001; 22: 11-16.
46. London SJ, Yuan JM, Coetzee GA, Gao YT, Ross RK and Yu MC. CYP1A1 I462V genetic polymorphism and lung cancer risk in a cohort of men in Shanghai, China. *Cancer Epidemiol Biomarkers Prev* 2000; 9: 987-991.
47. Alexandrie AK, Sundberg MI, Seidegard J, Tornling G and Rannug A. Genetic susceptibility to lung cancer with special emphasis on CYP1A1 and GSTM1: a study on host factors in relation to age at onset, gender and histological cancer types. *Carcinogenesis* 1994; 15: 1785-1790.
48. Nelson D. Cytochrome P450s in humans, available at *http://drnelson.utmem.edu/P450lect.html*. 2007.
49. Ingelman-Sundberg M, Oscarson M and McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *Trends Pharmacol Sci* 1999; 20: 342-349.
50. Ingelman-Sundberg M. Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. *Naunyn Schmiedebergs Arch Pharmacol* 2004; 369: 89-104.

51. Chen S, Xue K, Xu L, Ma G and Wu J. Polymorphisms of the CYP1A1 and GSTM1 genes in relation to individual susceptibility to lung carcinoma in Chinese population. *Mutat Res* 2001; 458: 41-47.
52. Garcia-Closas M, Kelsey KT, Wiencke JK, Xu X, Wain JC and Christiani DC. A case-control study of cytochrome P450 1A1, glutathione S-transferase M1, cigarette smoking and lung cancer susceptibility. *Cancer Causes Control* 1997; 8: 544-553.
53. Ohshima S and Xu Y. p53 gene mutations, and CYP1A1 and GSTM1 genotypes in pulmonary squamous cell carcinomas. *Mol Pathol* 1997; 50: 108-110.
54. Tefre T, Ryberg D, Haugen A, Nebert DW, Skaug V, Brogger A and Borresen AL. Human CYP1A1 (cytochrome P(1)450) gene: lack of association between the Msp I restriction fragment length polymorphism and incidence of lung cancer in a Norwegian population. *Pharmacogenetics* 1991; 1: 20-25.
55. Kawajiri K, Nakachi K, Imai K, Watanabe J and Hayashi S. Germ line polymorphisms of p53 and CYP1A1 genes involved in human lung cancer. *Carcinogenesis* 1993; 14: 1085-1089.
56. Huang CS, Chern HD, Chang KJ, Cheng CW, Hsu SM and Shen CY. Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. *Cancer Res* 1999; 59: 4870-4875.
57. Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J and Boffetta P. Metabolic polymorphisms and susceptibility to cancer. *IARC Sci Publ* 1999; 148.
58. Malaiyandi V, Sellers EM and Tyndale RF. Implications of CYP2A6 genetic variation for smoking behaviors and nicotine dependence. *Clin Pharmacol Ther* 2005; 77: 145-158.
59. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM and Ederer F. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med* 1993; 328: 1365-1371.
60. Ma Q. Induction of CYP1A1. The AhR/DRE paradigm: transcription, receptor regulation, and expanding biological roles. *Curr Drug Metab* 2001; 2: 149-164.
61. Schwarz D, Kisselev P, Cascorbi I, Schunck WH and Roots I. Differential metabolism of benzo[a]pyrene and benzo[a]pyrene-7,8-dihydrodiol by human CYP1A1 variants. *Carcinogenesis* 2001; 22: 453-459.
62. Marnett LJ. Aspirin and the potential role of prostaglandins in colon cancer. *Cancer Res* 1992; 52: 5575-5589.
63. Sachse C, Brockmoller J, Bauer S and Roots I. Functional significance of a C-->A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. *Br J Clin Pharmacol* 1999; 47: 445-449.
64. Han XM, Ouyang DS, Chen XP, Shu Y, Jiang CH, Tan ZR and Zhou HH. Inducibility of CYP1A2 by omeprazole in vivo related to the genetic polymorphism of CYP1A2. *Br J Clin Pharmacol* 2002; 54: 540-543.
65. van Schaik RH. Cancer treatment and pharmacogenetics of cytochrome P450 enzymes. *Invest New Drugs* 2005; 23: 513-522.
66. Aklillu E, Oscarson M, Hidestrand M, Leidvik B, Otter C and Ingelman-Sundberg M. Functional analysis of six different polymorphic CYP1B1 enzyme variants found in an Ethiopian population. *Mol Pharmacol* 2002; 61: 586-594.
67. Aklillu E, Ovrebo S, Botnen IV, Otter C and Ingelman-Sundberg M. Characterization of common CYP1B1 variants with different capacity for benzo[a]pyrene-7,8-dihydrodiol epoxide formation from benzo[a]pyrene. *Cancer Res* 2005; 65: 5105-5111.
68. Kamataki T, Nunoya K, Sakai Y, Kushida H and Fujita K. Genetic polymorphism of CYP2A6 in relation to cancer. *Mutat Res* 1999; 428: 125-130.
69. Cauffiez C, Lo-Guidice JM, Quaranta S, Allorge D, Chevalier D, Cenee S, Hamdan R, Lhermitte M, Lafitte JJ, Libersa C, Colombel JF, Stucker I and Broly F. Genetic polymorphism of the human cytochrome CYP2A13 in a French population: implication in lung cancer susceptibility. *Biochem Biophys Res Commun* 2004; 317: 662-669.

70. Shi WX and Chen SQ. Frequencies of poor metabolizers of cytochrome P450 2C19 in esophagus cancer, stomach cancer, lung cancer and bladder cancer in Chinese population. *World J Gastroenterol* 2004; 10: 1961-1963.
71. Uematsu F KH, Motomiya M, Abe T, Ishioka C, Kanamaru R, Sagami I, Watanabe M. Human cytochrome P450IIE1 gene: DraI polymorphism and susceptibility to cancer. *Tohoku J Exp Med* 1992; 168: 113-117.
72. Dally H, Edler L, Jager B, Schmezer P, Spiegelhalder B, Dienemann H, Drings P, Schulz V, Kayser K, Bartsch H and Risch A. The CYP3A4*1B allele increases risk for small cell lung cancer: effect of gender and smoking dose. *Pharmacogenetics* 2003; 13: 607-618.
73. Yeh KT, Chen JC, Chen CM, Wang YF, Lee TP and Chang JG. CYP3A5*1 is an inhibitory factor for lung cancer in Taiwanese. *Kaohsiung J Med Sci* 2003; 19: 201-207.
74. Fujieda M, Yamazaki H, Saito T, Kiyotani K, Gyamfi MA, Sakurai M, Dosaka-Akita H, Sawamura Y, Yokota J, Kunitoh H and Kamataki T. Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behavior and tobacco-related lung cancer risk in male Japanese smokers. *Carcinogenesis* 2004; 25: 2451-2458.
75. Tyndale RF and Sellers EM. Genetic variation in CYP2A6-mediated nicotine metabolism alters smoking behavior. *Ther Drug Monit* 2002; 24: 163-171.
76. Ariyoshi N, Miyamoto M, Umetsu Y, Kunitoh H, Dosaka-Akita H, Sawamura Y, Yokota J, Nemoto N, Sato K and Kamataki T. Genetic polymorphism of CYP2A6 gene and tobacco-induced lung cancer risk in male smokers. *Cancer Epidemiol Biomarkers Prev* 2002; 11: 890-894.
77. Oscarson M, McLellan RA, Gullsten H, Yue QY, Lang MA, Bernal ML, Sinues B, Hirvonen A, Raunio H, Pelkonen O and Ingelman-Sundberg M. Characterisation and PCR-based detection of a CYP2A6 gene deletion found at a high frequency in a Chinese population. *FEBS Lett* 1999; 448: 105-110.
78. Lin YS, Dowling AL, Quigley SD, Farin FM, Zhang J, Lamba J, Schuetz EG and Thummel KE. Co-regulation of CYP3A4 and CYP3A5 and contribution to hepatic and intestinal midazolam metabolism. *Mol Pharmacol* 2002; 62: 162-172.
79. The Allele Frequency Database (ALFRED), available at *http://alfred.med.yale.edu/alfred/sitesWithRsnumber.asp.* 2007.
80. Zhang X, Chen Y, Liu Y, Ren X, Zhang QY, Caggana M and Ding X. Single nucleotide polymorphisms of the human cyp2a13 gene: evidence for a null allele. *Drug Metab Dispos* 2003; 31: 1081-1085.
81. Kirchheiner J and Brockmoller J. Clinical consequences of cytochrome P450 2C9 polymorphisms. *Clin Pharmacol Ther* 2005; 77: 1-16.
82. Maliakal PP, Coville PF and Wanwimolruk S. Decreased hepatic drug metabolising enzyme activity in rats with nitrosamine-induced tumours. *Drug Metabol Drug Interact* 2002; 19: 13-27.
83. Desta Z, Zhao X, Shin JG and Flockhart DA. Clinical significance of the cytochrome P450 2C19 genetic polymorphism. *Clin Pharmacokinet* 2002; 41: 913-958.
84. Leclercq IA, Farrell GC, Field J, Bell DR, Gonzalez FJ and Robertson GR. CYP2E1 and CYP4A as microsomal catalysts of lipid peroxides in murine nonalcoholic steatohepatitis. *J Clin Invest* 2000; 105: 1067-1075.
85. Ronis M, Lindros K and Ingelman-Sundberg M. The cytochrome P4502E subfamily. In: Ioannidies I (ed) Cytochromes P450, pahrmacological and toxicological aspects. *CRC Press, Boca Ration* 1996; pp211-239.
86. Kato S, Shields PG, Caporaso NE, Sugimura H, Trivers GE, Tucker MA, Trump BF, Weston A and Harris CC. Analysis of cytochrome P450 2E1 genetic polymorphisms in relation to human lung cancer. *Cancer Epidemiol Biomarkers Prev* 1994; 3: 515-518.
87. Hu Y, Hakkola J, Oscarson M and Ingelman-Sundberg M. Structural and functional characterization of the 5'-flanking region of the rat and human cytochrome P450 2E1 genes: identification of a polymorphic repeat in the human gene. *Biochem Biophys Res Commun* 1999; 263: 286-293.

88. Amirimani B, Walker AH, Weber BL and Rebbeck TR. RESPONSE: re: modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J Natl Cancer Inst* 1999; 91: 1588-1590.
89. Rebbeck TR, Jaffe JM, Walker AH, Wein AJ and Malkowicz SB. Modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J Natl Cancer Inst* 1998; 90: 1225-1229.
90. Plummer SJ, Conti DV, Paris PL, Curran AP, Casey G and Witte JS. CYP3A4 and CYP3A5 genotypes, haplotypes, and risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2003; 12: 928-932.
91. Miyoshi Y and Noguchi S. Polymorphisms of estrogen synthesizing and metabolizing genes and breast cancer risk in Japanese women. *Biomed Pharmacother* 2003; 57: 471-481.
92. Nakachi K, Hayashi S, Kawajiri K and Imai K. Association of cigarette smoking and CYP1A1 polymorphisms with adenocarcinoma of the lung by grades of differentiation. *Carcinogenesis* 1995; 16: 2209-2213.
93. Nakachi K, Imai K, Hayashi S and Kawajiri K. Polymorphisms of the CYP1A1 and glutathione S-transferase genes associated with susceptibility to lung cancer in relation to cigarette dose in a Japanese population. *Cancer Res* 1993; 53: 2994-2999.
94. Xu X, Kelsey KT, Wiencke JK, Wain JC and Christiani DC. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 1996; 5: 687-692.
95. Yang XR, Wacholder S, Xu Z, Dean M, Clark V, Gold B, Brown LM, Stone BJ, Fraumeni JF, Jr. and Caporaso NE. CYP1A1 and GSTM1 polymorphisms in relation to lung cancer risk in Chinese women. *Cancer Lett* 2004; 214: 197-204.
96. Sugimura H, Hamada GS, Suzuki I, Iwase T, Kiyokawa E, Kino I and Tsugane S. CYP1A1 and CYP2E1 polymorphism and lung cancer, case-control study in Rio de Janeiro, Brazil. *Pharmacogenetics* 1995; 5 Spec No: S145-148.
97. Ozturk O, Isbir T, Yaylim I, Kocaturk CI and Gurses A. GST M1 and CYP1A1 gene polymorphism and daily fruit consumption in Turkish patients with non-small cell lung carcinomas. *In Vivo* 2003; 17: 625-632.
98. Loriot MA, Rebuissou S, Oscarson M, Cenee S, Miyamoto M, Ariyoshi N, Kamataki T, Hemon D, Beaune P and Stucker I. Genetic polymorphisms of cytochrome P450 2A6 in a case-control study on lung cancer in a French population. *Pharmacogenetics* 2001; 11: 39-44.
99. Miyamoto M, Umetsu Y, Dosaka-Akita H, Sawamura Y, Yokota J, Kunitoh H, Nemoto N, Sato K, Ariyoshi N and Kamataki T. CYP2A6 gene deletion reduces susceptibility to lung cancer. *Biochem Biophys Res Commun* 1999; 261: 658-660.
100. Wang H, Tan W, Hao B, Miao X, Zhou G, He F and Lin D. Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK. *Cancer Res* 2003; 63: 8057-8061.
101. Tan W, Chen GF, Xing DY, Song CY, Kadlubar FF and Lin DX. Frequency of CYP2A6 gene deletion and its relation to risk of lung and esophageal cancer in the Chinese population. *Int J Cancer* 2001; 95: 96-101.
102. Kato S, Shields PG, Caporaso NE, Hoover RN, Trump BF, Sugimura H, Weston A and Harris CC. Cytochrome P450IIE1 genetic polymorphisms, racial variation, and lung cancer risk. *Cancer Res* 1992; 52: 6712-6715.
103. Hirvonen A, Husgafvel-Pursiainen K, Anttila S, Karjalainen A and Vainio H. The human CYP2E1 gene and lung cancer: DraI and RsaI restriction fragment length polymorphisms in a Finnish study population. *Carcinogenesis* 1993; 14: 85-88.
104. Hamada GS, Sugimura H, Suzuki I, Nagura K, Kiyokawa E, Iwase T, Tanaka M, Takahashi T, Watanabe S, Kino I and et al. The heme-binding region polymorphism of cytochrome P450IA1 (CypIA1), rather than the RsaI polymorphism of IIE1 (CypIIE1), is associated with lung cancer in Rio de Janeiro. *Cancer Epidemiol Biomarkers Prev* 1995; 4: 63-67.

105. Watanabe J, Yang JP, Eguchi H, Hayashi S, Imai K, Nakachi K and Kawajiri K. An Rsa I polymorphism in the CYP2E1 gene does not affect lung cancer risk in a Japanese population. *Jpn J Cancer Res* 1995; 86: 245-248.
106. Sobti RC, Sharma S, Joshi A, Jindal SK and Janmeja A. Genetic polymorphism of the CYP1A1, CYP2E1, GSTM1 and GSTT1 genes and lung cancer susceptibility in a north indian population. *Mol Cell Biochem* 2004; 266: 1-9.
107. Wu X, Amos CI, Kemp BL, Shi H, Jiang H, Wan Y and Spitz MR. Cytochrome P450 2E1 DraI polymorphisms in lung cancer in minority populations. *Cancer Epidemiol Biomarkers Prev* 1998; 7: 13-18.
108. Persson I, Johansson I, Bergling H, Dahl ML, Seidegard J, Rylander R, Rannug A, Hogberg J and Sundberg MI. Genetic polymorphism of cytochrome P4502E1 in a Swedish population. Relationship to incidence of lung cancer. *FEBS Lett* 1993; 319: 207-211.
109. Wu X, Shi H, Jiang H, Kemp B, Hong WK, Delclos GL and Spitz MR. Associations between cytochrome P4502E1 genotype, mutagen sensitivity, cigarette smoking and susceptibility to lung cancer. *Carcinogenesis* 1997; 18: 967-973.
110. Wang SL, Lee H, Chen KW, Tsai KJ, Chen CY and Lin P. Cytochrome P4502E1 genetic polymorphisms and lung cancer in a Taiwanese population. *Lung Cancer* 1999; 26: 27-34.
111. Wang J, Deng Y, Li L, Kuriki K, Ding J, Pan X, Zhuge X, Jiang J, Luo C, Lin P and Tokudome S. Association of GSTM1, CYP1A1 and CYP2E1 genetic polymorphisms with susceptibility to lung adenocarcinoma: a case-control study in Chinese population. *Cancer Sci* 2003; 94: 448-452.
112. Wu MT, Lee JM, Wu DC, Ho CK, Wang YT, Lee YC, Hsu HK and Kao EL. Genetic polymorphisms of cytochrome P4501A1 and oesophageal squamous-cell carcinoma in Taiwan. *Br J Cancer* 2002; 87: 529-532.
113. Wang AH, Sun CS, Li LS, Huang JY and Chen QS. Relationship of tobacco smoking CYP1A1 GSTM1 gene polymorphism and esophageal cancer in Xi'an. *World J Gastroenterol* 2002; 8: 49-53.
114. Wang AH, Sun CS, Li LS, Huang JY, Chen QS and Xu DZ. Genetic susceptibility and environmental factors of esophageal cancer in Xi'an. *World J Gastroenterol* 2004; 10: 940-944.
115. Abbas A, Delvinquiere K, Lechevrel M, Lebailly P, Gauduchon P, Launoy G and Sichel F. GSTM1, GSTT1, GSTP1 and CYP1A1 genetic polymorphisms and susceptibility to esophageal cancer in a French population: different pattern of squamous cell carcinoma and adenocarcinoma. *World J Gastroenterol* 2004; 10: 3389-3393.
116. Morita S, Yano M, Shiozaki H, Tsujinaka T, Ebisui C, Morimoto T, Kishibuti M, Fujita J, Ogawa A, Taniguchi M, Inoue M, Tamura S, Yamazaki K, Kikkawa N, Mizunoya S and Monden M. CYP1A1, CYP2E1 and GSTM1 polymorphisms are not associated with susceptibility to squamous-cell carcinoma of the esophagus. *Int J Cancer* 1997; 71: 192-195.
117. Lu XM, Zhang YM, Lin RY, Arzi G, Wang X, Zhang YL, Zhang Y, Wang Y and Wen H. Relationship between genetic polymorphisms of metabolizing enzymes CYP2E1, GSTM1 and Kazakh's esophageal squamous cell cancer in Xinjiang, China. *World J Gastroenterol* 2005; 11: 3651-3654.
118. Lin D, Tang Y and Peng Q. [Genetic polymorphisms of cytochrome P450 2E1 and glutathione S-transferase P1 and susceptibility to esophageal cancer]. *Zhonghua Zhong Liu Za Zhi* 1998; 20: 94-97.
119. Tan W, Song N, Wang GQ, Liu Q, Tang HJ, Kadlubar FF and Lin DX. Impact of genetic polymorphisms in cytochrome P450 2E1 and glutathione S-transferases M1, T1, and P1 on susceptibility to esophageal cancer among high-risk individuals in China. *Cancer Epidemiol Biomarkers Prev* 2000; 9: 551-556.
120. Li D, Dandara C and Parker MI. Association of cytochrome P450 2E1 genetic polymorphisms with squamous cell carcinoma of the oesophagus. *Clin Chem Lab Med* 2005; 43: 370-375.
121. Dandara C, Ballo R and Iqbal Parker M. CYP3A5 genotypes and risk of oesophageal cancer in two South African populations. *Cancer Lett* 2005; 225: 275-282.

122. Gonzalez-Perez A, Garcia Rodriguez LA and Lopez-Ridaura R. Effects of non-steroidal anti-inflammatory drugs on cancer sites other than the colon and rectum: a meta-analysis. *BMC Cancer* 2003; 3: 28.
123. Roth MJ, Abnet CC, Johnson LL, Mark SD, Dong ZW, Taylor PR, Dawsey SM and Qiao YL. Polymorphic variation of Cyp1A1 is associated with the risk of gastric cardia cancer: a prospective case-cohort study of cytochrome P-450 1A1 and GST enzymes. *Cancer Causes Control* 2004; 15: 1077-1083.
124. Tsukino H, Kuroda Y, Qiu D, Nakao H, Imai H and Katoh T. Effects of cytochrome P450 (CYP) 2A6 gene deletion and CYP2E1 genotypes on gastric adenocarcinoma. *Int J Cancer* 2002; 100: 425-428.
125. Cai L, Zheng ZL and Zhang ZF. Cytochrome p450 2E1 polymorphisms and the risk of gastric cardia cancer. *World J Gastroenterol* 2005; 11: 1867-1871.
126. Wu MS, Chen CJ, Lin MT, Wang HP, Shun CT, Sheu JC and Lin JT. Genetic polymorphisms of cytochrome p450 2E1, glutathione S-transferase M1 and T1, and susceptibility to gastric carcinoma in Taiwan. *Int J Colorectal Dis* 2002; 17: 338-343.
127. Colombo J, Rossit AR, Caetano A, Borim AA, Wornrath D and Silva AE. GSTT1, GSTM1 and CYP2E1 genetic polymorphisms in gastric cancer and chronic gastritis in a Brazilian population. *World J Gastroenterol* 2004; 10: 1240-1245.
128. Slattery ML, Samowtiz W, Ma K, Murtaugh M, Sweeney C, Levin TR and Neuhausen S. CYP1A1, Cigarette Smoking, and Colon and Rectal Cancer. *Am J Epidemiol* 2004; 160: 842-852.
129. Sachse C, Smith G, Wilkie MJ, Barrett JH, Waxman R, Sullivan F, Forman D, Bishop DT and Wolf CR. A pharmacogenetic study to investigate the role of dietary carcinogens in the etiology of colorectal cancer. *Carcinogenesis* 2002; 23: 1839-1849.
130. Kiss I, Sandor J, Pajkos G, Bogner B, Hegedus G and Ember I. Colorectal cancer risk in relation to genetic polymorphism of cytochrome P450 1A1, 2E1, and glutathione-S-transferase M1 enzymes. *Anticancer Res* 2000; 20: 519-522.
131. Landi S, Gemignani F, Moreno V, Gioia-Patricola L, Chabrier A, Guino E, Navarro M, de Oca J, Capella G and Canzian F. A comprehensive analysis of phase I and phase II metabolism gene polymorphisms and risk of colorectal cancer. *Pharmacogenet Genomics* 2005; 15: 535-546.
132. Moonen H, Engels L, Kleinjans J and Kok T. The CYP1A2-164A-->C polymorphism (CYP1A2*1F) is associated with the risk for colorectal adenomas in humans. *Cancer Lett* 2005; 229: 25-31.
133. Tranah GJ, Chan AT, Giovannucci E, Ma J, Fuchs C and Hunter DJ. Epoxide hydrolase and CYP2C9 polymorphisms, cigarette smoking, and risk of colorectal carcinoma in the Nurses' Health Study and the Physicians' Health Study. *Mol Carcinog* 2005; 44: 21-30.
134. Martinez C, Garcia-Martin E, Ladero JM, Sastre J, Garcia-Gamito F, Diaz-Rubio M and Agundez JA. Association of CYP2C9 genotypes leading to high enzyme activity and colorectal cancer risk. *Carcinogenesis* 2001; 22: 1323-1326.
135. McGreavey LE, Turner F, Smith G, Boylan K, Timothy Bishop D, Forman D, Roland Wolf C and Barrett JH. No evidence that polymorphisms in CYP2C8, CYP2C9, UGT1A6, PPARdelta and PPARgamma act as modifiers of the protective effect of regular NSAID use on the risk of colorectal carcinoma. *Pharmacogenet Genomics* 2005; 15: 713-721.
136. van der Logt EM, Bergevoet SM, Roelofs HM, Te Morsche RH, Dijk YV, Wobbes T, Nagengast FM and Peters WH. Role of epoxide hydrolase, NAD(P)H: quinone oxidoreductase, cytochrome P450 2E1 or alcohol dehydrogenase genotypes in susceptibility to colorectal cancer. *Mutat Res* 2005.
137. Yu MW, Chiu YH, Yang SY, Santella RM, Chern HD, Liaw YF and Chen CJ. Cytochrome P450 1A1 genetic polymorphisms and risk of hepatocellular carcinoma among chronic hepatitis B carriers. *Br J Cancer* 1999; 80: 598-603.
138. Silvestri L, Sonzogni L, De Silvestri A, Gritti C, Foti L, Zavaglia C, Leveri M, Cividini A, Mondelli MU, Civardi E and Silini EM. CYP enzyme polymorphisms and susceptibility to HCV-related chronic liver disease and liver cancer. *Int J Cancer* 2003; 104: 310-317.

139. Mochizuki J, Murakami S, Sanjo A, Takagi I, Akizuki S and Ohnishi A. Genetic polymorphisms of cytochrome P450 in patients with hepatitis C virus-associated hepatocellular carcinoma. *J Gastroenterol Hepatol* 2005; 20: 1191-1197.
140. Agundez JA, Ledesma MC, Benitez J, Ladero JM, Rodriguez-Lescure A, Diaz-Rubio E and Diaz-Rubio M. CYP2D6 genes and risk of liver cancer. *Lancet* 1995; 345: 830-831.
141. Yu MW, Gladek-Yarborough A, Chiamprasert S, Santella RM, Liaw YF and Chen CJ. Cytochrome P450 2E1 and glutathione S-transferase M1 polymorphisms and susceptibility to hepatocellular carcinoma. *Gastroenterology* 1995; 109: 1266-1273.
142. Munaka M, Kohshi K, Kawamoto T, Takasawa S, Nagata N, Itoh H, Oda S and Katoh T. Genetic polymorphisms of tobacco- and alcohol-related metabolizing enzymes and the risk of hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2003; 129: 355-360.
143. Lee HS, Yoon JH, Kamimura S, Iwata K, Watanabe H and Kim CY. Lack of association of cytochrome P450 2E1 genetic polymorphisms with the risk of human hepatocellular carcinoma. *Int J Cancer* 1997; 71: 737-740.
144. Blum HE. Hepatocellular carcinoma: susceptibility markers. *IARC Sci Publ* 2001; 154: 241-244.
145. Haiman CA, Stram DO, Pike MC, Kolonel LN, Burtt NP, Altshuler D, Hirschhorn J and Henderson BE. A comprehensive haplotype analysis of CYP19 and breast cancer risk: the Multiethnic Cohort. *Hum Mol Genet* 2003; 12: 2679-2692.
146. Hefler LA, Tempfer CB, Grimm C, Lebrecht A, Ulbrich E, Heinze G, Leodolter S, Schneeberger C, Mueller MW, Muendlein A and Koelbl H. Estrogen-metabolizing gene polymorphisms in the assessment of breast carcinoma risk and fibroadenoma risk in Caucasian women. *Cancer* 2004; 101: 264-269.
147. Taioli E, Trachman J, Chen X, Toniolo P and Garte SJ. A CYP1A1 restriction fragment length polymorphism is associated with breast cancer in African-American women. *Cancer Res* 1995; 55: 3757-3758.
148. Huang CS, Shen CY, Chang KJ, Hsu SM and Chern HD. Cytochrome P4501A1 polymorphism as a susceptibility factor for breast cancer in postmenopausal Chinese women in Taiwan. *Br J Cancer* 1999; 80: 1838-1843.
149. Ishibe N, Hankinson SE, Colditz GA, Spiegelman D, Willett WC, Speizer FE, Kelsey KT and Hunter DJ. Cigarette smoking, cytochrome P450 1A1 polymorphisms, and breast cancer risk in the Nurses' Health Study. *Cancer Res* 1998; 58: 667-671.
150. da Fonte de Amorim L, Rossini A, Mendonca G, Lotsch P, de Almeida Simao T, de Moura Gallo C and Pinto L. CYP1A1, GSTM1, and GSTT1 polymorphisms and breast cancer risk in Brazilian women. *Cancer Lett* 2002; 181: 179-186.
151. Ambrosone CB, Freudenheim JL, Graham S, Marshall JR, Vena JE, Brasure JR, Laughlin R, Nemoto T, Michalek AM, Harrington A and et al. Cytochrome P4501A1 and glutathione S-transferase (M1) genetic polymorphisms and postmenopausal breast cancer risk. *Cancer Res* 1995; 55: 3483-3485.
152. Miyoshi Y, Takahashi Y, Egawa C and Noguchi S. Breast cancer risk associated with CYP1A1 genetic polymorphisms in Japanese women. *Breast J* 2002; 8: 209-215.
153. Zheng W, Xie DW, Jin F, Cheng JR, Dai Q, Wen WQ, Shu XO and Gao YT. Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2000; 9: 147-150.
154. Kocabas NA, Sardas S, Cholerton S, Daly AK and Karakaya AE. Cytochrome P450 CYP1B1 and catechol O-methyltransferase (COMT) genetic polymorphisms and breast cancer susceptibility in a Turkish population. *Arch Toxicol* 2002; 76: 643-649.
155. Bergman-Jungestrom M, Gentile M, Lundin AC and Wingren S. Association between CYP17 gene polymorphism and risk of breast cancer in young women. *Int J Cancer* 1999; 84: 350-353.
156. Spurdle AB, Hopper JL, Dite GS, Chen X, Cui J, McCredie MR, Giles GG, Southey MC, Venter DJ, Easton DF and Chenevix-Trench G. CYP17 promoter polymorphism and breast cancer in Australian women under age forty years. *J Natl Cancer Inst* 2000; 92: 1674-1681.

157. Feigelson HS, Coetzee GA, Kolonel LN, Ross RK and Henderson BE. A polymorphism in the CYP17 gene increases the risk of breast cancer. *Cancer Res* 1997; 57: 1063-1065.
158. Miyoshi Y, Iwao K, Ikeda N, Egawa C and Noguchi S. Genetic polymorphism in CYP17 and breast cancer risk in Japanese women. *Eur J Cancer* 2000; 36: 2375-2379.
159. Kristensen VN, Andersen TI, Lindblom A, Erikstein B, Magnus P and Borresen-Dale AL. A rare CYP19 (aromatase) variant may increase the risk of breast cancer. *Pharmacogenetics* 1998; 8: 43-48.
160. Haiman CA, Hankinson SE, Spiegelman D, De Vivo I, Colditz GA, Willett WC, Speizer FE and Hunter DJ. A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. *Int J Cancer* 2000; 87: 204-210.
161. Baxter SW, Choong DY, Eccles DM and Campbell IG. Polymorphic variation in CYP19 and the risk of breast cancer. *Carcinogenesis* 2001; 22: 347-349.
162. Lee KM, Abel J, Ko Y, Harth V, Park WY, Seo JS, Yoo KY, Choi JY, Shin A, Ahn SH, Noh DY, Hirvonen A and Kang D. Genetic polymorphisms of cytochrome P450 19 and 1B1, alcohol use, and breast cancer risk in Korean women. *Br J Cancer* 2003; 88: 675-678.
163. Kristensen VN, Harada N, Yoshimura N, Haraldsen E, Lonning PE, Erikstein B, Karesen R, Kristensen T and Borresen-Dale AL. Genetic variants of CYP19 (aromatase) and breast cancer risk. *Oncogene* 2000; 19: 1329-1333.
164. Tanaka Y, Sasaki M, Kaneuchi M, Shiina H, Igawa M and Dahiya R. Polymorphisms of the CYP1B1 gene have higher risk for prostate cancer. *Biochem Biophys Res Commun* 2002; 296: 820-826.
165. Chang BL, Zheng SL, Isaacs SD, Turner A, Hawkins GA, Wiley KE, Bleecker ER, Walsh PC, Meyers DA, Isaacs WB and Xu J. Polymorphisms in the CYP1B1 gene are associated with increased risk of prostate cancer. *Br J Cancer* 2003; 89: 1524-1529.
166. Jakobsson J, Karypidis H, Johansson JE, Roh HK, Rane A and Ekstrom L. A functional C-G polymorphism in the CYP7B1 promoter region and its different distribution in Orientals and Caucasians. *Pharmacogenomics J* 2004.
167. Wadelius M, Andersson AO, Johansson JE, Wadelius C and Rane E. Prostate cancer associated with CYP17 genotype. *Pharmacogenetics* 1999; 9: 635-639.
168. Gsur A, Feik E and Madersbacher S. Genetic polymorphisms and prostate cancer risk. *World J Urol* 2004; 21: 414-423.
169. Suzuki K, Nakazato H, Matsui H, Koike H, Okugi H, Kashiwagi B, Nishii M, Ohtake N, Nakata S, Ito K and Yamanaka H. Genetic polymorphisms of estrogen receptor alpha, CYP19, catechol-O-methyltransferase are associated with familial prostate carcinoma risk in a Japanese population. *Cancer* 2003; 98: 1411-1416.
170. Suzuki K, Matsui H, Nakazato H, Koike H, Okugi H, Hasumi M, Ohtake N, Nakata S, Takei T, Hatori M, Ito K and Yamanaka H. Association of the genetic polymorphism in cytochrome P450 (CYP) 1A1 with risk of familial prostate cancer in a Japanese population: a case-control study. *Cancer Lett* 2003; 195: 177-183.
171. Murata M, Watanabe M, Yamanaka M, Kubota Y, Ito H, Nagao M, Katoh T, Kamataki T, Kawamura J, Yatani R and Shiraishi T. Genetic polymorphisms in cytochrome P450 (CYP) 1A1, CYP1A2, CYP2E1, glutathione S-transferase (GST) M1 and GSTT1 and susceptibility to prostate cancer in the Japanese population. *Cancer Lett* 2001; 165: 171-177.
172. Sasaki M, Tanaka Y, Kaneuchi M, Sakuragi N and Dahiya R. CYP1B1 gene polymorphisms have higher risk for endometrial cancer, and positive correlations with estrogen receptor alpha and estrogen receptor beta expressions. *Cancer Res* 2003; 63: 3913-3918.
173. Zeigler-Johnson C, Friebel T, Walker AH, Wang Y, Spangler E, Panossian S, Patacsil M, Aplenc R, Wein AJ, Malkowicz SB and Rebbeck TR. CYP3A4, CYP3A5, and CYP3A43 genotypes and haplotypes in the etiology and severity of prostate cancer. *Cancer Res* 2004; 64: 8461-8467.
174. Tayeb MT, Clark C, Sharp L, Haites NE, Rooney PH, Murray GI, Payne SN and McLeod HL. CYP3A4 promoter variant is associated with prostate cancer risk in men with benign prostate hyperplasia. *Oncol Rep* 2002; 9: 653-655.

175. Paris PL, Kupelian PA, Hall JM, Williams TL, Levin H, Klein EA, Casey G and Witte JS. Association between a CYP3A4 genetic variant and clinical presentation in African-American prostate cancer patients. *Cancer Epidemiol Biomarkers Prev* 1999; 8: 901-905.
176. Loukola A, Chadha M, Penn SG, Rank D, Conti DV, Thompson D, Cicek M, Love B, Bivolarevic V, Yang Q, Jiang Y, Hanzel DK, Dains K, Paris PL, Casey G and Witte JS. Comprehensive evaluation of the association between prostate cancer and genotypes/haplotypes in CYP17A1, CYP3A4, and SRD5A2. *Eur J Hum Genet* 2004; 12: 321-332.
177. Zhenhua L, Tsuchiya N, Narita S, Inoue T, Horikawa Y, Kakinuma H, Kato T, Ogawa O and Habuchi T. CYP3A5 gene polymorphism and risk of prostate cancer in a Japanese population. *Cancer Lett* 2005; 225: 237-243.
178. Stone A, Ratnasinghe LD, Emerson GL, Modali R, Lehman T, Runnells G, Carroll A, Carter W, Barnhart S, Rasheed AA, Greene G, Johnson DE, Ambrosone CB, Kadlubar FF and Lang NP. CYP3A43 Pro(340)Ala polymorphism and prostate cancer risk in African Americans and Caucasians. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 1257-1261.
179. Lunn RM, Bell DA, Mohler JL and Taylor JA. Prostate cancer risk and polymorphism in 17 hydroxylase (CYP17) and steroid reductase (SRD5A2). *Carcinogenesis* 1999; 20: 1727-1731.
180. Gsur A, Bernhofer G, Hinteregger S, Haidinger G, Schatzl G, Madersbacher S, Marberger M, Vutuc C and Micksche M. A polymorphism in the CYP17 gene is associated with prostate cancer risk. *Int J Cancer* 2000; 87: 434-437.
181. Yamada Y, Watanabe M, Murata M, Yamanaka M, Kubota Y, Ito H, Katoh T, Kawamura J, Yatani R and Shiraishi T. Impact of genetic polymorphisms of 17-hydroxylase cytochrome P-450 (CYP17) and steroid 5alpha-reductase type II (SRD5A2) genes on prostate-cancer risk among the Japanese population. *Int J Cancer* 2001; 92: 683-686.
182. Kittles RA, Panguluri RK, Chen W, Massac A, Ahaghotu C, Jackson A, Ukoli F, Adams-Campbell L, Isaacs W and Dunston GM. Cyp17 promoter variant associated with prostate cancer aggressiveness in African Americans. *Cancer Epidemiol Biomarkers Prev* 2001; 10: 943-947.
183. Habuchi T, Liqing Z, Suzuki T, Sasaki R, Tsuchiya N, Tachiki H, Shimoda N, Satoh S, Sato K, Kakehi Y, Kamoto T, Ogawa O and Kato T. Increased risk of prostate cancer and benign prostatic hyperplasia associated with a CYP17 gene polymorphism with a gene dosage effect. *Cancer Res* 2000; 60: 5710-5713.
184. Haiman CA, Stampfer MJ, Giovannucci E, Ma J, Decalo NE, Kantoff PW and Hunter DJ. The relationship between a polymorphism in CYP17 with plasma hormone levels and prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2001; 10: 743-748.
185. Ntais C, Polycarpou A and Ioannidis JP. Association of the CYP17 gene polymorphism with the risk of prostate cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2003; 12: 120-126.
186. Latil AG, Azzouzi R, Cancel GS, Guillaume EC, Cochan-Priollet B, Berthon PL and Cussenot O. Prostate carcinoma risk and allelic variants of genes involved in androgen biosynthesis and metabolism pathways. *Cancer* 2001; 92: 1130-1137.
187. Ko Y, Abel J, Harth V, Brode P, Antony C, Donat S, Fischer HP, Ortiz-Pallardo ME, Thier R, Sachinidis A, Vetter H, Bolt HM, Herberhold C and Bruning T. Association of CYP1B1 codon 432 mutant allele in head and neck squamous cell cancer is reflected by somatic mutations of p53 in tumor tissue. *Cancer Res* 2001; 61: 4398-4404.
188. Neuhaus T, Ko YD, Lorenzen K, Fronhoffs S, Harth V, Brode P, Vetter H, Bolt HM, Pesch B and Bruning T. Association of cytochrome P450 2E1 polymorphisms and head and neck squamous cell cancer. *Toxicol Lett* 2004; 151: 273-282.
189. Sasaki M, Tanaka Y, Okino ST, Nomoto M, Yonezawa S, Nakagawa M, Fujimoto S, Sakuragi N and Dahiya R. Polymorphisms of the CYP1B1 gene as risk factors for human renal cell cancer. *Clin Cancer Res* 2004; 10: 2015-2019.

190. Garner EI, Stokes EE, Berkowitz RS, Mok SC and Cramer DW. Polymorphisms of the estrogen-metabolizing genes CYP17 and catechol-O-methyltransferase and risk of epithelial ovarian cancer. *Cancer Res* 2002; 62: 3058-3062.
191. Lear JT, Smith AG, Bowers B, Heagearty AH, Jones PW, Gilford J, Alldersea J, Strange RC and Fryer AA. Truncal tumor site is associated with high risk of multiple basal cell carcinoma and is influenced by glutathione S-transferase, GSTT1, and cytochrome P450, CYP1A1 genotypes, and their interaction. *J Invest Dermatol* 1997; 108: 519-522.
192. Strange RC, Ellison T, Ichii-Jones F, Bath J, Hoban P, Lear JT, Smith AG, Hutchinson PE, Osborne J, Bowers B, Jones PW and Fryer AA. Cytochrome P450 CYP2D6 genotypes: association with hair colour, Breslow thickness and melanocyte stimulating hormone receptor alleles in patients with malignant melanoma. *Pharmacogenetics* 1999; 9: 269-276.
193. Cardon LR and Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001; 2: 91-99.
194. Ioannidis JP. Genetic associations: false or true? *Trends Mol Med* 2003; 9: 135-138.
195. Ioannidis JP, Ntzani EE and Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004; 36: 1312-1318.
196. Hirschhorn JN, Lohmueller K, Byrne E and Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; 4: 45-61.
197. Dekker MC and van Duijn CM. Prospects of genetic epidemiology in the 21st century. *Eur J Epidemiol* 2003; 18: 607-616.
198. Ioannidis JP, Bernstein J, Boffetta P, Danesh J, Dolan S, Hartge P, Hunter D, Inskip P, Jarvelin MR, Little J, Maraganore DM, Bishop JA, O'Brien TR, Petersen G, Riboli E, Seminara D, Taioli E, Uitterlinden AG, Vineis P, Winn DM, Salanti G, Higgins JP and Khoury MJ. A network of investigator networks in human genome epidemiology. *Am J Epidemiol* 2005; 162: 302-304.
199. Finta C and Zaphiropoulos PG. Intergenic mRNA molecules resulting from trans-splicing. *J Biol Chem* 2002; 277: 5882-5890.
200. Domanski TL, Finta C, Halpert JR and Zaphiropoulos PG. cDNA cloning and initial characterization of CYP3A43, a novel human cytochrome P450. *Mol Pharmacol* 2001; 59: 386-392.
201. Waxman DJ, Attisano C, Guengerich FP and Lapenson DP. Human liver microsomal steroid metabolism: identification of the major microsomal steroid hormone 6 beta-hydroxylase cytochrome P-450 enzyme. *Arch Biochem Biophys* 1988; 263: 424-436.
202. Waxman DJ, Lapenson DP, Aoyama T, Gelboin HV, Gonzalez FJ and Korzekwa K. Steroid hormone hydroxylase specificities of eleven cDNA-expressed human cytochrome P450s. *Arch Biochem Biophys* 1991; 290: 160-166.
203. Lee AJ, Conney AH and Zhu BT. Human cytochrome P450 3A7 has a distinct high catalytic activity for the 16alpha-hydroxylation of estrone but not 17beta-estradiol. *Cancer Res* 2003; 63: 6532-6536.
204. Miller KK, Cai J, Ripp SL, Pierce WM, Jr., Rushmore TH and Prough RA. Stereo- and regioselectivity account for the diversity of dehydroepiandrosterone (DHEA) metabolites produced by liver microsomal cytochromes P450. *Drug Metab Dispos* 2004; 32: 305-313.
205. Kittles RA, Chen W, Panguluri RK, Ahaghotu C, Jackson A, Adebamowo CA, Griffin R, Williams T, Ukoli F, Adams-Campbell L, Kwagyan J, Isaacs W, Freeman V and Dunston GM. CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? *Hum Genet* 2002; 110: 553-560.
206. Westlind A, Lofberg L, Tindberg N, Andersson TB and Ingelman-Sundberg M. Interindividual differences in hepatic expression of CYP3A4: relationship to genetic polymorphism in the 5'-upstream regulatory region. *Biochem Biophys Res Commun* 1999; 259: 201-205.
207. Amirimani B, Ning B, Deitz AC, Weber BL, Kadlubar FF and Rebbeck TR. Increased transcriptional activity of the CYP3A4*1B promoter variant. *Environ Mol Mutagen* 2003; 42: 299-305.

208. Ando Y, Tateishi T, Sekido Y, Yamamoto T, Satoh T, Hasegawa Y, Kobayashi S, Katsumata Y, Shimokata K and Saito H. Re: Modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J Natl Cancer Inst* 1999; 91: 1587-1590.
209. Spurdle AB, Goodwin B, Hodgson E, Hopper JL, Chen X, Purdie DM, McCredie MR, Giles GG, Chenevix-Trench G and Liddle C. The CYP3A4*1B polymorphism has no functional significance and is not associated with risk of breast or ovarian cancer. *Pharmacogenetics* 2002; 12: 355-366.
210. Hamzeiy H, Bombail V, Plant N, Gibson G and Goldfarb P. Transcriptional regulation of cytochrome P4503A4 gene expression: effects of inherited mutations in the 5'-flanking region. *Xenobiotica* 2003; 33: 1085-1095.
211. Floyd MD, Gervasini G, Masica AL, Mayo G, George AL, Jr., Bhat K, Kim RB and Wilkinson GR. Genotype-phenotype associations for common CYP3A4 and CYP3A5 variants in the basal and induced metabolism of midazolam in European- and African-American men and women. *Pharmacogenetics* 2003; 13: 595-606.
212. Rebbeck TR. More about: modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J Natl Cancer Inst* 2000; 92: 76.
213. Schirmer M, Toliat MR, Haberl M, Suk A, Kamdem LK, Klein K, Brockmoller J, Nurnberg P, Zanger UM and Wojnowski L. Genetic signature consistent with selection against the CYP3A4*1B allele in non-African populations. *Pharmacogenet Genomics* 2006; 16: 59-71.
214. Diagnostic Systems Laboratories (DSL), Webster, Texas, available at *http://www.dslabs.com.* 2007.
215. Sodergard R, Backstrom T, Shanbhag V and Carstensen H. Calculation of free and bound fractions of testosterone and estradiol-17 beta to human plasma proteins at body temperature. *J Steroid Biochem* 1982; 16: 801-810.
216. van den Beld AW, de Jong FH, Grobbee DE, Pols HA and Lamberts SW. Measures of bioavailable serum testosterone and estradiol and their relationships with muscle strength, bone density, and body composition in elderly men. *J Clin Endocrinol Metab* 2000; 85: 3276-3282.
217. Blauer KL, Poth M, Rogers WM and Bernton EW. Dehydroepiandrosterone antagonizes the suppressive effects of dexamethasone on lymphocyte proliferation. *Endocrinology* 1991; 129: 3174-3179.
218. May M, Holmes E, Rogers W and Poth M. Protection from glucocorticoid induced thymic involution by dehydroepiandrosterone. *Life Sci* 1990; 46: 1627-1631.
219. Shafagoj Y, Opoku J, Qureshi D, Regelson W and Kalimi M. Dehydroepiandrosterone prevents dexamethasone-induced hypertension in rats. *Am J Physiol* 1992; 263: E210-213.
220. Horton NJ and Laird NM. Maximum likelihood analysis of generalized linear models with missing covariates. *Stat Methods Med Res* 1999; 8: 37-50.
221. Smit P, van Schaik RH, van der Werf M, van den Beld AW, Koper JW, Lindemans J, Pols HA, Brinkmann AO, de Jong FH and Lamberts SW. A common polymorphism in the CYP3A7 gene is associated with a nearly 50% reduction in serum dehydroepiandrosterone sulfate levels. *J Clin Endocrinol Metab* 2005; 90: 5313-5316.
222. Legrain S and Girard L. Pharmacology and therapeutic effects of dehydroepiandrosterone in older subjects. *Drugs Aging* 2003; 20: 949-967.
223. de Ronde W, Hofman A, Pols HA and de Jong FH. A direct approach to the estimation of the origin of oestrogens and androgens in elderly men by comparison with hormone levels in postmenopausal women. *Eur J Endocrinol* 2005; 152: 261-268.
224. Bangsi D, Zhou J, Sun Y, Patel NP, Darga LL, Heilbrun LK, Powell IJ, Severson RK and Everson RB. Impact of a genetic variant in CYP3A4 on risk and clinical presentation of prostate cancer among white and African-American men. *Urol Oncol* 2006; 24: 21-27.
225. Tayeb MT, Clark C, Haites NE, Sharp L, Murray GI and McLeod HL. CYP3A4 and VDR gene polymorphisms and the risk of prostate cancer in men with benign prostate hyperplasia. *Br J Cancer* 2003; 88: 928-932.
226. Lamberts SW, van den Beld AW and van der Lely AJ. The endocrinology of aging. *Science* 1997; 278: 419-424.

227. Thomas HV, Reeves GK and Key TJ. Endogenous estrogen and postmenopausal breast cancer: a quantitative review. *Cancer Causes Control* 1997; 8: 922-928.
228. Roodi N, Bailey LR, Kao WY, Verrier CS, Yee CJ, Dupont WD and Parl FF. Estrogen receptor gene analysis in estrogen receptor-positive and receptor-negative primary breast cancer. *J Natl Cancer Inst* 1995; 87: 446-451.
229. Zuppan P, Hall JM, Lee MK, Ponglikitmongkol M and King MC. Possible linkage of the estrogen receptor gene to breast cancer in a family with late-onset disease. *Am J Hum Genet* 1991; 48: 1065-1068.
230. Andersen TI, Heimdal KR, Skrede M, Tveit K, Berg K and Borresen AL. Oestrogen receptor (ESR) polymorphisms and breast cancer susceptibility. *Hum Genet* 1994; 94: 665-670.
231. Iwase H, Greenman JM, Barnes DM, Hodgson S, Bobrow L and Mathew CG. Sequence variants of the estrogen receptor (ER) gene found in breast cancer patients with ER negative and progesterone receptor positive tumors. *Cancer Lett* 1996; 108: 179-184.
232. Wang M, Dotzlaw H, Fuqua SA and Murphy LC. A point mutation in the human estrogen receptor gene is associated with the expression of an abnormal estrogen receptor mRNA containing a 69 novel nucleotide insertion. *Breast Cancer Res Treat* 1997; 44: 145-151.
233. Cai Q, Gao YT, Wen W, Shu XO, Jin F, Smith JR and Zheng W. Association of breast cancer risk with a GT dinucleotide repeat polymorphism upstream of the estrogen receptor-alpha gene. *Cancer Res* 2003; 63: 5727-5730.
234. Cai Q, Shu XO, Jin F, Dai Q, Wen W, Cheng JR, Gao YT and Zheng W. Genetic polymorphisms in the estrogen receptor alpha gene and risk of breast cancer: results from the Shanghai Breast Cancer Study. *Cancer Epidemiol Biomarkers Prev* 2003; 12: 853-859.
235. Shou M, Korzekwa KR, Brooks EN, Krausz KW, Gonzalez FJ and Gelboin HV. Role of human hepatic cytochrome P450 1A2 and 3A4 in the metabolic activation of estrone. *Carcinogenesis* 1997; 18: 207-214.
236. Yamazaki H, Shaw PM, Guengerich FP and Shimada T. Roles of cytochromes P450 1A2 and 3A4 in the oxidation of estradiol and estrone in human liver microsomes. *Chem Res Toxicol* 1998; 11: 659-665.
237. Kadlubar FF, Berkowitz GS, Delongchamp RR, Wang C, Green BL, Tang G, Lamba J, Schuetz E and Wolff MS. The CYP3A4*1B variant is related to the onset of puberty, a known risk factor for the development of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2003; 12: 327-331.
238. Mugford CA and Kedderis GL. Sex-dependent metabolism of xenobiotics. *Drug Metab Rev* 1998; 30: 441-498.
239. Smith DA. Species differences in metabolism and pharmacokinetics: are we close to an understanding? *Drug Metab Rev* 1991; 23: 355-373.
240. Tanaka E. Gender-related differences in pharmacokinetics and their clinical significance. *J Clin Pharm Ther* 1999; 24: 339-346.
241. Wolbold R, Klein K, Burk O, Nussler AK, Neuhaus P, Eichelbaum M, Schwab M and Zanger UM. Sex is a major determinant of CYP3A4 expression in human liver. *Hepatology* 2003; 38: 978-988.
242. Gandhi M, Aweeka F, Greenblatt RM and Blaschke TF. Sex differences in pharmacokinetics and pharmacodynamics. *Annu Rev Pharmacol Toxicol* 2004; 44: 499-523.
243. Hunt CM, Westerkam WR and Stave GM. Effect of age and gender on the activity of human hepatic CYP3A. *Biochem Pharmacol* 1992; 44: 275-283.
244. Gillberg M, Skaanild MT and Friis C. Regulation of gender-dependent CYP2A expression in pigs: involvement of androgens and CAR. *Basic Clin Pharmacol Toxicol* 2006; 98: 480-487.
245. Forman BM, Tzameli I, Choi HS, Chen J, Simha D, Seol W, Evans RM and Moore DD. Androstane metabolites bind to and deactivate the nuclear receptor CAR-beta. *Nature* 1998; 395: 612-615.
246. Kawamoto T, Kakizaki S, Yoshinari K and Negishi M. Estrogen activation of the nuclear orphan receptor CAR (constitutive active receptor) in induction of the mouse Cyp2b10 gene. *Mol Endocrinol* 2000; 14: 1897-1905.

247. Levi F, Lucchini F, Negri E, Boyle P and La Vecchia C. Cancer mortality in Europe, 1995-1999, and an overview of trends since 1960. *Int J Cancer* 2004; 110: 155-169.
248. Leslie A, Carey FA, Pratt NR and Steele RJ. The colorectal adenoma-carcinoma sequence. *Br J Surg* 2002; 89: 845-860.
249. Gonzalez FJ and Gelboin HV. Role of human cytochromes P450 in the metabolic activation of chemical carcinogens and toxins. *Drug Metab Rev* 1994; 26: 165-183.
250. Ding X and Kaminsky LS. Human extrahepatic cytochromes P450: function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. *Annu Rev Pharmacol Toxicol* 2003; 43: 149-173.
251. Lee CR, Goldstein JA and Pieper JA. Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. *Pharmacogenetics* 2002; 12: 251-263.
252. Nebert DW and Russell DW. Clinical importance of the cytochromes P450. *Lancet* 2002; 360: 1155-1162.
253. Yokose T, Doy M, Taniguchi T, Shimada T, Kakiki M, Horie T, Matsuzaki Y and Mukai K. Immunohistochemical study of cytochrome P450 2C and 3A in human non-neoplastic and neoplastic tissues. *Virchows Arch* 1999; 434: 401-411.
254. Goldstein JA and de Morais SM. Biochemistry and molecular biology of the human CYP2C subfamily. *Pharmacogenetics* 1994; 4: 285-299.
255. Badawi AF, Stern SJ, Lang NP and Kadlubar FF. Cytochrome P-450 and acetyltransferase expression as biomarkers of carcinogen-DNA adduct levels and human cancer susceptibility. *Prog Clin Biol Res* 1996; 395: 109-140.
256. Daikh BE, Lasker JM, Raucy JL and Koop DR. Regio- and stereoselective epoxidation of arachidonic acid by human cytochromes P450 2C8 and 2C9. *J Pharmacol Exp Ther* 1994; 271: 1427-1433.
257. Rifkind AB, Lee C, Chang TK and Waxman DJ. Arachidonic acid metabolism by human cytochrome P450s 2C8, 2C9, 2E1, and 1A2: regioselective oxygenation and evidence for a role for CYP2C enzymes in arachidonic acid epoxygenation in human liver microsomes. *Arch Biochem Biophys* 1995; 320: 380-389.
258. Cao Y, Pearman AT, Zimmerman GA, McIntyre TM and Prescott SM. Intracellular unesterified arachidonic acid signals apoptosis. *Proc Natl Acad Sci U S A* 2000; 97: 11280-11285.
259. Michaelis UR, Fisslthaler B, Medhora M, Harder D, Fleming I and Busse R. Cytochrome P450 2C9-derived epoxyeicosatrienoic acids induce angiogenesis via cross-talk with the epidermal growth factor receptor (EGFR). *Faseb J* 2003; 17: 770-772.
260. Vondracek J, Stika JV, Soucek K, Minksova K, Blaha L, Hofmanova J and Kozubik A. Inhibitors of arachidonic acid metabolism potentiate tumour necrosis factor-alpha-induced apoptosis in HL-60 cells. *Eur J Pharmacol* 2001; 424: 1-11.
261. Watson AJ. Chemopreventive effects of NSAIDs against colorectal cancer: regulation of apoptosis and mitosis by COX-1 and COX-2. *Histol Histopathol* 1998; 13: 591-597.
262. Ricchi P, Zarrilli R, Di Palma A and Acquaviva AM. Nonsteroidal anti-inflammatory drugs in colorectal cancer: from prevention to therapy. *Br J Cancer* 2003; 88: 803-807.
263. Thun MJ, Namboodiri MM and Heath CW, Jr. Aspirin use and reduced risk of fatal colon cancer. *N Engl J Med* 1991; 325: 1593-1596.
264. Ingelman - Sundberg M DA, Nebert DW. CYP2C9 allele nomenclature. Vol. 2005 (Human CYP450 Allele Nomenclature Committee, 2005).
265. Xie HG, Prasad HC, Kim RB and Stein CM. CYP2C9 allelic variants: ethnic distribution and functional significance. *Adv Drug Deliv Rev* 2002; 54: 1257-1270.
266. Clayton D and M H. Time-varying explanatory variables. in *Statistical models in epidemiology* 307-318 (Oxford University Press, Oxford, England, 1993).
267. AppliedBiosystems, Foster City, California, USA, available at *http://www.appliedbiosystems.com/northernamerica*. 2007.

268. Hosmer DW and Lemeshow S. Confidence interval estimation of interaction. *Epidemiology* 1992; 3: 452-456.
269. Andersson T, Alfredsson L, Kallberg H, Zdravkovic S and Ahlbom A. Calculating measures of biological interaction. *Eur J Epidemiol* 2005; 20: 575-579.
270. Peleg, II, Lubin MF, Cotsonis GA, Clark WS and Wilcox CM. Long-term use of nonsteroidal antiinflammatory drugs and other chemopreventors and risk of subsequent colorectal neoplasia. *Dig Dis Sci* 1996; 41: 1319-1326.
271. Rosenberg L, Palmer JR, Zauber AG, Warshauer ME, Stolley PD and Shapiro S. A hypothesis: nonsteroidal anti-inflammatory drugs reduce the incidence of large-bowel cancer. *J Natl Cancer Inst* 1991; 83: 355-358.
272. Smalley W, Ray WA, Daugherty J and Griffin MR. Use of nonsteroidal anti-inflammatory drugs and incidence of colorectal cancer: a population-based study. *Arch Intern Med* 1999; 159: 161-166.
273. Thun MJ, Namboodiri MM, Calle EE, Flanders WD and Heath CW, Jr. Aspirin use and risk of fatal cancer. *Cancer Res* 1993; 53: 1322-1327.
274. Chan AT, Giovannucci EL, Meyerhardt JA, Schernhammer ES, Curhan GC and Fuchs CS. Long-term use of aspirin and nonsteroidal anti-inflammatory drugs and risk of colorectal cancer. *Jama* 2005; 294: 914-923.
275. Cook NR, Lee IM, Gaziano JM, Gordon D, Ridker PM, Manson JE, Hennekens CH and Buring JE. Low-dose aspirin in the primary prevention of cancer: the Women's Health Study: a randomized controlled trial. *Jama* 2005; 294: 47-55.
276. Sorensen HT, Friis S, Norgard B, Mellemkjaer L, Blot WJ, McLaughlin JK, Ekbom A and Baron JA. Risk of cancer in a large cohort of nonaspirin NSAID users: a population-based study. *Br J Cancer* 2003; 88: 1687-1692.
277. Iacopetta B. Are there two sides to colorectal cancer? *Int J Cancer* 2002; 101: 403-408.
278. Bigler J, Whitton J, Lampe JW, Fosdick L, Bostick RM and Potter JD. CYP2C9 and UGT1A6 genotypes modulate the protective effect of aspirin on colon adenoma risk. *Cancer Res* 2001; 61: 3566-3569.
279. Chan AT, Tranah GJ, Giovannucci EL, Hunter DJ and Fuchs CS. A prospective study of genetic polymorphisms in the cytochrome P-450 2C9 enzyme and the risk for distal colorectal adenoma. *Clin Gastroenterol Hepatol* 2004; 2: 704-712.
280. Samowitz WS, Wolff RK, Curtin K, Sweeney C, Ma KN, Andersen K, Levin TR and Slattery ML. Interactions between CYP2C9 and UGT1A6 polymorphisms and nonsteroidal anti-inflammatory drugs in colorectal cancer prevention. *Clin Gastroenterol Hepatol* 2006; 4: 894-901.
281. Lundblad MS, Ohlsson S, Johansson P, Lafolie P and Eliasson E. Accumulation of celecoxib with a 7-fold higher drug exposure in individuals homozygous for CYP2C9*3. *Clin Pharmacol Ther* 2006; 79: 287-288.
282. Steward DJ, Haining RL, Henne KR, Davis G, Rushmore TH, Trager WF and Rettie AE. Genetic association between sensitivity to warfarin and expression of CYP2C9*3. *Pharmacogenetics* 1997; 7: 361-367.
283. Takanashi K, Tainaka H, Kobayashi K, Yasumori T, Hosakawa M and Chiba K. CYP2C9 Ile359 and Leu359 variants: enzyme kinetic study with seven substrates. *Pharmacogenetics* 2000; 10: 95-104.
284. Waddell WR and Loughry RW. Sulindac for polyposis of the colon. *J Surg Oncol* 1983; 24: 83-87.
285. Jacobs EJ, Rodriguez C, Mondul AM, Connell CJ, Henley SJ, Calle EE and Thun MJ. A large cohort study of aspirin and other nonsteroidal anti-inflammatory drugs and prostate cancer incidence. *J Natl Cancer Inst* 2005; 97: 975-980.
286. Sabichi AL and Lippman SM. COX-2 inhibitors and other nonsteroidal anti-inflammatory drugs in genitourinary cancer. *Semin Oncol* 2004; 31: 36-44.
287. Moorman PG, Grubber JM, Millikan RC and Newman B. Association between non-steroidal anti-inflammatory drugs (NSAIDs) and invasive breast cancer and carcinoma in situ of the breast. *Cancer Causes Control* 2003; 14: 915-922.

288. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A and Willett WC. Aspirin use and the risk for colorectal cancer and adenoma in male health professionals. *Ann Intern Med* 1994; 121: 241-246.
289. Chan AT, Giovannucci EL, Schernhammer ES, Colditz GA, Hunter DJ, Willett WC and Fuchs CS. A prospective study of aspirin use and the risk for colorectal adenoma. *Ann Intern Med* 2004; 140: 157-166.
290. Garcia Rodriguez LA and Huerta-Alvarez C. Reduced incidence of colorectal adenoma among long-term users of nonsteroidal antiinflammatory drugs: a pooled analysis of published studies and a new population-based study. *Epidemiology* 2000; 11: 376-381.
291. Sandler RS, Halabi S, Baron JA, Budinger S, Paskett E, Keresztes R, Petrelli N, Pipas JM, Karp DD, Loprinzi CL, Steinbach G and Schilsky R. A randomized trial of aspirin to prevent colorectal adenomas in patients with previous colorectal cancer. *N Engl J Med* 2003; 348: 883-890.
292. Gupta RA and DuBois RN. Translational studies on Cox-2 inhibitors in the prevention and treatment of colon cancer. *Ann N Y Acad Sci* 2000; 910: 196-204; discussion 204-196.
293. Muto T, Bussey HJ and Morson BC. The evolution of cancer of the colon and rectum. *Cancer* 1975; 36: 2251-2270.
294. Giovannucci E and Chan A. NSAIDs: Role in prevention of colorectal cancer. in *Up-to-Date* (2006).
295. Garcia-Rodriguez LA and Huerta-Alvarez C. Reduced risk of colorectal cancer among long-term users of aspirin and nonaspirin nonsteroidal antiinflammatory drugs. *Epidemiology* 2001; 12: 88-93.
296. Giovannucci E, Egan KM, Hunter DJ, Stampfer MJ, Colditz GA, Willett WC and Speizer FE. Aspirin and the risk of colorectal cancer in women. *N Engl J Med* 1995; 333: 609-614.
297. Peleg, II, Maibach HT, Brown SH and Wilcox CM. Aspirin and nonsteroidal anti-inflammatory drug use and the risk of subsequent colorectal cancer. *Arch Intern Med* 1994; 154: 394-399.
298. Baron JA and Sandler RS. Nonsteroidal anti-inflammatory drugs and cancer prevention. *Annu Rev Med* 2000; 51: 511-523.
299. Sturmer T, Buring JE, Lee IM, Kurth T, Gaziano JM and Glynn RJ. Colorectal cancer after start of nonsteroidal anti-inflammatory drug use. *Am J Med* 2006; 119: 494-502.
300. Eberhart CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S and DuBois RN. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. *Gastroenterology* 1994; 107: 1183-1188.
301. Sullivan JL. Aspirin and colorectal cancer. *Ann Intern Med* 1995; 123: 390; author reply 391.
302. Tuynman JB, Peppelenbosch MP and Richel DJ. COX-2 inhibition as a tool to treat and prevent colorectal cancer. *Crit Rev Oncol Hematol* 2004; 52: 81-101.
303. Bertagnolli MM. The potential of non-steroidal anti-inflammatory drugs (NSAIDs) for colorectal cancer prevention. *J Surg Oncol* 2003; 84: 113-119.
304. Lichtenberger LM. Where is the evidence that cyclooxygenase inhibition is the primary cause of nonsteroidal anti-inflammatory drug (NSAID)-induced gastrointestinal injury? Topical injury revisited. *Biochem Pharmacol* 2001; 61: 631-637.
305. Piazuelo E, Jimenez P and Lanas A. COX-2 inhibition in esophagitis, Barrett's esophagus and esophageal cancer. *Curr Pharm Des* 2003; 9: 2267-2280.
306. Menezes RJ, Huber KR, Mahoney MC and Moysich KB. Regular use of aspirin and pancreatic cancer risk. *BMC Public Health* 2002; 2: 18.
307. Davies GL. Cyclooxygenase-2 and chemoprevention of breast cancer. *J Steroid Biochem Mol Biol* 2003; 86: 495-499.
308. Caliper LS, Mountain View, California, USA, available at *http://www.caliperls.com.* 2007.
309. Papafili A, Hill MR, Brull DJ, McAnulty RJ, Marshall RP, Humphries SE and Laurent GJ. Common promoter variant in cyclooxygenase-2 represses gene expression: evidence of role in acute-phase inflammatory response. *Arterioscler Thromb Vasc Biol* 2002; 22: 1631-1636.
310. ABgene, Epsom, UK, available at *http://www.abgene.com.* 2007.
311. Deerac Fluidics, Dublin, Ireland, available at *http://www.deerac.com.* 2007.

312. Harris RE, Chlebowski RT, Jackson RD, Frid DJ, Ascenseo JL, Anderson G, Loar A, Rodabough RJ, White E and McTiernan A. Breast cancer and nonsteroidal anti-inflammatory drugs: prospective results from the Women's Health Initiative. *Cancer Res* 2003; 63: 6096-6101.
313. Harris RE, Beebe-Donk J and Alshafie GA. Reduction in the risk of human breast cancer by selective cyclooxygenase-2 (COX-2) inhibitors. *BMC Cancer* 2006; 6: 27.
314. Terry MB, Gammon MD, Zhang FF, Tawfik H, Teitelbaum SL, Britton JA, Subbaramaiah K, Dannenberg AJ and Neugut AI. Association of frequency and duration of aspirin use and hormone receptor status with breast cancer risk. *Jama* 2004; 291: 2433-2440.
315. The Cancer Genome Anatomy Project, Virtual Northern Tissue gene expression, available at *http://cgap.nci.nih.gov/Tissues/VirtualNorthern*. 2007.
316. Campa D, Zienolddiny S, Maggini V, Skaug V, Haugen A and Canzian F. Association of a common polymorphism in the cyclooxygenase 2 gene with risk of non-small cell lung cancer. *Carcinogenesis* 2004; 25: 229-235.
317. Cox DG, Pontes C, Guino E, Navarro M, Osorio A, Canzian F and Moreno V. Polymorphisms in prostaglandin synthase 2/cyclooxygenase 2 (PTGS2/COX2) and risk of colorectal cancer. *Br J Cancer* 2004; 91: 339-343.
318. Koh WP, Yuan JM, van den Berg D, Lee HP and Yu MC. Interaction between cyclooxygenase-2 gene polymorphism and dietary n-6 polyunsaturated fatty acids on colon cancer risk: the Singapore Chinese Health Study. *Br J Cancer* 2004; 90: 1760-1764.
319. Wiesner GL, Platzer P, Buxbaum S, Lewis S, MacMillen M, Olechnowicz J, Willis J, Chakravarti A, Elston RC and Markowitz SD. Testing for colon neoplasia susceptibility variants at the human COX2 locus. *J Natl Cancer Inst* 2001; 93: 635-639.
320. Ulrich CM, Whitton J, Yu JH, Sibert J, Sparks R, Potter JD and Bigler J. PTGS2 (COX-2) -765G > C promoter variant reduces risk of colorectal adenoma among nonusers of nonsteroidal anti-inflammatory drugs. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 616-619.
321. Zhan J, Liu JP, Zhu ZH, Yao HR and Chen CY. Relationship between COX-2 expression and clinicopathological features of colorectal cancers. *Chin Med J (Engl)* 2004; 117: 1151-1154.
322. Ali IU, Luke BT, Dean M and Greenwald P. Allellic variants in regulatory regions of cyclooxygenase-2: association with advanced colorectal adenoma. *Br J Cancer* 2005; 93: 953-959.
323. Panguluri RC, Long LO, Chen W, Wang S, Coulibaly A, Ukoli F, Jackson A, Weinrich S, Ahaghotu C, Isaacs W and Kittles RA. COX-2 gene promoter haplotypes and prostate cancer risk. *Carcinogenesis* 2004; 25: 961-966.
324. Shahedi K, Lindstrom S, Zheng SL, Wiklund F, Adolfsson J, Sun J, Augustsson-Balter K, Chang BL, Adami HO, Liu W, Gronberg H and Xu J. Genetic variation in the COX-2 gene and the association with prostate cancer risk. *Int J Cancer* 2006; 119: 668-672.
325. Ridker PM, Cushman M, Stampfer MJ, Tracy RP and Hennekens CH. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N Engl J Med* 1997; 336: 973-979.
326. Waskewich C, Blumenthal RD, Li H, Stein R, Goldenberg DM and Burton J. Celecoxib exhibits the greatest potency amongst cyclooxygenase (COX) inhibitors for growth inhibition of COX-2-negative hematopoietic and epithelial cell lines. *Cancer Res* 2002; 62: 2029-2033.
327. Masferrer JL, Leahy KM, Koki AT, Zweifel BS, Settle SL, Woerner BM, Edwards DA, Flickinger AG, Moore RJ and Seibert K. Antiangiogenic and antitumor activities of cyclooxygenase-2 inhibitors. *Cancer Res* 2000; 60: 1306-1311.
328. Slattery ML, Samowitz W, Hoffman M, Ma KN, Levin TR and Neuhausen S. Aspirin, NSAIDs, and colorectal cancer: possible involvement in an insulin-related pathway. *Cancer Epidemiol Biomarkers Prev* 2004; 13: 538-545.
329. Coussens LM and Werb Z. Inflammation and cancer. *Nature* 2002; 420: 860-867.
330. Balkwill F and Mantovani A. Inflammation and cancer: back to Virchow? *Lancet* 2001; 357: 539-545.

331. Hashimoto K, Ikeda Y, Korenaga D, Tanoue K, Hamatake M, Kawasaki K, Yamaoka T, Iwatani Y, Akazawa K and Takenaka K. The impact of preoperative serum C-reactive protein on the prognosis of patients with hepatocellular carcinoma. *Cancer* 2005; 103: 1856-1864.
332. Chung YC and Chang YF. Serum C-reactive protein correlates with survival in colorectal cancer patients but is not an independent prognostic indicator. *Eur J Gastroenterol Hepatol* 2003; 15: 369-373.
333. Nikiteas NI, Tzanakis N, Gazouli M, Rallis G, Daniilidis K, Theodoropoulos G, Kostakis A and Peros G. Serum IL-6, TNFalpha and CRP levels in Greek colorectal cancer patients: prognostic implications. *World J Gastroenterol* 2005; 11: 1639-1643.
334. Danesh J and Pepys MB. C-reactive protein in healthy and in sick populations. *Eur Heart J* 2000; 21: 1564-1565.
335. Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GD, Pepys MB and Gudnason V. C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N Engl J Med* 2004; 350: 1387-1397.
336. Ito Y, Suzuki K, Tamakoshi K, Wakai K, Kojima M, Ozasa K, Watanabe Y, Kawado M, Hashimoto S, Suzuki S, Tokudome S, Toyoshima H, Hayakawa N, Kato K, Watanabe M, Ohta Y, Maruta M and Tamakoshi A. Colorectal cancer and serum C-reactive protein levels: a case-control study nested in the JACC Study. *J Epidemiol* 2005; 15 Suppl 2: S185-189.
337. Il'yasova D, Colbert LH, Harris TB, Newman AB, Bauer DC, Satterfield S and Kritchevsky SB. Circulating levels of inflammatory markers and cancer risk in the health aging and body composition cohort. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 2413-2418.
338. Zhang SM, Buring JE, Lee IM, Cook NR and Ridker PM. C-reactive protein levels are not associated with increased risk for colorectal cancer in women. *Ann Intern Med* 2005; 142: 425-432.
339. Platz EA, De Marzo AM, Erlinger TP, Rifai N, Visvanathan K, Hoffman SC and Helzlsouer KJ. No association between pre-diagnostic plasma C-reactive protein concentration and subsequent prostate cancer. *Prostate* 2004; 59: 393-400.
340. Erlinger TP, Platz EA, Rifai N and Helzlsouer KJ. C-reactive protein and the risk of incident colorectal cancer. *Jama* 2004; 291: 585-590.
341. Rifai N, Buring JE, Lee IM, Manson JE and Ridker PM. Is C-reactive protein specific for vascular disease in women? *Ann Intern Med* 2002; 136: 529-533.
342. Van Der Meer IM, De Maat MP, Hak AE, Kiliaan AJ, Del Sol AI, Van Der Kuip DA, Nijhuis RL, Hofman A and Witteman JC. C-reactive protein predicts progression of atherosclerosis measured at various sites in the arterial tree: the Rotterdam Study. *Stroke* 2002; 33: 2750-2755.
343. Immage® Immunochemistry System, Beckman Coulter, USA. *http://www.beckmancoulter.com*. *2*007.
344. Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, 3rd, Criqui M, Fadl YY, Fortmann SP, Hong Y, Myers GL, Rifai N, Smith SC, Jr., Taubert K, Tracy RP and Vinicor F. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation* 2003; 107: 499-511.
345. Seattle Program for Genomic Applications, available at *http://pga.gs.washington.edu.* 2007.
346. Stephens M, Smith NJ and Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; 68: 978-989.
347. Carlson CS, Aldred SF, Lee PK, Tracy RP, Schwartz SM, Rieder M, Liu K, Williams OD, Iribarren C, Lewis EC, Fornage M, Boerwinkle E, Gross M, Jaquish C, Nickerson DA, Myers RM, Siscovick DS and Reiner AP. Polymorphisms within the C-reactive protein (CRP) promoter region are associated with plasma CRP levels. *Am J Hum Genet* 2005; 77: 64-77.
348. Kardys I, de Maat MP, Uitterlinden AG, Hofman A and Witteman JC. C-reactive protein gene haplotypes and risk of coronary heart disease: the Rotterdam Study. *Eur Heart J* 2006; 27: 1331-1337.

349. Ockene IS, Matthews CE, Rifai N, Ridker PM, Reed G and Stanek E. Variability and classification accuracy of serial high-sensitivity C-reactive protein measurements in healthy adults. *Clin Chem* 2001; 47: 444-450.
350. Harris H. Enzyme and protein diversity in human populations. in *The Principle of Human Biochemical Genetics* 211-242 (American Elsevier Publishing Company, New York, 1970).
351. Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, Sheffield VC, Sunden S, Duyk GM and et al. A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). *Science* 1994; 265: 2049-2054.
352. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J and Weissenbach J. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996; 380: 152-154.
353. Newton-Cheh C and Hirschhorn JN. Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 2005; 573: 54-69.
354. van Duijn CM and Porta M. Good prospects for genetic and molecular epidemiologic studies in the European Journal of Epidemiology. *Eur J Epidemiol* 2003; 18: 285-286.
355. Horwitz RI and Feinstein AR. The problem of "protopathic bias" in case-control studies. *Am J Med* 1980; 68: 255-258.
356. Visser LE, van Schaik RH, van Vliet M, Trienekens PH, De Smet PA, Vulto AG, Hofman A, van Duijn CM and Stricker BH. The risk of bleeding complications in patients with cytochrome P450 CYP2C9*2 or CYP2C9*3 alleles on acenocoumarol or phenprocoumon. *Thromb Haemost* 2004; 92: 61-66.
357. Visser LE, van Vliet M, van Schaik RH, Kasbergen AA, De Smet PA, Vulto AG, Hofman A, van Duijn CM and Stricker BH. The risk of overanticoagulation in patients with cytochrome P450 CYP2C9*2 or CYP2C9*3 alleles on acenocoumarol or phenprocoumon. *Pharmacogenetics* 2004; 14: 27-33.
358. Botstein D and Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 2003; 33 Suppl: 228-237.
359. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ and Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002; 296: 2225-2229.
360. Greenspan RJ. The flexible genome. *Nat Rev Genet* 2001; 2: 383-387.
361. Scriver CR. After the genome--the phenome? *J Inherit Metab Dis* 2004; 27: 305-317.
362. Freimer N and Sabatti C. The human phenome project. *Nat Genet* 2003; 34: 15-21.
363. Jeong H, Tombor B, Albert R, Oltvai ZN and Barabasi AL. The large-scale organization of metabolic networks. *Nature* 2000; 407: 651-654.

6. Summary / Samenvatting

6.1. Summary

Improvement of medical care is one of the causes of aging of the population. As a consequence, the incidence of age-related diseases, like cancer, increases. The importance of studies on cancer is therefore widely recognized. Substantial efforts in cancer research over the past decades have resulted in new insights into the complex systems involved in cancer development, treatment and prognosis. Moreover, individualized approaches are becoming state of the art. Knowledge about genetic variation might significantly contribute to the insights into the complexity of this group of diseases.

The aim of this thesis was to give a deeper understanding of the role genetic variation in genes encoding for specific enzymes could play in cancer development and prognosis. As for the association with cancer multiple candidate genes can be studied, a selection of two potentially important topics was made. These are cytochrome P450 genes and inflammatory related genes.

Chapter 1 starts with a general introduction about the aim of this thesis (**Chapter 1.1.**), The Rotterdam Study as a source population for all analyses (**Chapter 1.2.**), the demographics of cancer in the world in general and in The Rotterdam Study in particular (**Chapter 1.3**) and the most important aspects of genetic variation (**Chapter 1.4.**).

Chapter 2 describes studies that point at the association between genetic variation in cytochrome P450 (CYP) genes and cancer. It starts with a summary of the present knowledge in which we restricted ourselves to genetic polymorphisms as a susceptibility factor, therefore excluding literature on the polymorphic expression of CYPs in (tumor) tissue. (**Chapter 2.1.**) The review reveals a great heterogeneity in respect to ethnicity studied, the composition of control groups, the phenotype of interest, and the variant type and reference type used. At the end of this chapter, the need for standardization of study designs and methods to enable meta-analyses is discussed. The following two chapters describe studies on the association between *CYP3A* genotypes, which include *CYP3A4*, *CYP3A5* and *CYP3A7*, steroid hormone levels and hormone related cancers. **Chapter 2.2.** studies associations in males and consequently focuses on prostate cancer. **Chapter 2.3.** studies associations in females and consequently focuses on breast cancer. In both studies, associations are found between *CYP3A* genotypes and hormone levels, although there is gender-specificity since the associations are most pronounced in males. Androgen dependent inhibition of CYP3A expression was proposed to explain these gender differences. In females, some of the genotypes were associated with the duration of reproductive live, calculated from the age at onset of menarche and menopause. However, no associations were observed for the hormone related cancers. Finally, **Chapter 2.4.** describes the importance of interaction between NSAID use and *CYP2C9* slow metabolizing variants on the prevention of colorectal cancer. Although previous studies showed that duration of NSAID use is more important than dose, others found significant drug-gene interactions *in vivo*. Since these final studies focused on specific parts of the association, the study presented here aimed at gaining more insight into the complete causal pathway. Both long-term NSAID use and *CYP2C9* gene variants were associated with a decreased risk of colorectal

cancer, especially in proximal parts. The combination of both determinants was associated with a further risk reduction but without synergy. This might indicate independent pathophysiological mechanisms and underlines the importance of duration over dose of NSAIDs.

Chapter 3 points at studies that describe associations between inflammation and cancer. The association between long-term NSAID use and colorectal cancer prevention is commonly known. Nevertheless, most previous studies included self-reported medication use. Therefore, in **Chapter 3.1.** the association between NSAID use on a day-to-day basis and colorectal cancer is studied. The study additionally reports information on anatomical subtypes of colorectal cancer and explores the effect of discontinuation of the drug. Results add to the large body of evidence that long-term NSAID use protects against colorectal cancer, especially in proximal parts. The association is further explored in the study presented in **Chapter 3.2.** that investigates the influence of COX-2 expression on the beneficial effects of NSAIDs. A *COX-2* gene polymorphism that leads to 30% lower expression levels of its encoding enzyme was included in the analyses. Although the beneficial effects of NSAIDs are mostly explained by the inhibition of COX enzymes, this study shows evidence for cyclooxygenase independent effects, since no synergy could be found. The final study of this thesis explored the relation between inflammation as inducer of cancer development by using baseline C-reactive protein levels and *CRP* gene haplotypes as a proxy for chronic inflammation. Both high levels of CRP at baseline as a *CRP* haplotype associated with low CRP levels were associated with an increased risk of lung cancer, even after a latent period of 5-years.

Finally, **Chapter 4** starts with a general discussion of the main findings of the studies described in this thesis and places them in a broader perspective. Furthermore, methodological aspects that were used to study different clinical stages of cancer were described. At the end of this chapter, it was concluded that studying genetic associations, despite the exponential increase of studies in the last years, is still in its infancy. After a short summary of the different genotyping techniques that consecutively were introduced in the recent years, the considerations of some scientists, who questioned the current approach of studying genetic variation, was followed. These scientists argue for studying genome-phenome associations rather than genotype-phenotype associations. Although this new approach is in line with the way in which we think of complex traits such as cancer, practical reasons will delay the introduction. Nevertheless, it is good to have this idea in mind when performing association studies on the association between genotypes and phenotypes.

6.2. Samenvatting

De verbetering van de gezondheidszorg is één van de factoren, die leidt tot vergrijzing van de populatie. Als gevolg hiervan stijgt de incidentie van leeftijdsgerelateerde ziekten zoals kanker. Het belang van onderzoek naar kanker wordt daarom algemeen onderkend. Een grote hoeveelheid inspanningen in de laatste tientallen jaren heeft geleid tot meer inzicht in de complexe systemen die bij de ontwikkeling, behandeling en prognose van kanker betrokken zijn. Bovendien wordt de persoonlijke aanpak van deze gebieden steeds meer gouden standaard. Kennis over genetische variatie zou mogelijk significant kunnen bijdragen aan de inzichten in de complexiteit van deze groep ziekten.

Het doel van dit proefschrift was om een gedetailleerd begrip te verkrijgen van de rol die genetische variaties in genen, die coderen voor specifieke enzymen, hebben in de ontwikkeling en prognose van kanker. Aangezien er vele genen bestaan die kandidaat kunnen zijn voor associatie studies met kanker, werden twee potentieel belangrijke gebieden geselecteerd. Dit betrof cytochroom P450 genen en genen die betrokken zijn bij inflammatie.

Hoofdstuk 1 begint met een algemene introductie over de doelstellingen van dit proefschrift (**Hoofdstuk 1.1.**), de Rotterdam Studie als bronpopulatie voor alle analyses (**Hoofdstuk 1.2.**), de epidemiologie van kanker in de wereld in het algemeen en in de Rotterdam Studie in het bijzonder (**Hoofdstuk 1.3.**) en de meest belangrijke aspecten van genetische variatie (**Hoofdstuk 1.4.**).

In **Hoofdstuk 2** worden studies beschreven die zich richten op de associatie tussen genetische variatie in cytochroom P450 (CYP) genen en kanker. Het hoofdstuk begint met een samenvatting van de huidige kennis, waarbij wij ons beperken tot genetische polymorfismen als factor voor vatbaarheid, waardoor literatuur over de polymorfische expressie van CYPs in (tumor)weefsel niet geïncludeerd is. (**Hoofdstuk 2.1.**) Het overzicht laat zien dat er een grote heterogeniteit tussen de studies bestaat als het gaat om de etniciteit die bestudeerd is, the samenstelling van de controle groepen, het fenotype waar men in geïnteresseerd is en de varianten en referentie typen die gebruikt zijn. Op het eind van dit hoofdstuk wordt het belang van standaardisatie van studie opzet en methodologie besproken die meta-analyses mogelijk zou moeten maken. De volgende twee hoofdstukken beschrijven studies die de associatie tussen *CYP3A* genotypen, welke *CYP3A4*, *CYP3A5* en *CYP3A7* bevatten, steroïd hormoon spiegels en hormoongerelateerde kankers bestuderen. **Hoofdstuk 2.2.** bestudeert de associatie voor mannen en richt zich als gevolg hiervan op prostaatcarcinomen. **Hoofdstuk 2.3.** bestudeert de associatie voor vrouwen en richt zich als gevolg hiervan op mammacarcinomen. In beide studies worden associaties gevonden tussen *CYP3A* genotypen en hormoonspiegels, hoewel dit wel geslachtsafhankelijk blijkt te zijn, aangezien de associaties voor mannen meer uitgesproken zijn. Deze geslachtsverschillen zouden veroorzaakt kunnen worden door androgeen afhankelijke remming van de CYP3A expressie. Sommige genotypen bleken geassocieerd te zijn met de duur van de vruchtbare periode van vrouwen. Deze werd berekend uit de leeftijd van het begin van de menarche en menopauze. Er worden echter geen associaties gevonden met de hormoongerelateerde kankers. Uiteindelijk wordt in **Hoofdstuk**

2.4. het belang van synergisme tussen NSAID gebruik en *CYP2C9* varianten, die geassocieerd zijn met een vertraagd metabolisme, op de preventie van colorectaalkanker bestudeerd. Ondanks dat eerdere studies hebben aangetoond dat de duur van NSAID gebruik belangrijker is dan de dosering, vonden andere studies een significante interactie tussen het medicijn en het genotype *in vivo*. Aangezien deze laatsten zich uitsluitend richtten op specifieke delen van de associatie, had de studie die hier gepresenteerd wordt als doel om meer inzicht van het totaalbeeld te geven. Zowel langdurig NSAID gebruik alsook *CYP2C9* gen varianten waren geassocieerd met een verlaagd risico op colorectaalkanker van voornamelijk de proximale delen van het colon. De combinatie van beide determinanten leidde tot een verdere risicoverlaging, maar er bleek geen sprake te zijn van synergisme. Dit kan duiden op onafhankelijke pathofysiologische mechanismen en onderstreept het belang van de duur van NSAID gebruik boven het belang van de dosering.

In **Hoofdstuk 3** worden studies beschreven die zich richten op de associatie tussen inflammatie en kanker. De preventieve werking van langdurig gebruik van NSAIDs op het risico van colorectaalkanker is wereldwijd geaccepteerd. De meeste studies die deze associatie hebben aangetoond maakten echter gebruik van vragenlijsten over het gebruik in het verleden. Daarom werd in **Hoofdstuk 3.1.** de associatie tussen NSAID gebruik, verkregen via apotheekgegevens, en colorectaalkanker bestudeerd. Verder geeft deze studie informatie over het effect op kanker verdeeld over de anatomische locaties en wordt het effect van stoppen van de medicatie bekeken. De resultaten, die gevonden werden, dragen bij aan de grote hoeveelheid bewijzen dat langdurig NSAID gebruik beschermt tegen colorectaalkanker van voornamelijk de proximale locaties. Deze associatie wordt verder bestudeerd in **Hoofdstuk 3.2.** waarin gekeken wordt naar de rol van COX-2 expressie. In de analyses wordt gebruik gemaakt van een polymorfisme in het *COX-2* gen dat leidt tot 30% lagere expressies van het corresponderende enzym. Ondanks dat de preventieve werking van NSAID meestal verklaard wordt door de remmende werking die het heeft op COX enzymen, laat deze studie een bewijs zien voor COX onafhankelijke effecten, aangezien er geen synergisme gevonden werd. De laatste studie beschreven in dit proefschrift bekijkt de relatie tussen inflammatie als veroorzaker van carcinogenese waarbij het gebruik maakt van 'baseline' CRP waarden en haplotypen van het *CRP* gen als marker voor chronische inflammatie. Zowel hoge 'baseline' waarden van CRP alsook een *CRP* haplotype dat gerelateerd is met lage CRP waarden waren geassocieerd met een verhoogd risico op longkanker, zelfs na includeren van een latentietijd van 5 jaar.

Uiteindelijk wordt er in **Hoofdstuk 4** begonnen met een algemene discussie over de belangrijkste bevindingen van de studies beschreven in dit proefschrift en worden deze in een breder perspectief geplaatst. Vervolgens worden er methodologische aspecten beschreven die gebruikt zijn om de verschillende stadia van de kliniek van kanker te beschrijven. Op het einde van dit hoofdstuk, werd er geconcludeerd dat het aspect van genetische associatie studies, ondanks dat er een zeer grote toename van het aantal is geweest de laatste jaren, nog steeds in de kinderschoenen staat. Na een korte samenvatting van de verschillende tech-

nieken van genotypering die achtereenvolgens geïntroduceerd zijn gedurende de laatste jaren, worden de overwegingen van een aantal die de huidige aanpak van genetische associatie studies in een ander daglicht stellen, gevolgd. Deze wetenschappers houden een betoog voor het bestuderen van genoom-fenoom associaties in plaats van genotype-fenotype associaties. Ondanks dat deze nieuwe aanpak in overeenstemming is met de manier waarop er over complexe aandoeningen zoals kanker gedacht wordt, zullen praktische redenen de introductie van deze manier van analyseren sterk vertragen. Niettemin is het goed om het idee van complexe netwerken in het achterhoofd te houden wanneer er naar genotype-fenotype associaties onderzoek wordt gedaan.

7. Een wonderbaarlijke reis…

Een promotieonderzoek is als het leven; het gaat er niet om waar je eindigt, maar hoe de weg ernaartoe is geweest. Op deze plaats laat ik graag nog eens de mensen voorbij komen die mij de laatste jaren vergezeld hebben tijdens deze wonderbaarlijke reis…

Succes dient niet alleen gemeten te worden aan de hoogten die zijn bereikt, maar ook aan de obstakels die zijn overwonnen.

Traditiegetrouw begin ik met het noemen van mijn twee promotoren, Prof. dr. B. (Bruno) H. Ch. Stricker, hoogleraar Farmaco-Epidemiologie en Prof. dr. H. (Huibert) A. P. Pols, hoogleraar Inwendige Geneeskunde/decaan Erasmus MC en mijn co-promotor Dr. L. (Loes) E. Visser, ziekenhuisapotheker.

Bruno, jij bent in meerdere opzichten een bijzonder persoon. Je laagdrempeligheid en doorzettingsvermogen zijn opmerkelijk. Daarnaast ben je echter ook het levende bewijs dat je niet rechtlijnig kunt denken in een ronde schedel (D. Ozon) Voor een controlefreak als ikzelf is het bijna ondoenlijk geen planning te hebben, te vertrouwen op eigen kennis zonder een beroep te doen op de expertise van anderen en ingewikkelde denkstructuren steevast te doorbreken met het aanhoren van anekdotes. De vraag blijft echter of al jouw promovendi die hun onderzoek hebben afgerond dit ondanks of juist dankzij deze manier van werken zo doeltreffend hebben kunnen doen.

Er is geen betere wijsheid dan eigenwijsheid.

Huib, als mijn tweede promotor was je slechts zijdelings betrokken bij mijn onderzoek. Niettemin wist je altijd in zeer korte tijd waar de pittfalls zaten en zorgde je iedere keer opnieuw voor een gezellige sfeer en bemoedigende woorden. Jouw passie voor de Interne Geneeskunde – die ik voor het eerst zag tijdens colleges zo'n 10 jaar geleden – heeft ertoe geleid dat ook ik mij al lange tijd verbonden voel met dit geweldige en veelzijdige specialisme. De laatste jaren heb je er tevens voor gezorgd dat ik de toegevoegde waarde van epidemiologische kennis voor een internist heb weten te erkennen. Bedankt voor alle leerzame momenten.

Artsen hebben in de praktijk altijd al epidemiologisch gedacht. (Feinstein)

Loes, wij begonnen als collega's. Na het afronden van je eigen promotieonderzoek bleef je naast je werk als ziekenhuisapotheker nog 1 dag per week betrokken bij het onderzoek op de afdeling Epidemiologie & Biostatistiek en werd ik jouw eerste promovendus die je in de hoedanigheid van co-promotor ondersteunde. Jouw no-nonsense mentaliteit heeft ertoe geleid dat er laatste versies van mijn artikelen bestaan en je wist als ervaringsdeskundige altijd wel een oplossing voor het doorbreken van de vele obstakels die je als promovendus tegenkomt.

De tijd die verstrijkt tussen het doen van onderzoek, het schrijven van een artikel en de uiteindelijke publicatie is positief gecorreleerd met het mannelijk geslacht en een hogere academische graad. (Blumenthal 1997)

Dit promotieonderzoek zou niet mogelijk zijn geweest zonder het initiatief van Prof. dr. A. (Albert) Hofman, die bijna twintig jaar geleden een bevolkingsonderzoek oprichtte waaruit al jarenlang gegevens gehaald worden die bijdragen aan veelzijdige en hooggewaardeerde publicaties. Bert, jouw passie voor de epidemiologie en vooral jouw visualisatie van prospectieve en retrospectieve cohort studies zal ik altijd blijven herinneren.

Regeren is vooruitzien. (Emile de Girardin)

Grote waardering gaat tevens uit naar alle ERGO-deelnemers, ondersteunend personeel en huisartsen in Ommoord die er al deze jaren al voor zorgen dat "The Rotterdam Study" een begrip is in de internationale wetenschappelijke literatuur.

Give us the tools and we will finish the job. (Sir Winston Churchill)

Alle co-auteurs wil ik hierbij bedanken voor hun intellectuele bijdrage aan mijn artikelen. Een speciaal woord van dank gaat uit naar Prof. dr. J.W. (Jan-Willem) W. Coebergh voor zijn gedegen en snelle commentaar, Prof. dr. T. (Theo) Stijnen voor zijn hulp bij straightforward en minder straightforward analyses, Prof. dr. T. (Ted) A.W. Splinter voor zijn belangrijke bijdrage aan mijn CRP stuk, Prof. dr. F. (Frank) H. de Jong voor het begrijpelijk maken van de vele resultaten die betrekking hadden op het hormoonmetabolisme, Dr. R. (Ron) N. van Schaik voor zijn pionierswerk binnen het begrip cytochroom P450 en niet te vergeten Dr. A. (André) G. Uitterlinden die voor mij het toonbeeld van een wetenschappelijk onderzoeker is en vooral een belangrijke bijdrage heeft geleverd aan mijn review.

Overtuiging is een luxe voor mensen aan de zijlijn.

Prof. dr. J.W.W. Coebergh, Prof. dr. E.J. Kuipers en Prof. dr. ir. F.E. van Leeuwen wil ik bedanken voor hun bereidheid zitting te nemen in de kleine commissie en dit proefschrift inhoudelijk te beoordelen.

Wetenschap is één van de weinige zaken die je vermenigvuldigd door haar te delen.

Belangrijke reisgenoten waren mijn collega's. Jullie aanwezigheid heeft mijn tijd op de afdeling Epidemiologie & Biostatistiek zeer verrijkt en ik ben blij dat ik er weer een aantal vrienden voor het leven bij heb gekregen. Mijn kamergenoten verdienen een bijzonder plaatsje. Albert-Jan,

samen in Dordt en toen samen in Rotterdam. Het was een hele verandering, maar gelukkig hoefde ik deze omschakeling niet alleen te doen. Wij zijn de laatste jaren steeds dichter naar elkaar toe gegroeid en hebben ondertussen aan een half woord genoeg. De dagen dat jij in Den Haag zat was het stil op de kamer en ik genoot ervan als er op de dinsdagen, woensdagen en donderdagen weer koffie werd gezet (of warm water voor de thee werd gehaald) en de hete post, dondervogels en cynische en relativerende opmerkingen door de kamer heen vlogen. Naast de mentale steun die ik telkens weer van jou kon verwachten, aarzelde je ook nooit om mij weer eens met analyses te helpen. Jij doorgrondde HaploView, HaploStat, en genetische databases in een verbazingwekkend tempo en vele aspecten in mijn artikelen zijn aan jou te danken. Heel veel succes nog bij het afronden van je eigen promotieonderzoek en ik wacht alweer op je terugkeer in Dordt! Cornelis, mijn andere soul-brother. Samen met Albert-Jan vormde je een geweldig team dat tegengas gaf aan mijn vurige en kleurrijke persoonlijkheid (zoals je dat zelf in je proefschrift noemde). Op het moment dat ik je leerde kennen was je al een aardige hoeveelheid obstakels gepasseerd en wist je mij van goede adviezen te voorzien die het beloop van mijn promotietraject zouden moeten vereenvoudigen. Op jouw grote dag mocht ik naast je staan en dat zal ik niet snel vergeten. Het was het begin van een lange vriendschap waarbij ook Antsje niet mag ontbreken. De afstand van de Randstad naar het Friese land moet dan ook makkelijk te overbruggen zijn. Dus: wanneer gaan we zeilen? Sabine, wat moet je zeggen van iemand die het voor elkaar krijgt een gezin draaiende te houden, een zeer belangrijke functie te bekleden en ook nog de tijd weet te vinden voor sociale contacten terwijl ze dit zelf relativeert met een ontegenzeggelijke zelfspot? Ik weet alleen maar dat jouw vele opbeurende en gezellige woorden, die waarschijnlijk het gevolg zijn van jouw levenswijsheid, een groot aandeel hebben gehad in het feit dat er nu van mij een proefschrift op tafel ligt. Ik wil je bedanken voor alle gezellige momenten als kamergenoot en wat miste ik je toen je weer volledig terug ging naar Den Haag. En je had gelijk: dat boekje komt er wel…

De belangrijkste dingen kosten vaak de minste tijd.

Een speciaal woord is ook gericht aan Mark E. en Ronald, 'mijn' studenten. Mark, wij begonnen samen te werken in het begin van mijn promotietraject. Jouw statistische en medische kennis leidde tot interessante discussies en tijd om over andere dingen te praten. Onze samenwerking is uitgegroeid tot een leuke vriendschap met jou en Marieke. Succes nog even met de co-schappen en dan mogen jullie ook het geweldigste beroep ter wereld uitvoeren…Ronald, jij was de tweede die onze kamer kwam versterken. Een jaar lang heb je alle ins en outs van promotieonderzoek doen mogen meebeleven en heb je hier een beeld van kunnen vormen voor een gedegen keuze in de toekomst. Jouw onderzoek over ACE en kanker was ingewikkeld, maar potentievol. De meeste gesprekken hebben we gehad over de kliniek. Wanneer je dit leest ben je begonnen met je co-schappen en ik hoop dat het je bevalt. Veel succes in de toekomst.

Wie geen vertrouwen stelt in anderen, zal ook nimmer het vertrouwen van anderen winnen. (Lao-tse)

Ook de andere collega's van de Farmaco-Epidemiologie, met wie elke stafbespreking op dinsdag weer een gezellig samenzijn werd, mogen hier niet vergeten worden: Annemiek (heel veel succes met je studie), Bert (laat het me weten als je OCD-clubje is opgericht), Betty (influenza en Indisch hebben ineens een andere betekenis gekregen), Charlotte (je bent een goede opvolgster), Dika (bedankt voor onze gesprekken en een kus voor Daantje), Eva (Lissabon was gezellig en je hebt een goede keuze gemaakt, zet 'm op!), Gianlucca (it was a pleasure meeting you, good luck for the future), Hedi (jij weet als geen ander hoe je niet-significante genetische associaties tot een proefschrift kunt verwoorden), Jeanne (wetenschap is jouw ding, heel veel succes in de toekomst), Katia (bedankt voor je Vlaamse inbreng, het was altijd erg gezellig) en haar man Guy (bedankt voor de zondagochtend overpeinzingen ten aanzien van mijn CRP stuk), Mariëtte (kort maar krachtig op de Epi, tot de volgende ROIG), Marissa (altijd druk, druk, druk, maar jij komt er wel), Martina (ineens weet je het: je wordt doctor!), Matthijs (en ik maar denken dat ik de meest serieuze was), Mendel (we keep in touch! Geniet van je moederschap), Miriam (knap heeft inderdaad meerdere betekenissen, ik wacht op de oratie!), Monique (binnen een aantal jaren heeft Nederland er weer een zeer veelzijdige ziekenhuisapotheker bij), Nathalie (het was altijd weer gezellig als je even kwam buurten en ook wij zien elkaar weer bij het volgende ROIG), Patty (stille wateren hebben diepe gronden), Rudolf (heb je maar kort mogen meemaken, maar je lijkt veelbelovend) en Yannick (nog meer Vlaamse inbreng, succes met je COPD en CRP, ben erg benieuwd).

Het werk van een AIO is te vergelijken met dat van een monnik qua bezieling, doorzetting en beloning.

Daarnaast wil ik alle anderen van de afdeling Epidemiologie & Biostatistiek met wie ik vele lunch-, koffie-/thee- en borrelmomenten heb ervaren hier noemen: Abbas (sounds like 'lots of fun'), Arfan (respect voor jouw kennis), Arlette (voor regeltjes kom ik voortaan naar jou toe), Christiane (we definitely like the same designs), Dominiek (the roots of education are bitter, but the fruit is sweet), Frank-Jan (bedankt voor de momenten van de laatste maanden), Isabella (een goede buur is beter dan een verre vriend), Julia (slaaponderzoek doen is alles behalve slaapverwekkend), Lonneke en Marieke (hard werken en gezelligheid gaan prima samen), Mark Sie (what's in a name, wij zullen elkaar altijd tegen blijven komen), Meike (gezellig dat ik een studiegenootje weer tegenkwam), Michiel (samen op het Alfrink en samen op de Epi), Miranda (postdocs zijn net mensen), Monika (je had bijna dezelfde reis gemaakt), Sharmila (jij was altijd in voor Coenen of andere uitjes), Simone (de ogen zijn de poort naar de wereld), Sjoerd (jij bent een SigmaPlot-kenner), Susanne (Viel Glück in Deutschland), Suzette

(heerlijk actief) en Ylian (bedankt voor je tip van de bloedbank, het heeft mijn promotietijd verrijkt)

Voor een promotieonderzoek kunnen Neil Armstrong's woorden beter verdraaid worden tot "that's one giant leap for man, but a small step for mankind".

Liong-Kie Ko en Tanneke Marijt, hoofd donorartsen, wil ik bedanken voor de mogelijkheid die zij mij hebben geboden als donorarts mijn onderzoeksweek op de woensdagen te onderbreken met praktijkgerichte werkzaamheden. Ik vond het heerlijk om weer even deel uit te maken van een team en met mensen bezig te zijn. Door de jaren heen heb ik het erg gezellig gehad met alle donorassistenten in Breda en Gouda en alle andere plaatsen waar ik wel eens ingevallen heb. Jullie hebben er vanaf het begin voor gezorgd dat ik me meer dan welkom voelde op 'jullie' bloedbank. Bovendien heb ik veel geleerd van bloedproducten; het geven van een zakje packed cells, trombocyten of plasma zal nooit meer hetzelfde zijn…

Het verschil tussen theorie en praktijk is dat er in theorie geen verschil is tussen beiden maar in praktijk wel degelijk.

Het mooie aan een reis maken is dat je altijd weer thuiskomt en zo voelde het ook toen ik weer terug kwam in het Albert Schweitzer ziekenhuis in Dordrecht. De spontaniteit, glimlachen en goede bedoelingen voelden weer aan als een warm bad en ik ben blij dat ik mijn nieuwe reis tot specialist voor een groot gedeelte mag afleggen met mensen die mij stimuleren en een lach op mijn gezicht toveren. In het bijzonder ben ik Dr. A. (Adrie) C.M. van Vliet en Dr. E. (Eric) F.M. van Bommel, opleiders Interne geneeskunde, dank verschuldigd voor het vertrouwen dat zij altijd in mij gehad hebben. Adrie, jij hebt vanaf het moment dat ik co-assistent op jouw afdeling was mij altijd gestimuleerd tot een combinatie van onderzoek en patiëntenzorg. Bedankt voor je vertrouwen de laatste jaren en ik hoop nog veel van je te leren. Eric, door jou ben ik de band met het Albert Schweitzer nooit helemaal kwijtgeraakt. Jij laat mij van dichtbij meegenieten van je paradepaardje, de retroperitoneale fibrose. De manier waarop je met een populatie uit de dagelijkse praktijk een belangrijke bijdrage levert aan de wetenschappelijke kennis is voor mij het voorbeeld van een 'echte' clinicus. Yes…

Wie oude kennis koestert en voortdurend nieuwe vergaart, mag een leraar van anderen zijn. (Confucius)

Alhoewel onderzoek doen soms een eenzame aangelegenheid is, sta ik tijdens de verdediging op 9 mei zeker niet alleen. Ik prijs me gelukkig dat twee van mijn vriendinnen al vanaf het begin van deze hele reis bereid zijn geweest naast me te staan tijdens deze bijzondere dag. Jolanda, wij zijn ondertussen al alles van elkaar geweest: je begon als mijn co, maar al

snel groeide dit uit naar vriendinnen, mede-promovendus en nu zelfs directe collega's. De vriendschap met jou is goud waard. Suzanne, ook jou leerde ik als co-assistent kennen. Jij volgde echter een andere weg en ik hoop dat je je doel mag gaan bereiken. Jouw zachte en liefdevolle karakter zal dit zeker niet in de weg staan. Lieve Jo en Suus, we maken er een mooi dagje van en ik hoop dat we gezamenlijk nog vele mooie momenten in de toekomst mogen beleven. Ik ben blij dat jullie mijn vriendinnen en paranimfen zijn. Daarnaast wil ik jullie mannen, Johan en Rens, bedanken voor alle gezellige momenten samen en dat ik de dames af en toe eens mag 'lenen'.

Een vrouw moet twee keer zo goed zijn als een man, om half zo goed gevonden te worden. Gelukkig is dit niet moeilijk. (Charlotte Whitton)

Alle familieleden, vrienden en collega's die ik hier niet met naam en toenaam noem, maar die wel altijd geïnteresseerd zijn geweest in mijn werkzaamheden, wil ik bedanken voor hun interesse en bemoedigende woorden.

Men vraagt zich ten aanzien van het eigen promotieonderzoek zo nu en dan wel af met welk superdetail van de evolutie men bezig is wanneer men dit onderzoek aan een gedwee luisterende buitenstaander tracht uit te leggen.

De belangrijkste plaats in mijn hart wordt ingenomen door mijn ouders. Lieve pap en mam, jullie stonden aan de basis van mijn leven en hebben ervoor gezorgd dat ik geworden ben wat ik nu ben. Pap, op momenten dat het nodig is ben je er altijd voor me. Je hebt me 17 jaar geleden samen met Yolanda ook het mooiste cadeau geschonken: mijn zus(je) Karen. Yolanda, wij kunnen heerlijk over alle ziekenhuis-perikelen spreken en groeien mede daardoor steeds meer naar elkaar toe. Kaatje, jij bent mijn alles. Jouw spontaniteit, vrijheid en innemende lach zijn aanstekelijk. Heel veel succes met je toekomstige carrière en ik zal er altijd voor je zijn. Mam, moeki, jij kent mij waarschijnlijk beter dan dat ik mezelf ken. Jij hebt alles in mijn leven van dichtbij meegemaakt en weet als geen ander wat de laatste jaren voor mij betekent hebben. Je weet altijd de juiste woorden te vinden om het maximale uit mijn leven te halen. Als ik ook maar half jouw doorzettingsvermogen en positiviteit heb geërfd komt het allemaal zeker goed. Lieve mam, ik hou van je en hoop nog lang vriendinnen te blijven…

Een bloem kan pas groeien op vruchtbare grond.

Claire

8. Curriculum Vitae

Claire Siemes was born on October 17th 1977 in Leidschendam, The Netherlands. In 1996, she completed secondary school at the "Alfrink College" in Zoetermeer, and started Medical School at the "Erasmus University" in Rotterdam. During her study, she worked in the students' team of Hematology and for the Traumacenter of the South-West region of The Netherlands, both at the Erasmus Medical Center, Rotterdam. In addition, she worked for several general practices in the surroundings of Rotterdam. From January 1998 until December 1999 she was a member of the "Mentorencommissie". From January until December 1999, she was chairman of this committee and first ambassador of the Erasmus University, department of Health Sciences. To complete her Master of Science in Medicine, she participated in research on the psychosocial evaluation and quality of life of Dutch women with Turner Syndrome after long-term growth hormone therapy in combination with low-dose estrogen replacement therapy. This research was performed at the department of Pediatric Endocrinology of the Sophia Children's Hospital, Rotterdam (Supervisor: Dr. S.M.P.F. de Muinck Keizer-Schrama). In March 2003, she obtained her medical degree. For fourteen months she worked at the Albert Schweitzer Hospital in Dordrecht as a resident of Internal Medicine (Head: Dr. A.C.M. Van Vliet). In June 2004, she started the work presented in this thesis at the pharmaco-epidemiology unit (Head: Prof. dr. B.H.Ch. Stricker) of the department of Epidemiology & Biostatistics (Head: Prof. dr. A. Hofman) in close collaboration with the department of Internal Medicine (Head: Prof. dr. H.A.P. Pols) of the Erasmus University in Rotterdam. A part of this thesis was presented at a pharmaco-genetic seminar in Utrecht and at a congress of the International Society of Pharmaco-Epidemiology (ISPE 2006) in Lisbon, Portugal. As of August 2005, she was a member of the ERGO-management team, representing PhD-students. She obtained a Master of Science degree in Clinical Epidemiology at The Netherlands Institute for Health Sciences (NIHES) in June 2006. During her PhD-project, she worked as a medical examiner (keuringsarts) at Sanquin Blood facility for one day per week in Breda and Gouda (Heads: Drs. L.K. Ko and Drs. T. Marijt). In January 2007 she started her specialist training in Internal Medicine at the Albert Schweitzer Hospital in Dordrecht (Head: Dr. A.C.M. Van Vliet / Dr. E.F.B. van Bommel).

9. List of manuscripts

Siemes C, van der Meulen J, van Vliet AC. Comment on: [Tuberculosis spondylodiscitis in three older patients] Spondylodiscitis door tuberculose bij 3 oudere patienten. *Ned Tijdschr Geneeskd* 2003;147(16):783.

Siemes C, Frenay HME, van Vliet ACM, Cleophas TJ. Relapsing bacteremias and about forgetting the aorta, a case report. *Angiology* 2003;41:259-62.

Siemes C, Frenay HME, van Vliet ACM. Abdominal aortic aneurysm as source of infection (abstract). *Neth J Intern Med – Abstract Book* 2003;15:77.

Siemes C, van der Meulen J, de Jong GH. An artistically acquired acute renal failure (abstract). *Neth J Intern Med – Abstract Book* 2004;16:33.

Siemes C, Govaert GAM, Avontuur JAM, Hendriksz TR, van Bommel EFH. Retroperitoneal fibrosis and mycotic abdominal aneurysm – a challenging combination (abstract). *Neth J Intern Med – Abstract Book* 2005;17:39.

Gonzalez-Zuloeta Ladd AM, Arias-Vasquez A, **Siemes C**, Yazdanpanah M, Coebergh JW, Hofman A, Stricker BHCh, van Duijn CM. Differential Roles of Angiotensinogen and Angiotensin Receptor type 1 Polymorphisms in Breast Cancer Risk. *Breast Cancer Res Treat* 2007: Mar;101(3):299-304..

Siemes C, Visser LE, Coebergh JWW, van Schaik RHN, Hofman A, Pols HAP, Stricker BHCh. Long-term NSAID use and CYP2C9 variant alleles protect against colorectal cancer (abstract). *Pharmacoepidemiology and Drug Safety* 2006;15:S63.

van der Hooft CS, Dielemans JP, **Siemes C**, Aarnoudse AJ, Verhamme KMC, Stricker BHCh. Ziekenhuisopnames door bijwerkingen van geneesmiddelen – een inventarisatie -. *Rapport College voor Zorgverzekeringen*. 2006.

van der Hooft CS, Dielemans JP, **Siemes C**, Aarnoudse AJ, Verhamme KMC, Stricker BHCh. ADR-related hospitalisations: A population-based cohort study (abstract). *Pharmacoepidemiology and Drug Safety* 2006;15:S99.

Gonzalez-Zuloeta Ladd AM, Arias-Vasquez A, **Siemes C**, Coebergh JW, Hofman A, Witteman JC, Uitterlinden AG, Stricker BHCh, van Duijn CM. Transforming-growth factor beta(1) Leu10Pro polymorphism and breast cancer morbidity. *Eur J Cancer* 2007:Jan;43(2):371-374.

Siemes C, Visser LE, Coebergh JW, Splinter TA, Witteman JC, Uitterlinden AG, Hofman A, Pols HAP, Stricker BHCh. C-reactive protein levels, variation in the C-reactive protein gene, and cancer risk: The Rotterdam Study. *J Clin Oncol.* 2006 Nov 20; 24(33):5216-22.

Gonzalez-Zuloeta Ladd AM, Arias-Vasquez A, **Siemes C**, Rivadeneira F, Hofman A, Uitterlinden AG, Stricker BHCh, van Duijn CM. ESR1 polymorphism and breast cancer risk. *Breast Cancer Res Treat*. 2007 ahead of publication.

van Bommel EFH, **Siemes C**, van der Veer SJ, Han SH, Huiskes AWLC, Hendriksz TR. Clinical value of gallium-67 SPECT scintigraphy in the diagnostic and therapeutic evaluation of retroperitoneal fibrosis. A prospective study. *Neth J of Med.* 2007 ahead of publication.

van Bommel EFH, **Siemes C**, van der Veer SJ, Han SH, Huiskes AWLC, Hendriksz TR. Clinical value of gallium-67 SPECT scintigraphy in the diagnostic and therapeutic evaluation of retroperitoneal fibrosis. A prospective study. *Neth J Intern Med – Abstract Book* 2007.

van Bommel EFH, **Siemes C**, Hak AE, van der Veer SJ, Hendriksz TR. Long-term renal and patient outcome in non-malignant retroperitoneal fibrosis treated with prednisone. *Am J Kidney Diseases* 2007 ahead of publication.

van Bommel EFH, **Siemes C**, Hak AE, van der Veer SJ, Hendriksz TR. Long-term renal and patient outcome in non-malignant retroperitoneal fibrosis treated with prednisone. *Neth J Internal Med. – Abstract Book* 2007.

Siemes C, Visser LE, Van Schaik RHN, Uitterlinden AG, Coebergh JWW, Hofman A, Pols HAP, Stricker BHCh. A critical appraisal of association studies of cytochrome P450 gene polymorphisms and susceptibility to cancer. *Submitted*.

Siemes C, Visser LE, de Jong FH, Coebergh JWW, Uitterlinden AG, Hofman A, Pols HAP, Stricker BHCh, Van Schaik RHN. *CYP3A* gene variation, steroid hormone serum levels and prostate cancer risk: The Rotterdam Study. *Submitted*.

Siemes C, Visser LE, de Jong FH, Coebergh JWW, Van Schaik RHN, Uitterlinden AG, Hofman A, Pols HAP, Stricker BHCh. *CYP3A* gene variation, steroid hormone serum levels and breast cancer risk: The Rotterdam Study. *Submitted*.

Siemes C, Visser LE, Coebergh JWW, Van Schaik RHN, Uitterlinden AG, Hofman A, Pols HAP, Stricker BHCh. Cytochrome P450 (CYP) *2C9*2* and **3* variant alleles do not increase the beneficial effects of NSAIDs on colorectal cancer. *Submitted*.

Siemes C, Visser LE, Coebergh JWW, Hofman A, Pols HAP, Stricker BHCh. A population-based study on NSAID use and colorectal cancer risk. *Submitted*.

Siemes C, Visser LE, Coebergh JWW, Van Schaik RHN, Uitterlinden AG, Hofman A, Pols HAP, Stricker BHCh. The protective effect of NSAIDs on cancer risk and influence by COX-2 expression. *Submitted*.

Van der Knaap R, **Siemes C**, Visser LE, Coebergh JWW, van Duijn CM, Hofman A, Stricker BHCh. ACE insertion/deletion polymorphism, ACE inhibitors and cancer risk. *Submitted*.

Van der Hooft CS, Verhamme KMC, **Siemes C**, Aarnoudse AJ, Dieleman JP, Stricker BHCh, Sturkenboom MCJM. ADR-related hospitalisations: a population-based cohort study. *Submitted*.

Siemes C, van der Meulen J, de Jong GH. An artistically acquired acute renal failure. *Submitted*.

Acknowledgement:
van Bommel EFH, Hendriksz TR, Huiskes AWLC, Zeegers AGM. Brief communication: Tamoxifen Therapy for nonmalignant retroperitoneal fibrosis. *Ann Intern Med* 2006;144:101-106.