Meta-analysis of the population and phenotypic expression of CYP2C9/2C19 polymorphism on drug metabolism in different ethnicities

Yingxin Xu

Doctor of Philosophy

ASTON UNIVERSITY

March 2006

Aston University

Meta-analysis of the population distribution and phenotypic expression of CYP2C9/2C19 polymorphism on drug metabolism in different ethnicities.

Yingxin Xu
Doctor of Philosophy 2006

The term "pharmacogenetics" has been defined as the scientific study of inherited factors that affect the human drug response. Many pharmacogenetic studies have been published since 1995 and have focussed on the principal enzyme family involved in drug metabolism, the cytochrome P450 family, particularly cytochrome P4502C9 and 2C19. In order to investigate the pharmacogenetic aspect of pharmacotherapy, the relevant studies describing the association of pharmacogenetic factor(s) in drug responses must be retrieved from existing literature using a systematic review approach. In addition, the estimation of variant allele prevalence for the gene under study between different ethnic populations is important for pharmacogenetic studies.

In this thesis, the prevalence of *CYP2C9/2C19* alleles between different ethnicities has been estimated through meta-analysis and the population genetic principle. The clinical outcome of *CYP2C9/2C19* allelic variation on the pharmacotherapy of epilepsy has been investigated; although many new antiepileptic drugs have been launched into the market, carbamazepine, phenobarbital and phenytoin are still the major agents in the pharmacotherapy of epilepsy. Therefore, phenytoin was chosen as a model AED and the effect of *CYP2C9/2C19* genetic polymorphism on phenytoin metabolism was further examined.

An estimation of the allele prevalence was undertaken for three *CYP2C9/2C19* alleles respectively using a meta-analysis of studies that fit the Hardy-Weinberg equilibrium The prevalence of *CYP2C9*1* is approximately 81%, 96%, 97% and 94% in Caucasian, Chinese, Japanese, African populations respectively; the pooled prevalence of *CYP2C19*1* is about 86%, 57%, 58% and 85% in these ethnic populations respectively. However, the studies of association between *CYP2C9/2C19* polymorphism and phenytoin metabolism failed to achieve any qualitative or quantitative conclusion. Therefore, mephenytoin metabolism was examined as a probe drug for association between *CYP2C19* polymorphism and mephenytoin metabolic ratio. Similarly, analysis of association between *CYP2C9* polymorphism and warfarin dose requirement was undertaken.

It was confirmed that subjects carrying two mutated *CYP2C19* alleles have higher S/R mephenytoin ratio due to deficient CYP2C19 enzyme activity. The studies of warfarin and *CYP2C9* polymorphism did not provide a conclusive result due to poor comparability between studies.

The genetic polymorphism of drug metabolism enzymes has been studied extensively, however other genetic factors, such as multiple drug resistance genes (MDR) and genes encoding ion channels, which may contribute to variability in function of drug transporters and targets, require more attention in future pharmacogenetic studies of antiepileptic drugs.

Key words: CYP2C9, CYP2C19, population distribution and pharmacogenetics

Acknowledgement

I am very grateful to my supervisor, Professor Helen Griffiths, for her invaluable advice, helpful guidance, and encouragement throughout the period of this thesis.

The scholarship provided by the Centre of Evidence-based Pharmacotherapy (CEBP) from Professor Alain Li Wan Po is also faithfully appreciated. Without it, I would not have been able to accomplish this thesis.

I would also like to extend my special thank to Dr. Dan Cornford and Dr. Weiya Zhang for their invaluable advice on the statistical methods applied in this thesis.

Thanks also to my colleagues, Katerina, Sumiah, Shamyla, Lili, Cynthia, Rita, Heyi, Shereen and others who encouraged me and were always there to give me advice and help.

Finally, my deepest thanks to my dearest parents and husband for their endless patience, understanding, consistent support and encouragement throughout the whole period of my PhD and study abroad.

List of content

List of table

List of figure

Abbreviation

AEDs – anti-epileptic drugs

CI – confidence interval

CYP – cytochrome P450

FDA – Food and Drug Administration

GABA – gamma aminobutyric acid

HPPH – 5-(4-hydroxyphenyl)-5-phenyhydantoin

HWE– Hardy-Weinberg equilibrium

MDR – multiple drug resistance

PCR – polymerase chain reaction

PM – poor metabolizer

S/R ratio – ratio of racemic (S)-mephenytoin and (R)-mephenytoin in urine

SD – standard deviation

SNP – single nucleotide polymorphism

WHO – world health organization

Glossary in the thesis

Allele – one of the several alternative forms of a gene or DNA sequence that can exist at a single locus

Candidate genes – Genes that are thought to be more likely to have polymorphisms that influence response to a given drug compared with a random gene from the genome

Exon – any non-intron section of the coding sequence of a gene; together, the exons constitute the mRNA and are translated into protein

Ethnic/Ethnicity – a term referring human population whose members identify with each other on the basis of a presumed common genealogy or ancestry, which are usually united by common cultural, behavioural, linguistic or religious practices

Epistatic – a situation in which the differential phenotypic expression of genotypes at one locus depends on the genotype at another locus

F-statistics – a general statistical tool for analyzing variances (variation in gene frequencies)

Genetic polymorphism – the occurrence in a population or among populations of several phenotypic forms associated with alleles of one gene or homologs of one chromosome

Genetic variation – phenotypic variance resulting from the presence of different genotypes in the population

Genotype – the specific allelic components for gene or a set of genes

Hardy-Weinberg Equilibrium – the stable frequency distribution of genotypes A/A, A/a, and a/a, in the proportions p^2, $2pq$, and q^2, respectively (where p and q are the frequencies of the allele A and a), that is consequence of random mating in the absence of mutation, migration, natural selection, or random drift

Intron – alternative name of intervening sequence, a segment of largely unknown function within a gene. This segment is initially transcribed, but the transcript is not found in mRNA

LINEs – long interspersed element; a type of large repetitive DNA segment found throughout the genome

Linkage disequilibrium – a condition in which the haplotype frequencies in a population deviate from the value they would have if the genes at each locus were combined at random

Markov Chain – a usually discrete stochastic process (e.g. a random walk) in which the probabilities of occurrence of various future states depend only on the present state of the system or on the immediately preceding state and not on the path by which that present state was achieved

Meta-analysis – a statistical technique for combining the finding from many independent studies

Missense mutation – a mutation that alters a codon so that is encodes a different amino acid

Mutation – (1) the process that produces a gene or a chromosome set differing from the wild type. (2) the gene or chromosome set that results from such a process

Nonsense mutation – a mutation converts a codon that encodes an amino acid into a stop codon, one that specifies the termination of translation

Pharmacogenetics –a study of genetic variation underlying differential response to drugs, particularly genes involved in drug metabolism

Phenotype – the detectable outward manifestations of a specific genotype

Phylogenetic tree – a tree showing the evolutionary interrelation ships among various species or other entities that are believed to have a common ancestor

Polygenes – a number of genes

Population –a group of individuals who share one or more characteristics, often defined by the demographical and geographical features.

Pseudogene – a defective gene that does not produce a functional protein. Pseudogenes are relics of gene duplications where one of the copies had degenerated and lost its function.

Race – a term commonly used to distinguish a population of humans from other populations

Regulator – regulatory genes that have roles in turning on or off the transcription of structural genes

SINE – short interspersed element; a type of small repetitive DNA sequence found throughout a eukaryotic genome

Single nuclear polymorphism (SNPs) – a single nucleotide change in a sequence of DNA

Splicing – the reaction that removes introns and joins together exons in mRNA

Systematic review – a review with a clearly formulated question has been prepared using a systematic approach and explicit methods to identify, select and critically appraise relevant research, and to collect and analysis data from the studies in order to minimize biases and random errors that are documented in a materials and methods section. It may or may not apply statistical analysis to the results from independent studies

Transcription – the synthesis of RNA with the use of a DNA template

Chapter 1 Introduction

1.1 Population Genetics

The need for study of Population genetics

As it is widely accepted that variable drug response is highly associated with patient ethnic or geographic origin (Evans *et al.* 2001), the analysis of population based pharmacogenetic factors would be useful for further improvement of pharmacotherapy. In this thesis, the theory of population genetics will be applied to pharmacogenetic data analysis.

Before analyzing the data from different studies, the population genetic theory should be applied to ensure that each of the studies can be considered as a fair sample from the population under study. Therefore the further comparisons will be both valid and meaningful.

What is population genetics?

Population genetics is the study of the distribution of and change in allele frequencies under the influence of the five evolutionary forces: natural selection, genetic drift, mutation, migration and non-random mating. It also takes population subdivision and population structure into account. As such, it attempts to explain and analyse such phenomena as adaptation and speciation using mathematical principles (Ewens 1979, Weiss 1993, Durbin 1998).

Population genetics has changed dramatically with advances in molecular technology. The focus of polymorphism is switching from observation of polymorphic phenotypes, such as Mendelian experiments, to molecular polymorphism, *i.e.* DNA polymorphism, or protein polymorphism, which is typically presented in the data that consist of a sample of aligned sequences. The polymorphism of the aligned sequences provides information about genealogical or ancestral relationships between genomes and is usually demonstrated in a phylogenetic tree. There are a few essential theoretical concepts used in population genetics for analysing genetic polymorphisms (Hartl and Clark 1999), such as Hardy-Weinberg Equilibrium (HWE), Linkage Disequilibrium (LD), Random Genetic Drift (RGD), Hudson-Kreitman Aguade test (HKA), Poisson Random Field (PRF) and Quantitative Train Locus (QTL).

The well developed and ever growing theoretical knowledge currently allows population genetics to make quantitative prediction of human evolution by using the information of 'genetic polymorphism', which underlies those variations, (Carvalli-Sforza 1998, Chaabani 2002). Population genetics is not only a vital ingredient in modern evolution research, but also a bridge from human evolutionary history to genetic medicine (Jorde *et al.* 2001). In pharmacogenetics, it is well known that allelic variations affect the human response to many drugs, especially the genetic polymorphism in Drug Metabolism Enzymes (DMEs). This is the main concern in pharmacogenetics research and probably contributes significantly to the drug response (Wilson *et al.* 2001). For instance, the poor-metabolizer phenotype of the antiepileptic drug Mephenytoin is observed in 2-6% of Caucasians, but in 14-22% of populations in far Eastern countries such as Japan, China and India (Meyer 1994a). Therefore, it would be possible to understand more clinically about the variation in response among populations if the population genetic basis of variable drug response was available.

Hardy-Weinberg Equilibrium

In population genetics, the Hardy-Weinberg equilibrium (HWE) is a simple description of how genotype and allele frequency change in a population. It predicts expected genotype frequency using allele frequency in a diploid Mendelian population.

There are five assumptions in the Hardy-Weinberg equilibrium.

1. The population includes a very large (virtually infinite) number of individuals.
2. Individuals in the population mate at random.
3. There is no migration into or out of the population.
4. No new mutations appear in the population gene pool.
5. There is no genotype-dependent difference in the ability to survive to reproductive age and hence to transmit genes to next generation.

With these assumptions, the genotype frequencies for a gene with two alleles (A and a) would fit the following equation: (p as frequency of allele A, q as frequency of allele a)

Allele frequencies: $p + q = 1$

Expected genotype frequencies: $p^2 + 2pq + q^2 = 1$

Although assumptions of Hardy-Weinberg equilibrium (HWE) cannot hold true in reality – populations are finite, mutations occur constantly, migration is common, and many genotypes affect the survival of the individual - most natural populations show

approximate agreement with the Hardy-Weinberg expectation. The genotype frequencies typically conform to Hardy-Weinberg expectation from the knowledge of allele frequencies (Crow 1988). In population genetics, the deviations from HWE usually arise from one of four sources: (Cavalli-Sforza *et al.* 1996)

1. The genetic model (postulated alleles and dominance relationships) is incorrect.
2. Laboratory procedures or testing reagents are not always adequate.
3. Natural selection eliminates some phenotypes preferentially (an example is the effect of malaria mortality when testing the distribution of phenotypes for sickle-cell anemia among adults).
4. The population sample is heterogeneous, being made of socio-economic and geographic strata that do not mate randomly with each other and differ in their gene frequencies.

So the test of deviation from HWE is practically and frequently used for testing or identifying genotype analysis problems (Xu *et al.* 2002, Lewis 2002). In this thesis, the population genetics theory, the HWE, has been applied in the initial analysis of pharmacogenetic data in order to have a clear picture of diversity of the genetic polymorphisms of *CYP2C9&2C19* between populations.

Ethnicity & Race

In population genetics, any analysis begins with the definition of population. As described earlier, there is variation of certain phenotypes or genotypes between defined populations, such as Caucasian, Japanese or Chinese. It is clear that ethnicity or race classifications are broadly used in scientific research, especially in medicine or biomedical research. Consequently, the racial or ethnic differences, which are important for scientists to understand better the variations in the prevalence and severity of diseases and in the response to treatments, often bring social and political issues into inequality of health care (Pfeffer 1998, Burchard *et al.* 2003). However, it is necessary to give some information about the terms "Ethnicity", "Race" and "Population" that are frequently used in the current project, especially as these terms are often used interchangeably in the relevant publications.

In ecology, a population is defined as a group of organisms, all of the same species, which occupies a given area or ecosystem. In genetics, a population refers the Mendelian

population, which is defined as those individuals generally randomly mating within a population. The Mendelian population is very hard to define in reality as all populations or population clusters are derived by considering that single genes are overlapping, and in almost all populations, all alleles are present but at different frequencies (Cavalli-Sforza *et al.* 1996). As a result, in the study of population genetic structure of variable drug response (Wilson *et al.* 2001), genetic cluster analysis is only powerful to separate the individuals whose ancestors diverged many millennia ago; there are many debates about whether the Ethnicity/Race classifications have a biological basis (Mcleod and Evans 2001, Goldstein *et al.* 2003a), however these terms are still useful in biomedical research particularly at present when there is only limited knowledge about the genetic contribution to variations in complex traits available (Mountain and Risch 2004).

Ethnicity is the identity in an ethnic group where the individual share same ancestry, history, and cultural heritage. Race is a sensitive word, which is avoided in society, although the study of human race dates to antiquity. The classification of race is primarily based on superficial physical characteristics such as skin colour and facial features, for instance, Caucasian, Mongolian, Ethiopian, American, and Malay, as stated by J. F. Blumenbach (1724-1804) who is considered as the father of physical anthropology.

In medical and biomedical research, the ethnic/racial differences are widely considered. Dating back to 1988, the Food and Drug Administration of United States (FDA) published guidelines and regulations to emphasize the importance of including analysis of demographic subset data in new drug applications, which promoted the collection of race related data during research and even recommended the analysis of the data for race effects (http://www.fda.gov/cber/gdlns/racethclin.htm).

Through the selection of references in the project, many studies have been found sampling predefined populations that are usually based on the culture or geographical characteristics of recruited subjects. While there are many systems, geographic, anthropological, linguistic and ethnographic theorem, applied in the categorisation of human variation, the phylogenetic tree of world populations that was established with DNA markers by Cavalli-Sforza *et al.* in 1994, which took into account many aspects of those theories systematically, has not been found to be in disagreement with other current molecular genetic markers, such as micosatellites or SNPs. The population clusters studied in the thesis follow the phylogenetic tree principally, which is presented in fig 1.1.1 here.

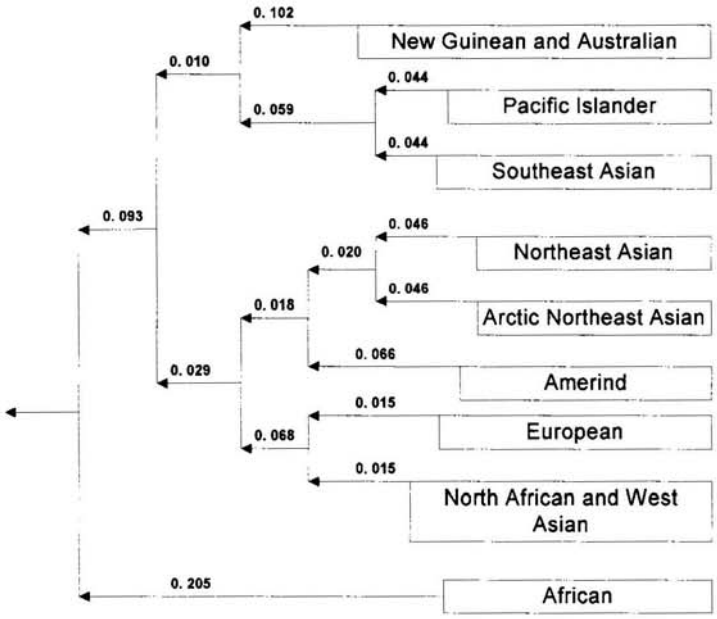


Figure 1.1.1: phylogenetic tree of major populations in the world (Cavalli-Sforza *et al.* 1994) (The number on each arrow means the genetic distance, which is calculated from the results of polymorphic DNA markers)

The types of Genetic Polymorphism in populations

Apart from monozygotic twins, in pharmacogenetic research, the natural populations always show diverse characteristics, either the physical appearance/status of individuals (skin colour, eye and hair colours, height, weight, body conformation and so on) or various responses to medical treatment (such as extensive metabolizer of given drug). Most of the traits show phenotypic diversity, which are observable phenotypes, and can be categorized into many phenotypic variations such as morphologic polymorphism, immunologic polymorphism and protein polymorphism.

Furthermore, a large amount of phenotype diversity is caused by the difference in molecular genetics among the individuals of the population. Genetic polymorphism exists in different forms when the DNA molecule is presented at high resolution or lower resolution; the polymorphism can be demonstrated from individual base pairs to the larger restriction fragments or amplified fragments of DNA. In the studies of protein polymorphism, the study of the molecular basis of variation has been carried down to the level of the polypeptides encoded by the genes, which is often presented as the variation of single base or SNP (figure 1.1.2).

Not all of the polymorphisms essentially cause or induce polymorphism in phenotypes. For instance the synonymous polymorphism is a single-nucleotide polymorphism in a coding region that alters the amino acid codon; but the codon produced by this SNP is a

synonymous codon that does not result in different amino acid replacement in the consequent protein product, consequently the function of the protein would not be altered in the person who carries this synonymous SNP.

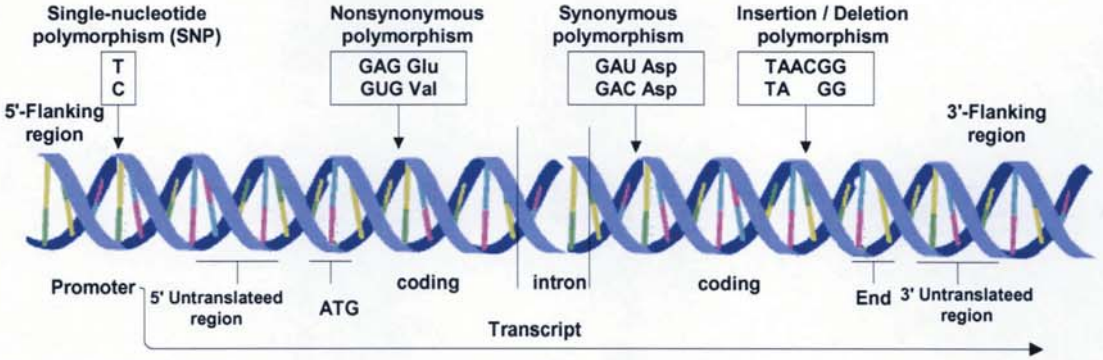


Figure 1.1.2: The DNA molecular polymorphisms at high resolution

Association of genetic variation and polymorphism phenotype

In population genetics, many complex mathematical models are applied to explain and analyse the association between alleles or genes and related phenotypes, such as the Quantitative Genes model, the Maximum likelihood estimation, and the Bayes theorem (Durbin 1998). The frequency of a genetically related phenotype, such as poor drug metabolizer, or disease, depends on the relevant allele and genotype frequency.

For instance, if a phenotype trait is encoded by a single diallelic locus, the phenotype distribution curve is shown in figure 1.1.3 with the phenotype frequency for the known relevant genotypes AA, Aa, and aa. In graph a and b, the phenotype difference among three genotypes is relatively small, and particularly in graph b, the genetic contribution to the phenotype is minor. In population genetics both of them could be categorised as quantitative genes and require certain mathematical principles to analyse (Weiss 1993). But in graph c and d, the phenotypes according to each genotype are obviously discrete variables, and the genetic contribution to phenotype is clearly separated by observed population data. In graph d, the phenotype distribution in the population is obviously bimodal, which is like most of biomedical variables, such as normal and affected, or positive and negative. Therefore the association of gene and phenotype can be estimated by classifying the population into discrete phenotypes, such as normal and affected, and ignoring the quantitative variability within each category (Weiss 1993). Even without complex mathematical application in the candidate gene approached of pharmacogenetic research, the consequent result may still be useful for improvement of clinical safety or

efficacy of medical treatment. For instance the dosage or regime could be adjusted if the patient has mutated alleles of metabolism enzymes that would categorise himself as a poor metabolizer relevant of certain drug.

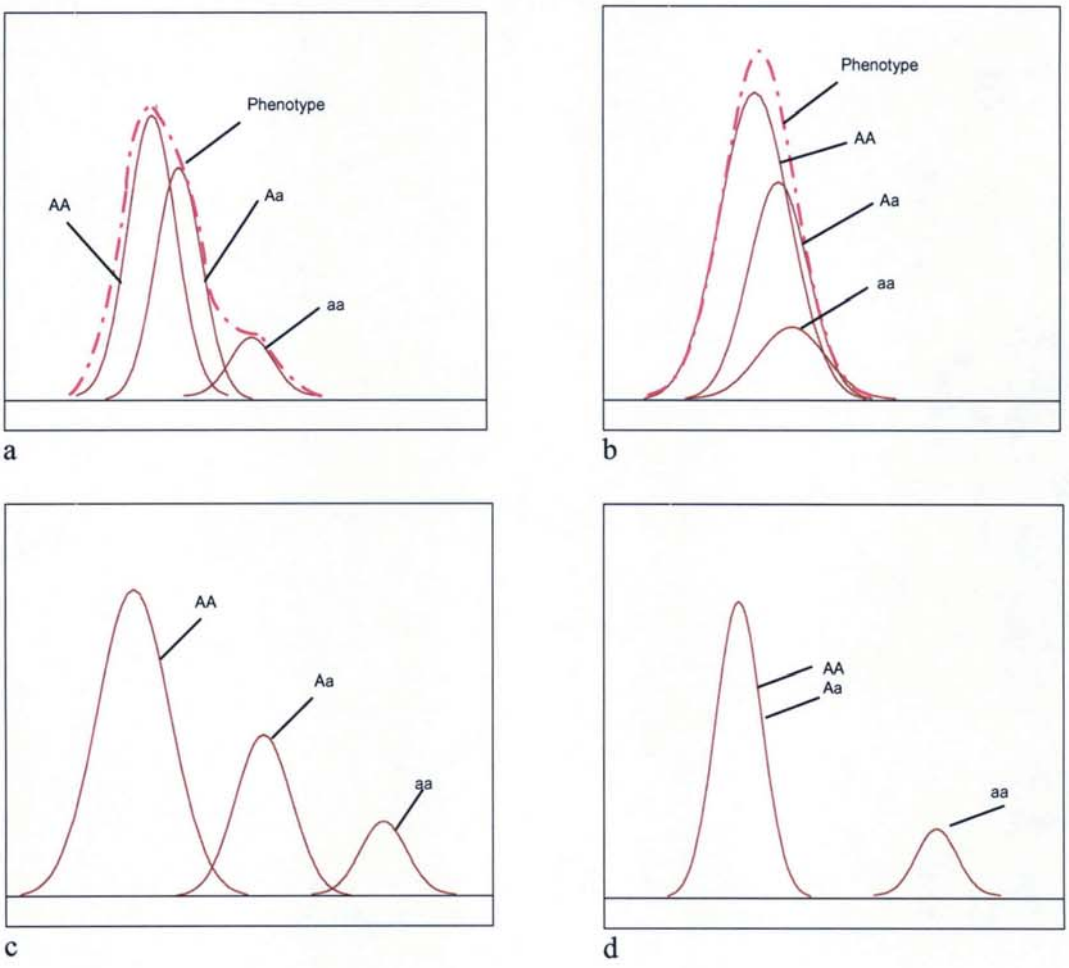


Figure 1.1.3: Illustration of different phenotype distributions related to a single diallelic locus. (The abscissa in each graph represents the phenotype variable, for instance plasma concentration, the ordinate present the frequency of phenotype. The genotype of each curve is indicated by label beside.)

1.2 Pharmacogenetics

Background

Pharmacogenetics is not a new term in the health sciences field. However, pharmacogenetics has become one of the most popular biomedical sciences, which has been invigorated with current advances in genomics. Motulsky (1957) first referred to it when he stated that genetic variation in enzyme activity could be associated with certain adverse drug reactions. Today, pharmacogenetics is defined as the scientific study of inherited factors that affect human drug response. It aims to improve drug design and possibly tailor medicines according to a person's unique genetic make-up; thereby making treatment safer and more effective. (http://www.nigms.nih.gov/pharmacogenetics).

Pharmacogenomics is often used interchangeably with pharmacogenetics. However, it is generally accepted that Pharmacogenomics encompasses a broader meaning (Bailey *et al.* 1998); implying an interaction between many genes or an entire expressed genome and drug action, whilst pharmacogenetics entails the study of genetic variation responsible for an individual's response to drug therapy. With the shift of focus from Mendelian examples to more complex modes of genetic causation, these two terms have become less distinguishable from each other (Goldstein *et al.* 2003a, b). However, it is not within the scope of this thesis to discuss the differences between the two terms. For purpose of this thesis, the term 'pharmacogenetics' will be consistently used to reduce any misunderstanding or confusion.

Prospects of Pharmacogenetics

When prescribing drugs under appropriate regimes, physicians can always find variations of the drug response among individual patients with similar syndromes, which are revealed as the difference in either efficacy or safety of given drugs. It is well known that drug response is dependent on various factors, both endogenous (age, gender, physiological status, genetics etc.); and exogenous (diet, environment etc.). However, only the genetic aspect remains stable throughout a person's lifetime. Pharmacogenetics can be seen as the study of the genotype of any proteins involved in drug actions, thus including drug-metabolizing enzymes, drug transporters and drug targets. With the continuous rise of information about human genomics and the latest biotechnology, pharmacogenetics could

potentially bring clinicians and scientists together in order to practice and achieve personalised medicine-'the right medicine, for individual patients, at the right dose under right regime' (Figure 1.2.1: the prospect of pharmacogenetics research).

Figure 1.2.1: The prospect of Pharmacogenetic research (Johnson J. A. 2003)

Despite current knowledge concerning the association of human genetic variation and polymorphic drug response, there is still a long way to go to realise personalised medicine. Nevertheless, pharmacogenetics has already identified many clinically relevant genetic polymorphisms (Evans and Relling 1999).

History of Pharmacogenetic research

Variability in human drug responses, especially the variation of metabolism has been widely studied since the early 1950s. A well-known example is the slow and rapid acetylation of the antituberculous drug isoniazid (INH) first described in 1960. Evans *et al.* studied INH blood concentration of 267 subjects 6 hours after receiving a standard dose of 9.8mg per kg. The histogram showing the frequency distribution of the data exhibited bimodality, which led to the later definition of the terms, poor metabolizers and extensive metabolizers as illustrated in figure 1.2.2.

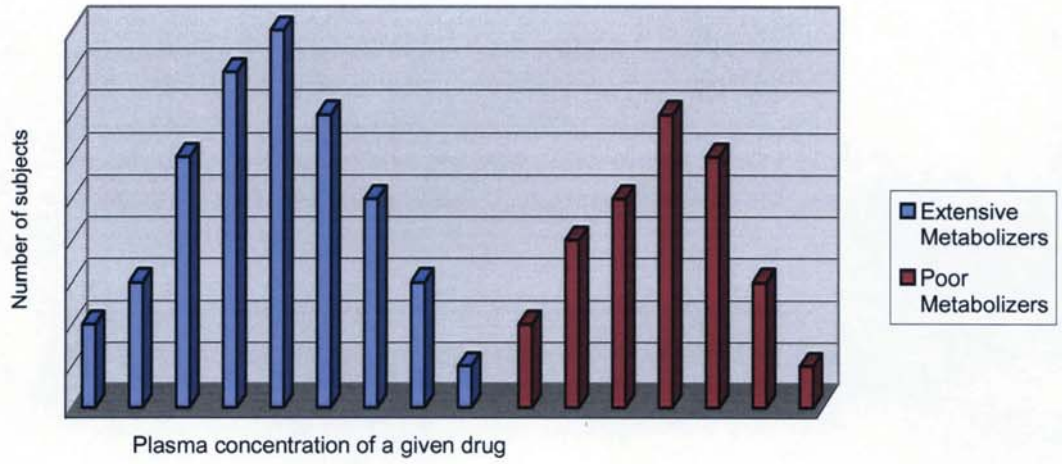


Figure 1.2.2: Illustration of earlier clinical observation studies: the bimodal frequency distribution of plasma concentrations of a given drug among a population of individuals who received the same dosage.

The major metabolic pathway of INH is the acetylation to acetylisoniazid by the N-acetyltransferase present in the liver and the small intestine. The patients phenotyped as slow acetylators normally have higher blood lever of INH and longer elimination half-lives, therefore appear as being at increased risk of toxic reactions resulting from higher drug plasma concentration, such as peripheral neuropathy (Evans *et al.* 1960, Weber and Hein 1979, Kalow 1982). In 1973, the World Health Organization (WHO) first reported that genetic variations of acetyltransferase were responsible for the isoniazid adverse effects resulting from differing acetylation rates (WHO technical report series NO. 524, 1973). Pharmacogenetic factors influencing drug metabolism were considered more widely from 1972 onwards, when Vesell and his colleagues demonstrated that identical twins showed higher similarity in plasma half-lives of numerous drugs when compared to fraternal twins (Vesell 1972).

Besides the concentration or half-lives of drugs, *i.e.* pharmacokinetic characteristics, early clinical observation studies also indicated polymorphic pharmacodynamic features. Specifically, the efficacy and safety of drugs, such as coumarin, anticoagulants (e.g. warfarin) and the antimalarial drug ((e.g. primaquine) was shown to vary between individuals (La Du 1972). Therefore, a regular dosage regime, which is defined as the therapeutic range from pharmacodynamic research, cannot have a similar clinical outcome among individual patients of different genotypes. As illustrated in the dose-effect curve shown in figure 1.2.3, most individuals would have an optimal effect when the prescribed

drug was given within the therapeutic range, and show less efficacy or even toxicity when the drug dose was lower or higher than the therapeutic range; however, others that have different genotypes may require lower dose or higher dose for the same expected drug response.

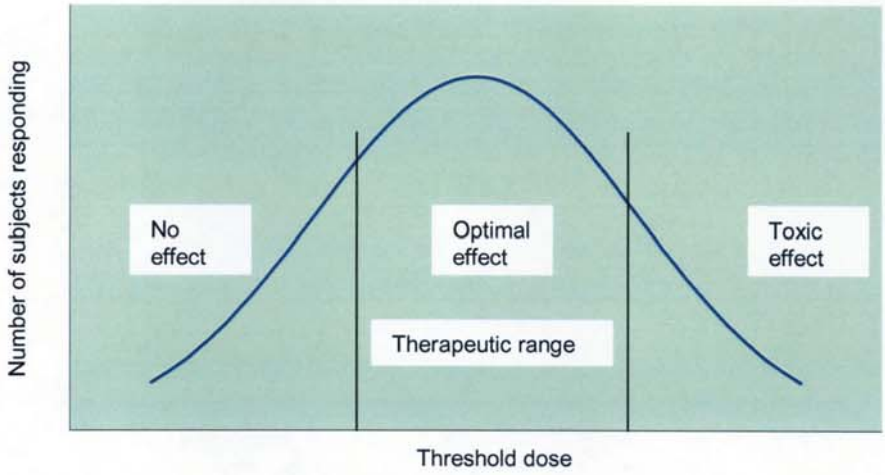


Figure 1.2.3: Illustration of clinical outcome studies: the frequency distribution of individuals clinical response under or over the threshold dose compared with the regular therapeutic range of a given drug

Current Pharmacogenetic research

Pharmacogenetics began with the observations of variable drug metabolism. Most of the pharmacogenetic traits were identified many years ago as monogenic inherited traits, such as glucose-6-phosphate dehydrogenase (G6PD) deficiency, acetyltransferase polymorphism, and methemoglobin reductase deficiency (WHO technical report, 1972). However, for many drugs, a simple monogenic trait did not fit the clinical observations of drug response, therefore other complex models were investigated. Currently, polygenic models of drug response are gaining more attention than ever. The pharmacogenetic factors, which have to be considered concurrently in polymorphism studies of drug responses, can be summarised into genetic polymorphism related to drug disposition, genetic polymorphism related to drug targets, and other indirect genetic polymorphism related to drug response. Those genes associated with polymorphism of drug responses are frequently called candidate genes (Goldstein *et al.* 2003c, Evans and McLeod 2003) in pharmacogenetic research.

In considering the polymorphism of drug disposition, especially the polymorphism of phase I metabolism enzymes, cytochrome P4502D6 (CYP2D6) is an outstanding example that demonstrates the potential clinical implication of pharmacogenetics and shows how

pharmacogenetic research can proceed from the phenotype to an understanding of molecular mechanisms at the genotype level (Weinshilboum 2003). CYP2D6 is one of the most important enzymes in the cytochrome P450 subfamily II. It is responsible for the metabolism of approximately 20% of all drugs, such as anti-arrhythmics, tricyclic antidepressants, neuroleptics, codeine, and debrisoquine. Mahgoub and colleagues (1977) firstly identified the inherited polymorphism of CYP2D6 by studying the metabolism of debrisoquine; one of Mahgoub's colleagues, Smith and his relatives had suffered severe side effects after taking debrisoquine due to deficient CYP2D6 enzyme activity. With current advances in molecular cloning technology, which enable the sequencing and comparison the protein encoding genes by single base pairs, more than 75 CYP2D6 alleles have been described in various publications (http://www.imm.ki.se/CYPalleles/ accessed on 30th Apr. 2004). These alleles are a series of genetic variants; they include deleted genes, single mutated genes and duplicated or multi-duplicated genes, which are consequently responsible for the absence/lack of the CYP2D6 enzyme, low level of CYP2D6 activity or highly increased CYP2D6 activity as shown in figure 1.1.4 (Ingelman-Sundberg 1999a).

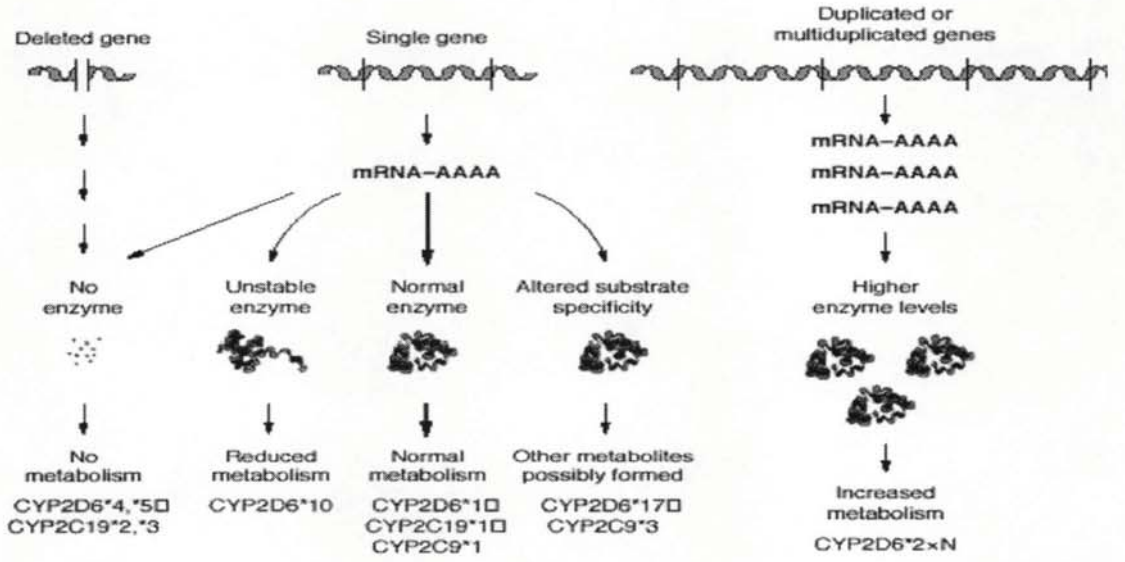


Figure 1.2.4: Genetic variants and relevant phenotypic consequence of *CYP2D6* (Ingelman-Sundberg 1999a)

In addition to the polymorphism of metabolism enzymes, which represents polymorphism with respect to pharmacokinetic factors, there exists similar polymorphism in relation to pharmacodynamic factors, e.g. genes encoding transporters and receptors. For instance, the multiple drug resistant gene 1 (MDR1), which produces P-glycoprotein, was found to be associated with multiple drug resistance to cancer chemo-therapeutic agents (Ambudkar *et al.* 1999). Another example of a mutated gene for a drug target is the angiotensin-converting enzyme (ACE) gene, which affects the response of ACE inhibitors, such as

enalapril, which target ACE in order to reduce blood pressure and left ventricular mass (Kohno *et al.* 1999).

Complexity of genetic polymorphism in Pharmacogenetic research

Since most drug effects are determined by the interplay of several gene products that influence the pharmacokinetics and pharmacodynamics of medications, consideration of polygenic determinants of drug effects has become increasingly important. For instance, CYP3A4 and the P-glycoprotein (MDR1), which is an ATP-dependent drug efflux transporter, are co-expressed in human tissues such as the intestine, liver and kidney, where both have broad substrate specificities. The spatial relationship of P-glycoprotein traversing the plasma membrane and CYP3A4 inside the cell on the endoplasmic reticulum suggests that P-glycoprotein may act to control exposure of substrates to metabolism by CYP3A4 enzyme and their cooperative activity can be a key determinant governing the extent of oral drug bioavailability (Zhang and Benet 2001, Kim 2002). Therefore, both the polymorphisms of CYP3A4 and the MDR1 should be considered for the response of drugs that are dual substrates of the two proteins. This may also inform on which polymorphism is more important in governing clinical effect.

Furthermore, it has to be emphasized here that the concept of candidate genes relevant to pharmacogenetics is broader than pharmacokinetic and pharmacodynamic aspects, which are more often considered as the two major relevant aspects in pharmaceutical field. In pharmacogenetics, it is also important to consider genes that are involved in the underlying disease condition or intermediate phenotypes (Goldstein *et al.* 2003a,c). The polymorphic pathogenesis of disease is an essential component that is tightly associated with the variant drug response. For instance, long QT syndrome is a rare condition in which people have a slower repolarisation of the myocardium after depolarisation. This rare disorder can result in various cardiac pathologies. There are at least five genes associated with this syndrome. A mutation in LQT2 affects potassium channels, while a mutation in LQT3 affects sodium channels (Anantharam *et al.* 2003). What LQT syndromes tell us, is that despite a similar disease phenotype, the large variation in underlying molecular pathology means that LQT symptom cannot be treated by a single drug (Shimizu 2005).

Moreover, studies of pharmacogenetic traits can be classified according to the strategies employed, such as approaches for elucidating known candidate genes and approaches for

elucidating uncertain numbers of polygenic conditions. Results from various research strategies are diverse, and have varying potential value and limitations (McLeod and Evans 2001, Essionx *et al.* 2002). As a candidate gene is normally identified from existing knowledge about a drug's pharmacology and mechanisms of action, results of studies based on candidate genes can be easily linked with clinical outcome. Whereas the results from the polygenes approach, which applies anonymous single-nucleotide polymorphism (SNP) maps to perform genome-wide searches, cannot be easily interpreted without further gene-expression profiles and the aid of proteomic studies. These are still not fully available or characterised by the current knowledge of the human genome.

Difficulties in Pharmacogenetics: genetic facts

Genes essentially encode proteins that play multiple roles in the living organisms. All proteins work in a highly regulated and extremely complex dynamic equilibrium. Variation in genes will contribute to protein variability. Since the worldwide publication of the human genome in 2001, scientists have estimated that there might be fewer than 30,000 protein-coding genes with many non-coding and repeated sequences in the human genome (figure 1.2.5).

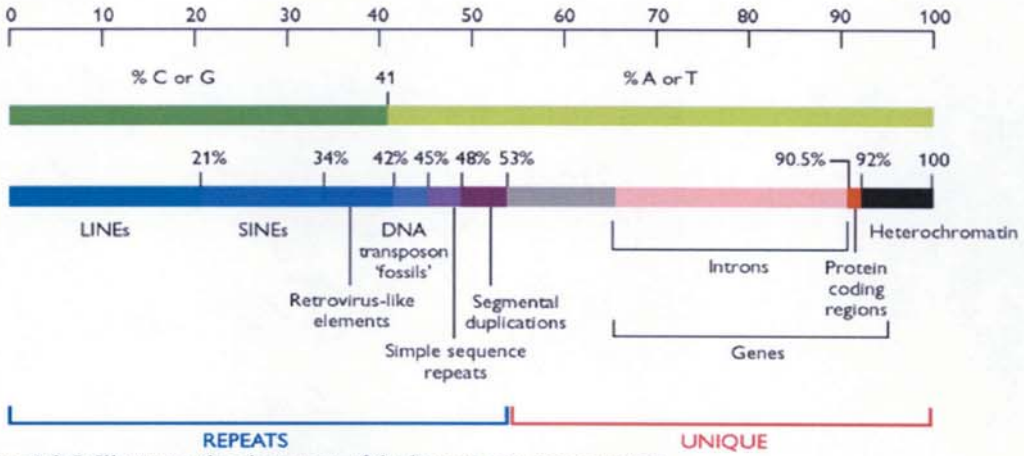


Figure 1.2.5: Illustrates the characters of the human genome sequence
Note: LINEs: long interspersed elements; SINEs: short interspersed elements

However, recent studies have re-estimated this number to approximately be 25,000, with the genome also representing a large number of non-coding RNA molecules. Comparison of the humane genome with other unsophisticated living organisms, such as the nematode (*Caenorhadbitis elegans*), shows that there is merely one-quarter more genes in the human genome, where a nematode has about 20,000 genes (Venter *et al.* 2001, Claverie 2001, 2005). On the other hand, there is an ever-increasing numbers of polyadenylated transcripts with 3'-ends that have been identified by single-pass sequencing of complementary DNA

(cDNA) library, and many of the transcriptions appear to be developmentally regulated not to encode protein (Mattick 2005). This paradox raises more and more questions, 'Are the gene numbers sufficient for humans to perform with such a biological complexity?' 'What role does the large number of non-coding RNAs play in this complexity?' and 'What does the frequent appearance of repeated sequence in the human genome mean in the biological complexity?'

Venter (2001) proposed that combinatorial diversity generated at the levels of protein architecture, transcriptional and translational control, post-translational modification of proteins, or post-transcriptional regulation may be compensating for the fewer than predicted evident number of genes in the human genome. Recently Willingham (2005) proposed a strategy for probing the function of non-coding RNAs (ncRNAs), and found that an ncRNA is a repressor of the nuclear factor of activated T cells (NFAT). Scientists believed that there are potentially thousands of RNA regulators within the large numbers of ncRNAs, which may effectively amplify the complexity of the human genome with a limited number of protein-coding genes by RNA-RNA, RNA-DNA, or RNA-protein interactions, which may explain the long-standing discrepancy between the small numbers of protein-coding genes and the ever-increasing numbers of polyadenylated transcripts. These studies indicate that the posttranscriptional control of eukaryotic gene expression is much more elaborate and extensive than previously thought (Claverie 2005, Moore 2005, Willingham *et al.* 2005).

In 1958, Crick first proposed the central dogma that established how genetic information within DNA molecules was transferred to RNA and expressed as protein. It has been followed in molecular biology research since then. However, the Central Dogma is no longer a sufficient organizational paradigm when numerous RNAs are found to interfere with the expression of genetic information within the DNA or even within proteins. There are many uncertain mechanisms behind the process of gene expression from DNA to protein. Moreover, there are many unknown mechanisms about the modification of synthesized polypeptide chain before the functional and highly folded proteins in three dimensions are actually created.

It may be still too early for scientists to elucidate how the gene sequences manipulate the production of numerous proteins and influence the performance of their function in such a dynamic equilibrium (figure 1.2.6).

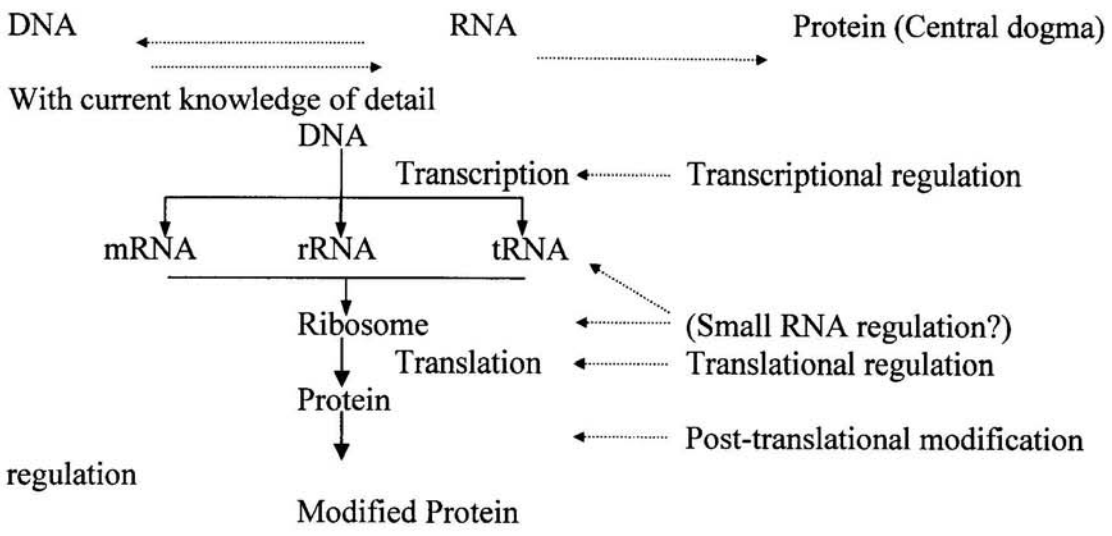


Fig 1.2.6: The known process from gene sequence to protein creation with uncertain regulator networks

Conceptually, polymorphism can have more than one meaning. Defined in rigid terms, polymorphism denotes base-pair substitution at the level of DNA. However, polymorphism can also be expressed at the level of post-transcriptional, modification and onwards through translational process. In summary, despite the completion of human genome sequencing, the dynamic equilibrium from gene sequence into transcribed protein appear to be more complex than previously anticipated. At present, the candidate genes in pharmacogenetic research are defined as the proteins involved in drug disposition and which act as drug targets (Goldstein *et al.* 2003a,c). Thus, the candidate genes approach can only be applicable to the drugs with clearly defined mechanisms of the bioprocesses in human body. However, often there are unknown pathways or factors that may exist for many drugs. The range of candidate genes will be more difficult to estimate if the possible unknown effects on transcriptional control, and post-transcriptional regulation are taken into account.

Data from human genome research has shown that any two unrelated individuals are more than 99.9% identical in their DNA sequences, with the remaining 0.1% sequences contributing to the differences observed amongst individuals, such as the physical appearance, the predisposition to disease and the variable response to drugs. From a genome-wide examination of sequence variation, about 10 million single nucleotide polymorphisms (SNPs) have been identified, with a rate of ~1 SNP per 1200 to 1500 base pair (bp). Moreover, 75% of the SNPs are located at the intergenic region. The SNPs in gene coding regions are either frequently silent or missense, whereas only a very small

proportion (<1%) potentially leads to dysfunctional alterations in proteins (Venter *et al.* 2001, Winter *et al.* 1998). Therefore many of the intergenic SNPs are likely to have a regulatory function.

The SNP studies can establish a range of haplotypes; a set of associated SNPs that are located closely together on a chromosome with strong linkage disequilibrium and are inherited together as a unit. As a result, a recent international haplotype map project, which has been initiated the research of associated SNPs in individuals of different ethnic or geographic origin, will provide more valuable information about how the mere genetic variation contribute to the magnitude of human diversity, especially for the variable risk of disease and drug response. (http://www.hapmap.org/)

Therefore, with the advance of the haplotype map, the polygenes approach in pharmacogenetic studies, which applies anonymous single-nucleotide polymorphism maps to perform genome-wide searches, can provide further important clarification for pharmacogenetic traits when the candidate genes are unknown. For instance, Kim (2001) reported the presence of multiple SNPs in the MDR1 genes, which suggested that consideration of haplotypes will be more important than single polymorphism in future studies.

The necessity for systematic review in pharmacogenetics

Publication of pharmacogenetics research has been increasing steadily over the past three decades, with a particularly marked rise since the 1990s. Pharmacogenetic research has mainly focused on the frequency of particular genes or genotypes, or on outcome/phenotype associations, and has primarily applied the candidate gene approach. As previously mentioned, the most well studied pharmacogenetic variations are those relevant to drug metabolism enzymes (DMEs). Nearly one-third of the polymorphisms in candidate gene studies are on DMEs (Goldstein *et al.* 2003c), where polymorphism of the cytochrome P450 families apparently result from the significant different allele frequencies among people from different ethnic or geographic origin (Bertilsson and Kalow 2001).

Although there are large numbers of publications on pharmacogenetics research, only a few have applied a systematic review or a meta-analysis to synthesise the results from polymorphism studies (Xie *et al.* 1999, Lee *et al.* 2002).

Meanwhile, most pharmacogenetic studies have employed distinctive study design methodologies according to the purpose of the research. Therefore the data provided by those studies vary extensively. For instance, in the Pharmacogenetics and Pharmacogenomics Knowledge base (PharmGKB), data are grouped according to genotype and phenotype. In terms of the phenotype, four different categories are used, clinical outcome, pharmacodynamics & drug response, pharmacokinetics, and molecular & cellular function assays, and presents a very clear picture of the current available pharmacogenetic research (http://www.pharmgkb.org/index.jsp).

For a clinical physician, the most practical information would be the relevance of a polymorphism to certain disease treatment or drug application, which could be useful for enhancement of medical treatment or personalized medicine. However, few genome or genome relevant databases exist. Whilst Online Mendelian Inheritance in Man (OMIM), UniGene, Nucleotide database offer enormous information on genes and genetic polymorphism, none of them is compatible with the requirements for pharmacogenetic research yet, and lack a user-friendly interface. Therefore, in this thesis, methods in evidence-based medicine are initially applied in the studies relevant to the medical treatment of epilepsy in order to produce a broad overview of the pharmacogenetic aspects of antiepileptic drugs. Following on from this, the thesis focuses on two genetic polymorphisms of drug metabolism enzymes, CYP2C9/2C19, that have comprehensive pharmacogenetic evidence available to further study the general implications of pharmacogenetics on drug treatment.

1.3 Genetic polymorphism of cytochrome P450

General information about CYP450s

The cytochrome P450s are a superfamily of enzymes that are found in all forms of living organisms. They are responsible for the metabolism of many endogenous compounds, and participated in the activation and deactivation of many carcinogens/xenobiotics, considered as major phase I enzymes in the drug metabolism system. There are 18 families, 41 subfamilies and 57 proteins in cytochrome P450s (Ortiz de Montellano and De Voss 2004, Nebert and Russell 2002). Ninety five percent of drug metabolism is relevant to cytochrome P450s (Rendic 2002). Bertz and his colleague had found that 56% of 315 drugs are predominantly metabolized by CYP450s (Bertz and Granneman 1997). Therefore, the cytochrome P450 families have received intensive interest from pharmacists and biomedical scientists since they were discovered.

CYP450s history

Axelrod and Brodie discovered the first experimental evidence relating to the existence of cytochrome P450 in 1955 and located the enzyme system in the endoplasmic reticulum of the liver (Brodie *et al.* 1955). In 1958, Garfinkel *et al.* detected a carbon monoxide (CO) binding pigment in liver microsomes that had an absorption spectrum (Garfinkel and Klingenberd 1958). In 1962, Omura and Sato coined the pigment as Cytochrome P450 as its absorption peak at 450nm (Omura and Sato 1962). With the feasibility of various new instruments and in *vitro* technology in 1950s, there appeared reports of the molecular mechanism of action of CYP450s in drug metabolism (for reviews see Estabrook, 2003).

As individual CYP450s were identified in various laboratories, several diverse CYP450 nomenclatures were adopted. In 1987, Nerbet and his colleagues developed the nomenclature system based on the amino acid sequence comparison of cytochrome P450. Several updates of the new CYP450 gene super-family have been published since. According to the nomenclature, the cytochrome P450s are classified into families and subfamilies according to the homology of their amino acid sequences. Families are based on at least a 40% sequence homology. Members of each subfamily must share at least at 55% identity (Nebert *et al.* 1987a, 1989, 1991). This nomenclature system is founded on the evolution of CYP450s and their structure relationships. Many of the cytochrome P450s

in same family have significantly different substrate groups and physiological functions (Guengerich 2004).

Biological characters of CYP450s

Cytochrome P450s are expressed primarily in the endoplasmic reticulum of human liver and only 6 are located in the mitochondria. In the liver microsomes, families 1, 2 and 3 appear to account for 70% of total amount of CYP450s. Among these, CYP3A accounts for 30% and CYP2 accounts for 20% (Shimada *et al.* 1994). Based on the substrate classes, the cytochrome P450 family 2 and 3 are responsible for most metabolism of xenobiotics (Guengerich *et al.* 2005). Therefore, the polymorphism of these two cytochrome P450 families would be highly relevant to clinical outcomes in terms of drug metabolism. Currently, individual forms of CYP450s have been found in a variety of extrahepatic tissues, such as kidney, lung, and small intestine, and brain. It is noted that the uneven distribution of CYPs in human brain regions, may create a micro-environment within the central nervous system, which would explain the inter-individual variability in response to centrally acting drugs as well as the risk for neurological diseases and pathogens (McFadyen *et al.* 1998, Gervasini *et al.* 2004, Guengerich 2004, Miksys and Tyndale 2004).

Through the adverse drug reaction (ADR) studies published between Jan 1995 and Jun 2000, Phillips and his colleagues found that 59% of the drugs cited in ADR studies are metabolized by at least one enzyme with a variant allele known to cause poor metabolism. Their results also demonstrated that the genetic variability of drug metabolizing enzymes is likely to be an important contributor to the incidence of ADRs (Phillips *et al.* 2001).

Since cytochrome P450s take part in the metabolism of varied endogenous compounds, such as steroids, this super-family has been considered as a cancer susceptibility factor especially in certain hormone sensitive cancers (Nebert *et al.* 1996, Brockmoller *et al.* 2000)

Cytochrome p450s have substrate specificities, but multiple or alternative pathways exist in CYP450s involved in the metabolism of any given drug (Hasler 1999, Nebert and Russell 2002, Ortiz de Montellano and De Voss 2004). Cytochrome P450s may also be induced or inhibited by their own substrate or substrates of other CYP450s (Hasler 1999). In drug metabolism, the cytochrome P450s work together with other proteins, for instance P-glycoprotein was found associated with metabolism that involved CYP3A4. Therefore,

the CYP450 super-family is important in drug metabolism, responsible for many variant drug responses, but should not be studied alone without considering other protein components that are involved in the drug metabolism pathway (Kim 2002).

Genetic aspects of CYP450s

In the early 21st century, a new era of CYP450s research work was revealed with the development of genetic technologies, such as genome sequences, genetic databases, and SNPs microarrays. There are more than 2,500 cytochrome P450 genetic sequences available now. Among them, the human has 57 CYP genes and 47 pseudogenes published (http://www.imm.ki.se/CYPalleles/ accessed on 30th Apr. 2004). Human Cytochrome P450s are approximately 500 amino acid long proteins. The genes encoding those amino acids are normally under complex and distinct control during development, either following exposure of organisms to various foreign compounds or in response to important endogenous signals (Nebert and Gonzalez 1987b, Lewis *et al.* 1998). In clinics, it has been reported that same cytochrome P450s have a polymorphic impact on drug metabolism from person to person, in addition to the genetics polymorphism of certain cytochrome P450s.

As pharmacogenetics has expanded into many fields, published literature now covers the genetic of a variety of molecules associated with drug disposition and drug action. A variety of factors, from drug metabolizing enzymes, drug transporters and drug receptors, contribute to the genetic polymorphism in drug response (Evans and Relling 1999, Johnson and Evans 2002, Pirmonhamed and Park 2001). However, the knowledge of inter-individual differences in drug transporters and drug receptors are still lacking compared with the extensive available information of the polymorphism of drug metabolizing enzymes, especially the polymorphism of Cytochrome P450 family (Tanaka 1998a, 1999a, Nebert and Russell 2002, Johnson and Evans 2002, Ingelman-sundberg 2004, van der Weide and Steijns 1999).

The cytochrome P450 family 1, 2 and 3 are responsible for the metabolism of most drugs. All genes encoding these families are found to be polymorphic (Nebert and Russell 2002). The variant genes are variably distributed among different ethnic groups (Ingelman-sundberg 2004, Kim *et al.* 2004, Solus *et al.* 2004). The enzymes in the three families play significantly different roles in drug metabolism with different pharmacological

mechanisms. The relative importance of the polymorphisms varies according to drug response studies (Ingelman-sundberg 2004, Ma *et al.* 2004). In a genetic variation study of 11 cytochrome P450 enzymes, the members in CYP2 subfamily show the highest level of genetic diversity in the exons and flanking intronic regions (Solus *et al.* 2004).

Prospects of CYP450s pharmacogenetic research

Genetic polymorphism means there is at least a 1% frequency of a gene variant variable existing in a population (Meyer 1991). When 1% of the population had the variant allele/mutant, there will be 1% of the population who may have different drug responses from the remaining population. Under same medication regime, 1% of the population either may have less desired therapeutic effects (efficacy), or more undesired adverse effects, even a high risk of toxicity. This is especially important when the medication has relative narrow therapeutic window, such as, phenytoin, tolbutamide and Warfarin. Consequently, there are more potentially uncontrolled seizures, elevated blood sugar, and bleeding issues. In pharmacogenetics, these actually represent two aspects of the polymorphism of cytochrome P450, the polymorphism of genotype and the polymorphism of phenotype.

Before the polymerase chain reaction (PCR) technology was developed, several different approaches were applied for polymorphic studies of cytochrome P450, in s*ilico*, *in vitro*, *in vivo* (animal, human) (Rodrigues and Rushmore 2002, Smith *et al.* 1998). All of them aimed to define the important factors that govern the significance of cytochrome P450 polymorphisms. However, the outcomes of these studies are actually phenotypes, in terms of pharmacokinetic characters or parameters of certain drugs or probe drugs. In pharmacogenetic studies, those approaches are still applicable. Furthermore, with available genotypes or alleles, it is possible to obtain the association of genotypes and phenotypes. Based on the association result, patients with known genotypes of cytochrome P450s could be prescribed an adjusted dose or regime before they actually experience poor control of disease, or serious adverse reactions or drug interactions due to genetic polymorphism of cytochrome P450s.

Since the genetic polymorphism of *CYP2D6* have been studied thoroughly, there are already some clinical diagnostic chips available in the market, such as the AmpliChip CYP450 test from Roche, which can predict the possible drug response relevant to genetic polymorphism of *CYP2D6* and can be used effectively to prevent associated adverse

reaction or drug interaction due to polymorphism of *CYP2D6*. In this thesis, the available studies that are related with the genetic polymorphism of *CYP2C9/2C19* and the variable response to phenytoin or probe drugs mephenytoin and warfarin are selected from PubMed database and applied with further analysis in order to evaluate the phenotype and genotype association.

Genetic features of CYP2C: 2C9 and 2C19

Cytochrome P450 proteins have been categorized into families and subfamilies by their sequence similarities. Humans have 18 families of cytochrome P450 genes and 43 subfamilies. Now, there are more than 2500 cytochrome P450 sequences published in relevant database (http://www.imm.ki.se/CYPalleles/ accessed on 30th Apr. 2004). As previously mentioned, CYP2C9, CYP2C19, CYP3A4 play a major role in the metabolism of antiepileptic drugs among those cytochrome P450 enzymes. In this thesis, the genotypes of *CYP2C9*, *CYP2C19* that are responsible for majority of phenytoin metabolism have been the genes under study.

CYP2C9 and *CYP2C19* are located in the CYP2C gene cluster on chromosome 10; the LocusLink cytogenetic band is 10q24.1-q24.3 and the Ensembl cytogenetic band is 10q23.33 (figure 1.3.1) (Gray *et al.* 1995).

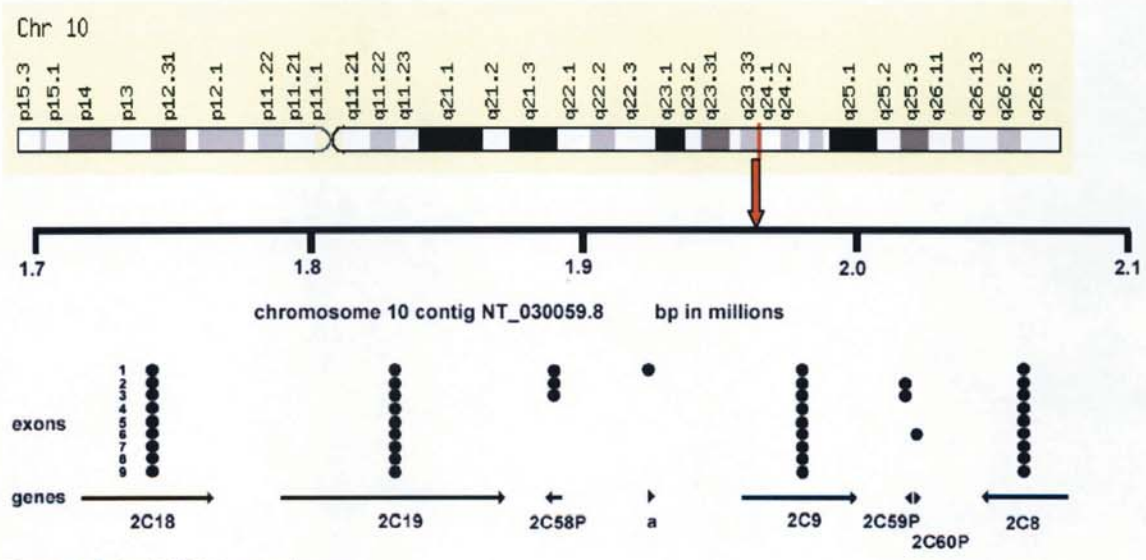


Figure 1.3.1: CYP2C gene cluster

As in figure 1.3.1, both of the *CYP2C9, CYP2C19* have nine exons. Meanwhile they encode CYP2C9 or CYP2C19 enzyme, each of which has 490 amino acids. The few well-

known mutant alleles of *CYP2C9, CYP2C19* consequently produce enzymes with decreased activity or without activity at all (table 1.3.1).

Table 1.3.1: Published descriptions of common mutated *CYP2C9, CYP2C19* alleles

Allele nomenclature	Nucleotide change in cDNA [a]	Other form used before		Amino acid change	Phenotype	Reference
		Nucleotide or substitutions	Trivial name			
*CYP2C19*2A*	681G>A	G-A	m1, m1A	N/A*	Mephenytoin hydroxylase, poor metabolism	de Morais SM *et al.*, 1994a
*CYP2C19*2B*	276G>C 681G>A	GAGg-GAC G-A	m1B	Glu92Asp	Mephenytoin hydroxylase, poor metabolism	Ibeanu GC *et al.*, 1998b
*CYP2C19*3*	636G>A	TGGa-TGA	m2	Trp212Term	Mephenytoin hydroxylase, poor metabolism	de Morais SM *et al.*, 1994b
*CYP2C19*4*	911A>G	aATG-GTG	m3	Met1Val	Mephenytoin hydroxylase, poor metabolism	Ferguson RJ *et al.*, 1998
*CYP2C19*5A*	1297C>T	aCGG-TGG	m4	Arg433Trp	Mephenytoin hydroxylase, poor metabolism	Ibeanu GC *et al.*, 1998a
*CYP2C19*6*	395G>A	CGG-CAG	m5	Arg132Gln	Mephenytoin hydroxylase, poor metabolism	Ibeanu GC *et al.*, 1998b
*CYP2C19*7*	IVS5+2T>A	T-A	N/A	N/A*	Mephenytoin hydroxylase, poor metabolism	Ibeanu GC *et al.*, 1999
*CYP2C9*2*	430C>T	cCGT-TGT	N/A	Arg144Cys	Poor phenytoin metabolism	Aynacioglu AS *et al.*, 1999
*CYP2C9*3*	1075A>C	cATT-CTT	N/A	Ile359Leu	Poor tolbutamide metabolism	Sullivan-klose TH *et al.*, 1996
*CYP2C9*5*	1080C>G	GACc-GAG	N/A	Asp360Glu	Poor substrate metabolism	Dickmann LJ *et al.*, 2001
*CYP2C9*6*	818delA	AAAATG^GA GAaGGAAAAGCAC	N/A	N/A*	Phenytoin toxicity	Kidd RS *et al.*, 2001

Note: a standard form for mutated alleles according to nomenclature system.
N/A stands for not applicable. * splicing defect.

Before the nomenclature system of cytochrome P450 families was established, there were many alternative names generally used in publications. As presented in table 1.3.1, those alternative names often appeared in the articles selected for the thesis. Before 1995, the mutants had first been named as 'm' after the gene names, in the sequence of discovery, for example m1, m2 and so on. According to the Human Cytochrome P450 Allele Nomenclature Community, currently the genes of cytochrome P450 are displayed as the

same name as the corresponding protein but in italics and the wild type of allele is named as allele 1 with star* in front, consequently m1 as allele *2. For instance, *CYP2C9* wild type allele is named to be *CYP2C9*1*, and *CYP2C9m1* is nominated as *CYP2C9*2*. Detail of this nomenclature can be found in the Human Cytochrome P450 Allele Nomenclature Community website (http://drnelson.utmem.edu/Nomenclature.html).

By 3rd May 2005, there were already 20 and 16 alleles identified for *CYP2C9/2C19* separately each with a different genetic mutation. (http://www.imm.ki.se/CYPalleles/)
Among the identified *CYP2C9/2C19* genetic variations, the alleles, **1*, **2* and **3*, exhibit significant polymorphism, and contribute to major inter-individual and intra-population variations in the response to medication by consequent deficient or inactive enzyme functions (Ingelman-Sundberg *et al.* 1999b, Ma *et al.* 2002, Pirmohamed and Park 2003). Thus, the current thesis focuses on the polymorphism studies of *CYP2C9*1*, **2*, **3* and *CYP2C19*1*, **2*, **3*.

Furthermore, early family or twin studies show that cytochrome P450 genetic polymorphisms are autosomally inherited in populations with the phenotypes of poor metabolizers (PM) and extensive metabolizers (EM), while further classification also has identified ultra-extensive metabolizers for some CYPs, such as CYP2D6 (Nebert *et al.* 1987a,b, 1999). This has been confirmed as monogenetic trait by the latest biochemistry and molecular studies of genetic polymorphism of Cytochrome P450 (Meyer 1990, 1994a,b, 1997; Goldstein and de Morais 1994a, Ingelman-Sundberg 2004). In most references about *CYP2C19* polymorphisms, the subject without *CYP2C19*1* allele was frequently denoted as poor metabolizer (**PM**), whose genotype is either *CYP2C19*2/*2* or *CYP2C19*2/*3* or *CYP2C19*3/*3*. Although those subjects without *CYP2C9*1* allele are not named as **PM**, they frequently present with deficient CYP2C9 enzyme activity in the metabolism of drugs, such as phenytoin, or tolbutamide.

While varied prevalence of *CYP2C9, CYP2C19* polymorphic alleles was found in different ethnic groups with relevant polymorphism of drug metabolism, it was believed that the genetic polymorphism of CYP2C9, CYP2C19 enzyme had potentially important impacts on the metabolic features of many drugs, especially when those cytochrome P450 enzymes were responsible for major metabolic pathways of the drug. Meanwhile, the polymorphic

drug response among different ethnic populations is considered as the subsequent clinical outcome of dissimilar allele prevalence among those people with different ethnic origin.

Although many factors may affect the expression of a given genotype, which can include epistatic genes, modifiers of the gene-protein translational process, and an adapting effect of the environment, as previously mentioned, there are two distinguished phenotype traits associated with drug metabolism, poor metabolizer and extensive metabolizer, which present in clinical medical treatment when the variable drug response is associated to polymorphic drug-metabolism enzymes (figure 1.2.2). In this thesis, pharmacogenetic polymorphism of anti-epileptic drugs (AEDs) response was focused. Since the few AEDs are mainly metabolized by cytochrome P450s, and variable response is significantly related with deficient cytochrome P450s enzymes. It would be valuable to see whether the variable genes of cytochrome P450s could be helpful in the prediction of the variable drug response in epilepsy pharmacotherapy and pre-diagnose patients as poor metabolizer or extensive metabolizer before setting the medication regime. Therefore it would be possible to prevent any low/reduced efficacy or high toxicity due to the impact of mutated cytochrome P450 genotypes.

1.4 Epilepsy and anti-epileptics

General introduction about epilepsy

Epilepsy is a common neurological disorder that affects people in every country throughout the world. It can be identified as far back as ancient Indian medicine records during the Vedic period of 4500-1500BC. Epilepsy was firstly described as *'apasmara'*, which means *'loss of consciousness'*. At present, epilepsy is defined as a condition of recurrent spontaneous seizures by International League Against Epilepsy (ILAE), which was founded in 1909 and became a professional and international organization with over 90 divisions covering about 60 countries from every continent by the end of 20th century. Epilepsy is characterized by repeated seizures, and is the result of bursting and excessive electrical activity in the brain. However, the fundamental mechanisms of epilepsy remain unknown. The known causes are categorized into genetic predisposition, brain damage due to infections, tumours, alcohol or other toxic substance, and some infectious diseases, such as aspergillosis, meningococcal meningitis, schistosomiasis, malaria and encephalitis (Berg *et al.* 1996, WHO factsheet 2000). Moreover, the commonest cause of epilepsy differs from country to country. For instance, cerebrovascular disease and head injuries were identifed as the commonest cause in Bradford, UK (Wright *et al.* 2000); however, neurocysticercosis is the most common cause of epilepsy in rural Zambia (Birbeck and Kalichi 2004).

According to the classification proposed by ILAE (1981, 1989), epilepsy can be categorised by means of the aetiology or seizure types (http://www.ilae-epilepsy.org/). Based on aetiology, epilepsy is classified as symptomatic epilepsy, which implies a known underlying cause; idiopathic epilepsy, which entails unknown cause; and cryptogenic epilepsy, which has suspected symptomatic cause but is not confirmed. Classification by seizure types is dichotomous and is described as: (1) focal epilepsy with localised seizure (2) generalised epilepsy with either localised or generalised seizure.

The different seizure types and syndromes create complexities in establishing the incidence, prevalence and prognosis of epilepsy. In the United Kingdom, the age-standardised prevalence of epilepsy in 1998 was reported at 7.4 per 1,000 in men, and 7.2 per 1,000 in women, which was based on a total size of 1.4 million patients from 211

general practices in England and Wales (Purcell *et al.* 2002). However, the reported prevalence ratios vary remarkably between and even within countries. By reviewing available door-to-door surveys of epilepsy prevalence in sub-Saharan African, Prenux and Druet-Cabanac (2005) found the prevalence of epilepsy varied from 5.3 per 1,000 to 74.4 per 1,000 across rural and urban region of sub-Saharan African, while the median prevalence was 15 per 1,000 people. According to the latest international survey report by the WHO (2005), the mean prevalence of active epilepsy, which presents as refractory seizures and requires essential medical treatment, is approximately 8.2 per 1000 of the general population. However, this estimation is even higher in the developing countries, where the prevalence is more than 10 per 1000. Thus, the lifetime prevalence of epilepsy is approximately 100 million people, which is the number of people presently in the world who have epilepsy now or have had it in the past or will experience it in the future (figure 1.4.1)
(http://www.who.int/mental_health/management/globaepilepsycampaign/en/).

Figure 1.4.1: Worldwide incidence of epilepsy (N: number of countries responding to the survey in corresponding region, which are indicated by colours)

The obscure aetiology and intermittent occurrence of epilepsy has both social and physiological consequences that can impact on individuals' life more intensively than the seizures that patients experienced. Such consequences include damaged employment prospects, personal development, mental health and difficulty in personal relationships. Furthermore, a recent study has shown that epilepsy is associated with an increase risk of mortality (Bruce *et al.* 2004). This is supported by the latest international report entitled 'Global campaign again Epilepsy: out of the shadow', which is a worldwide project launched by the World Health Organization (WHO), International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Meanwhile, Gaitatzis (2004) revisited epidemiological studies that have described the mortality of patients with epilepsy and found that the mortality of the epilepsy population was 2-3 times higher than that of the general population, where the mortality was measured by the standardized mortality ratio (SMR) (Gaitatzis and Sander 2004, WHO 2005). However, despite the adverse prognosis for particular groups of epilepsy patient, Gaitatzis (2004) suggested that optimising treatment and care for epilepsy patients could prevent those seizure-related deaths effectively.

Available treatments for epilepsy

Currently, pharmacotherapy and surgery are two major medical treatments available for epilepsy patients (Aiken and Brown 2000, WHO 2005).

Surgery is principally considered as an important treatment for intractable epilepsy patients, where resection of the temporal lobe is informed by localisation of the seizure focus. Surgery was first introduced in the late 19th century and has became more effective with the great improvement in neuroimaging, specifically magnetic resonance imaging (MRI), positron emission tomography (PET) and single photo emission computed tomography (SPECT), which further improves the ability to identify epileptogenic brain lesions that were not apparent with previous diagnostic tools (Cascino 2004). Intractable epilepsy is defined when the patient has true epileptic seizures of unacceptable frequency and severity after a trial of at least two drugs up to tolerance level for a period of two years. In order to achieve a successful surgical treatment, certain evaluations of the epileptic syndrome and possible outcomes after surgery have to be considered, such as seizure onset, likelihood of neurological deficit afterwards, and destruction or isolation of the focus (Kemeny 2001). Currently, as Schmidt (2003) proposed, the "cure rate" from

epilepsy surgery requires further study, involving a long-term randomised control trial to evaluate effect on specific surgical treatment on well-defined epilepsy symptoms. Moreover, surgical therapy is only available in 40.9% of the responding countries (155 countries cross the world) as stated by the latest WHO survey. Thus it is believed that at present pharmacotherapy is the main medical treatment for epilepsy patients across the world. The pharmacotherapy of epilepsy will be focused on in this thesis, with particular attention paid to those pharmacogenetic factors relevant to drug response.

The pharmacotherapy of epilepsy can be traced back 1856, when the first medicine, Bromide, was introduced for epilepsy syndromes. Until the early of 1990s, there were six major compounds; carbamazepine (CBZ); ethosuximide (ETS); phenobarbital (PBB); phenytoin (PHT); primidone (PRM) and valproic acid (VPA), widely used for medical control of seizure. With these drugs, about 50% of patients with newly diagnosed epilepsy can achieve seizure control immediately after drug treatment, and 20-30% of patients can achieve remission after one ore more changes of the daily dose or after switching to another compound. Pharmacotherapy can effectively control seizure in up to 70% epilepsy patients by either mono-therapy or poly-therapy. After 2-5 years of successful treatment, the drugs can be withdrawn in about 70% of children and 60% of adults without relapse of seizure (Brodie and Dichter 1996, Lima 2000, WHO report 2005). Before the 1990's, two well-known medical treatment strategies for epilepsy were established, which were related to AEDs profile and seizure types or epilepsy classifications (table 1.4.1).

Table 1.4.1: Epilepsy pharmacotherapy strategy with early AEDs (Lima 2000)

	Seizure types	**1st line**	**2nd line**
Classical strategy	Partial Seizures	VPA	CBZ, PHT
	Generalised convulsive seizures		CBZ, PHT
	Myoclonic, tonic, atonic-astatic seizures		BDZs
	Absence seizures		ESM
Specialist strategy	**Epilepsy type or syndrome**	**1st line**	**2nd line**
	Cryptogenic, Symptomatic partial Epilepsies	CBZ or PHT	VPA
	General, Idiopathic partial Epilepsies	VPA	CBZ or PHT BDZs (Myoclonic) ESM (Absences)
	Lennox-Gastaut syndrom	VPA	BDZs, CBZ, PHT, ESM
	West syndrom	ACTH	VPA, CBZ

Note: BDZs: Benzodiazepines are used as antiepileptic drugs primarily to treat status epilepcus and to terminate serial seizures, but they are also used clinically for antianxiety, as muscle relaxants, and for their hypnotic activity. The target of BDZs is a receptor for the neutral amino acid ɣ-aminobutyric acid (GABA).
ACTH: Adrenocorticotropic hormone produced in the anterior pituitary gland, often used for stopping certain epilepsy seizure, especially for infant spasms.

In the late 1990s, few new antiepileptic drugs (AED) were approved and launched into the market, such as gabapentin (GPT); Lamotrigine (LTG); felbamate (FLB); clobazam(CLB); vigabatrin (VGB); oxcarbazepine (OXC); zonisamide (ZNS); tiagabine (TGB); and topiramate (TPM) (Beghi 2004).

The common issues in Pharmacotherapy of epilepsy

Although, antiepileptic drugs (AEDs) can prevent the occurrence of seizures, attenuate or eliminate the seizure-related risks, such as mortality and morbidity, and improve patients' life quality, the other risks, especially due to side effects of AEDs, have to be considered thoroughly. Most of the AEDs have a relatively narrow therapeutic index, and have higher likelihood of drug interactions when the poly-therapy strategy is applied or other medication is required (Tanaka 1999a, Aiken and Brown 2000, Klassen and Sadler 2001).

There are a few common issues that should be emphasized in the pharmacotherapy of epilepsy (Lima 2000, Stephen and Brodie 2002, Loscher 2002b, Perucca *et al.* 2000).

1. Mono-therapy is preferred in the pharmacotherapy of epilepsy, but about 30% of epilepsy patients, especially those with multiple seizure types or refractory epilepsy, generally require various combinations of AEDs. Withdrawing or altering the dose of an exiting AED can induce seizures and status epilepticus.
2. The pharmacotherapy of epilepsy is not necessarily continuous for a lifetime, but long-term treatment is very common. The AED tolerance of patients is a major concern in epilepsy pharmacotherapy.
3. The cause or mechanism of epilepsy is commonly unknown; many epilepsy patients have concomitant disease. Therefore the co-admission of other drugs with AEDs is commonplace.

Additionally the mechanism of most AEDs is currently not known with certainty. Many of them appear to act through multiple complementary mechanisms and target critical brain systems, to achieve the protection against seizure activity. The pharmacokinetic and pharmacodynamic characters of often-used AEDs are summarized in table 1.4.2.

Table .1.4.2: Summary of Antiepileptic drugs' pharmacokinetics and pharmacodynamics characteristics and relevant interactions (Rogawski 2002, Patsalos *et al.* 2002, Anderson 1998, Dichter and Brodie 1996, Brodie and Dichter 1996, Asconape 2002, Tanaka 1999a)

Drug generic name		Pharmacokinetics						Pharmacodynamics or Mechanisms.
		Metabolism (major pathway)				Relative Elimination (%)	Protein Binding (%)	
		CYP450s			Others			
		Substrate	Inducer.	Inhibitor				
1st generation	CBZ	Yes	Yes			98	75	Na^+
	ETS	Yes				85	0	Na^+ Ca^{2+}
	PBB	Yes	Yes			75	50	GABAa
	PHT	Yes	Yes			95	90	Na^+
	PRM	Yes	Yes			80	25	
	VPA	Yes		Yes		97	90	
2nd generation	FBM	Yes	Yes	Yes		50	25	Na^+ GABAa NMDAr
	GBP		N/A	N/A	Not*	0	0	Ca^{2+} GABAb
	LEV		Not	Not	NoHep	34	0	
	LTG		N/A	N/A	UGT	80	56	Na^+ Ca^{2+}
	OXC	Minor*	Yes	Yes	UGT	50	40	Na^+
	TGB	Yes	N/A			98	98	GABAtr
	TPM		Yes		UGT	20	15	Na^+ GABAa
	VGB		Not	Not	Not*	4	0	GABAts
	ZNS	Yes				65	60	Na^+ Ca^{2+}

Note: N/A: Not available data; Not*: not metabolised; Minor*: minor metabolism of its primary pharmacologically active metabolite; NoHep: non-hepatic hydrolysis; UGT: uridine diphosphate glucuronosyl transferase; Na^+ Ca^{2+}s: Na^+ Ca^{2+} channel; GABA: gamma aminobutyric acid; GABAa: GABA receptor type A, GABAb: GABA receptor type B, GABAtr: GABA transporter, GABAts: GABA transaminase, NMDAr: N-methyl-D-asparate receptor; Relative elimination presented in the relative contribution ratio of hepatic metabolism and renal excretion for AEDs, for instance 85% means 85% elimination is contributed by hepatic metabolism, 15% by renal excretion.

From the table, it is apparently that the older AEDs have different pharmacokinetic characteristics from the new AEDs, where the first generation of AEDs is primarily eliminated through hepatic metabolism catalyzed by cytochrome P450s and uridine diphosphate glucuronosytransferase (UGT) enzymes. These AEDs cause the induction or inhibition of cytochrome P450s, which are major causes of drug interactions related to AED therapeutics (Levy 1995, Anderson 1998, Tanaka 1999a,b).

Following the development of the second generation of AEDs, the pharmacotherapy of epilepsy had been expected to be easier due to their linear kinetics and the less frequent effect on cytochrome P450 metabolism. However, except LTG and OXC, most new AEDs are licensed for adjunctive therapy only (BNF, 2006). Furthermore, each of the new AEDs provides a unique profile of pharmacokinetics, adverse effects, and mechanism of actions, making an appreciation of how these agents are best utilized even more difficult.

Meanwhile, there remains no well-established treatment guidelines or strategy for choosing new AEDs (LaRoche and Helmers 2004). According to currently available clinical trials, Wilby and his colleagues (2005) found that little good-quality evidence can support that the use of the newer mono-therapy or adjunctive therapy new AEDs are no more effective

compared to older drugs. Also there is insufficient evidence to prove the use of one new AED in preference to another. Furthermore, for achieving equivalent effectiveness, the treatment with a newer AED was more costly than with older AEDs in monotherapy of epilepsy (Wilby *et al.* 2005). Therefore, in most countries, the financial problem related to the new AEDs prevents their wide usage in clinics. According to the latest WHO survey (WHO 2005), the older AEDs, such as carbamazepine (CBZ), phenobarbital (PBB), phenytoin (PHT), are still the major agents in the pharmacotherapy of epilepsy.

With a lack knowledge of epilepsy aetiology, the selection of AEDs is mainly based on their efficacy for specific type of seizures, tolerability, and safety (Perucca *et al.* 2000). Monitoring of plasma concentration of AEDs is frequently required because of their narrow therapeutic windows, the varied inter-individual responses under identical treatment regimes, and the high prevalence of side effects.

Additionally, significant clinical interactions between AEDs, or AEDs and non-AEDs, are widely observed in the pharmacotherapy of epilepsy (Patsalos *et al.* 2002). The major cause of these interactions as previously mentioned is the metabolism and induction or inhibition of cytochrome P450 enzymes, especially in those patients who have genetic polymorphisms in genes that are coding for relevant cytochrome P450s.

How Pharmacogenetics could improve treatment with AEDs

Pharmacogenetics aims to find inherited factors relevant to drug responses. It is possible to explore the variant genes that cause variable AED responses, particularly in the mutated cytochrome P450 genes, and this may help a physician to predict those patients with deficient metabolism, potential interaction with other AEDs or non-AEDs, and possible side effects due to deficient metabolism. Meanwhile, the genes associated with AED targets, such as the ion channel, GABA receptors, and GABA transaminase also contribute to the various AED responses, polymorphism of these genes would be useful to explain the variability of AED responses in pharmacodynamic aspects.

In addition to the identification of genetic polymorphisms related to the therapeutic effects of AEDs, pharmacogenetic research may help to elucidate some serious adverse reaction of AEDs in the future with the discovery of genetic variation regarding the side effects after taking AEDs. For instance, the genetically determined fetal or maternal susceptibility to teratogenic effect of AEDs may explain the different prevalence of birth defects, which is a

commonly seen serious side-effect of AEDs, among female patients who had exposed fetuses to certain AEDs, such as CBZ, and PHT (Lindhout and Omtzigt 1992). Moreover, pharmacogenetic research of multi-drug transporters could possibly explain the multidrug resistance of some epilepsy patients based on recent studies of P-glycoprotein expression and members of the multiple drug resistant-associated proteins (MDRs) families. In the future, these genetic variations related to pharmacotherapy with AEDs may contribute to the effective and safe application of pharmacotherapy in epilepsies (Loscher 2002a, b).

In the current thesis, the antiepileptic drug, phenytoin, has been a focus for study. The pharmacogenetic factors related to the major enzymes responsible for the metabolism of phenytoin in particular, the genetic polymorphisms of cytochrome P450, *CYP2C9*, *CYP2C19*, are examined.

1.5 Objectives of the thesis

At present, there is no systematic review or meta-analysis undertaken for studies of population distribution of *CYP2C9/2C19* polymorphism. Apart from reviews of pharmacogenetics in antiepileptic drugs, no any systematic review or meta-analysis has been employed to evaluate the genetic effect of *CYP2C9/2C19* polymorphism on the antiepileptic drug responses. Therefore, the first objective of this thesis is to undertake a meta-analysis method to assess the distribution of *CYP2C9/2C19* alleles/genotypes in different ethnic populations. The second objective is to examine the influence of *CYP2C9/2C19* genetic polymorphism on phenytoin metabolism. The association studies between *CYP2C9/2C19* polymorphism and phenytoin metabolism were explored using systematic searching strategy. Since the current available studies cannot provide sufficient evidence to quantify the effect of *CYP2C9* and *CYP2C19* allele/genotype polymorphism on phenytoin metabolism, association studies of *CYP2C9/2C19* polymorphism and relevant probe drugs, warfarin/mephenytoin, were further investigated in order to quantify the genetic effect of *CYP2C9/2C19* polymorphism on drug metabolism.

Chapter 2 Methodology

2.1 Preface

With the deciphering of the human genome, genetic polymorphism is being recognized with increasing frequency (http://www-hto.usc.edu/~cbmp/2001/SNP/index.html; Nebert *et al*, 2001). Extensive work has been carried out on the polymorphism of drug-response related genes. PubMed searches retrieved 69892 publications by using the keyword (MeSH) 'genetic polymorphism', covering the period 1964-2003; 2043 publications with pharmacogenetics as the keywords, and 598 publications with both 'polymorphism and pharmacogenetics' as keywords over the same period. Out of 598 papers, 476 were published in or after 1995.

Pharmacogenetic research has been conducted utilising different approaches across many biomedical scientific fields (Meyer 1990, 2000). The valuable data relating to gene-drug relationships remains scattered throughout the published literature. None of the available databases have comprehensively aggregated the work on pharmacogenetics research. Extracting and characterizing the gene-drug relationships from the literature requires new methods to enable scientists to keep up to date with the ever-growing amount of published literature, and to create a useful information source for future Pharmacogenetic discoveries (Chang 2004). The search methods employed in Evidence-based Pharmacotherapy (EBP) cannot be applied easily with success until a clear picture about the pharmacogenetics of AEDs is obtained. Meanwhile, the quality of Pharmacotherapy evidence differs significantly from one study to another, depending on the original study design (Alain Li Wan 1996). Unlike clinical Randomised Controlled Trials (RCTs), pharmacogenetics studies lack a gold standard.

Most drugs have multiple pharmacological effects and affect multiple organs. Therefore, the overall drug response is a complex process. Polymorphism may be associated with any proteins including those involved in signal transduction, and these proteins may lead to altered drug response. In practice, identification, measurement of drug response and hence genotyping may be very difficult particularly for diseases with poorly defined aetiologies or related symptoms. Moreover, single gene polymorphism is not expected to capture the

gene-drug relationship for multi-factorial diseases very well, such as in the case of epilepsy pharmacotherapy.

Antiepileptic drugs have multiple-targets (Macdonald *et al*, 1995; Rogawski *et al*, 2000), and complex mechanisms of action. Moreover the effect of genetic polymorphism on response to antiepileptic drugs (AED) is not yet well defined.

In order to have an efficient searching method for studies describing polymorphic drug response to AEDs, a combination of general terms, such as antiepileptic/anticonvulsant drug and polymorphism/pharmacogenetics, was chosen for preliminary search. All relevant articles including some pharmacogenetics reviews were read through. Then the most frequently appearing terms in those publications were chosen for further search. After comparing results from searches with different combined terms, a searching strategy was defined with aim of gaining as complete as possible relevant publication list.

An initial search in MEDLINE and EMBASE, the two most comprehensive biomedical databases, retrieved 29 articles with combined terms of 'antiepileptic/anticonvulsant drug' and 'polymorphism/pharmacogenetics'. However, only one relevant article was caught by search in EMBASE, which also appeared in MEDLINE's result. Due to a lack of database subscriptions in the University, searches for the project were restricted to PubMed, a search engine managed by the National Library of Medicine in the United State, and which includes over 15 million citations from MEDLINE with the addition of other life science journals for medical research.

Genomics information is widely accessible through many electronic databases. A hand search of the key journal, Epilepsia, confirmed the fact that any reference relating polymorphic genes and antiepileptic drug response could be retrieved by searching the electronic database PubMed.

Due to the wealth of information on genetic polymorphism, bibliographic software was used to organize and manage relevant references. EndNote was found to be the better user-friendly program to retrieve and organize references when compared with another available software, ProCite. EndNote was employed to retrieve relevant references from PubMed for the project.

In addition to searches in electronic databases, three key journals were chosen according to the aims of current project. They were:

1. Pharmacogenetics and Genomics
2. Clinical Pharmacology and Therapeutics
3. Epilepsia.

Pharmacogenetics and Genomics (previously known as Pharmacogenetics) is one of the few key-journals in Pharmacogenetics research. It started publishing in 1991 and became one of the most valuable journals in the Pharmacogenetic and Genomic research area. According to the journal citation report in the Web of Science, the impact factor of this journal increased from 4.371 to 6.406 from 2001 to 2004, which indicated this journal was receiving recognition for publication of valuable papers. Clinical Pharmacology and Therapeutics had been available since 1960, and issues from 1995 are available in full-text from the electronic database. It is also a well-respected journal for pharmacology research in clinical therapy. Epilepsia is the official publication of the International League Against Epilepsy. It covered most of the research aspects relevant to epilepsy.

The searches in the key-journals aim to supplement the major searches in PubMed, which were performed with a combination of general terms, such as antiepileptic/anticonvulsant drug with polymorphism/pharmacogenetics in Epilepsia, and CYP2C9/2C19 polymorphism with phenytoin in the other two journals.

2.2 Search strategy

Search terms always included 'Polymorphism' or 'Pharmacogenetics'. MeSH keywords were always chosen when they were available for searching aim. The searching field was limited to the abstract or keywords, except for overviews of gene polymorphism of human response to antiepileptics, which was searched in any field. Each of the terms was searched in both the singular and plural forms. The common searching process is displayed in the flowchart. Under the subheading of each aim, searching terms or MeSH keywords and brief criteria of reference selection is presented.

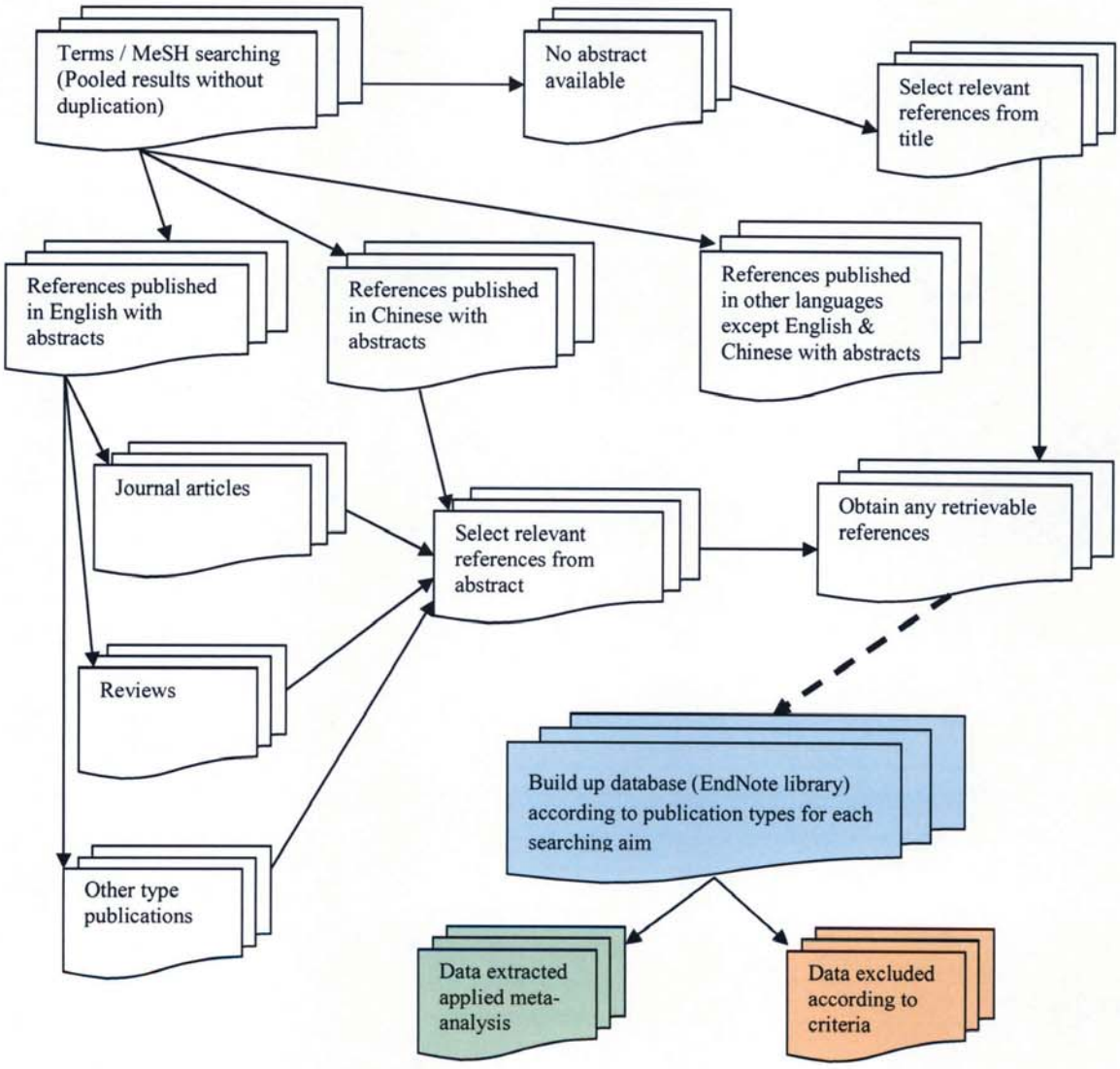


Figure 2.2.1: The flowchart of searching strategy

The search terms are listed in table 2.2.1 to table 2.2.5 according to each aim.

Table 2.2.1: Search terms for the review of gene polymorphism of human response to AED

Combined Terms	Antiepileptics/anticonvulsants Polymorphism Pharmacogenetics
Searching Field	Any field
Criteria	Gene polymorphism of human response to AED

Table 2.2.2: Search terms for assessment of the population distribution of CYP2C9, CYP2C19 alleles/genotypes

Combined Terms with CYP2C9/19	Allele/s Variant Mutation/s Ethnicity/ties/ethnic Frequency Polymorphism Polymorphic metabolism Pharmacogenetics Population/s
Searching Field	Abstract
Criteria	CYP2C9/2C19*1 *2 *3 genotyped. Clearly defined ethnicity of subjects.

Table 2.2.3: Search terms for the estimation/valuation of association studies of CYP2C9/2C19 polymorphism and phenytoin metabolism

Combined Terms	CYP2C9/CYP2C19 Phenytoin Polymorphism
Searching Field	Keywords (any field)
Criteria	CYP2C9/2C19*1 *2 *3 genotyped. Clearly defined ethnicity of subjects. Phenytoin phenotype data reported.

Table 2.2.4: Search terms for estimation/valuation of association studies of the CYP2C9/2C19 polymorphism and warfarin/mephenytoin metabolism

Combined Terms	CYP2C9/CYP2C19 Polymorphism Warfarin (Mephenytoin)
Searching Field	Abstract (Keywords)
Criteria	CYP2C9/2C19*1 *2 *3 genotyped. Clearly defined ethnicity of subjects. Mephenytoin or Warfarin phenotype data reported.

Table 2.2.5: Search terms using in key journals

Journal	Combined Terms
Pharmacogenetics and Genomics Clin Pharmacol & Ther	Polymorphism Pharmacogenetics (Keywords) CYP2C9 CYP2C19 (Abstract)
Epilepsia (Drug response)	Polymorphism (Keyword) Pharmacogenetics (Any Field)

2.3 Approach for retrieving data

2.3.1 References for the overview of gene polymorphism relevant to the human response to AED

All articles that studied the gene polymorphism of the human response to AED were included.

For consistency, the references were categorized according to the data or relevant aspects of pharmacogenetic studies. As mentioned in Chapter 1, relevant papers are clinical outcome studies, pharmacodynamics & drug response studies, and pharmacokinetics studies. In addition to those pharmacogenetic aspects of AEDs, some aetiological studies about inheritance of epilepsy are also included, which provided useful supplemental information about other gene polymorphisms associated with the human response to AED.

2.3.2 References for CYP2C9, CYP2C19 population distributions

In the selected references, general information for each study was retrieved and recorded into summary tables, which included the ethnicity or geographic location of recruited subjects (*i. e.* American Chinese, Taiwan Chinese), the approach of subject recruitment (*i. e.* randomly selection from local clinics, random selection of healthy volunteers), physiological status of subjects studied, and the PCR based genotype techniques or methods. The geographic area from which subjects were recruited (*i. e.* Tamilian, Israeli Arab) was recorded separately when no clear ethnicity was defined.

The frequencies of *CYP2C9, CYP2C19* alleles or genotypes were retrieved and recorded into individual tables for each study. Essential calculation was performed when the allele or genotype frequencies (the number of subjects carrying certain alleles or genotypes) were not reported clearly or directly. For instance, few studies reported the proportion (percentage) of alleles or genotypes instead of the frequency (number) of alleles or genotypes in the studied subjects.

Furthermore, in those studies where the data was presented in graphical form rather than text or table, the graph digitizer program, Grafula, was used for retrieving data. After inputting the graph into Grafula, a calibration curve was created using the original graph axes presented in the publication. The data was established from the average of three times readings.

2.4 Applied Statistical analysis

2.4.1 Studies of CYP2C9, CYP2C19 population distributions

Original data presentation for each study:

Based on the Binomial distribution, the 95% confidence interval of the proportion of six possible genotypes and three alleles of *CYP2C9* and *CYP2C19* was calculated in Minitab, using the Exact Method (Johnson N.L. 1969) by default.

The ethnicity of subjects in each study was identified from the original description in the publication. Therefore all studies were categorized into different ethnicity groups, as Caucasian, Chinese, Japanese, African and miscellaneous groups. In the miscellaneous group, the ethnicity of each study was defined as original publication and categorized as the closest ethnicity group using the phylogenic tree of Carvalli-Sforza (1994). Subsequently in each category, a Forest plot was produced as a visual result of the allele or genotype proportions and their 95% confidence interval in Statistica 6.0 for each of the studies selected for the *CYP2C9, CYP2C19* population distribution.

Hardy-Weinberg equilibrium test for each study:

For those studies with genotype data available, the Chi-square goodness of fit test has been applied within each of the genotype studies in order to test whether the population's genotype follows the Hardy-Weinberg Equilibrium principle (HWE); this was performed in the GENEPOP program. As the alleles *CYP2C9*3* and *CYP2C19*3* have less than 1% prevalence in quite a few of the studies, the Fisher Exact test was performed using the GENEPOP program (Louis and Dempster1987). (http://wbiomed.curtin.edu.au/genepop/) Studies were retained for the next analysis stage when they had probability over 0.05 in the result of HWE test. The data was excluded when the probability of fitting HWE was less than 0.05.

Estimation for the prevalence of three *CYP2C9, CYP2C19* alleles in Caucasian, Chinese, Japanese and African

R*C contingency table analysis was performed for the estimation of heterogeneity or population differentiation among those studies that fitted with HWE. For each ethnicity, the unbiased estimate of the exact test P-value was achieved using the Markov Chain method as described by Raymond and Rousset (1995) in the GENEPOP program. When

the P-value was greater than 0.05, the null hypothesis was accepted that those studies had no statistically significant heterogeneity or population differentiation.

When there was not any heterogeneity or population differentiation observed among studies, The pooled prevalence was calculated as the estimated prevalence of *CYP2C9, CYP2C19* alleles in each ethnicity in the fixed model of meta-analysis method. Otherwise, the random model for meta-analysis was applied.

Trinomial plots were produced for the three alleles of *CYP2C9* and *CYP2C19* based on ethnicity categories, which are presented as visual results of the separation or overlap of the proportions or prevalence of three alleles appeared in different ethnicity groups. The plotting was performed in the Simfit program, which was developed by Manchester University.

(http://www.simfit.man.ac.uk/default.htm)

2.4.2 Studies of the genetic polymorphism of CYP2C9, CYP2C19 effect on AED or probe drug metabolism

The departure from the Hardy-Weinberg Equilibrium principle was tested for each study using the GENEPOP program (Louis and Dempster1987).

The approach of standardized difference in Meta-analysis was applied for the antiepileptic drug (Whitehead 1991), phenytoin, and the probe drugs, mephenytoin and warfarin according to each of the genotypes of *CYP2C9* or *CYP2C19*, where the difference of plasma concentrations or metabolic ratios or daily doses was compared between subjects or patients with *CYP2C9*1* and *CYP2C9*2* or **3* alleles, otherwise with *CYP2C19*1* and *CYP2C19*2* or **3* alleles.

According to whether there was heterogeneity between the studies, the fixed or random model analysis was applied to estimate the association of the different *CYP2C9, CYP2C19* allele and variable phenytoin, mephenytoin and warfarin pharmacokinetic characters.

Chapter 3 Population Distribution of CYP2C9&2C19 Alleles and Genotypes

3.1 Preface

The *CYP2C9* and *CYP2C19* genes are located in the CYP2C gene cluster on chromosome 10 with nine exons each (figure 1.4.1). The different alleles, *CYP2C9*1, *2, *3* and *CYP2C19*1, *2, *3,* contribute to the major inter-individual and intra-population variations in the response to medication by the consequent deficient or inactive enzyme functions (Ingelman-Sundberg *et al.* 1999,b Ma *et al.* 2002, Pirmohamed and Park 2003).

Earlier pharmacological studies have proved that 2-6% of Caucasian populations, and 14-22% of Asian populations, such as Japanese, Chinese and India, are poor metabolizers of the antiepileptic drug, mephenytoin, which is mainly metabolised by enzyme CYP2C19 (Meyer 1994a). Later, Xie *et al.* (1997, 1999) reported that there was good agreement between the population distribution of the CYP2C19 poor metaboliser phenotypes and *CYP2C19* alleles/genotypes in healthy Caucasians living in different geographical areas. It is believed that the prevalence of mutated *CYP2C19* alleles could give a better estimation of the occurrence of mephenytoin poor metabolizers, and may also be applied to other CYP2C19 substrates (drugs). Furthermore, by reviewing in-vitro data of *CYP2C9* polymorphism, Lee *et al.* (2002) proposed that individuals carrying *CYP2C9*2* or **3* appeared to be more susceptible to adverse events with narrow therapeutic index agents, such as phenytoin; meanwhile from population distribution data he found that around 33~51% Caucasian individuals have the mutated allele **2* and **3*, while only 3~13% individuals from African and 2~10% from Asian populations had those mutated alleles. As with other *CYP2C9/219* polymorphism studies, the prevalence of variant *CYP2C9* or *CYP2C19* alleles for different ethnic populations was found to exist over a wide range. None of studies had applied meta-analysis to estimate the prevalence of *CYP2C9/2C19* for each ethnic population. For the drugs primarily metabolised by CYP2C9/2C19 enzymes, such as phenytoin, a good estimation of the variant alleles' prevalence would give better understanding about the variation in drug response among people with different ethnic

origins regarding the distinctive genetic makeup of the CYP2C9 and CYP2C19 enzymes. Therefore, in the thesis the prevalence of *CYP2C9*1, *2, *3* and *CYP2C19*1, *2, *3* was initially derived from the included studies for people with different ethnic origins. Following the population genetic theory, the Hardy-Weinberg equilibrium was applied to test the validation of genotype data from the retrieved studies. The meta-analysis method was employed to estimate the prevalence of **1, *2* and **3* for *CYP2C9* and *CYP2C19* respectively, which was presented as the proportion of each allele present in different ethnicity populations.

General information about the three alleles of *CYP2C9, CYP2C19* is reiterated here to aid interpretation of results. *CYP2C9*1* and *CYP2C19*1* are often called wild alleles and consequently the encoded enzymes perform normal functions in relevant substrate metabolism, while *CYP2C9*2, *3* and *CYP2C19*2, *3* encode enzymes with decreased capability in substrate metabolism. As with most CYP450s enzymes, both of the normal CYP2C9/2C19 enzymes are proteins with 490 amino acids (figure 3.1.1.).

Instead of *CYP2C9*2, *3* and *CYP2C19*2, *3,* quite few earlier publications used alternative names, nucleotide changes or amino acid changes to describe the mutated alleles, which are shown in table 3.1.1 and table 3.1.2. By 3rd May 2005, 20 *CYP2C9* variant alleles and 16 *CYP2C19* variant alleles were reported respectively in the cytochrome P450 database (http://www.imm.ki.se/CYPalleles/).

Table 3.1.1: General information about *CYP2C9*1, *2* and **3*

Alleles	Nucleotide Changes	Location	Amino Acid changed	Enzyme function	Alternative names
*CYP2C9*1*	---		---	*Normal*	*CYP2C9 wt*
*CYP2C9*2*	430C>T	Exon 3	Arg144cys	*Decreased*	*CYP2C9 m1*
*CYP2C9*3*	1075A>C	Exon 7	Ile359Leu	*Decreased*	*CYP2C9 m2*

Table 3.1.2: General information about *CYP2C19*1, *2* and **3*

Alleles	Nucleotide Changes	Location	Effect of change	Enzyme function	Alternative names
*CYP2C19*1*	---		---	*Normal*	*CYP2C19 wt*
*CYP2C19*2*	681G>A	Exon 5	Early stop codon	*Decreased*	*CYP2C19 m1*
*CYP2C19*3*	636G>A	Exon 4	Early stop codon	*Decreased*	*CYP2C19 m2*

3.2 Studies of CYP2C9 & CYP2C19 Population distribution

3.2.1 The EndNote library for genotype studies of CYP2C9 & CYP2C19

A search of the PubMed database was performed in the specific field "abstract" by employing EndNote software with combined terms of CYP2C9/2C19 and others, such as allele, variant mutation, ethnicity, ethnic, and frequency in singular and plural forms respectively.

Whilst, most publications in PubMed database provided abstracts, when any articles had no abstracts available, the search by default would be performed in free-text, such as keywords and titles. Retrieved publications were pooled together for *CYP2C9* and *CYP2C19* respectively. After deleting duplicated articles by using the EndNote organizing function, each library was initially arranged into two sub-libraries for articles with abstracts and without abstracts where the library of articles with abstracts was sub-categorized according to publication languages and publication types as described in figure 2.2.1. The further selection of articles was performed using English and Chinese publications.

In the beginning, population-based studies with genotyping of *CYP2C9/2C19*1, *2, *3* were selected by reading the abstracts. When abstracts were not available, the selection was based on the publication titles. As figure 2.2.1 shows, the selected publications were pooled together into one EndNote library for CYP2C9 and CYP2C19 correspondingly for further assortment and data retrieval. Then, the original publication of each selected article was gained from university libraries or other source.

Besides the major search in PubMed, a few key journals are defined for supplementary purposes with CYP2C9/2C19 combined with general search terms, such as polymorphism and pharmacogenetics.

Searches for *CYP2C19* population distribution retrieved 474 articles up until July 2004. By reading through the abstract, 130 articles were selected according to the research aim. The same search was repeated in August 2005. Seven further articles were selected from the new publications.

The EndNote library was built up for the *CYP2C19* population distribution by inputting the

combined 137 articles. In addition, two articles were selected from other search results, which also have *CYP2C19* genotype data available. One was found from 'Pharmacogenetics and Genomics' key-journal search. One was from a search of CYP2C9 population distribution. This provided a total of 139 relevant articles in the CYP2C19 EndNote library.

The same methods were applied to studies of CYP2C9; 329 articles were retrieved by July 2004. 104 of them were selected according to the criteria. 11 articles were selected from a repeated search result in August 2005. The EndNote library of CYP2C9 had 116 references citations, which also included one article from search of CYP2C19 population distribution (104+11+1).

In order to estimate the prevalence of *CYP2C9/2C19 *1, *2* and **3* in different ethnicity populations, the desired studies have genotyped each of the three alleles, and provided the ethnicity of recruited subjects. After primary reading of each article in the two EndNote libraries, 150 articles fell into this criterion, where 74 and 76 articles are for *CYP2C9*, *CYP2C19* respectively.

In general, case reports or review articles do not provide original data, however two case reports and three review articles of *CYP2C19* provided unpublished relevant data, which were included for further analysis. A few highly relevant review articles were read carefully to check whether any relevant data was missing in the initial EndNote library, however none of them had cited references that were not retrieved by the searching strategy. The exact procedure is presented in figure 3.2.1 and figure 3.2.2.

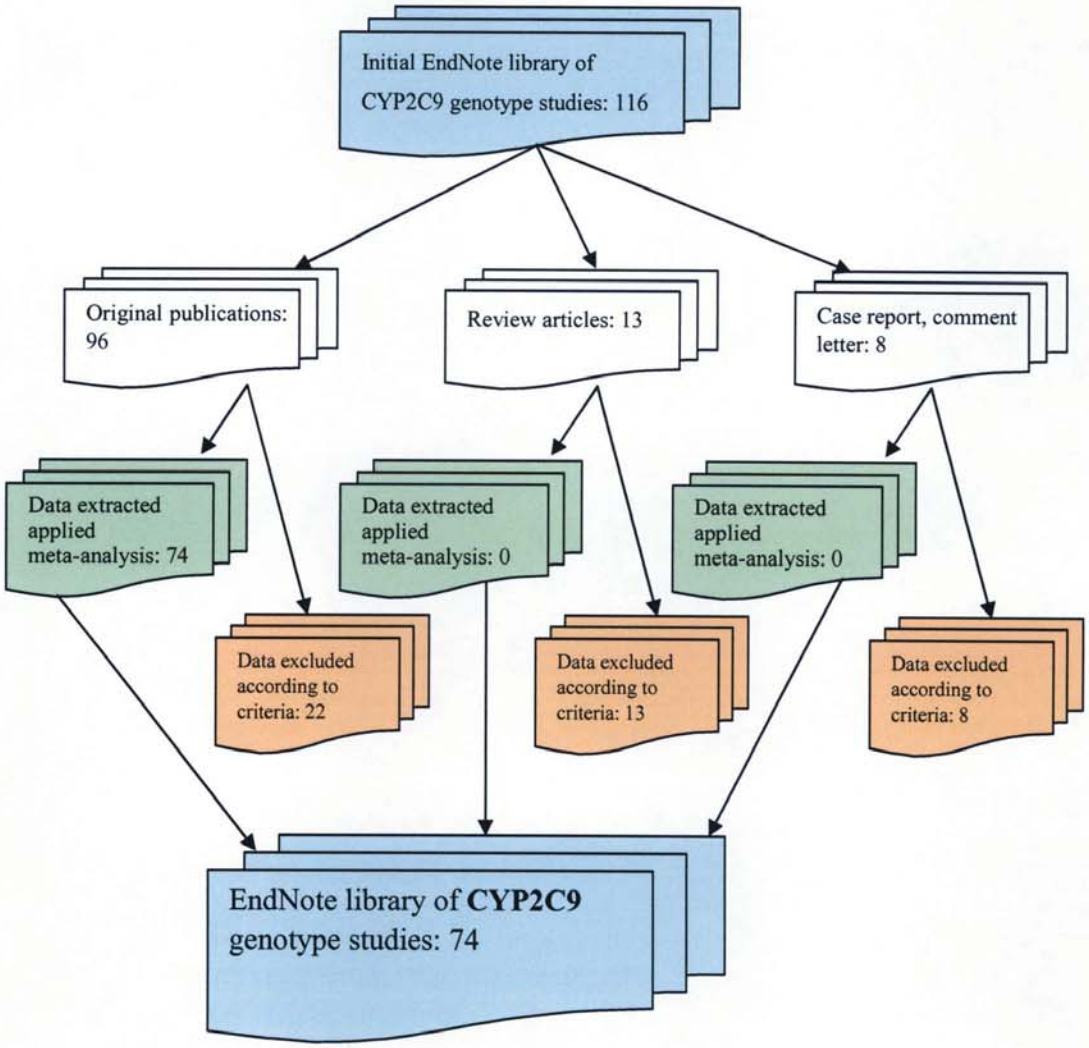


Figure 3.2.1 The EndNote library of CYP2C9 references

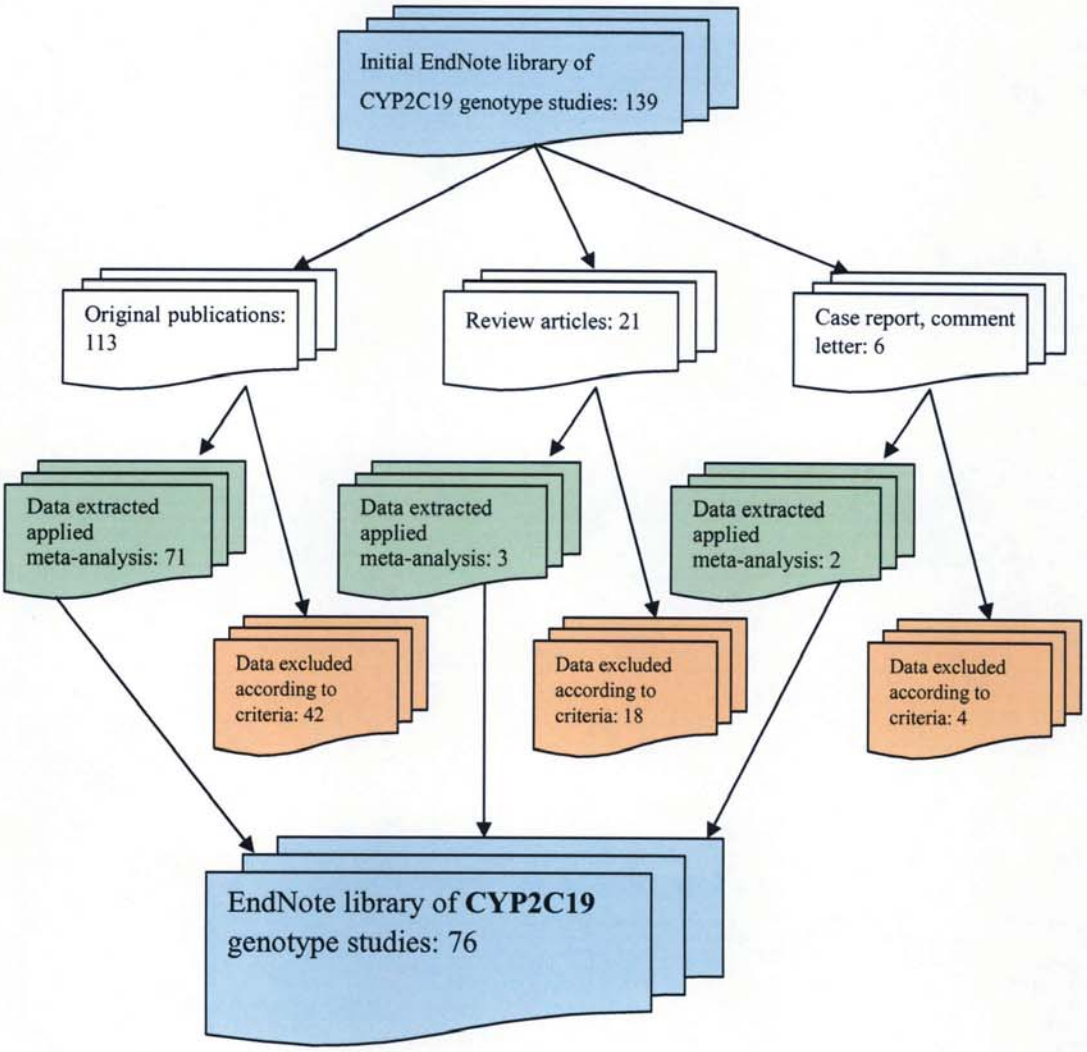


Figure 3.2.2 The EndNote library of CYP2C19 references

3.2.2 Summary of studies describing the population genetic distribution of CYP2C9&2C19

As different alleles of *CYP2C9/2C19* are found in Caucasian, Chinese, and Africans, determination of the genotype prevalence or mutant allele prevalence is valuable for further *CYP2C9/2C19* related pharmacogenetic estimation. This can provide information on the possible proportion of poor metabolizers in each population, particularly for those drugs which are primarily metabolised by *CYP2C9/2C19*.

Although Pharmacogenetics related studies currently have no gold standard available, most pharmacogenetic studies aim to find the genetic factors that are associated with the variable drug response, and are often performed with a candidate gene or genes. Similar to gene-disease association studies, pharmacogenetic studies can be considered as association studies between candidate genes and drug response or generally as one of the genetic association studies. Recently, Little *et al.* (2002) proposed a checklist for reporting and appraising studies of genotype prevalence and gene-associations, and this was followed for assessing the quality of selected studies in this thesis. According to the checklist, the purpose of the study, the analytical validity of genotype determination, and the selection of study subjects are very important in gene-disease association studies especially for further statistical analysis; therefore, these aspects of the selected articles were summarized before analysis was applied.

For the population distribution of *CYP2C9/2C19*, *i. e.* the population genotype prevalence, the characteristics of each study were derived for evaluation of the quality, which included the genotype method, the reported data format, the sampling methods, the study designs, and ethnicity identification.

The methods used for determining *CYP2C9/2C19* are referenced back to original publications. The data format describes whether the publication reported the genotype in number (frequency) or proportion (percentage); the rounding-up error introduced by calculating from proportion may give an apparent total larger than one (sum of proportions >100%). Some publications only provided the number or proportion of *CYP2C9/2C19*

alleles for recruited subjects. Without an assumption of fitness with Hardy-Weinberg equilibrium, the number or proportion of *CYP2C9/2C19* genotypes cannot be precisely calculated for these subjects. Therefore, apart from the allele proportion plots, further analysis cannot be undertaken for those studies.

Study design is very difficult to define in many studies. Genetic association studies could recruit subjects by family or recruit subjects without affiliation, where the desired data for estimation of genotype prevalence would be achieved from the studies of unrelated subjects who were randomly selected from same ethnic population. Among studies retrieved in this thesis, some recruited subjects from patient populations with specific diseases, some recruited healthy volunteers as subjects. Of the studies reported in patient populations, some studies divided the subjects into case and control groups separately according to Cohort study design. The genotype results were accepted for further analysis when the patients were recruited from multiple clinics randomly. Those issues are summarized into "study aims".

Furthermore, ethnicity contributes to the population prevalence of particular alleles, so the method of ethnicity identification is reviewed from the studies. In many publications, the authors reported the ethnicity without further detail of the methodology of how they define which ethnic group the subjects belong to. Few of them had applied a survey or questionnaire to track back two more generations from the subject under study. Some had identified the ethnicity for parents and four grandparents of the subject. The detail for each study provided in the 'Ethnicity identification' column in table 3.2.1. Meanwhile, the 'Ethnicity' column in table 3.2.1 provided the subjects' ethnicity as original prescription in those studies.

All the information has been summarized into the table 3.2.1 (the notes for abbreviations used in the table are attached); the studies were arranged by alphabetical sequence of first authors' name. Few studies reported allele or genotype data for both of *CYP2C9* and *CYP2C19*, which were indicated in the 'CYP' column; moreover the table is divided into three sections for the studies of *CYP2C9, CYP2C19* respectively, and the studies of both *CYP2C9* and *CYP2C19*.

Table 3.2.1 Summary of selected references for population distribution of ***CYP2C9***

Author	Year	Ethnicity	CYP	Genotype Method [a]	Data Formats [b]	Sampling Populations [c]	Study Aims [d]	Ethnicity Identification [e]
A Llerena	2004	Mexican	CYP2C9	1, 2, 3	AN, AP, GN, GP	Un	Genotype	3G
Aithal GP	2000	Caucasian	CYP2C9	5	GN, GP	Ra, Un	Ass. Disease	AuN
Aithal, G.P.	1999	Caucasian	CYP2C9	5	GN, GP	Ra, Un	Ass. Drug	N/A
Aynacioglu AS	1999	Turkish	CYP2C9	G	GN, GP	Un	Ass. Drug	AuN, residing
Babaoglu MO	2004	Turkish	CYP2C9	1, 3, 5	GN, GP	Un, Ad	Ass. Drug	AuN
Burian M	2002	German	CYP2C9	1, *	AP, GN**	Un	Genotype	AuN
Caraco	2001	Caucasian	CYP2C9	G, 4	GN	Un, Ad, Stu	Ass. Drug	AuN
Chen K	2005	Chinese (Han)	CYP2C9	*	GN	Un, Vol	Ass. Drug	AuN
Dickmann L	2001	African American, European American	CYP2C9	1, 3	GN, GP	Un	Genotype	Se, residing
Dorado P	2003	Spanish (white)	CYP2C9	1, 2, 3	AP, GN, GP	Un, Sta	Ass. Drug	AuN
Freeman BD	2000	White, African American	CYP2C9	2	GN, GP	Ra, Pa, 1999, 6-1999, 8**	Ass. Drug	AuN
Gaedigk A	2001	Inuit, Caucasian, Chinese, Canadian Native Indian (CNI)	CYP2C9	1, 6, M	AP, GN, GP	Pa, Un	Genotype DifEthnic	Se (Caucasian), 3G (Chinese, CNI), residing (Inuit)
Garcia ME	2001	Spanish (white)	CYP2C9	1, 3	AN, AP, GN, GP	Vol, Sta, Stu	Genotype	AuN
Garcia ME	2004	Spanish (white)	CYP2C9	1, 3	GN, GP	Vol, Sta, Stu	Ass. Drug	AuN
Hallberg	2002	Caucasian (White Swedish)	CYP2C9	*	GN	other	Ass. Drug Disease	AuN
Herman, D.	2005	Caucasian (Slovenian)	CYP2C9	3	GN	Pa, Cohort	Ass. Drug	AuN
Higashi, M.K.	2002	Mixture (96.2%Caucasian, 3.85 Hispanic)	CYP2C9	*	GN	Pa, Mult	Ass. Drug	AuN
Hillman WA	2004	Caucasian	CYP2C9	kit	GN, GP	Pa, Cohort	Ass. Drug	AuN
Ho PC et al	2003	Caucasian	CYP2C9	7	GN	Liver Sample from tissue bank	Ass. Drug	AuN
Hummers-Pradier,	2003	Caucasian (German)	CYP2C9	6	GN*	Pa, 1999.8-2000.2** Mult	Ass. Drug	N/A
Joffe, H.V.	2004	Mixture (89%Caucasian, 33%African-American, 3other races)	CYP2C9	3	GN, GP	Pa, 1996-2001**	Ass. Drug	AuN
Lee SS	2005	Vietnamese	CYP2C9	pyrosequence	AP, GN	Un	Genotype	AuN, residing
Lee, S.	2003	Korean	CYP2C9	5	GN	Pa, 1999-2000**	Ass. Drug	AuN
Linder, M.W.	2002	Mixture (98%Caucasians with 1 Hispanic, 2% Africa-American)	CYP2C9	1, 2	GN	Pa, 2000-2001**	Ass. Drug	AuN
Loebstein, R.	2001	Israel?	CYP2C9	5, 6	AN, GN*	Pa, Cohort	Ass. Drug	N/A
London SJ	1996	Caucasian, African-American	CYP2C9	*	AN	Un, Com	Ass. Disease	AuN, residing
Martin J H	2001	Caucasian	CYP2C9	1, 3, 10	GP	Pa, 1997,9-1999,11**	Ass. Disease	AuN
Martinez C	2001	Spanish (white)	CYP2C9	11	GN, GP	Un, Pa, Mult, Sta	Ass. Disease	AuN, residing
Nasu K	1997	Japanese	CYP2C9	5	GN	Un	Genotype	AuN
Ozawa S	1999	Hungarians	CYP2C9	5, 13	GN, GP	Pa	Ass. Disease	AuN
Pchelina, SN	2005	Caucasian (Slavic)	CYP2C9	3	GN, GP	Un, Pa	Ass. Drug	AuN, residing
Pederson RS	2004	Caucasian (Danish)	CYP2C9	6	AN, AP, GN, GP	Vol, Stu	Genotype	AuN, residing

Author	Year	Ethnicity	CYP	Genotype Method [a]	Data Formats [b]	Sampling Populations [c]	Study Aims [d]	Ethnicity Identification *
Continuous...								
Peyvandi, F.	2004	Caucasian (Italian)	CYP2C9	1, 3	GN*	Pa, Cohort	Ass. Drug	AuN
Pirmohamed M	2000	Caucasian (5.4 or 6.4%mixture)	CYP2C9	1, 13, 14	GN, GP	Pa	Ass.Drug Disease	AuN
Sandberg M	2004	no defined (Swedish?)	CYP2C9	kit	GN	Sta, Stu	Ass. Drug	AuN
Scordo M	2001	Italian, Ethiopian	CYP2C9	3	GN	Un	Genotype DifEthnic	AuN
Si D	2004	Chinese	CYP2C9	1, 3, 5	unclear	Un	Ass. Drug	AuN
Stubbins MJ	1996	Caucasian	CYP2C9	original	AP, GN	Ra	Genotype	AuN
Sullivan-Klose TH	1996	European American, African American, Chinese (Taiwanese)	CYP2C9	original	AP	Ra, Un, Ad, Pa	Genotype	AuN
Tabrizi AR	2002	African American, Caucasian	CYP2C9	5	AN, AP, GN	Pa, Cohort	Ass. Drug	AuN
Takahashi H	2003	Caucasians, Japanese	CYP2C9	2, A, A'	GN, GP	Pa	Ass. Drug	AuN
Topic E	2004	Croatians	CYP2C9	2, 3	GN, GP	Un, Blood Donor	Ass. Drug	AuN
van der Weide	2001	Duntch	CYP2C9	1, 3	GN, GP	Pa	Ass. Drug	AuN
Vianna JR	2004	White (Brazil), Intermediate (Brazil), Black (Brazil)	CYP2C9	1	AP, GN, GP	Un, Blood Bank	Ass. Drug	questionnaire
Visser LE	2004	Caucasian	CYP2C9	G	GN, GP	Pa, 1985,1-1998,12**, Mult	Ass. Drug	AuN
Visser LE	2004	Caucasian	CYP2C9	G	GN	Pa, 1990-1993**	Ass. Drug	AuN
Wang SL	1995	Chinese Aborigines, Chinese (Han)	CYP2C9	original	unclear	Un	Genotype	AuN, residing
Yang JQ	2003	Caucasian (French), Chinese (Han)	CYP2C9	6	GP	Un	Genotype	AuN, residing
Yasar	1999	Swedish	CYP2C9	1, 2, 5	AN, AP, GN, GP	Un, Sta, Stu	Genotype	AuN
Yasar	2002	Swedish	CYP2C9	kit	GN	other	Genotype	AuN
Yasar U	2003	Caucasian (Swedish)	CYP2C9	*	AP, GP	Ra, Pa	Ass. Disease	AuN
Yilmaz N	2001	Turkish	CYP2C9	*	GN, GP	Ra, Pa	Ass. Drug	AuN, residing
Yoon	2001	Korean	CYP2C9	2	GN, GP	Un	Genotype	AuN

? There is no description of ethnicity in original publication

* Genotype method did not referred to any original publication

Index note of the table found on page 72-73

Table 3.2.1 Summary of selected references for population distribution of ***CYP2C19***

Author	Year	Ethnic	CYP	Genotype Method [a]	Data Formats [b]	Sampling Populations [c]	Study Aims [d]	Ethnicity Identification [e]
Adithan	2003	Tamilians	CYP2C19	4	AP, GP	Un	Genotype	AuN, residing
Aklillu E	2002	Ethiopian (in Sweden), Ethiopian (in Ethiopia)	CYP2C19	A, B, C, D	GP	Un, Sta	Genotype&Enviroment	AuN, residing
Aynacioglu S	1999	Turkish, German	CYP2C19	A, A', C, F	AN, AP, GN, GP	Un	Genotype DifEthnic	AuN, residing
Bathum	1998	Danish	CYP2C19	A, A', H	GN	Un, Vol	Genotype&Longevity	AuN, residing
Bathum	1999	Negroid (Tanzania)	CYP2C19	A, A', H	GN*	Un	Ass. Drug	AuN, residing
Bramness JG	2003	Norwegian	CYP2C19	C, I	AN, AP, GN, GP	*	Ass. Drug	AuN
Brosen K	1995	Danish	CYP2C19	A, A'	graph	multiple families	Ass. Drug	AuN
Chang M	1995	Swedish	CYP2C19	A, A'	unclear	Vol, UnRa	Ass. Drug	AuN
Dandara	2001	Tanzanian, Venda, Zimbabwean	CYP2C19	A, A'	AP, GP	Vol, Pa	Genotype	AuN
de Morais	1994	American, Japanese, Swiss	CYP2C19	original	GN	Pa	Genotype*	AuN
de Morais	1994	American, Japanese, Swiss	CYP2C19	original	GN	Pa	Genotype*	AuN
de Morais	1995	Chinese	CYP2C19	A, A'	GN	Un	Ass. Drug	AuN
Edeki TI	1996	African -American	CYP2C19	A, A'	GN*	Un	Ass. Drug	AuN, residing
Ferguson RJ	1998	French	CYP2C19	I	AP	other	Ass. Drug	AuN
Fu LQ	2004	Chinese (Han)	CYP2C19	A, A'	GN, GP	Vol	Genotype DifEthnic	3G, AuN
Furuta T	2001	Japanese	CYP2C19	A. A'	GN	Pa, 1998, 8- 2000, 1	Ass.Drug Disease	AuN
Gaedigk A	2000	Inuit, Caucasian, Chinese,	CYP2C19	1, 6, M	AP	Un	Genotype DifEthnic	AuN
Garcia-Barcelo M	1999	Chinese	CYP2C19	C, F, L	AP	Un	Genotype	AuN
Goldstein JA	1997	Caucasian (European American),	CYP2C19	A, A'	AP, GP*	Un, Sta, Pa	Ass. Drug	AuN, residing
Griese EU	2001	Australian	CYP2C19	M	AP, GP	Com	Genotype	AuN, residing
He N	2002	Chinese (Dai)	CYP2C19	A, A'	GN, GP	Un, Stu	Ass. Drug	3G, AuN
Herrlin K	1998	Tanzanian	CYP2C19	A', O	AN, AP, GN, GP	Un, Sta, Stu	Ass. Drug	AuN
Hoskins	1998	Caucasian	CYP2C19	A, A'	GN, GP	Un	Ass. Drug	Se, 2G
Hu YR	2004	Chinese (Han)	CYP2C19	A, A', I	GN*	Un	Ass. Drug	AuN
Hung CC	2004	Chinese (Taiwanese)	CYP2C19	5, A, A'	GN	Pa	Ass. Drug	AuN
Ieiri I	1996	Japanese	CYP2C19	A, A'	GN	Un	Ass. Drug	AuN
Ieiri I	1997	Japanese	CYP2C19	5	GN*	Un, Sta	Ass. Drug	AuN
Jurima-Romet M	1996	Canadians	CYP2C19	A, A'	GN	Com	Ass. Drug	AuN, residing
Kaneko A	1997	Tanna, Malakula	CYP2C19	A	GN	Com	Ass. Drug	AuN, residing
Kaneko A	1999	Vanuatu (subgroups)	CYP2C19	A, A'	GN**, GP	Com	Genotype DifEthnic	AuN, residing
Kubota T	1996	Japanese	CYP2C19	A, A'	GN, GP	Un	Ass. Drug	AuN
Kubota T	1998	Japanese	CYP2C19	A, A'	GN	Un	Genotype	AuN
Lamba	2000	North Indian	CYP2C19	A, I	AN, GN	Un	Ass. Drug	AuN
Lee J	2004	Korean	CYP2C19	N/A	AP, GP	N/A	Ass. Drug	AuN
Marandi T	1997	Russian	CYP2C19	A, A', O	GN*	Un	Ass. Drug	AuN, 3G
Marinac JS	1996	Black American	CYP2C19	P	GN	Un	Ass. Drug	AuN

Author	Year	Ethnic	CYP	Genotype Method [a]	Data Formats [b]	Sampling Populations [c]	Study Aims [d]	Ethnicity Identification *
Continuous...								
Martin DE	1998	African American, Caucasian	CYP2C19	N/A	GP	Vol	Genotype	AuN
Masimirembwa C	1995	Zimbabweans	CYP2C19	A'	GN	Un, Sta, Stu	Ass. Drug	AuN
Masta A	2003	Jawia, Kinambu, Witupe	CYP2C19	A, A'	GP	Com	Genotype	AuN, residing
Nowak MP	1998	Chinese, CNI	CYP2C19	A, A'	AP	Com	Genotype	Se, 2G, 3G
Persson I	1996	Ethiopian	CYP2C19	A, A'	GN, GP	Un	Ass. Drug	AuN
Roddam PL	2000	Caucasian	CYP2C19	*	GN, GP	Pa, Mult	Ass. Disease	Se
Roh HY	1996a	Korean	CYP2C19	A, A'	GN*	Un, Sta, Stu	Ass. Drug	AuN
Roh HY	1996b	Korean	CYP2C19	A, A'	AN, AP, GNg	Vol, Sta, Stu	Ass. Drug	AuN
Ruas JL	1997	Portuguese (Caucasian)	CYP2C19	A, A'	AN, AP, GN, GP	Ra, Un	Genotype	AuN
Sviri S	1999	Jewish	CYP2C19	original	GN*	Stu	Ass. Drug	AuN
Takakubo F	1996	Japanese	CYP2C19	A, A'	GN	Un, Fa	Ass. Drug	AuN
Tamminga WJ	2001	Caucasian (Dutch)	CYP2C19	A, A'	AP, GP*	Un	Ass. Drug	AuN
Tassaneeyakul W	2002	Thai	CYP2C19	I	GN	Un	Genotype	AuN, 2G, Se
Tsuneoka	1996	Japanese	CYP2C19	A, A'	GN	Ra, Pa, Mult	Genotype	AuN
Xiao ZS	1997	Chinese (Han), Chinese (Bai)	CYP2C19	A, A'	GN, GP	Un, Sta, Stu	Ass. Drug	AuN
Xie HG	1997	Chinese	CYP2C19	A, A'	GNg	Ra, Com	Ass. Drug	AuN, residing
Yamada H	1998	Swedish	CYP2C19	A, A', O	AP, GN	Elder, Un	Genotype	AuN, residing
Yamada S	2001	Chinese, Thai, Vietnamese, Japanese	CYP2C19	A, A'	AN, AP, GP	Un	Genotype DifEthnic	AuN
Yao TW	2001	Chinese	CYP2C19	*	GN	Vol	Ass. Drug	AuN

* Genotype method did not referred to any original publication

Index note of the table found on page 72-73

Table 3.2.1 Summary of selected references for population distribution of ***CYP2C9*** and ***CYP2C19***

Author	Year	Ethnic	CYP	Genotype Method [a]	Data Formats [b]	Sampling Populations [c]	Study Aims [d]	Ethnicity Identification [e]
Alliabi A	2003	Belgian, Beninese	CYP2C9, 19	E	GP	Un	Genotype DifEthnic	AuN, residing
Bozina	2003	Croatians	CYP2C9, 19	5, E	AN, AP, GN, GP	Un	Genotype	AuN, residing
Bravo-Villalta	2005	Bolivian	CYP2C9, 19	2, A, A', J	AP, GP	Un	Genotype DifEthnic	3G, AuN, residing
Brockmoller J	1995	Caucasian (German)	CYP2C9, 19	A', orginal(2c9)	GN, GP	Vol	Ass. Drug	AuN
Gaikovitch	2003	Russian	CYP2C9, 19	A', G, K	GN, GP	Un, Pa	Genotype	AuN
Halling J	2005	Caucasian (Nordic, Faroese)	CYP2C9, 19	A, A', 2	AP, GP	Ra	Ass. Drug	AuN
Hamdy	2002	Egyptian	CYP2C9, 19	N	AP, GN, GP	Un, Sta, Stu	Genotype	AuN
Hashimoto Y	1996	Japanese	CYP2C9, 19	A, A', 5	GN	Pa	Ass. Drug	AuN
Huang Y	2004	Chinese	CYP2C9, 19	8, 9	AN, GN	Pa	Ass.Drug Disease	AuN
Inoue K	1997	Japanese, Caucasian	CYP2C9, 19	5, A, A'	GN	Liver Sample	Genotype and Enzyme	AuN
Inoue K	1998	Japanese, Caucasian	CYP2C9, 19	5, A, A'	individual	Pa, liver donnor	Genotype	AuN
Jose R	2005	Andhra Pradesh, Karnataka,	CYP2C9, 19	4,	AP, GP	Un	Genotype	3G, residing
Kerb R	2001	Turkish	CYP2C9, 19	A, G	GN, GP	Vol	Ass. Drug	AuN, residing
Kimura M	1998	Japanese	CYP2C9, 19	5, A, A'	GN	Un	Genotype	AuN
Mamiya K	1998	Japanese	CYP2C9, 19	5, A, A'	GN	Pa	Ass. Drug	AuN
Odani A	1997	Japanese	CYP2C9, 19	5, A, A'	GN	Pa	Ass. Drug	AuN
Ogawa K	2003	Japanese	CYP2C9, 19	5, 12, A, A'	GN, GP	N/A	Genotype	AuN
Scordo M	2002	Caucasian (Italian)	CYP2C9, 19	3, A'	AP, GN, GP	Pa	Ass. Drug	AuN
Scordo M	2004	Caucasian (Italian)	CYP2C9, 19	3	AP, GN, GP	Ra, Un, Sta, Stu	Genotype	AuN
Thijssen HH	2000	Caucasian	CYP2C9, 19	5, A, A'	GN	Pa, Cohort	Ass. Drug	N/A
Takahashi H	1998	Japanese	CYP2C9, 2C19	2, A, A'	GN, GP	Pa	Ass. Drug	AuN

Index note of the table found on page 72-73

The index note for table 3.2.1:

a: index note for genotype method. The references provided the genotyping of CYP2C9 and CYP2C19 are listed by numbers and alphabet sequence respectively (Appendix 1b).

1. Sullivan-Klose 1996;
2. Nasu K.*et al* 1997
3. Yasar U. *et al* 1999
4. Sviri S. *et al* 1999
5. Wang S.L. *et al* 1995
6. Stubbins M.J. *et al* 1996
7. Coller J.K. *et al* 1999
8. Kristensen V.N. *et al* 2001
9. Ninomiya H. *et al* 2000
10. Veronese M.E. *et al* 1993
11. London S.J. *et al* 1997.
12. Suno M, *et al* 2002
13. Furuya H., *et al* 1995
14. Gill H.J. *et al* 1999
15. Steward D.J. *et al* 1997

A. de Morais S.M., *et al.*1994a
A'. de Morais, S.M., *et al.*, 1994b
B. Persson I. *et al* 1996
C. Ferguson R.J. *et al* 1998
D. Ibeanu G.C. *et al* 1998
E. Hersberger M. *et al* 2001
F. Xiao Z.S. *et al* 1997.
G. Brockmoller J. *et al* 199
H. Bathum L. *et al* 1998
I. Goldstein J.A. and Blaisdell J.1996
J. Kimura M. *et al* 1998
K. Borlak J. and Thum T 2002
L. Ibeanu G.C. *et al* 1998
M. Griese E.U *et al* 1999
N. Hiratsuka M. *et al* 2000
O. Chang M. *et al* 1995
P. Goldstein J.A. *et al* 1994

b: index note for Data format

AN: the number of subjects for each allele
AP: the percentage of subjects carring certain alleles in the study
GN: the number of subjects for each genotype
GN*: the number of subjects for each genotype described in text
GN**: the number of subjects for each genotype provided by author
GNg: the number of subjects for genotype reported in groups, for instance CYP2C19*1/*2 with CYP2C19*1/*3 together or CYP2C19*2/*3 with CYP2C19*3/*3 together
GP: the percentage of subjects for each genotype
GP*: the genotype frequency reported in combined genotypes' group, for instance CYP2C19*1/*2 with CYP2C19*1/*3 together or CYP2C19*2/*3 with CYP2C19*3/*3 together.
Unclear: data reported in text but impossible to derive unambiguous figure for each genotype.
Individual genotypes: each subject's genotype reported

c: index note for sampling population

*: Subjects are suspected driver with drug consumotion
N/A: no detail available
Un: unrelated subject recruited
Sta: Staff of certain university, where author/s is/are working
Stu: Student of certain university, where the author/s is/are working
Ad: advertisement

Vol: volunteers
Ra: randomly selected subjects
Pa: patient from one local clinic or university hospital
Mult: patients from more than 3 clinics or hospitals
Com: subjects recruited from local community, either living in local residence or within county boundary
Cohort: patients selected according to Cohort study design
Other: subjects from previous double blind randomised trial
Graph: genotype prescribed in the pedigree tree graph
UnRa: un-random sample; parts of subjects from previous study
**: the time duration when patient were recruited

d: index note for study aims
Genotype: genotype prevalence studies
Genotype DifEthnic: genotype prevalence studies among different ethnicity/populations
Ass.: genetic association studies
Ass. Drug: association studies about polymorphic genes and variable drug metabolism
Ass. Disease: gene-disease association studies

e: index note for the Ethnicity identification method in the studies
G: Generation, for instance, 3G stands for three generations, 2G stand for two generations
AuN: Author reported ethnicity without clear description of the methodology applied in the ethnicity identification
Residing: subjects are actually residence of certain geographic region
N/A: no precise description of the ethnicity by author, but there is ambiguous indication of ethnicity
Se: subjects reported their ethnicity by themselves based on family history

From the table, it is noticeable that the information about subject recruitment, ethnicity determination, and even allele or genotype data in those studies is limited. Subjects recruited from the same community could be related in some studies or unrelated in other studies, and it is hard to find specific descriptions in many studies. The genotyping methods can be traced back to the first publication of the *CYP2C9/2C19* mutant alleles, but it is usual that studies have applied the method with minor modification. Whilst the *CYP2C9* sequence is similar to *CYP2C19*, there are other very similar CYP2Cs such as *CYP2C18*, *CYP2C8* located closely in the CYP2C gene cluster. It is difficult to amplify the region of interest only with polymerase chain reaction (PCR). In many studies, different PCR-based genotype assays were applied to improve specificity, such as PCR-RFLP (Restriction Fragment Length Polymorphism), PCR-SSLP (Single-Sequence Length Polymorphism), which were derived from early publications of de Morais *et al.* (1994, *CYP2C19*) and Wang *et al.* (1995, *CYP2C9*). Recently, commercial products became available for the analysis of CYP2Cs, such as the *CYP2C9* Mutation Detection Kit (Roche Diagnostics Corp., Basel, Switzerland), and unspecified oligonucleotide based DNA arrays. The comparison of the detail between the original method, modified methods or newly developed technology has not been undertaken, although it would be useful to evaluate the results based on the accuracy and specification of the different methods.

Many authors have defined the ethnicity without further detail; it is not clear whether the subjects defined themselves as Caucasian or any ethnicity by reporting family history, whether the authors actually define it by the geographical location where the subjects were recruited from or whether they applied any other methodology, such as questionnaire to define it. With rapid economy globalization in the world, intermarriage becomes more and more common as results of communication and sharing of different social and cultural lives. Therefore, it is acknowledged that the populations ethnicity in the country with high proportion immigrated people from all over the world will become increasingly difficult to identify.

The table 3.2.1 has been used to further sort the studies according to ethnicity for subsequent analysis of the population distribution of *CYP2C9/2C19*.

3.2.3 Primary data records for the populations distribution of CYP2C9&2C19

The genotype and allele numbers were retrieved from each article for further analysis. In addition to ethnicity, the health status of recruited subjects was recorded for each article. Meanwhile, if any drug or probe drug had been applied in subjects, the drug name was recorded. The details were presented in table 3.2.2 and table 3.2.3 for *CYP2C9* and *CYP2C19* respectively, where sub-tables were divided according to populations' ethnicity.

A number of studies did not provide subjects genotype in detail, which is particularly common in study of polymorphic *CYP2C19*. For instance, few of them grouped subjects with genotype *CYP2C19*1/*2* and *CYP2C19*1/*3* together as heterozygous for mutated allele and wild-type allele; meanwhile, *CYP2C19*2/*2*, *CYP2C19*2/*3*, *CYP2C19*3/*3* was presented as heterozygous and homozygous mutated genotype. The combined description of genotype brings difficulty in the accurate calculation of allele number detected in the population. The further analysis cannot be taken unless the author provided allele frequency (number) independently. Additionally, some studies provided allele data only without any detail of genotype; these were excluded for testing the goodness of fit to the Hardy-Weinberg equilibrium.

Furthermore, among three African population studies of *CYP2C9*, in addition to the mutant alleles **1*, **2* and **3*, two studies had found other variants at relatively high frequency. Dickmann *et al* (2001) found 4 subjects with one CYP2C9*5 allele within 120 subjects. Meanwhile, Allabi *et al* (2003) found that 4 subjects were carrying one *CYP2C9*5* and 6 subjects were carrying one *CYP2C9*11* allele in a study of 111 healthy subjects. Their results indicate that the genetic model of C*YP2C9*1*, **2* and **3* may not be appropriate for African populations. However, there were no more relevant studies of *CYP2C9* genotype in African populations found before this thesis was finished.

Table 3.2.2a: The retrieved distribution of ***CYP2C9*** polymorphism for **Caucasian** populations

General information for each study					Genotype data								Allele data				
Author	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C9 *1/*1**	**CYP2C9 *1/*2**	**CYP2C9 *1/*3**	**CYP2C9 *2/*2**	**CYP2C9 *2/*3**	**CYP2C9 *3/*3**	**Other**	**Total**	**CYP2C 9*1**	**CYP2C9 *2**	**CYP2C9 *3**	**Other**	**Total**
Aithal GP	1999	Caucasian	patients from anticoagulant clinic	warfarin	32	9	10	1	0	0		52	83	11	10		104
Aithal GP	2000	Caucasian	patients with diclofenac hepatotoxicity	diclofenac	17	5	2	0	0	0		24	41	5	2		48
Aithal GP	2000	Caucasian	healthy	N/A	60	20	17	0	2	1		100	157	22	21		200
Allabi A	2003	Caucasian (Belgian White)	not certain	N/A	81	22	14	0	2	1	1(*1/11)	121	199	24	18	1(*11)	242
Bozina	2003	Caucasian (Croatians)	not certain	N/A	148	45			7			200	341	59	0		400
Brockmoller J	1995	Caucasian	healthy	mephenytoin S-	95	30	?	2	?	?		127	220	34	0		254
Burian M	2002	Caucasian (German)	not certain	N/A	75	30	11	1	1	0		118	191	33	12		236
Caraco	2001	Caucasian	healthy	phenytoin	18	7	4	1	1	0		31	47	10	5		62
Dickmann L	2001	Caucasian (European American)	not certain	warfarin, diclofenac,	94	31	12	3	0	0		140	231	37	12		280
Dorado P	2003	Caucasian (Spanish white)	healthy	diclofenac	59	20	14	4	4	1		102	152	32	20		204
Gaedigk A	2001	Caucasian (White)	patients under warfarin treatment before orthopaedic surgery	N/A	196	74	41	9	5	0		325	507	97	46		650
Gaikovitch	2003	Caucasian (Russian)	71 healthy 219 out-patients	N/A	197	53	33	2	4	1		290	480	61	39		580
Garcia ME	2004	Caucasian (Spanish white)	healthy	ibuprofen	69	34	11	4	7	5		130	183	49	28		260
Garcia ME	2001	Caucasian (White Spanish)	healthy	N/A	78	25	37	3	14	0		157	218	45	51		314
Hallberg	2002	Caucasian (White Swedish)	patient with hypertension and LVH	Irbesartan atenolol	68	20	10	2	2	0		102	166	26	12		204
Halling J	2005	Caucasian (Nordic, Faroese)	not certain	mephenytoin (Racemic)	221	53	32	0	5	0		311	527	58	37		622
Herman D	2005	Caucasian (Slovenian)	outpatients on warfarin maintenance therapy	warfarin	118	32	27	2	6	3		188	295	42	39		376
Hillman WA	2004	Caucasian	patients under warfarin treatment	warfarin	295	86	55	7	8	2		453	731	108	67		906
Ho PC et al	2003	Caucasians	not certain	valproic acid	12	9	9	3	5	1		39	42	20	16		78
Hummers-Pradier E [a]	2003	Caucasian (German)	patients under anticoagulant treatment	phenprocoumon	132	32	14	0	1	0		179	310	33	15		358
Inoue K	1997	Caucasian	not certain	tobutamide	32	10	3	0	0	0		45	77	10	3		90

General information for each study					Genotype data								Allele data				
Author	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C9 *1/*1**	**CYP2C9 *1/*2**	**CYP2C9 *1/*3**	**CYP2C9 *2/*2**	**CYP2C9 *2/*3**	**CYP2C9 *3/*3**	**Other**	**Total**	**CYP2C 9*1**	**CYP2C9 *2**	**CYP2C9 *3**	**Other**	**Total**
Continuous...																	
London SJ	1996	Caucasian (non-hispanic)	patients with lung cancer	N/A	127	47	N/A	3	N/A	N/A		177	301	53	N/A		354
London SJ	1996	Caucasian (non-hispanic)	not certain	N/A	372	86	N/A	3	N/A	N/A		461	830	92	N/A		922
Martin J H [b]	2001	Caucasian	Rheumatology out-patient	NSAIDs	16	4	3	0	0	0		23	39 (35)	4	3		46 (42)
Martin J H	2001	Caucasian	patients with gastric ulcers	N/A	18	9	4	0	0	0		31	49	9	4		62
Martinez C	2001	Caucasian (White Spanish)	healthy	N/A	75	24	35	3	13	0		150	209	43	48		300
Martinez C	2001	Caucasian (White Spanish)	patients with colorectal cancer	N/A	80	19	19	4	7	0		129	198	34	26		258
Ozawa S	1999	Caucasian (Hungarians)	patients with lung disease	N/A	Not available								212	47	27		286
Pchelina SN	2005	Caucasian (Slavic, Russian)	out-patients under initial warfarin therapy	warfarin	46	10	5	0	1	0		62	107	11	6		124
Pchelina SN	2005	Caucasian (Slavic)	186 healthy male, 112 out-patients	62/289 warfarin	202	54	33	4	4	1		298	491	66	39		596
Pedersen RS	2004	Caucasian (Danish)	healthy	N/A	190	53	23	4	0	6		276	456	67	29		552
Peyvandi F	2004	Caucasian (Italian)	patients under initial oral anticoagulant treatment	warfarin	75	29	17	0	4	0		125	196	33	21		250
Scordo MG	2001	Caucasians (Italian)	healthy	N/A	102	24	22	4	2	3		157	250	34	30		314
Scordo MG	2002	Caucasians (Italian)	patients with cardiovascular disease	warfarin	54	15	16	2	4	2		93	139	23	24		186
Scordo MG	2004	Caucasians (Italian)	healthy	N/A	223	62	52	10	8	5		360	560	90	70		720
Stubbins MJ	1996	Caucasian	healthy	N/A	62	19	15	3	0	1		100	158	25	17		200
Sullivan-Klose TH	1996	Caucasian (European American)	healthy	tolbutamide warfarin	Not available								172	16	12		200
Takahashi H	2003	Caucasians	patients with stable warfarin dose	warfarin	26	12	4	4	1	0		47	68	21	5		94
Thijssen HH [a]	2000	Caucasian	patients under anticoagulant treatment	acenocoumarol	20	7	7	0	1	0		35	54	8	8		70
Topic E	2004	Caucasian (Croatians)	patient with thromboembolism	warfarin	104	49	5	18	4	1		181	262	89	11		362
Topic E	2004	Caucasian (Croatians)	healthy	warfarin	122	41	1	12	1	0		177	286	66	2		354
van der Weide et al	2001	Caucasian (Dutch)	patients with epilepsy	phenytoin	37	9	9	3	2	0		60	92	17	11		120
Visser LE	2004	Caucasian	patients under anticoagulant treatment	acenocoumarol phenprocoumon	771	239	73	23	18	0		1124	1854	303	91		2248

General information for each study					Genotype data								Allele data				
Author	Year	Ethnicity	Health status	Drug	CYP2C9 *1/*1	CYP2C9 *1/*2	CYP2C9 *1/*3	CYP2C9 *2/*2	CYP2C9 *2/*3	CYP2C9 *3/*3	Other	Total	CYP2C 9*1	CYP2C9 *2	CYP2C9 *3	Other	Total
Continuous...																	
Visser LE	2004	Caucasian	patients under anticoagulated treatment	acenocoumarol phenprocoumon	685	210	63	23	15	0		996	1643	271	78		1992
Yang JQ	2003	Caucasian (French)	healthy	N/A	94	30	16	4	6	1		151	234	44	24		302
Yasar	1999	Caucasian (Swedish)	healthy	N/A	287	80	50	2	8	3		430	704	92	64		860
Yasar U	2002	Caucasian (Swedish)	Control group [c, d]	N/A	1000	262	163	17	22	4		1468	2425	318	193		2936
Yasar U	2003	Caucasian (Swedish)	Control group [c]	N/A	1025	269	165	18	21	5		1503	2484	326	196		3006
Yasar U	2003	Caucasian (Swedish)	patients survived from first-time acute myocardial infarction	N/A	762	223	129	23	23	12		1172	1876	292	176		2344

Note: a. author didn't describe subjects' ethnicity directly; b. this study included on Caucasian subjects with genotype *CYP2C9*1*1*, they didn't account in further analysis; c. the control group was from the Stockholm Heart Epidemiology Program, a population based case-control study aiming to investigate the different risk factors for myocardial infarctions in men and women in Sweden; d. the study's data was excluded due to high possibility of duplicating sample in a later study; ? indication that allele *CYP2C9*3* was not studied or provided. N/A: not available.

Table 3.2.2b: The retrieved distribution of ***CYP2C9*** polymorphism for **Chinese** populations

General information for each study					Genotype data								Allele data				
Author	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C9 *1/*1**	**CYP2C9 *1/*2**	**CYP2C9 *1/*3**	**CYP2C9 *2/*2**	**CYP2C9 *2/*3**	**CYP2C9*3/*3**	**Other**	**Total**	**CYP2C9 *1**	**CYP2C9 *2**	**CYP2C9 *3**	**Other**	**Total**
Chen Kun	2005	Chinese	healthy	tolbutamide	159	0	9	0	0	1		169	327	0	11		338
Gaedigk A	2001	Chinese	not certain	N/A	91	0	11	0	0	0		102	193	0	11		204
Huang Y	2004	Chinese	patients with epilepsy	phenytoin	28	0	4	0	0	0		32	60	4	0		64
Hung CC	2004	Chinese	patients with epilepsy	phenytoin	151	0	18	0	0	0		169	320	0	18		338
Sullivan-Klose TH	1996	Chinese (Taiwanese)	maternity patients	tolbutamide warfarin	Not available								191	0	5		196
Wang SL	1995	Chinese Han	healthy	N/A	111	4	Not available					115	226	4	0		230
Yang JQ	2003	Chinese Han	helicobacter pylori-related children patients, healthy children, healthy adult	N/A	366	1	26	0	0	1		394	759	1	28		788

N/A: not available

Table 3.2.2c: The retrieved distribution of ***CYP2C9*** polymorphism for **Japanese** populations

General information for each study					Genotype data								Allele data				
Author	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C9 *1/*1**	**CYP2C9 *1/*2**	**CYP2C9 *1/*3**	**CYP2C9 *2/*2**	**CYP2C9 *2/*3**	**CYP2C9 *3/*3**	**Other**	**Total**	**CYP2C9 *1**	**CYP2C9 *2**	**CYP2C9 *3**	**Other**	**Total**
Inoue K [a]	1997 1998	Japanese	not certain	tolbutamide	36	0	3	0	0	0		39	75	0	3		78
Kimura M	1998	Japanese	N/A	N/A	135	0	5	0	0	0		140	275	0	5		280
Mamiya K	1998	Japanese	patients with epilepsy	phenytoin	131	0	3	0	0	0		134	265	0	3		268
Nasu K	1997	Japanese	healthy	N/A	209	0	9	0	0	0		218	427	0	9		436
Ogawa K	2003	Japanese	healthy	N/A	186	0	10	0	0	0		196	382	0	10		392
Takahashi H	2003	Japanese	patients under stable warfarin dose	warfarin	85	0	4	0	0	1		90	174	0	6		180

Note: a. Two publications of the author provided high possibility of duplicating data from same recruited subjects

Table 3.2.2d: The retrieved distribution of ***CYP2C9*** polymorphism for **African** populations

General information for each study					Genotype data								Allele data				
Author	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C9 *1/*1**	**CYP2C9 *1/*2**	**CYP2C9 *1/*3**	**CYP2C9 *2/*2**	**CYP2C9 *2/*3**	**CYP2C9 *3/*3**	**Other**	**Total**	**CYP2C9 *1**	**CYP2C9 *2**	**CYP2C9 *3**	**Other**	**Total**
Allabi A	2003	African (Beninese)	healthy	N/A	101	0	0	0	0	0	10 (4: *1/*5, 6: *1/*11)	111	212	0	0	10 (4: *5, 6: *11)	222
Dickmann L	2001	African (African American)	not certain	warfarin, diclofenac, Lauric Acid	107	6	3	0	0	0	4 (*1/*5)	120	227	6	3	4 (*5)	240
London SJ [a]	1996	African (African American)	lung cancer	N/A	143	9	N/A	0	N/A	N/A		152	295	9	N/A		304
London SJ [a]	1996	African (African American)	not certain	N/A	222	17	N/A	0	N/A	N/A		239	461	17	N/A		478
Scordo MG	2001	African (Ethiopian)	healthy	N/A	130	13	7	0	0	0		150	280	13	7		300
Sullivan-Klose TH	1996	African (African American)	healthy	tolbutamide warfarin	Not available								197	2	1		200

Note: a. allele *3 was not studied

Table 3.2.2e: The retrieved distribution of ***CYP2C9*** polymorphism for **miscellaneous** populations

General information for each study							Genotype data							Allele data			
Close to	phylogenetic tree	Author	Year	Ethnicity	Health status	Drug	CYP2C9 *1/*1	CYP2C9 *1/*2	CYP2C9 *1/*3	CYP2C9 *2/*2	CYP2C9 *2/*3	CYP2C9 *3/*3	Total	CYP2C9 *1	CYP2C9 *2	CYP2C9 *3	Total
Caucasian	West Asian	Jose R	2004	Andhra Pradesh	healthy	N/A	86	8	21	0	1	0	116	201	9	22	232
Caucasian	West Asian	Jose R	2004	Karnataka	healthy	N/A	82	7	17	3	1	0	110	188	14	18	220
Caucasian	West Asian	Jose R	2004	Kerala	healthy	N/A	97	5	17	0	0	1	120	216	5	19	240
Caucasian	European	Vianna JR	2004	White (Brazil)	healthy	tenoxicam	86	27	18	3	0	2	136	217	33	22	275
Caucasian	European	Sandberg M	2004	no defined (Swedish?)	healthy	losartan	81	19	20	2	2	2	126	201	25	26	252
Caucasian	West Asian	Hamdy	2002	Egyptian	not certain	N/A	164	47	29	6	0	1	247	404	59	31	494
Caucasian	West Asian	Aynacioglu AS et al	1999	Turkish	outpatients, healthy	phenytoin	308	90	86	5	6	4	499	792	106	100	998
Caucasian	West Asian	Babaoglu MO	2004	Turkish	healthy	losartan	58	10	12	3	1	1	85	138	17	15	170
Caucasian	West Asian	Kerb R et al	2001	Turkish	healthy	phenytoin	64	13	15	3	0	1	96	156	19	17	192
Caucasian	West Asian	Yilmaz N	2001	Turkish	healthy	N/A [a]	41	9	10	2	0	2	64	101	13	14	128
Chinese	Southeast Asian	Lee SS	2005	Vietnamese	healthy	N/A	150	0	7	0	0	0	157	307	0	7	314
Japanese	Northeast Asian	Lee S	2003	Korean	patients with anticoagulant	warfarin	90	0	0	0	0	0	90	180	0	0	180
Japanese	Northeast Asian	Yoon	2001	Korean	patients with epilepsy, healthy	N/A	561	0	13	0	0	0	574	1135	0	13	1148
African	African	Vianna JR [b]	2004	black (Brazil)	healthy	tenoxicam	67	5	3	0	2	0	77	142	7	5	154
Amerind	Amerind	Gaedigk A	2001	Canadian Native Indian (CNI)	healthy	N/A	94	7	13	0	0	0	114	208	7	13	228
Special	Arctic	Gaedigk A	2001	Inuit	not certain	N/A	151	0	0	0	0	0	151	302	0	0	302
Special		Loebstein R [c]	2001	Israel	patients with maintenance warfarin dose	warfarin	108	28	18	0	2	0	156	262	30	20	312
Special		Llerena A	2004	Mexican	healthy	N/A	72	15	10	0	1	0	98	169	16	11	196
Mixture	50% Canadian Native Indian (CNI)	Gaedigk A	2001	50% CNI	healthy	N/A	18	3	2	0	1	0	24	41	4	3	48
Mixture	75% Canadian Native Indian (CNI)	Gaedigk A	2001	75% CNI	healthy	N/A	10	2	1	1	0	1	15	23	4	3	30
Mixture	European & Amerind	Bravo-Villalta	2005	Bolivian	healthy	N/A	659	72	44	0	3	0	778	1434	75	47	1556
Mixture	African & European	Vianna JR	2004	Intermediate (Brazil)	healthy	tenoxicam	86	16	15	0	1	0	118	203	17	16	236

General information for each study							Genotype data							Allele data			
Close to	phylogenetic tree	Author	Year	Ethnicity	Health status	Drug	CYP2C9 *1/*1	CYP2C9 *1/*2	CYP2C9 *1/*3	CYP2C9 *2/*2	CYP2C9 *2/*3	CYP2C9 *3/*3	Total	CYP2C9 *1	CYP2C9 *2	CYP2C9 *3	Total
Continuous...																	
Mix-subjects (African-American and Caucasians)		Tabrizi AR	2002	N/A	patients with stable warfarin dose	warfarin	107	22	21	1	1	1	153	257	25	24	306
Mix-subjects (White/African American)		Freeman BD	2000	N/A	cardiovascular inpatients under Warfarin therapy	warfarin	27	8	2	1	0	0	38	64	10	2	76
Mix-subjects (93.2% Caucasian)		Pirmohamed M	2000	N/A	HIV patients [d]	Co trimoxazole	61	19	8	0	1	0	89	149	20	9	178
Mix-subjects (94.6% Caucasian)		Pirmohamed M	2000	N/A	HIV patient [e]	Co trimoxazole	33	10	8	0	5	0	56	84	15	13	112
Mix-subjects (89%Caucasian, 33%African-American, 3other races)		Joffe HV	2004	N/A	patients under warfarin treatment	warfarin	42	16	9	2	2	2	73	109	22	15	146
Mix-subjects (96.2%Caucasian, 3.85 Hispanic)		Higashi MK	2002	N/A	patients under warfarin treatment	warfarin	127	28	18	4	3	5	185	300	39	31	370
Mix-subjects (98%Caucasians with 1 Hispanic, 2% African-American)		Linder MW	2002	N/A	Patients on stable anticoagulation therapy	warfarin	34	13	4	2	3	0	56	85	20	7	112

Note: a. author did not study any allele except *CYP2C9*1, *2* and **3*, but found the genotype data did not fit Hardy-Weinberg equilibrium; b. serum tumour markers, cytokine; c. author did not provide the ethnicity directly; d. HIV patients without sulphamethoxazole hypersensitivity; e. HIV patients with sulphamethoxazole hypersensitivity

Table 3.2.3a: The retrieved distribution of ***CYP2C19*** polymorphism for **Caucasian** populations

General information for each study					Genotype data								Allele data				
AUTHOR	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C19 *1/*1**	**CYP2C19 *1/*2**	**CYP2C19 *1/*3**	**CYP2C19 *2/*2**	**CYP2C19 *2/*3**	**CYP2C19 *3/*3**	**Other**	**SUM**	**CYP2C 19*1**	**CYP2C 19*2**	**CYP2C 19*3**	**Other**	**Sum**
Alliabi A	2003	Caucasian (Belgian White)	not certain	N/A	101	18	0	2	0	0		121	220	22	0		242
Aynacioglu S	1999	Caucasian (German)	healthy	N/A	237	76	1	14	0	0		328	551	104	1		656
Bathum	1998	Caucasian (Danish)	healthy	N/A	43	19	0	2	0	0		64	105	23	0		128
Bathum	1998	Caucasian (Danish)	not certain	N/A	171	59	0	0	9	0		239	401	68	9		478
Bozina	2003	Caucasian (Croatians)	not certain	N/A	146	48			6			200	340	60	0		400
Bramness JG	2003	Caucasian (Norwegian)	not certain	N/A	70	19	0	5	0	0		94	159	29	0		188
Brockmoller J	1995	Caucasian	healthy	mephenytoin (racemic)	102	34	?	4	?	?		140	238	42	?		280
Brosen K	1995	Caucasian (Danish)	healthy	mephenytoin	12	25	1	15	1	0		54	50	56	2		108
Ferguson RJ	1998	Caucasian (French)	Healthy smoker	mephenytoin	99	22	1	7	0	0		130	221	36	1		258
Gaikovitch	2003	Caucasian (Russian)	outpatients, healthy	N/A	228	55	1	5	1	0		290	512	66	2		580
Goldstein JA	1997	Caucasian (European American)	healthy	mephenytoin (racemic)	79	24			2			105	183	27	0		210
Halling J	2005	Caucasian (Nordic, Faroese)	not certain	mephenytoin (racemic)	205	97	0	9	0	0		311	507	115	0		622
Hoskins	1998	Caucasian	healthy	proguanil	73	23	0	3	0	0	*	99	169	29	0	*	198
Marandi T	1997	Caucasian (Russian*)	healthy	mephenytoin S-		32		1	1	0	1 [d]	35	Not available				
Martin DE	1998	Caucasian	healthy	N/A	327	118	0	6	0	0		451	772	130	0		902
Roddam PL	2000	Caucasian	healthy	N/A	713	211	?	28	?	?		952	1637	267	?		1904
Ruas JL	1997	Caucasians (Portuguese)	healthy	N/A	115	36	*0*	2	0	0		153	266	40	0		306
Scordo M	2002	Caucasian (Italian)	cardiovascular outpatients	warfarin	72	19	?	2	?	?		93	163	23	?		186
Scordo M	2004	Caucasians (Italian)	healthy	N/A	286	68	0	6	0	0		360	640	80	0		720
Tamminga WJ	2001	Caucasian (Dutch)	healthy	cocktail [a]	554	163			19			736	Not available				
Yamada H	1998	Caucasian (Swedish)	healthy (elder)	N/A	59	22	1	1	0	0		83	141	24	1		166
Yamada H	1998	Caucasian (Swedish)	healthy	N/A	117	42	0	3	0	0		162	276	48	0		324
de Morais [b]	1994	Caucasian (American)	not certain	mephenytoin S-	8	2	0	2	0	0		12	18	6	0		24
de Morais [b]	1994	Caucasian (Swiss)	not certain	mephenytoin S-	5	5	0	5	0	0		15	15	15	0		30
Chang M [b,c]	1995	Caucasian (Swedish)	healthy	omeprazole	*113*	*41*	*0*	*6*	*0*	*0*	*2*	*160*	267	53	0		320

Note: a. mephenytoin, caffeine, dextromethorphan; b. studies intended to recruit more poor metabolizers to emphasize the mutated alleles, which were excluded for further analysis; c. those data in italics were obtained from other resource, such as reviews, instead of original publications; d. unknown allele; ? the data is not available since authors assumed that *CYP2C19*3* is rare in Caucasian. Some studies did not report genotype number individually, therefore no allele number was available; * author proposed other variants may exist in studied population.

Table 3.2.3b: The retrieved distribution of ***CYP2C19*** polymorphism for **Chinese** populations

General information for each study					Genotype data								Allele data				
AUTHOR	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C19 *1/*1**	**CYP2C19 *1/*2**	**CYP2C19 *1/*3**	**CYP2C19 *2/*2**	**CYP2C19 *2/*3**	**CYP2C19 *3/*3**	**Other**	**SUM**	**CYP2C 19*1**	**CYP2C 19*2**	**CYP2C 19*3**	**Other**	**Sum**
de Morais	1995	Chinese	healthy	mephenytoin S-	25	25	5	15	4	0		74	80	59	9		148
Fu LQ	2004	Chinese Han	healthy	N/A	54	62	5	15	4	0		140	175	96	9		280
Garcia-Barcelo M[a]	1999	Chinese	depression patients, healthy	N/A	Not available								155	73	9	1	238
Goldstein JA	1997	Chinese (Taiwan)	maternity patients	mephenytoin (racemic)	48	52		18				118	148	75	13		236
He N	2002	Chinese Dai	healthy	mephenytoin	81	85	9	15	2	1		193	256	117	13		386
Hu YR	2004	Chinese Han	healthy	lansoprazole	22	30	7	8	3	0		70	81	49	10		140
Huang Y	2004	Chinese	epilepsy patients	phenytoin	11	13	4	3	1	0		32	39	20	5		64
Hung CC	2004	Chinese (Taiwan)	epilepsy patients	phenytoin	53	88	11	10	7	0		169	205	115	18		338
Nowak MP	1998	Chinese	healthy	N/A	Not available								94	40	4		138
Xiao ZS	1997	Chinese Bai	healthy	mephenytoin	*102*	*64*	*9*	*26*			*1(*2*5)*	*202*	*Not available*				
Xiao ZS	1997	Chinese Han	healthy	mephenytoin	*32*	*42*	*7*	*20*				*101*	*Not available*				
Xiao ZS[b]	1997	Chinese Bai	healthy	mephenytoin	102	64	9	15	10	1	1(*2*5)	202	277	105	21	1 (*5)	404
Xiao ZS[b]	1997	Chinese Han	healthy	mephenytoin	32	42	7	13	6	1		101	113	74	15	0	202
Xie HG	1997	Chinese	healthy	mephenytoin (racemic)	23	30		19		0		72	Not available				
Yamada S	2001	Chinese	healthy	no certain	32	58	2	21	7	1		121	121	110	11		242
Yao TW	2001	Chinese	healthy	mephenytoin	6	7	3	9	1	0		26	22	26	4	0	52

Note: a the allele number was calculated from corresponding proportion of each allele; no genotype data was available; b the data was obtained from other publications of same research group instead of the selected article. The genotype data provided in this article is shown in above rows with italic font

Table 3.2.3c: The retrieved distribution of ***CYP2C19*** polymorphism for **Japanese** populations

General information for each study					Genotype data								Allele data				
AUTHOR	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C19 *1/*1**	**CYP2C19 *1/*2**	**CYP2C19 *1/*3**	**CYP2C19 *2/*2**	**CYP2C19 *2/*3**	**CYP2C19 *3/*3**	**Other**	**SUM**	**CYP2C 19*1**	**CYP2C 19*2**	**CYP2C 19*3**	**Other**	**Sum**
Furuta T	2001	Japanese	patients with *H.pylori*-positive gastritis	rabeprazole	34	35	16	8	6	2		101	119	57	26		202
Goldstein JA	1997	Japanese	healthy	mephenytoin	26	19		3	4	1		53	71	24	11		106
Ieiri I	1996	Japanese	healthy	omeprazole	10	5	5	5	2	0		27	30	17	7		54
Ieiri I	1997	Japanese	healthy	mephenytoin	2	0	1	2	1	0		6	5	5	2		12
Kimura M	1998	Japanese	not certain	N/A	44	48	15	12	19	2		140	151	91	38		280
Kubota T	1996	Japanese	healthy	mephenytoin	65	63	23	20	12	3		186	216	115	41		372
Kubota T	1998	Japanese	healthy	N/A	44	57	22	17	11	3		154	167	102	39		308
Mamiya K	1998	Japanese	epilepsy patients	phenytoin	55	47	17	7	5	3		134	174	66	28		268
Ogawa K	2003	Japanese	healthy	N/A	74	61	29	15	16	1		196	238	107	47		392
Takahashi	1998	Japanese	patients with heart disease	warfarin	23	21	12	5	3	2		66	79	34	19		132
Takakubo F[a]	1996	Japanese	healthy	N/A	84	74	26	13	16	4		217	268	119	47		434
Tsuneoka	1996	Japanese	healthy	N/A	30	16	9	3	6	0		64	85	28	15		128
Tsuneoka[b]	1996	Japanese	patients with different disease	N/A	56	60	18	18	16	1		169	190	112	36		338
Yamada S	2001	Japanese	healthy	no certain	35	25	17	7	9	1		96	110	52	30		192

Note: a. study did not report the genotype data directly; the data was retrieved from the text; b: subjects were recruited from patients with various diseases, their genotypes according to different disease group were provided separately

Table 3.2.3d: The retrieved distribution of ***CYP2C19*** polymorphism for **African** populations

General information for each study					Genotype data								Allele data				
AUTHOR	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C19 *1/*1**	**CYP2C19 *1/*2**	**CYP2C19 *1/*3**	**CYP2C19 *2/*2**	**CYP2C19 *2/*3**	**CYP2C19 *3/*3**	**Other**	**SUM**	**CYP2C 19*1**	**CYP2C 19*2**	**CYP2C 19*3**	**Other**	**Sum**
Aklillu E	2002	African (Ethiopian (Ethiopia))	healthy	mephenytoin (racemic)	85	22	1	3	3	0		114	193	31	4		228
Aklillu E	2002	African (Ethiopian (Sweden))	healthy	mephenytoin (racemic)	49	17	4	0	0	0		70	119	17	4		140
Alliabi A	2003	African (Beninese)	healthy	N/A	82	29	0	0	0	0		111	193	29	0		222
Bathum	1999	African (Tanzania)	healthy	mephenytoin (racemic)	160	32	0	2	0	0		194	352	36	0		388
Dandara	2001	African (Tanzanian)	healthy	N/A	72	30	0	4	0	0		106	174	38	0		212
Dandara	2001	African (Tanzanian)	psychiatric patients	N/A	56	27	0	2	1	0		86	139	32	1		172
Dandara	2001	African (Zimbabwean)	not certain	N/A	65	16	0	3	0	0		84	146	22	0		168
Dandara	2001	Venda	not certain	N/A	47	25	0	4	0	0		76	119	33	0		152
Edeki TI	1996	African (African-American)	healthy	mephenytoin (racemic)	48	27	0	1	0	0		76	123	29	0		152
Goldstein JA	1997	African (African-American)	healthy	mephenytoin (racemic)	60	40		8				108	162	54	0		216
Herrlin K	1998	African (Tanzanian)	healthy	mephenytoin omeprazole	166	75	2	7	1	0		251	409	90	3		502
Marinac JS	1996	African (Black American)	healthy	omeprazole	70	28	0	2	0	0		100	168	32	0		200
Martin DE	1998	African (African-American)	healthy	N/A	164	60	1	8	0	0		233	389	76	1		466
Masimirembwa C	1995	African (Zimbabweans)	healthy	mephenytoin	65	16	?	3	?	?		84	146	22	?		168
Persson I	1996	African (Ethiopians)	healthy	mephenytoin	85	22	1	3	3	0		114	193	31	4		228

Note: ? data is not available; the study was excluded for further analysis

Table 3.2.3e: The retrieved distribution of ***CYP2C19*** polymorphism for **Pacific Island** populations

General information for each study					Genotype data								Allele data				
Close to	Phylogenetic tree	AUTHOR	Year	Ethnicity	CYP2C19 *1/*1	CYP2C19 *1/*2	CYP2C19 *1/*3	CYP2C19 *2/*2	CYP2C19 *2/*3	CYP2C19 *3/*3	Other	SUM	CYP2C19 *1	CYP2C19 *2	CYP2C19 *3	Other	Sum
Northern-central Vanuatu	Pacific Islander	Kaneko A	1999	A-Gaua	26	138	36	143	45	11		399	226	469	103		798
	Pacific Islander	Kaneko A	1999	B-Gaua	9	43	14	43	27	13		149	75	156	67		298
	Pacific Islander	Kaneko A	1999	C-Santo	10	44	14	66	49	7		190	78	225	77		380
	Pacific Islander	Kaneko A	1999	D-Maewo	20	65	12	83	33	10		223	117	264	65		446
	Pacific Islander	Kaneko A	1999	E-Maewo	3	34	10	56	54	1		158	50	200	66		316
	Pacific Islander	Kaneko A	1999	F-Santo	2	40	16	120	69	11		258	60	349	107		516
	Pacific Islander	Kaneko A	1999	G-Malo	7	18	17	15	27	9		93	49	75	62		186
	Pacific Islander	Kaneko A	1999	H-Pentecost	44	185	35	263	101	9		637	308	812	154		1274
	Pacific Islander	Kaneko A	1999	I-Malakula	20	107	36	106	95	22		386	183	414	175		772
	Pacific Islander	Kaneko A	1999	J-Malakula	3	19	9	46	18	1		96	34	129	29		192
	Pacific Islander	Kaneko A	1999	K-Tongoa	11	55	15	108	49	5		243	92	320	74		486
	Pacific Islander	Kaneko A	1999	L-Emae	7	44	15	79	34	4		183	73	236	57		366
Polynesian	Pacific Islander	Kaneko A	1999	M-Emae*	3	24	8	33	21	0		89	38	111	29		178
Northern-central Vanuatu	Pacific Islander	Kaneko A	1999	N-Emae	5	20	2	47	10	0		84	32	124	12		168
	Pacific Islander	Kaneko A	1999	O-Nguna	4	26	12	29	27	19		117	46	111	77		234
Polynesian	Pacific Islander	Kaneko A	1999	P-Efate*	15	89	21	101	42	7		275	140	333	77		550
	Pacific Islander	Kaneko A	1999	Q-Ifira*	11	36	14	64	33	7		165	72	197	61		330
Southern Vanuatu	Pacific Islander	Kaneko A	1999	R-Erromango	8	72	16	274	77	7		454	104	697	107		908
Polynesian	Pacific Islander	Kaneko A	1999	S-Aniwa*	35	112	16	145	36	2		346	198	438	56		692
	Pacific Islander	Kaneko A	1999	T-Futuna*	46	111	23	84	26	2		292	226	305	53		584
Southern Vanuatu	Pacific Islander	Kaneko A	1999	U-Tanna	6	48	6	148	46	4		258	66	390	60		516
	Pacific Islander	Kaneko A	1999	V-Tanna	6	15	4	25	7	0		57	31	72	11		114
Polynesian	Pacific Islander	Kaneko A	1999	W-Aneityum*	7	31	0	23	8	0		69	45	85	8		138
Southern Vanuatu	Pacific Islander	Kaneko A	1999	X-Aneityum	16	89	3	199	10	0		317	124	497	13		634
New Guinean	Papua New Guinea	Masta A	2003	Jawia	8	33	9	12	13	0		75	58	70	22		150
New Guinean	Papua New Guinea	Masta A	2003	Kiniambu	21	37	26	25	22	2		132	106	108	50		264
New Guinean	Papua New Guinea	Masta A	2003	Witupe	29	74	17	39	33	4		196	147	182	54		384
Vanuatu	Pacific Islander	Kaneko A	1997	Malakula	6	49	18	103	49	2		227	79	304	71		454
Vanuatu	Pacific Islander	Kaneko A	1997	Tanna	6	60	6	144	46	4		266	78	394	60		532
Australian Aborigine	Australian	Griese EU	2001	Australian Aborigine	59	83	27	26	26	6		227	228	161	65		454

Note: the data was presented in two sections, where the first section presented ***CYP2C19*** data of 24 sub-grouped populations from Kaneko (1999), and the second section presented ***CYP2C19*** data of similar populations from other studies; only Masta (2003) studied the metabolism of proguanil in recruited subjects, the other studies did ***CYP2C19*** genotype only; the health status of subjects was not provided.

Table 3.2.3f: The retrieved distribution of ***CYP2C19*** polymorphism for **miscellaneous** populations

General information for each study							Genotype data							Allele data			
Close to	**Phylogenetic tree**	**AUTHOR**	**Year**	**Ethnicity**	**Health status**	**Drug**	**CYP2C19 *1/*1**	**CYP2C19 *1/*2**	**CYP2C19 *1/*3**	**CYP2C19 *2/*2**	**CYP2C19 *2/*3**	**CYP2C19 *3/*3**	**SUM**	**CYP2C 19*1**	**CYP2C 19*2**	**CYP2C 19*3**	**Sum**
Caucasian	West Asian	Jose R	2004	Andhra Pradesh	healthy	N/A	53	48	0	14	0	0	115	154	76	0	230
Caucasian	West Asian	Jose R	2004	Karnataka	healthy	N/A	42	47	0	18	1	0	108	131	84	1	216
Caucasian	West Asian	Jose R	2004	Kerala	healthy	N/A	55	50	1	11	1	0	118	161	73	2	236
Caucasian	West Asian	Adithan	2003	Tamilians	healthy	N/A	33	65	3	9	2	0	112	134	85	5	224
Caucasian	West Asian	Lamba	2000 1998	North Indian	healthy	omeprazole	58	54	9	0	0	0	121	179	54	9	242
Caucasian	West Asian	Hamdy	2002	Egyptian	not certain	N/A	194	50	1	2	0	0	247	439	54	1	494
Caucasian	West Asian	Aynacioglu S	1999	Turkish	272outpatient 132healthy	N/A	307	90	3	4	0	0	404	707	98	3	808
Caucasian	West Asian	Kerb R	2001	Turkish	healthy	phenytoin	75	16	0	3	0	0	94	166	22	0	188
Chinese	Southeast Asian	Goldstein JA	1997	Filipinos	healthy	mephenytoin	16	24		7	0	5	52	56	40	8	105
Chinese	Southeast Asian	Tassaneeyakul W	2002	Thai	healthy	omeprazole	51	47	3	5	0	1	107	152	57	5	214
Chinese	Southeast Asian	Yamada S	2001	Thai	healthy	no certain	45	49	8	16	3	0	121	145	85	12	242
Chinese	Southeast Asian	Yamada S	2001	Vietnamese	healthy	no certain	36	28	8	6	8	4	90	105	50	25	180
Japanese	Northeast Asian	Lee J	2004	Korean	not certain	omeprazole	118	100	24	25	11	4	282	360	161	43	564
Japanese	Northeast Asian	Roh HY	1996	Koreans*	healthy	omeprazole	48	43		12			103	139	43	24	206
Amerind	Amerind	Nowak MP	1998	Canadian Native Indian	healthy	N/A	Not available							186	44	0	230
Special	Arctic	Jurima-Romet M	1996	Inuit	healthy	mephenytoin	121	28	0	3	0	0	152	270	34	0	304
Special		Goldstein JA	1997	Saudi Arabian	healthy	mephenytoin	70	25		2	0	0	97	165	29	0	194
Special		Sviri S	1999	Jewish	healthy	mephenytoin	99	35	2	4			140	235	43	2	280
Special	Papua New Guinea	Masta A [a]	2003	Jawia	not certain	proguanil	8	33	9	12	13	0	75	59	71	23	152
Special	Papua New Guinea	Masta A	2003	Kiniambu	not certain	proguanil	21	37	26	25	22	2	132	106	108	50	264
Special	Pacific Islander	Kaneko A	1997	Malakula	not certain	N/A	6	49	18	103	49	2	227	79	304	71	454
Special	Pacific Islander	Kaneko A	1997	Tanna	not certain	N/A	6	60	6	144	46	4	266	78	394	60	532
Special	Papua New Guinea	Masta A [a]	2003	Witupe	not certain	proguanil	29	74	17	39	33	4	196	147	182	54	384
Special	Australian	Griese EU	2001	Australian Aborigine	not certain	N/A	59	83	27	26	26	6	227	228	161	65	454
Mixture	European & Amerind	Bravo-Villalta	2005	Bolivian	healthy	N/A	664	105	1	8	0	0	778	1434	121	1	1556
Mixture	25% CNI	Nowak MP	1998	25% Canadian Native Indian	healthy	N/A	Not available							33	3	0	36
Mixture	50% CNI	Nowak MP	1998	50% Canadian Native Indian	healthy	N/A	Not available							39	9	0	48
Mixture	75% CNI	Nowak MP*	1998	75% Canadian Native Indian	healthy	N/A	Not available							4	0	0	4

Note: a. rounding-up error was obvious after data was calculated from proportion (percentage);

3.3 Statistical analysis

3.3.1 Allele/genotype Proportion in each study

In the selected studies, there are allele numbers (frequencies) or genotype numbers (frequencies) reported. For three alleles, **1, *2*, and **3*, there are six possible combinations of genotypes, **1*1, *1*2, *1*3, *2*2, *2*3, *3*3*. The data of allele/genotype frequency was retrieved from every study. Then, the data of allele frequency and genotype were categorized into Caucasian, Chinese, Japanese, Negroid, and Miscellaneous groups for further analysis as shown in table 3.2.2, where the Miscellaneous group included the small number of studies that provided *CYP2C9, 2C19* alleles or genotype data for people with different ethnic origin from Caucasian, Chinese, Japanese and African.

Of the studies describing *CYP2C9* genotype, there are 49 Caucasian populations, 7 Chinese populations, 5 Japanese populations, 6 African populations and 29 miscellaneous populations.

Of the studies describing *CYP2C19* genotype, there are 25 Caucasian populations, 14 Chinese populations, 16 Japanese populations, 16 African populations and 28 miscellaneous populations that were identified in the period to July 2005. In order to estimate the prevalence of *CYP2C9/2C19*1, *2* and **3* in different ethnic populations, the numbers of each allele and genotype in recruited populations were required for further analysis. However, primary data from each study were not always completed (table 3.3.1)

Table 3.3.1: The availability of data in different populations

Ethnicity	CYP2C9		*CYP2C19*	
	Allele data	Genotype data	Allele data	Genotype data
Caucasian	49	47	24	25
Chinese	7	6	14	10
Japanese	5	5	14	15
African	6	5	16	15
Pacific Islander	--	--	24	24
Miscellaneous	29	29	28	21

For the purpose of allele prevalence estimation, the data of six *CYP2C9* and *CYP2C19* genotypes are very essential for further analysis; while the allele data can be easily calculated when genotype data is fully available for each of *CYP2C9* and *CYP2C19* genotypes. Most studies retrieved in the thesis provided data of six possible *CYP2C9/2C19* genotypes and three *CYP2C9/2C19* alleles. However, as shown in table 3.3.1, some studies did not provide completed data sets of genotypes. For instance, in the 49 Caucasian studies of *CYP2C9*, two had provided the allele data but no detail of genotype data. In such case, it is impossible to derive the genotype data without presuming the population in the study fitting HWE. Therefore, such studies without retrievable genotype data were excluded for further analysis. Furthermore, in 25 Caucasian studies of *CYP2C19*, one study reported genotype data with combining *CYP2C19*1/*2* and *CYP2C19*1/*3* or *CYP2C19*2/*2*, *CYP2C19*2/*3* and *CYP2C19*3/*3* together; consequently data of three *CYP2C19* alleles could not be retrieved for this study. In such case, no any further analysis could be applied. Therefore, the study was excluded due to lack of required data.

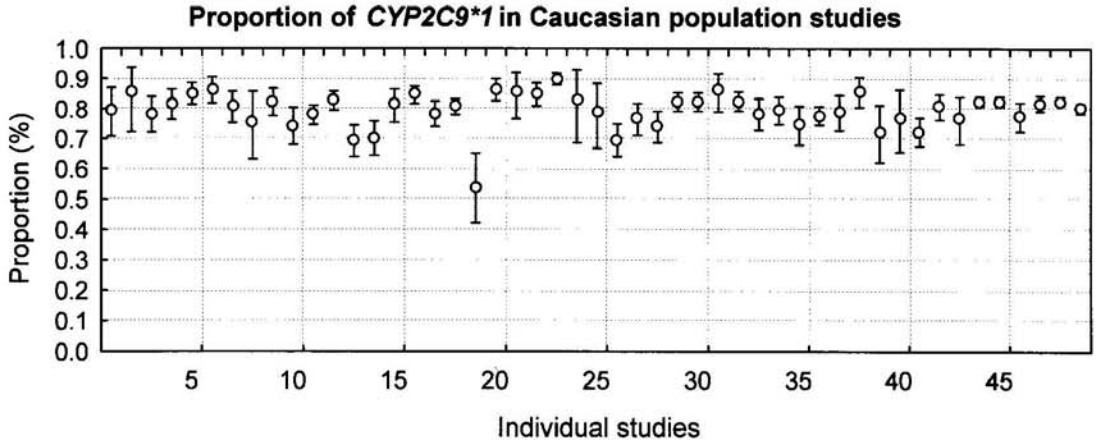
Overall, studies provided retrievable genotype frequency for six *CYP2C9/2C19* genotypes were valid for the Hardy-Weinberg equilibrium test and further statistical analysis. Those studies with only allele frequency available were presented in relevant plots, but they were excluded from any further analysis.

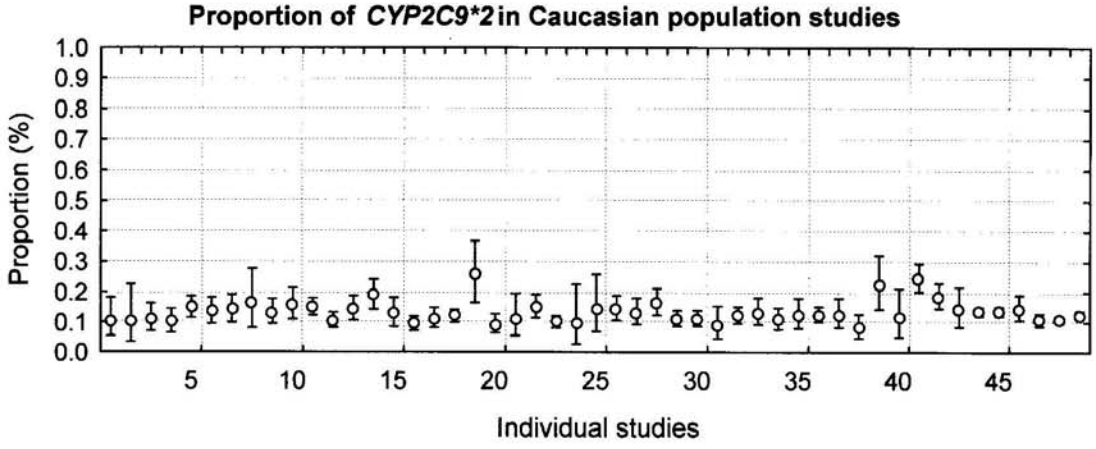
For any individual from any population, there are two possible results that either the person has the allele or genotype or he/she has not. This is considered as Binomial distribution. Therefore, the alleles/genotype proportion and 95% confidence interval (CI) are calculated in statistics software MinTab13.1 based on a Binomial distribution and under summarized data option.

The result of allele/genotype proportion is summarized in figure 3.3.1 and figure 3.3.2 for 74 studies of *CYP2C9*, and 76 studies of *CYP2C19* respectively. The data has been grouped as ethnic categories, such as Caucasian, Chinese, Japanese and African. Those studies not falling into the four ethnic groups were combined as the miscellaneous group.

Furthermore, some *CYP2C19* studies provided data for Pacific Islander populations, and they have been categorized separately and shown in figure 3.3.3. In study of Kaneko *et al.* (1999), one of Pacific Islander populations, Vanuatu population, was further divided into 24

subpopulations according to their language and geographical origin, which are presented in figure 3.3.3a. Meanwhile, Kaneko *et al.* (1999) have also combined these 24 subpopulations into three groups according to their geographical origin only, for instance Polynesian subjects, North-central Vanuatu subjects, and Southern Vanuatu subject; which were separately presented with entire Vanuatu subjects of this study in figure 3.3.2b. Additionally, Masta *et al.* (2003) and Kaneko *et al.* (1997) also provided *CYP2C19* genotypes for these Pacific Islander populations, which were shown in figure 3.3.4 with Kaneko *et al.* (1999) together.





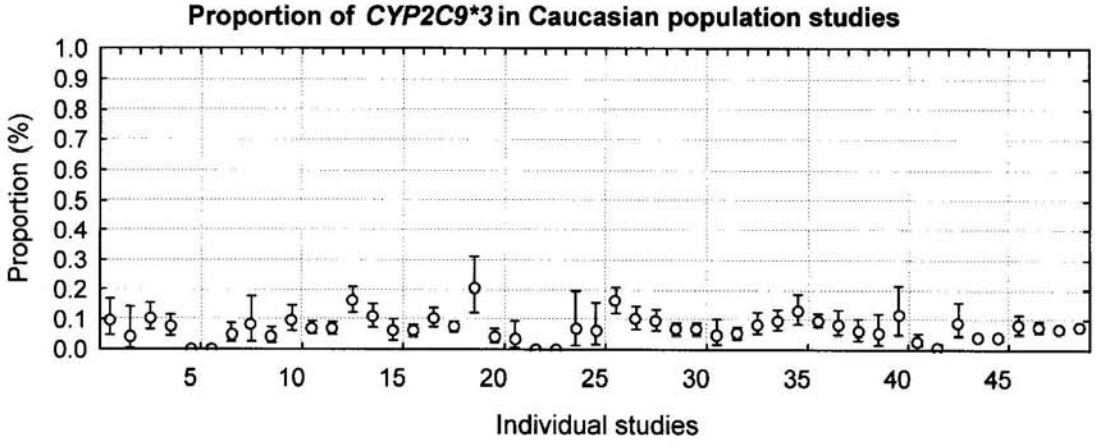
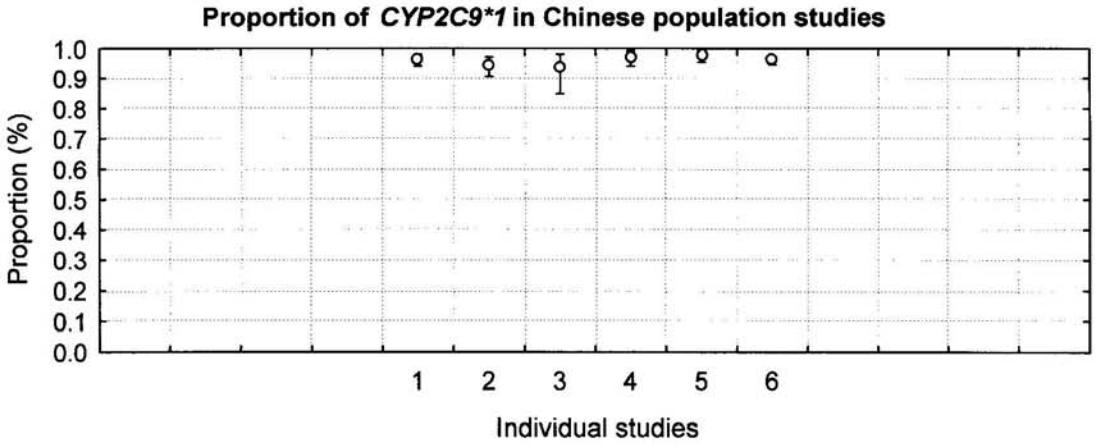


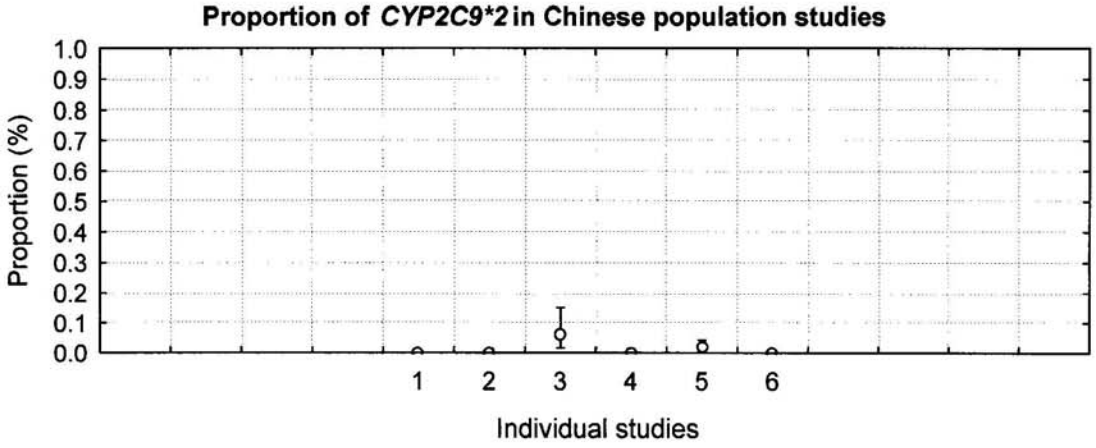
Figure 3.3.1: Proportion of CYP2C9 allele in different ethnic populations

O: Proportion of *CYP2C9* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





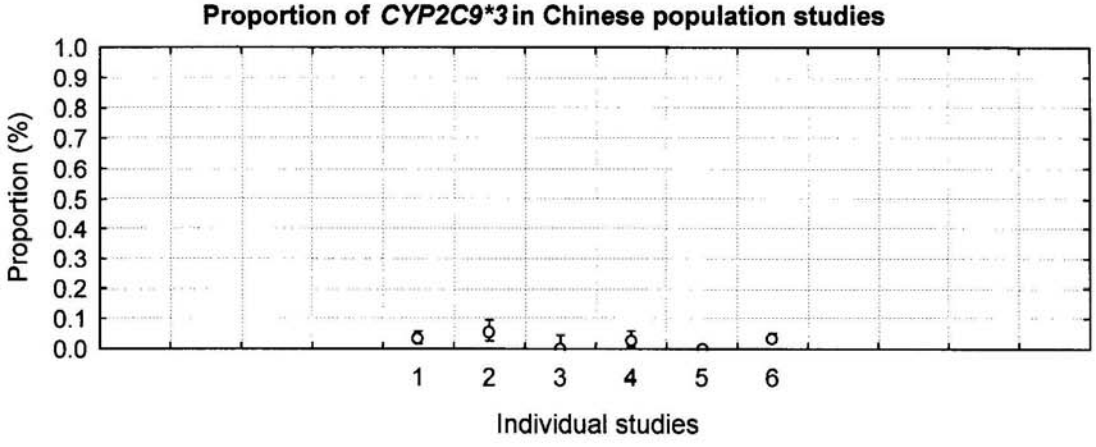
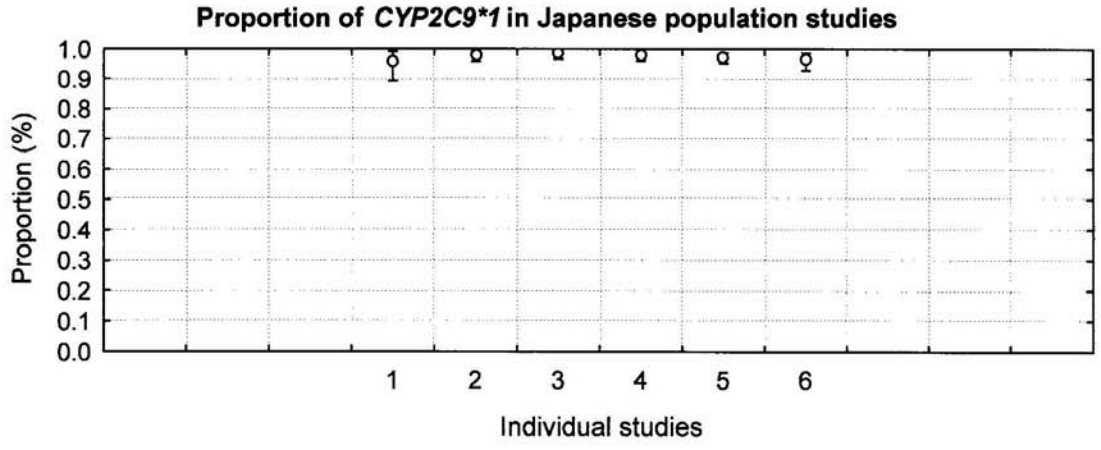


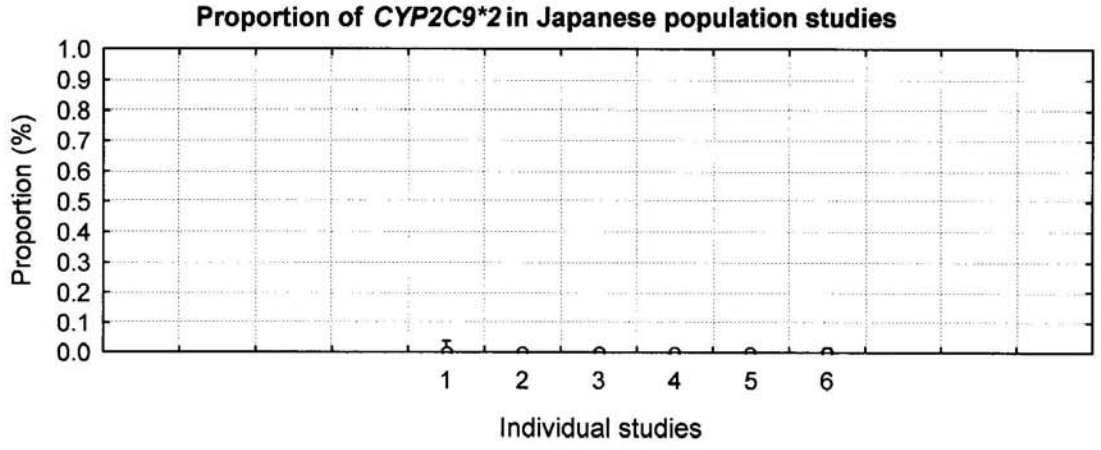
Figure 3.3.1: Proportion of CYP2C9 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C9* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





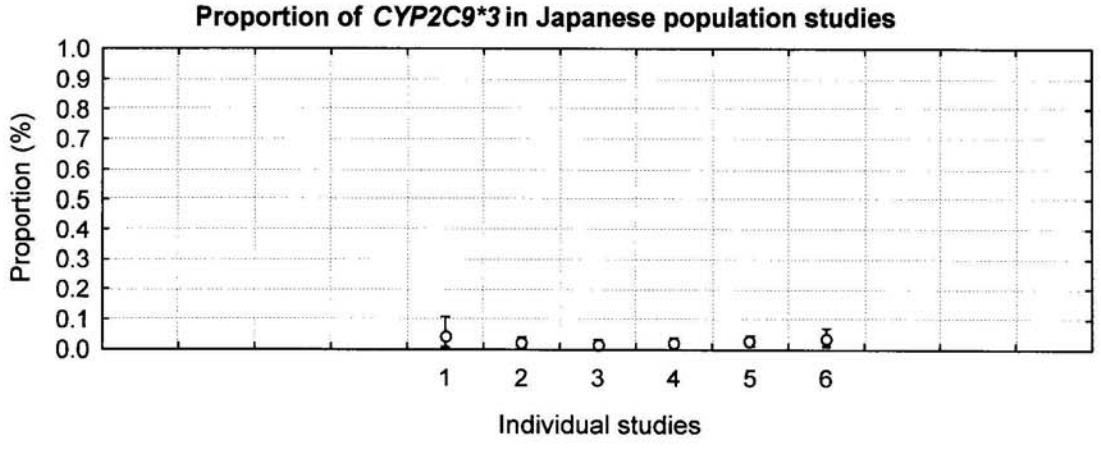
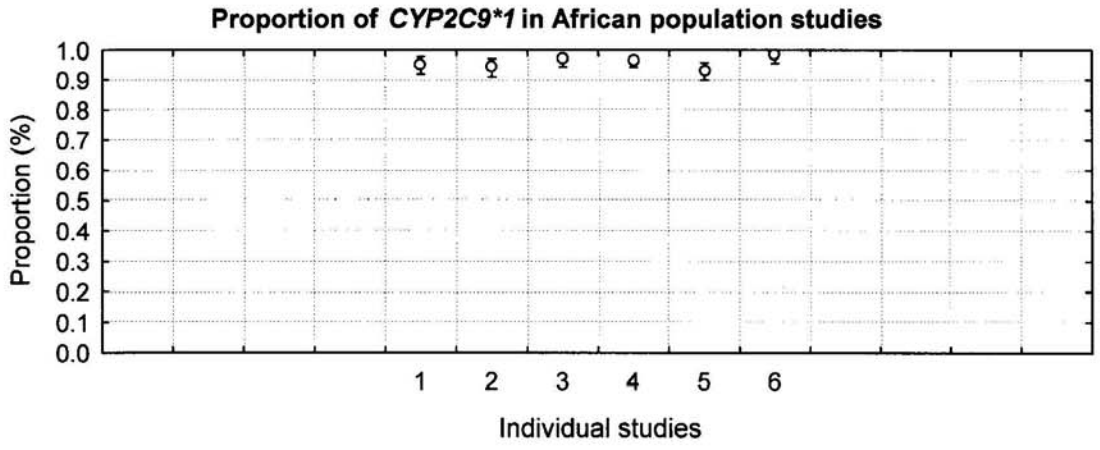


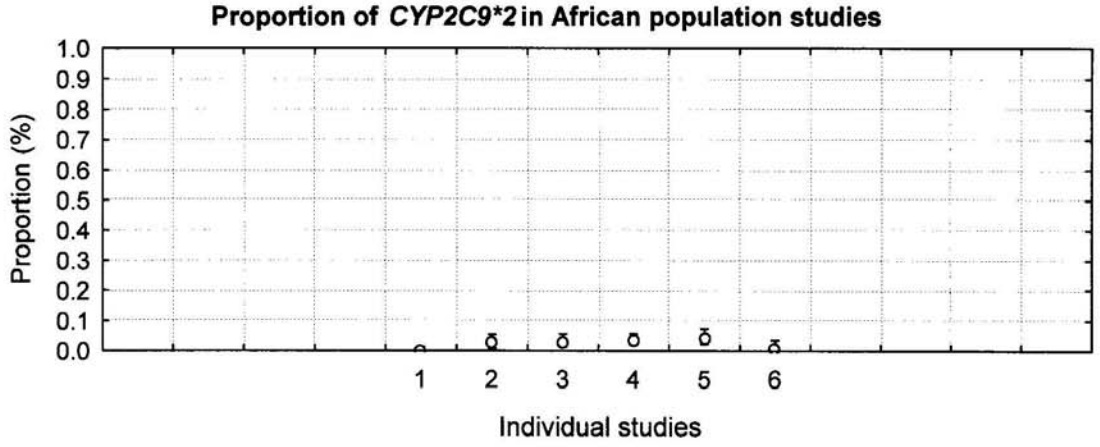
Figure 3.3.1: Proportion of CYP2C9 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C9* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





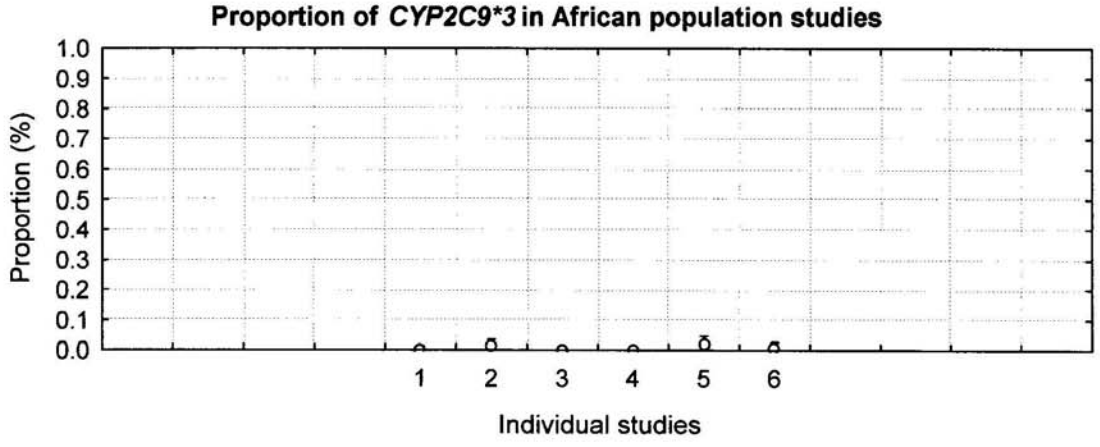


Figure 3.3.1: Proportion of CYP2C9 in different ethnic populations (Continuous)

O: Proportion of *CYP2C9* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%

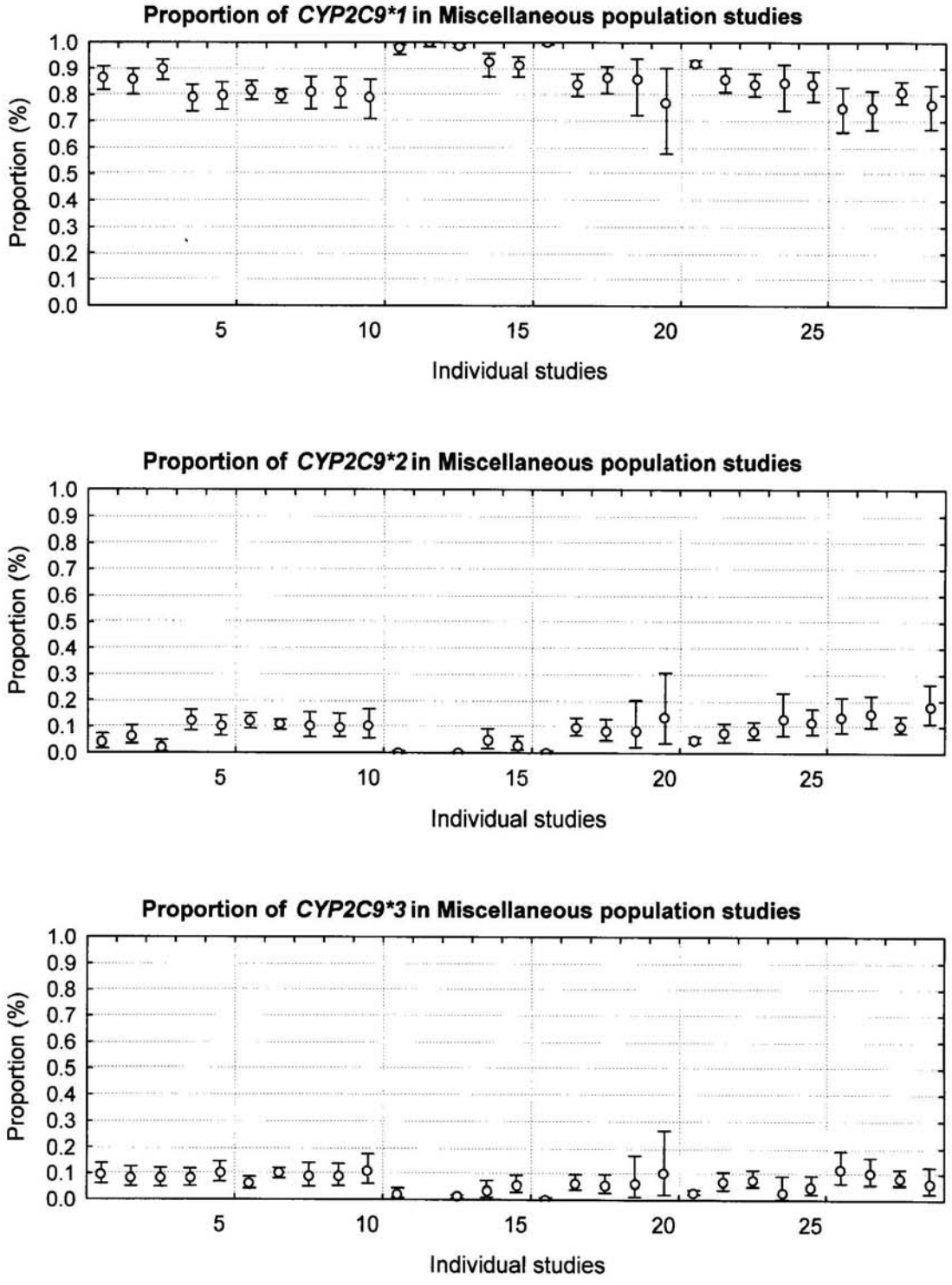


Figure 3.3.1: Proportion of CYP2C9 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C9* allele; ⊥: 95% confidence interval of the proportion; 1.0 stands for 100%

According to the phylogenetic tree (Carvalli-Sforza *et al.* 1994), populations of No.1 to 8 are close to Caucasian population; populations from No.11 to 13 are close to Chinese and Japanese populations. Studies of No.14 to 18 investigated some specific populations that have distinguished characteristics from other ethnicities (No.14: African descants living in Brazil; No.15: Canadian native Indian (CNI); No.16: Inuit in Arctic region; No.17: Israel; No.18: Mexican). Studies of No. 19 to 22 provided data of some admixed populations (No.19: 50% CNI; No.20: 75% CNI; No.21: Bolivian; No22: Intermediate population in Brazil). The remaining are studies with subjects from different ethnicities.

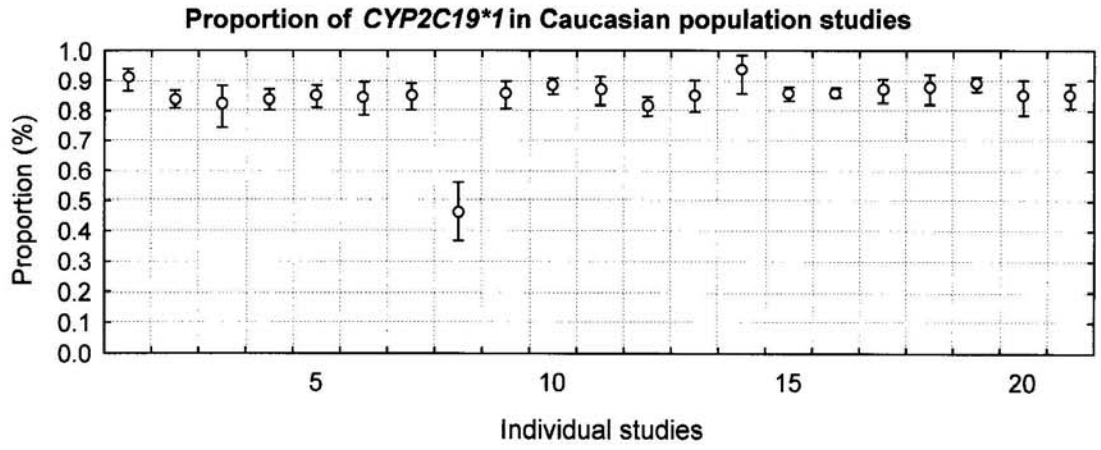
It can be seen in figure 3.3.1, around 70-90% Caucasians have *CYP2C9*1* allele, whereas over 90 percent of Chinese, Japanese populations have it; the variant of *CYP2C9*2* appear in Caucasians are about 10%, however it is very rare in Chinese and Japanese populations; the prevalence of *CYP2C9*3* among studies of Caucasian populations are various, and it is also rare allele in Chinese and Japanese populations. Although African people seem also have over 90% *CYP2C9*1* in most studies, two of the six studies found other *CYP2C9* variants, *CYP2C9*5* and *CYP2C9*11* with over 1% proportion among recruited subjects. Therefore, further studies are required to determine whether African populations have different *CYP2C9* genetic model, *i. e. CYP2C9*1*, **2* and **3* could be most important genetic variants contributed to major inter-individual polymorphic drug response in Caucasian, Chinese and Japanese, however in African populations other variants different from *CYP2C9*2* or **3* might exist and be responsible for major *CYP2C9* relevant polymorphism of drug response.

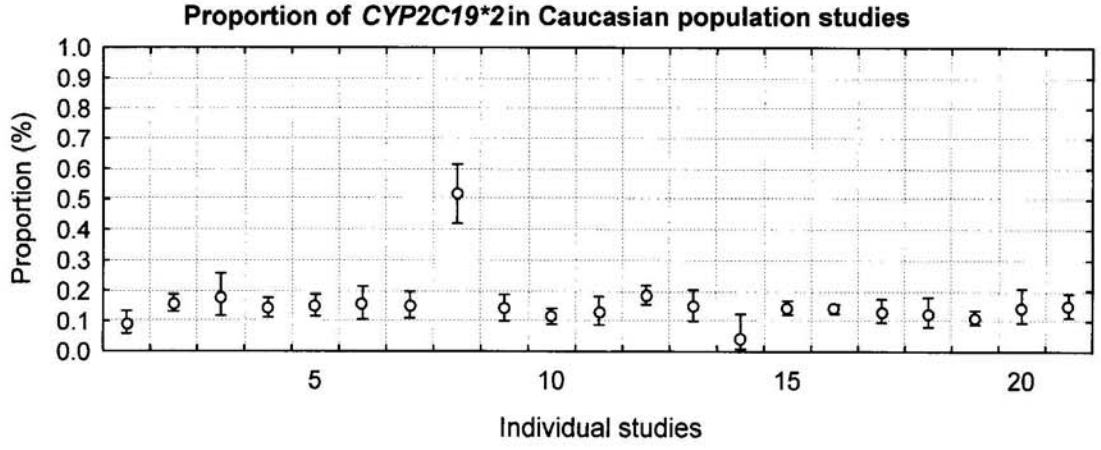
Referring table 3.2.2e, the first ten studies in miscellaneous populations are close to Caucasian in the phylogenetic tree (Cavalli-Sforza *et al.* 1994). *CYP2C9*1* proportions of these ten population were similar as Caucasian, *i. e.* around 70-90%. However, the proportion of *CYP2C9*2* or **3* of these populations shows different distributions from Caucasian. One Vietnamese and two Korean populations in the miscellaneous group have showed a similar distribution of the three *CYP2C9* alleles compared with Chinese and Japanese populations. In one of the Korean population, only the *CYP2C9*1* allele was found in the 90 subjects, which may be the result of low prevalence of *CYP2C9*2* or **3* alleles in the population. However, no statistical analysis has been undertaken in these populations due to insufficient studies.

Furthermore, in the miscellaneous group, some studies investigated the three CYP2C9 alleles in specific populations, which are either aboriginal populations, such as Canadian native Indian and Inuit, or populations influenced by complicated human immigration and evolution history, such as a black population living in Brazil, and Mexican. Due to a lack of replicate studies, it is difficult and inappropriate to conclude anything about the

differentiation of *CYP2C9* prevalence between these populations and other ethnic populations. Meanwhile, some studies recruited people with different ethnic origins. Although, a large proportion of subjects are Caucasians, the data of *CYP2C9* genotypes and alleles was not provided separately for each ethnicity. Therefore these studies were excluded for population distributions of CYP2C9 polymorphism.

For these studies in the miscellaneous group, no statistical analysis was undertaken apart from trinomial contours, which were expected to provide the visual representation of the diverse population distribution of the three CYP2C9 alleles.





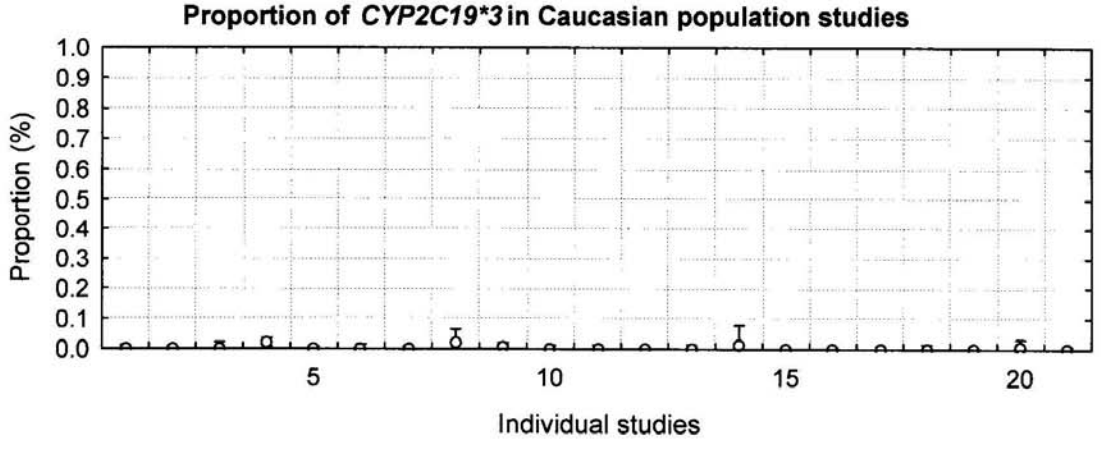
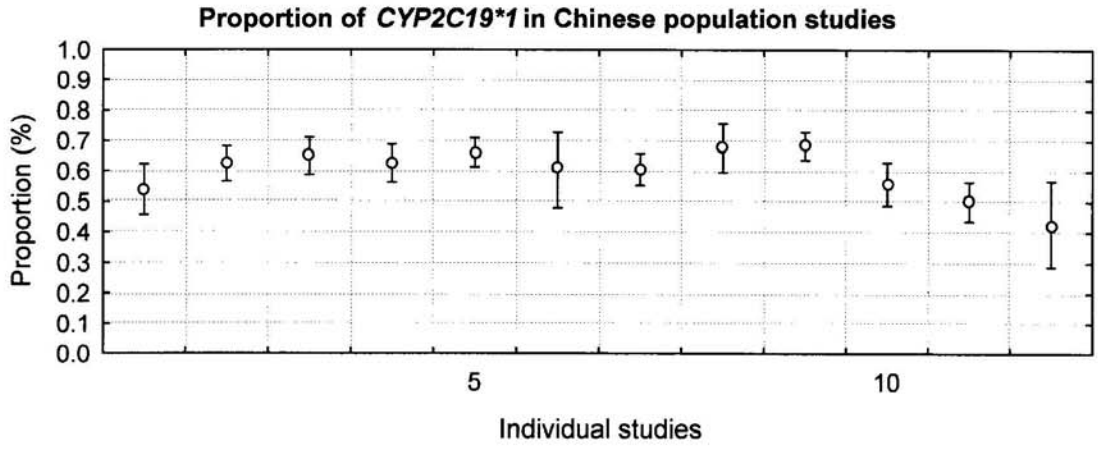


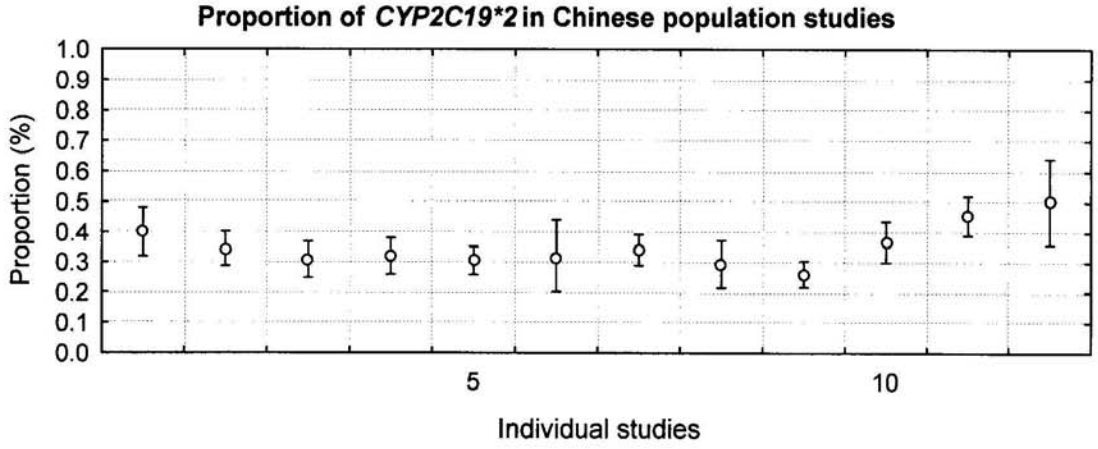
Figure 3.3.2: Proportion of CYP2C19 allele in different ethnic populations

O: Proportion of *CYP2C19* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





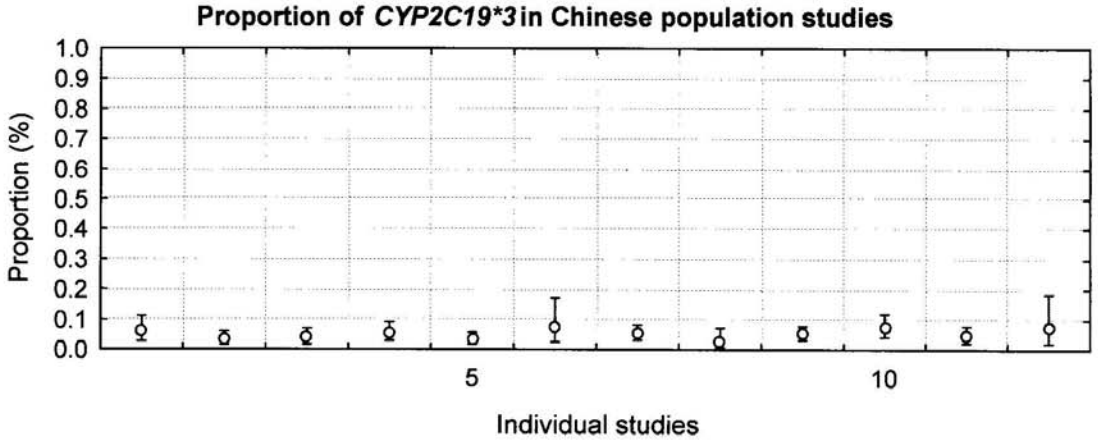
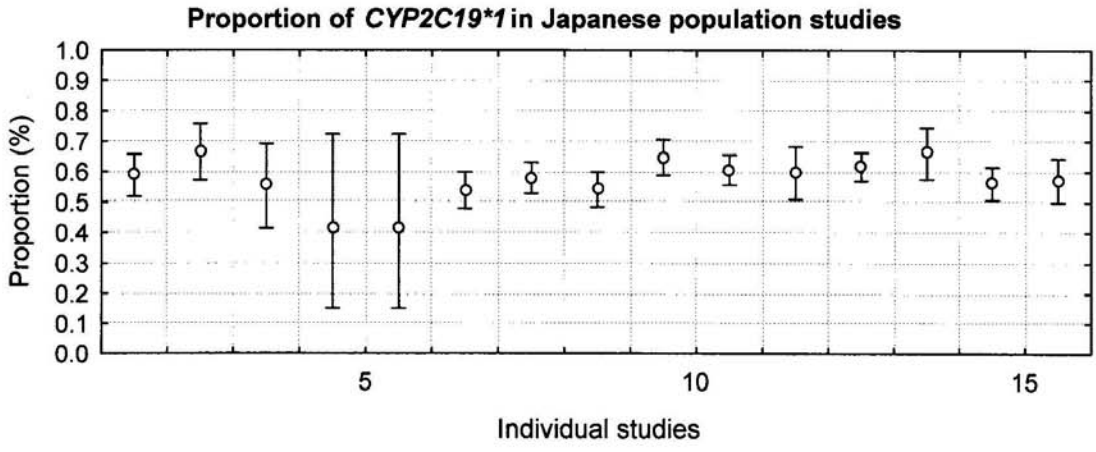


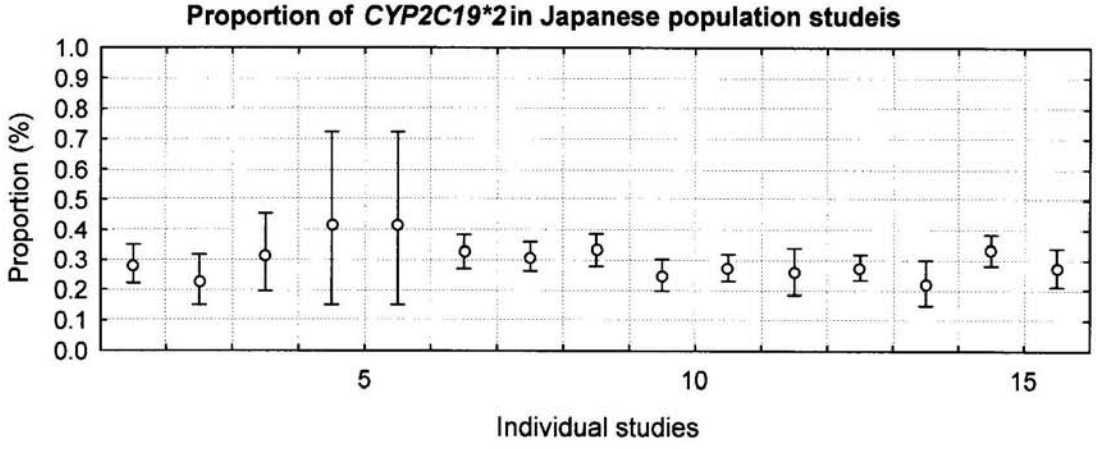
Figure 3.3.2: Proportion of CYP2C19 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C19* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





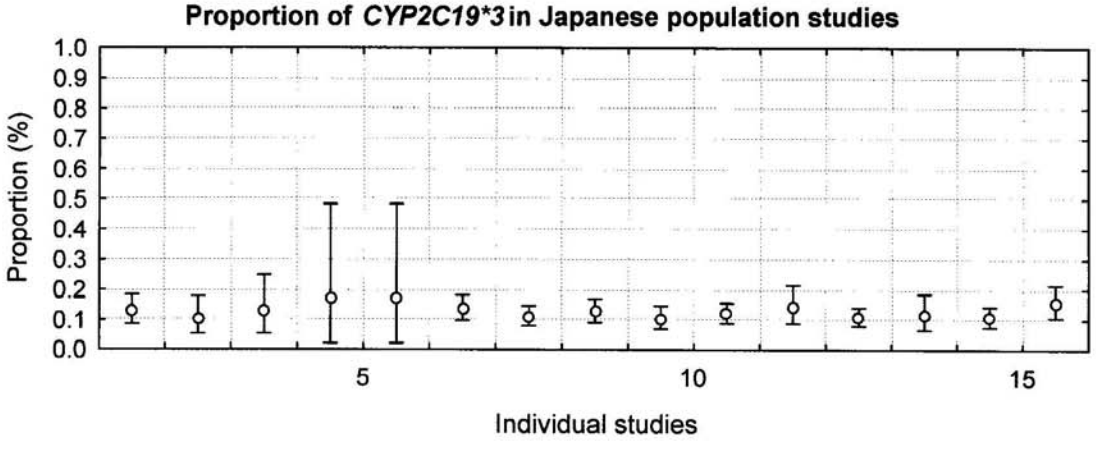
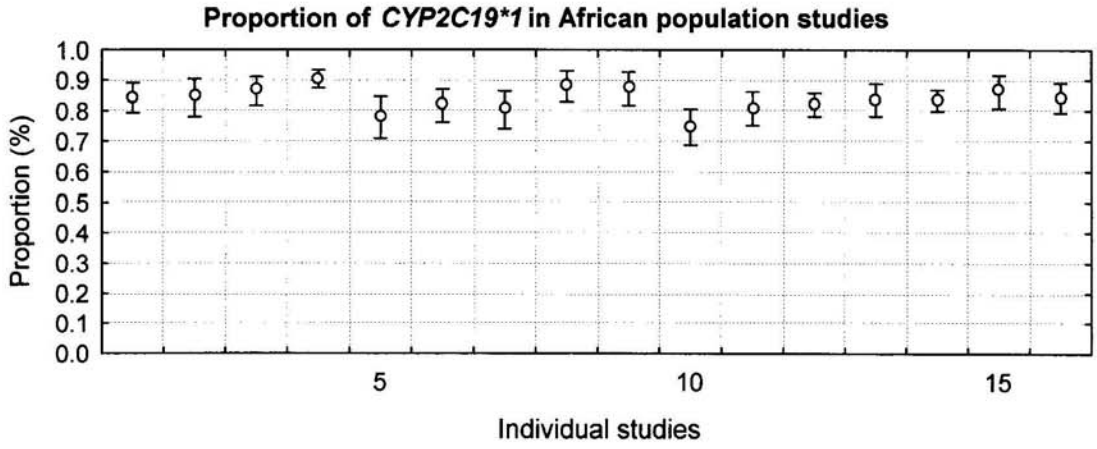


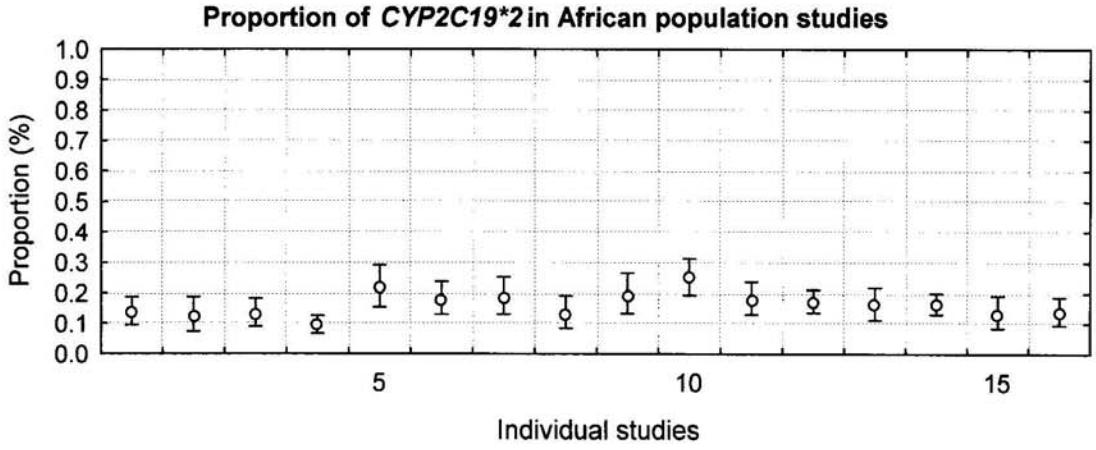
Figure 3.3.2: Proportion of CYP2C19 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C19* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%





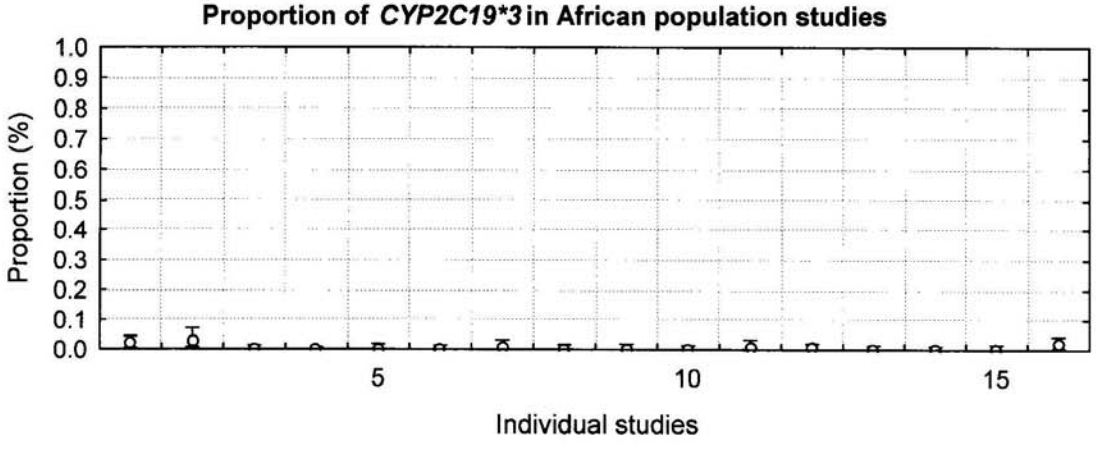


Figure 3.3.2: Proportion of CYP2C19 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C19* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%

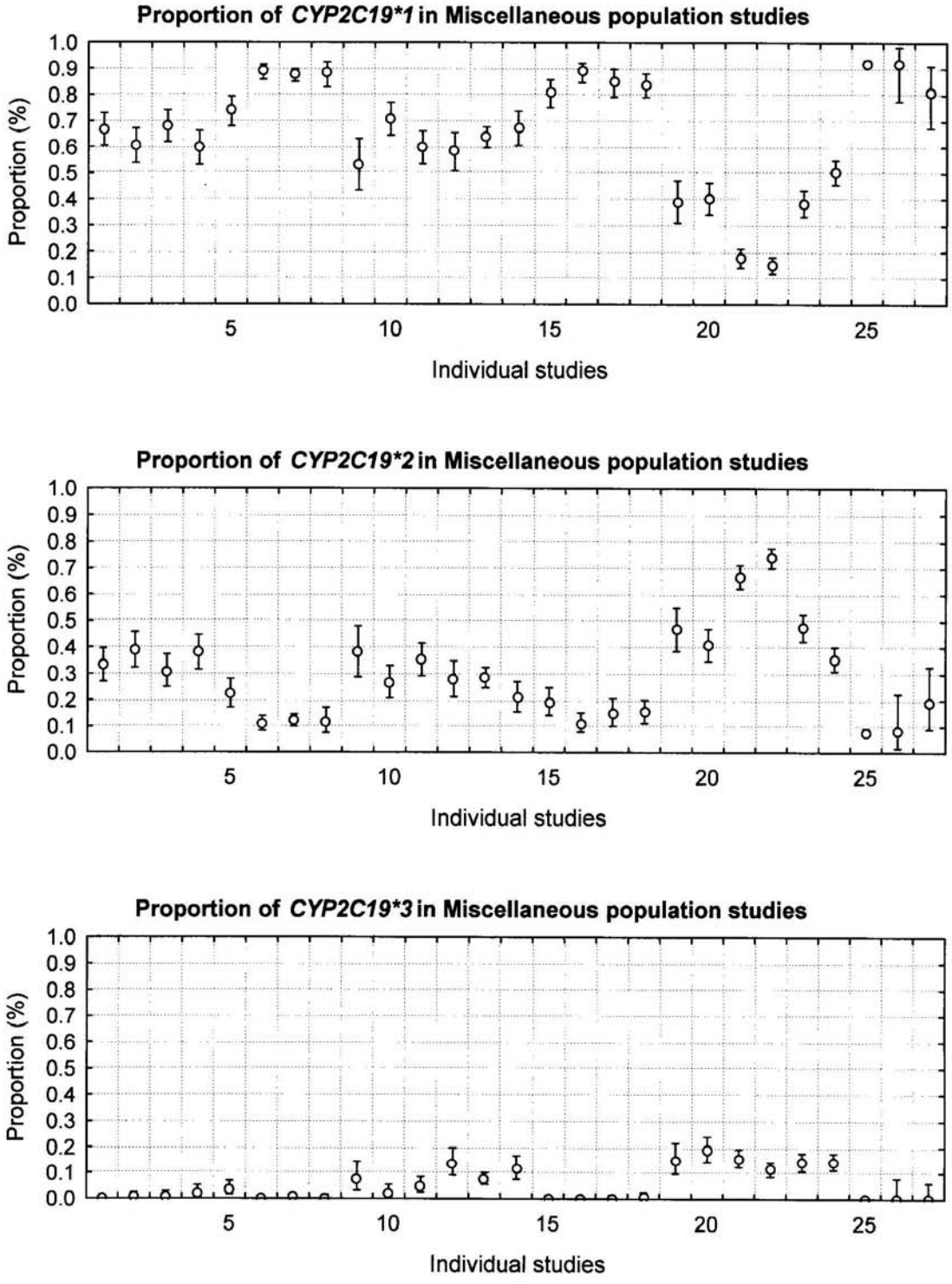


Figure 3.3.2: Proportion of CYP2C19 allele in different ethnic populations (Continuous)

O: Proportion of *CYP2C19* allele; ⊥: 95% confidence interval of the proportion; 1.0 stands for 100%

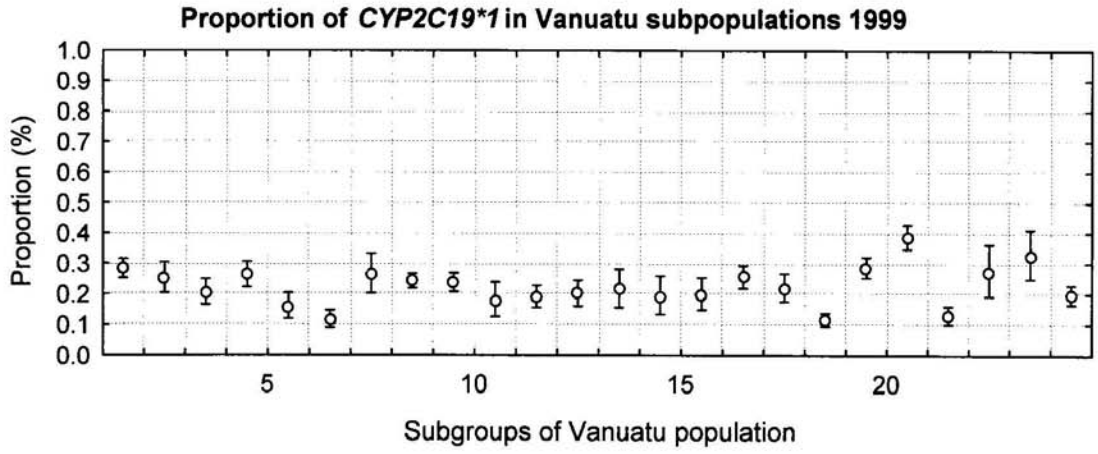
According to the phylogenetic tree (Carvalli-Sforza *et al.* 1994), populations of No.1 to 8 are close to Caucasian population; populations from No.9 to 14 are close to Chinese and Japanese populations. Studies of No.14 to 24 investigated some specific populations that have distinguished characteristics from other ethnicities (No.14: Canadian native Indian (CNI); No.15: Inuit in Arctic region; No.16: Inuit; No.17: Saudi Arabian; No.18: Jewish; No.19-23: Pacific Islander; No.24: Aborigine Australian). Studies of No. 25 to 28 provided data of some admixed populations (No.25: Bolivian; No.26: 25% CNI; No.27: 50% CNI; No.28: 75% CNI).

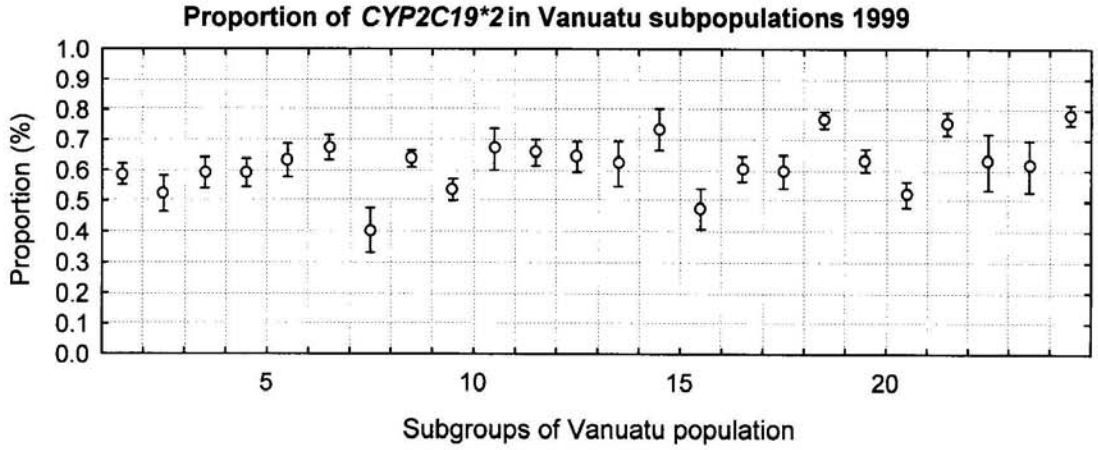
The proportions of *CYP2C19* among Caucasian, Chinese, Japanese and African populations appeared to show quite different characteristics compared with *CYP2C9*. Proportion of *CYP2C19*1*, **2* and **3* in Caucasian and African populations is very close, whereas Chinese and Japanese populations have similar proportion of three *CYP2C19* alleles. Furthermore, 50-60% Chinese and Japanese populations have *CYP2C9*1*, whereas around 80-90% Caucasian and African people have *CYP2C19*1*. The variant *2 appears to be doubled in Chinese and Japanese when comparing with Caucasian and African, 30-40% *vs* 10-20%. Meanwhile, *CYP2C19*3* is lower among Caucasian and African population with less than 1%, however it appears in about 10% of Japanese population. These differentiations in *CYP2C19* population distributions could be major contributors in the inter-individual and intra-population polymorphism of *CYP2C19* metabolizing drugs, which is further explored in the following chapters.

Referring to table 3.2.3f, the first eight studies in miscellaneous populations are close to Caucasian in the phylogenetic tree (Cavalli-Sforza *et al.* 1994). However, not all of their *CYP2C19* proportions are similar to Caucasians. The five Indian populations appear to have similar *CYP2C19* proportions to the Chinese, which is different from *CYP2C9* distributions. Meanwhile, as with *CYP2C9* alleles, three alleles of *CYP2C19* in Thai, Vietnamese and Korean populations have similar features to Chinese and Japanese populations.

However, comparing the CYP2C9/2C19 distribution in figure 3.3.1 and figure 3.3.2, the polymorphism of *CYP2C19* among different ethnic populations seems more significant than the polymorphism of *CYP2C9*, which is also presented in a trinomial graph in the last section of this chapter.

The proportion of the three *CYP2C19* alleles studied in the miscellaneous group was not undertaken for any statistical analysis, however, their data are represented in trinomial contours for a visual results of diverse *CYP2C19* population distributions.





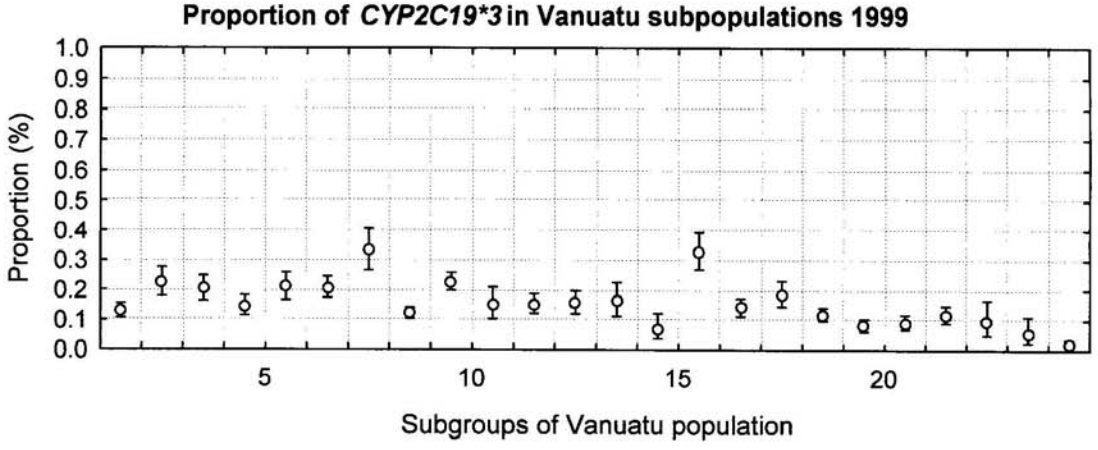
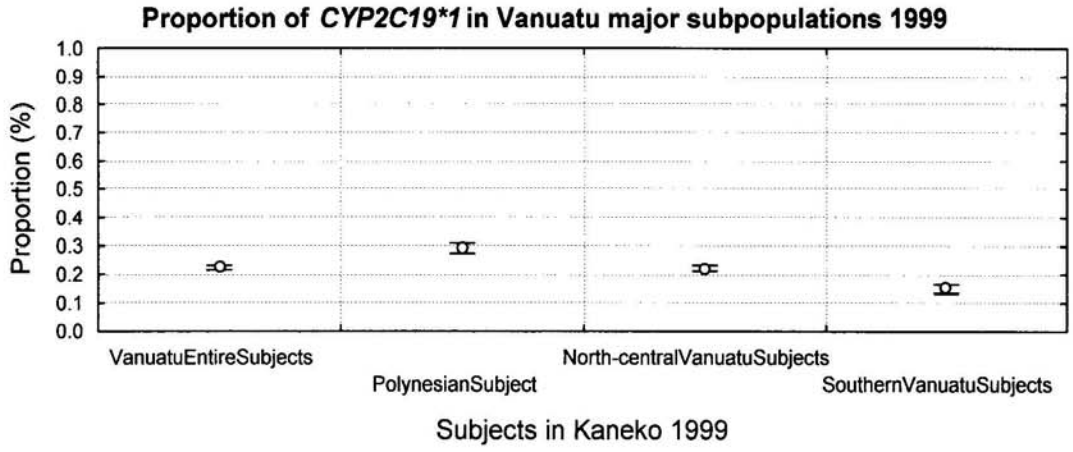
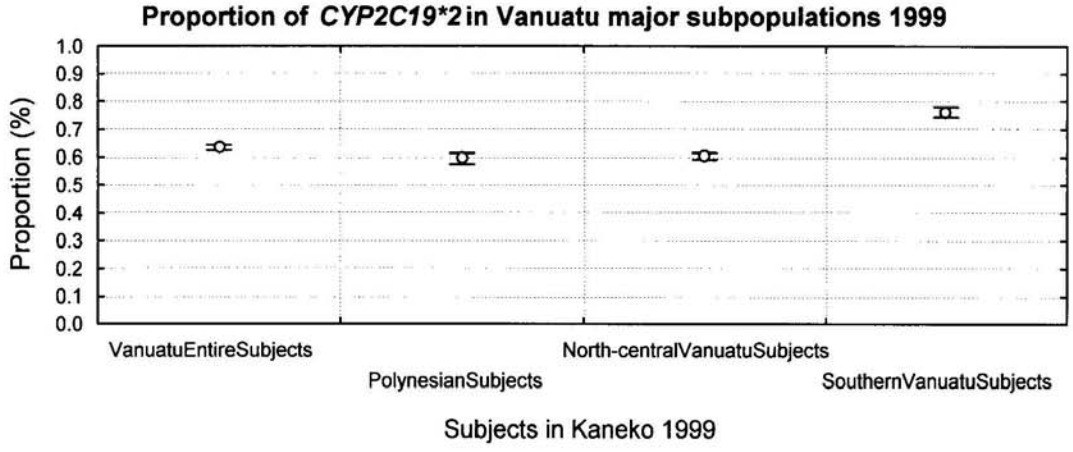


Figure 3.3.3a: Proportion of CYP2C19 allele in 24 Vanuatu populations (Kaneko *et al.* 1999)

O: Proportion of *CYP2C19* allele; ⊥: 95% confidence interval of the proportion; 1.0 stands for 100%

These 24 subpopulations of Pacific Islander are represented as categories of Austronesian languages in study of Kaneko *et al.* (1999). No.1 to 12, No. 14 and No.15 are Northern-central Vanuatu populations. No. 13, 16, 17, 19, 20 and 23 are Polynesian populations. No.18, No. 22, No. 23 and No. 24 are Southern Vanuatu populations.





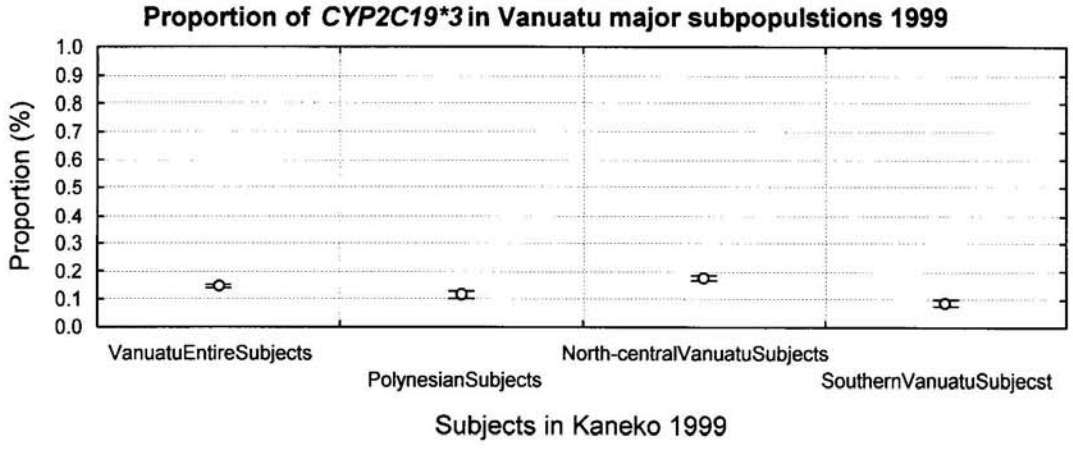


Figure 3.3.3b: Proportion of CYP2C19 allele in Vanuatu populations (Kaneko *et al.* 1999)

O: Proportion of *CYP2C19* allele

⊥: 95% confidence interval of the proportion

1.0 stands for 100%

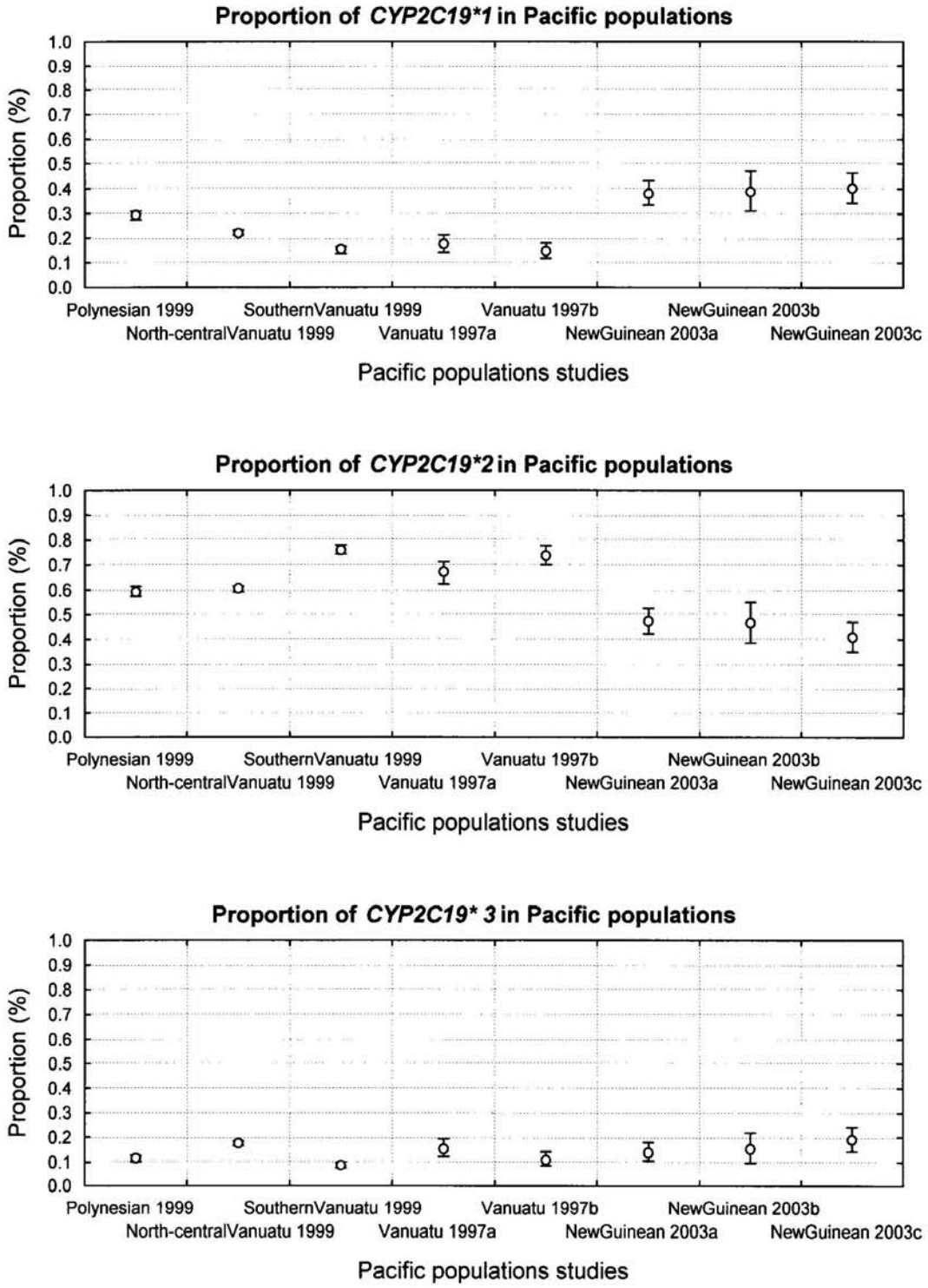


Figure 3.3.4: Proportion of CYP2C19 allele in Pacific Islander populations

O: Proportion of *CYP2C19* allele; ⊥: 95% confidence interval of the proportion; 1.0 stands for 100%

Polynesian 1999, Northern-central Vanuatu 1999, and Southern Vanuatu 1999 are represented from the study of Kaneko *et al.* (1999). Vanuatu 1997a and b are represented Malakula and Tanna populations from the study of Kaneko *et al.* (1997). New Guinean 2003a, b and c are represented Jawia, Kinambu and Witupe populations from the study of Masta *et al.* (2003).

In figure 3.3.3a, proportions of *CYP2C19* in 24 Vanuatu populations (Kaneko *et al.* 1999) are various, although they can be categorized into same ethnic group, Pacific Islander, according to phylogenetic tree (Cavalli-Sforza *et al.* 1994).

Regarding the geographic location and spoken language of these twenty-four Vanuatu populations, Kaneko *et al.* (1999) further categorized them into three groups, such as Polynesian subjects, Northern-central Vanuatu subjects and Southern Vanuatu subjects, which are presented in figure 3.3.3a with data of entire 24 Vanuatu populations. The difference of CYP2C19 variant proportions is obvious shown in the graph.

Referring to the phylogenetic tree shown in figure 1.3.1, New Guinean and Pacific Islander are unique from other populations in retrieved studies of this thesis; the genetic distance of these two population are smaller than the genetic distance between them and other populations. However, the prevalence of the three *CYP2C19* alleles exhibits apparent polymorphism between these populations. Moreover, two of *CYP2C19* variants, particularly *CYP2C19*2* exhibit a higher prevalence compared with any of the other populations investigated in this thesis. The prevalence of *CYP2C19*2* is approximate 40-75% in these Pacific Islander populations, whereas it was found in about 10-20% of Caucasian or African populations, and approximately 25-45% of Chinese or Japanese populations. Therefore, Kaneko *et al.* (1999) proposed that the majority of pacific islanders are CYP2C19 poor metabolizers, which suggests that great caution must be taken when drugs metabolized by CYP2C19 are prescribed to these pacific Islander populations. The visual presentation of *CYP2C19* distribution is shown in the figure 3.3.3a, b and 3.3.4.

3.3.2 Analysis of deviation from Hardy-Weinberg equilibrium for each study

Initially, the genotype and allele numbers from all studies for the six *CYP2C9/2C19* genotypes and three *CYP2C9/2C19* alleles were transformed into the required format for the GENEPOP program. Then, the Hardy-Weinberg equilibrium (HWE) test was performed in the GENEPOP program for each study. Meanwhile, the allele proportions for each study were calculated in GENEPOP program in order to check the accuracy of the transformed data. The results are presented in appendix 2. Studies were accepted as fitting HWE when their probability of fitting HWE was over 0.05 ($P>0.05$). Furthermore, population differentiation analysis was employed for those studies fitting HWE, which is considered as a similar analytical approach for the heterogeneity test in meta-analysis. Both of them employed a non-parametric test, using Chi-square statistics, to identify whether there was significant differentiation among populations in respect of their allele or genotype prevalence. However, the Fisher exact test was required in the analysis of this thesis because of low prevalence (<5%) of variant CYP2C9*2/*3 and CYP2C19*3 in Chinese, Japanese, and African populations. Controversially, CYP2C9*1 and CYP2C19*1 have a higher prevalence (>80%) in those populations, therefore the computation time was long for the Fisher exact test using the available statistical software (SAS 8.1). However, the GENEPOP program employed Markov Chain principles to perform the Fisher exact test, which provided acceptable results ($SE< 0.01$) within a very short computation time (Raymond M and Rousset F 1995). Therefore, the GENEPOP program has employed to test for HWE fitness of studies in the thesis, which provided valuable information for further meta-analysis in the estimation of allele prevalence.

3.3.3 Heterogeneity/population differentiation test of studies fitting HWE

According to ethnicity groups, such as Caucasian, Chinese, Japanese, and African, the heterogeneity/population differentiation test was performed in GENEPOP program for those studies fitting HWE.

In order to achieve a precise estimation of allele prevalence, the characteristics of subjects recruited in each study were re-checked. Ideally, allele prevalence for particular ethnic populations should be derived from studies where unrelated subjects with clearly defined ethnicity were randomly recruited from the target ethnic populations; however very few studies applied this criterion. Consequently, the data was accepted when the study had recruited unrelated healthy subjects or patients by "relatively" random methods; for instance, healthy subjects were volunteers or recruited from community or organization by advertisement or patients were recruited from few clinics with different clinical conditions. Therefore, the test results of population differentiation was performed by blinding the identification of each publications in order to check the selection of appropriate studies for the estimation of each CYP2C9/2C19 allele prevalence.

Eventually, the allele proportions from 14 Caucasian, 3 Chinese, 5 Japanese and 3 African populations studies were identified as falling into this criteria for the estimation of *CYP2C9 *1, *2* and **3* prevalence, which included 2709 Caucasian people, 665 Chinese people, 683 Japanese people and 381 African people. A further 10 Caucasian, 7 Chinese, 9 Japanese and 11 African population studies provided data for the prevalence estimation of *CYP2C19 *1, *2* and **3*, which included 2094 Caucasian people, 725 Chinese people, 1249 Japanese people, and 1363 African people.

As "Chinese/Japanese" defines more specific population groups than Caucasian or African, the higher heterogeneity among Caucasian and African population studies was expected. These results were presented in appendix 3. However, it was consistent with the Cochran Q test in the meta-analysis of allele prevalence. The Q value of Cochrane Q test for each ethnicity is shown individually in meta-analysis plots.

3.3.4 Prevalence estimation of *CYP2C9/2C19* allele for different ethnic populations

Studies that recruited subjects with clearly defined ethnic origin were selected again for estimation of prevalence of three major *CYP2C9/2C19* alleles. Unrelated subjects recruited randomly from the community or residential homes were preferred. In order to avoid population stratification, studies were excluded when subjects were recruited for genetic association observation with a particular aim, such as patients requiring the same clinical treatment.

The pooled proportions was calculated under both fixed and random models for alleles *CYP2C9*1, *2,* and **3 CYP2C19*1, *2,* and **3* independently. The fixed model was employed when the Cochrane Q value was less than the Chi-square table value or when the estimated variance of pooled studies, τ, was less than zero (Whitehead A and Whitehead J, 1991). Otherwise, the random model of meta-analysis was employed. Results are presented in figure 3.2.6. to figure 3.2.15 for *CYP2C9* and *CYP2C19* respectively according to ethnicity groups.

The estimated prevalence of CYP2C9/2C19 alleles in Caucasian, Chinese, Japanese and African populations is summarized in Table 3.3.2a and b for *CYP2C9* and *CYP2C19* respectively.

Table 3.3.2a: Summary of estimated *CYP2C9* allele prevalence with 95%CI in different ethnicity populations

Ethnicity	*CYP2C9*1*	*CYP2C9*2*	*CYP2C9*3*	Number of alleles	Number of individuals	Number of studies
Caucasian	0.8053 (0.7901~0.8201)	0.1183 (0.1087~0.1282)	0.0759 (0.0667~0.0857)	5418	2709	14
Chinese	0.9610 (0.9498~0.9706)	< 0.00127 [a] (0.0001~0.0042)	0.0383 (0.0287~0.0493)	1330	665	3
Japanese	0.9745 (0.9655~0.9822)	0 [a] (0.0000~0.0031)	0.0255 (0.0178~0.0345)	1366	683	5
African	0.9422 (0.9245~0.9576)	0.0327 [b] (0.0012~0.0576)	0.0169 [b] (0.0011~0.0312)	762	381	3

Note: a. The allele was rare in those populations. The value is presented for data compliance only. b. The value was inconclusive, presented for data compliance only.

The *CYP2C9*2* allele was not found in two of three Chinese population studies nor in any of five Japanese population studies. Therefore, the weight of the *CYP2C9*2* proportion cannot be calculated for these populations. However, the *CYP2C9*2* obviously is a rare

allele in Chinese and Japanese populations. Furthermore, among three African population studies, one study did not report either *CYP2C9*2* or *CYP2C9*3* allele. The prevalence of *CYP2C9*2* or *CYP2C9*3* was calculated by using the weighted proportion mean of the other two studies. Therefore, the estimated prevalence is inconclusive for African population. More studies are required for African populations.

Table 3.3.2b: Summary of estimated *CYP2C19* allele prevalence with 95%CI in different ethnicity populations

Ethnicity	*CYP2C19*1*	*CYP2C19*2*	*CYP2C19*3*	Number of alleles	Number of individuals	Number of studies
Caucasian	0.8611 (0.8421~0.8791)	0.1380 (0.1199~0.1573)	< 0.0038 [a] (0.0003~0.0022)	4188	2094	10
Chinese	0.5660 (0.5102~0.6210)	0.3783 (0.3312~0.4264)	0.0522 (0.0383~0.0681)	1450	725	7
Japanese	0.5824 (0.5630~0.6016)	0.2972 (0.2751~0.3197)	0.1211 (0.1086~0.1342)	2498	1249	9
African	0.8489 (0.8235~0.8727)	0.1555 (0.1320~0.1806)	< 0.0056 [a] (0.0010~0.0062)	2726	1363	10

Note: a. The allele was rare in those populations. The value was obtained for data compliance

As shown in the table, prevalence of *CYP2C19*3* appears very low in Caucasian and African populations. Among ten Caucasian studies, eight studies did not find any *CYP2C19*3*, whilst the other two found 0.34-0.60% of the population carrying this allele. Among African population studies, *CYP2C19*3* was only identified in four studies. However, for data compliance, the prevalence of *CYP2C19*3* is calculated from those studies that had found the allele.

As previously mentioned, among those reported *CYP2C9/2C19* mutant alleles, variants **2* and **3* contribute to major inter-individual and intra-population variations in the response to medication by consequent deficient or inactive enzyme functions. Therefore, considering the administration of a drug metabolized by CYP2C9, the unexpected toxicity due to deficient enzyme activity would be likely to occur more frequently in Caucasian populations than in Chinese and Japanese populations. Meanwhile, for drugs primarily metabolized by CYP2C19, Chinese and Japanese people would have a higher occurrence of ineffective metabolism and correlated adverse effects than Caucasians and Africans; consequently, these drugs should be prescribed with caution.

*Population distribution of CYP2C9*1, *2 and *3 in Caucasian, Chinese, Japanese and African:*

1. ***Caucasian*** population studies:

Nineteen of forty studies have been used in a meta-analysis for estimation of the CYP2C9 prevalence after checking the recruitment method and compliance of genotype data. Five studies were excluded because the genotype data dif not fit the Hardy-Weinberg equilibrium.

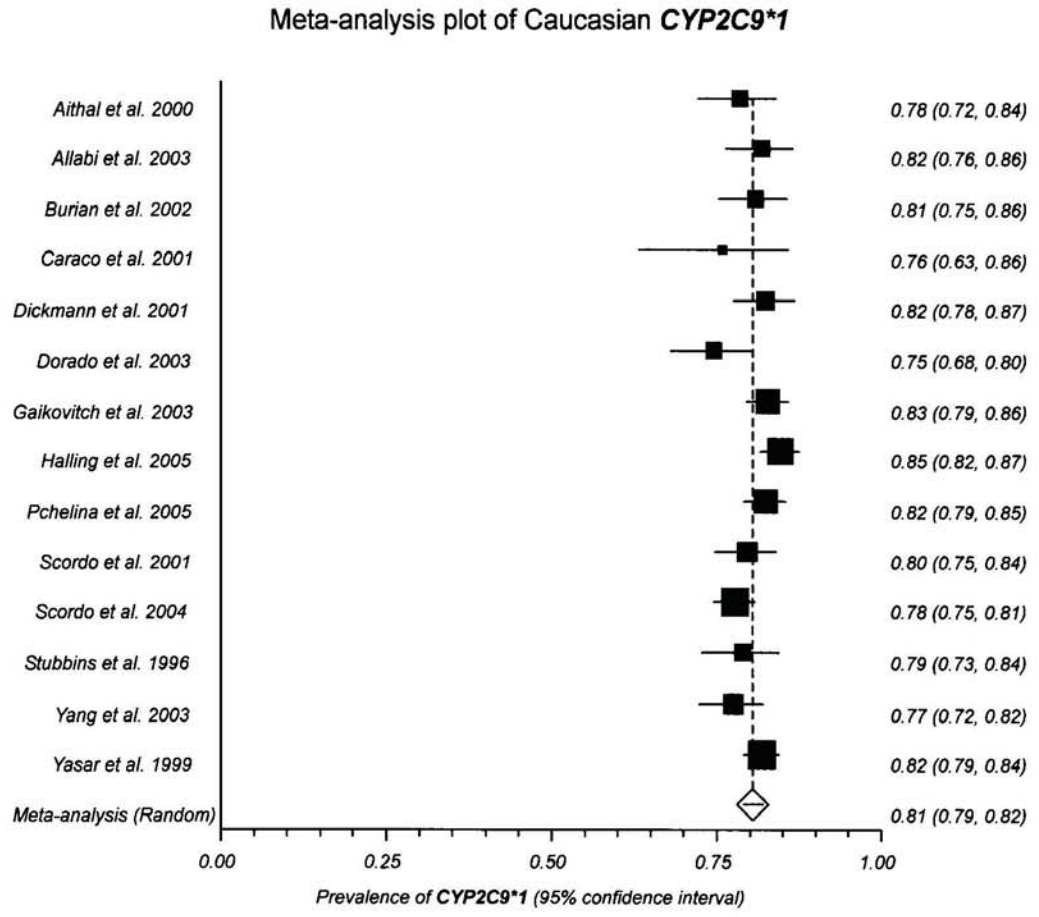


Figure 3.3.5a: Meta-analysis of ***CYP2C9*1*** prevalence in Caucasian populations. (Q = 23.80, df = 13)

The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence is represented by diamond.

The pooled prevalence for *CYP2C9*1* in Caucasian population was 81%, with 95% confidence interval 79 to 82%.

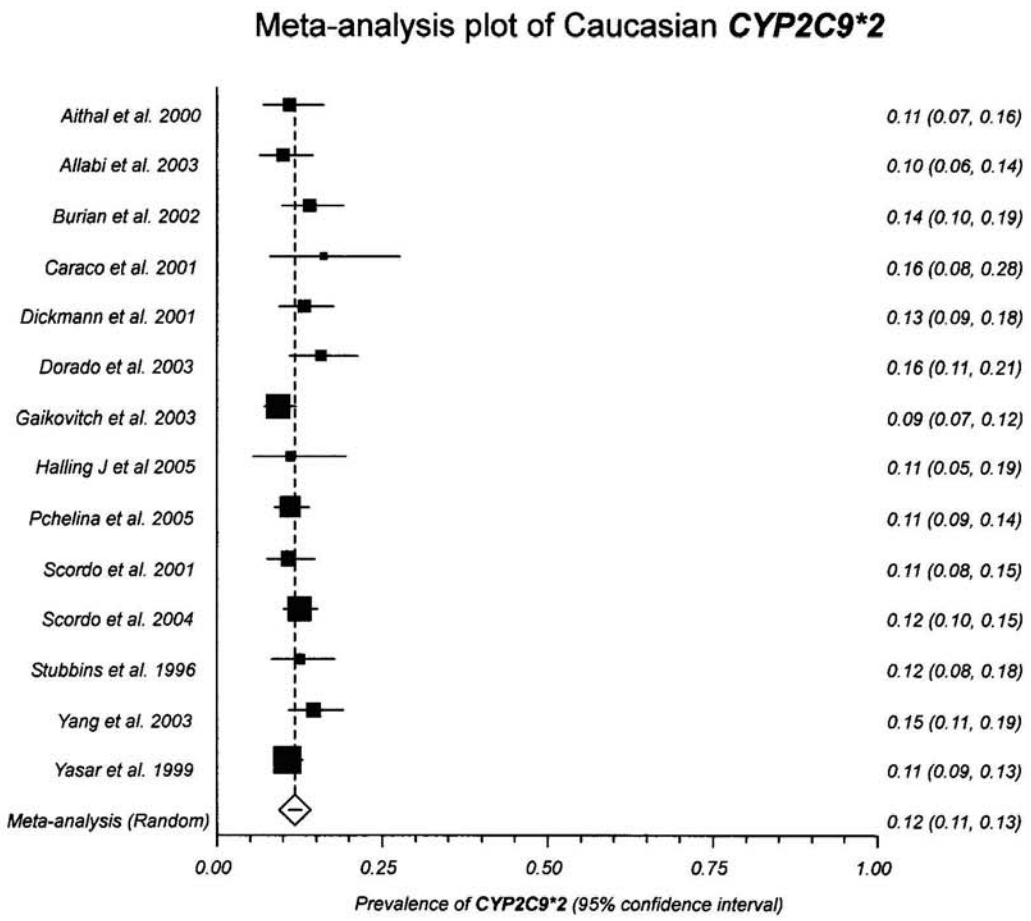


Figure 3.3.5b: Meta-analysis of ***CYP2C9*2*** prevalence in Caucasian populations. (Q = 14.64, df = 13)

The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence is represented by diamond.

The pooled prevalence for *CYP2C9*2* in Caucasian population was 12%, with 95% confidence interval 11 to 13%.

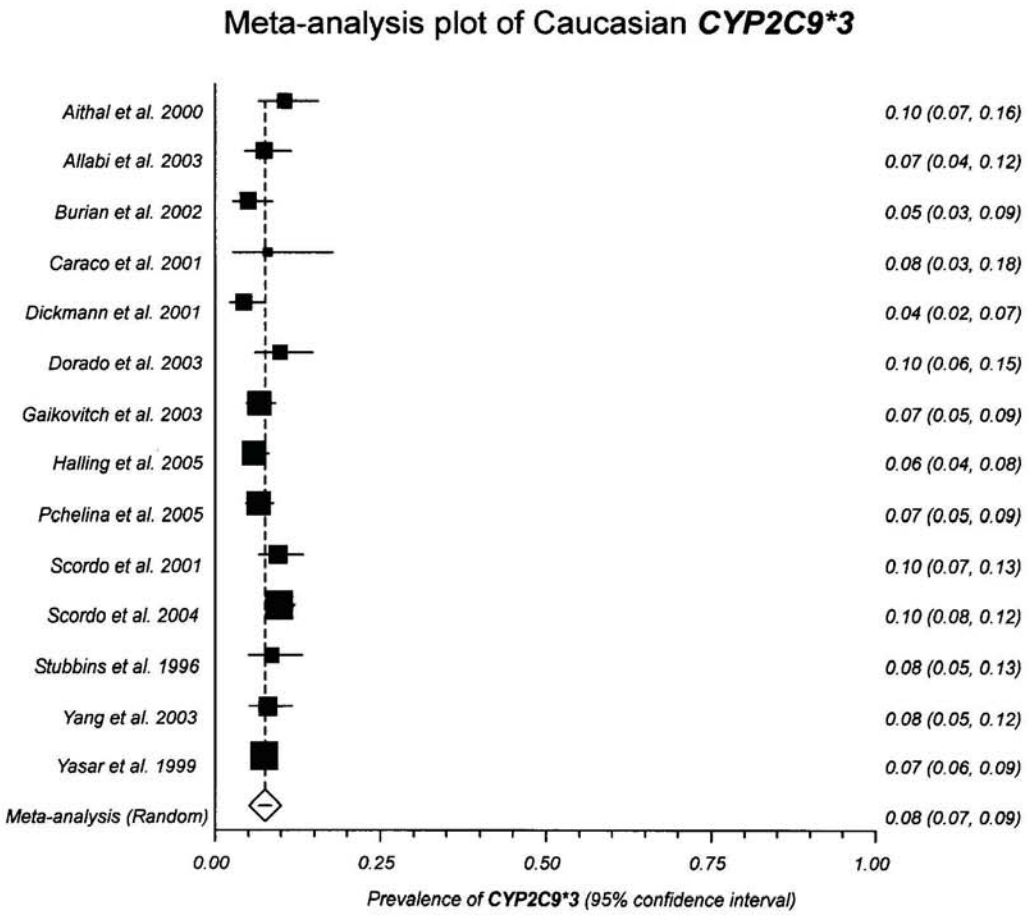


Figure 3.3.5c: Meta-analysis of ***CYP2C9*3*** prevalence in Caucasian populations. (Q = 21.47, df = 13)

The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence is represented by diamond.

The pooled prevalence for *CYP2C9*3* in Caucasian population was 8%, with 95% confidence interval 7 to 9%.

2. ***Chinese*** population studies:

Seven studies reporting *CYP2C9* polymorphism in Chinese populations have been retrieved. Two of them did not provide completed genotype. Two of the studies recruited epilepsy patients only, therefore they were excluded from meta-analysis for estimation of the CYP2C9 prevalence. However, the genotype data of all studies fitted the Hardy-Weinberg equilibrium.

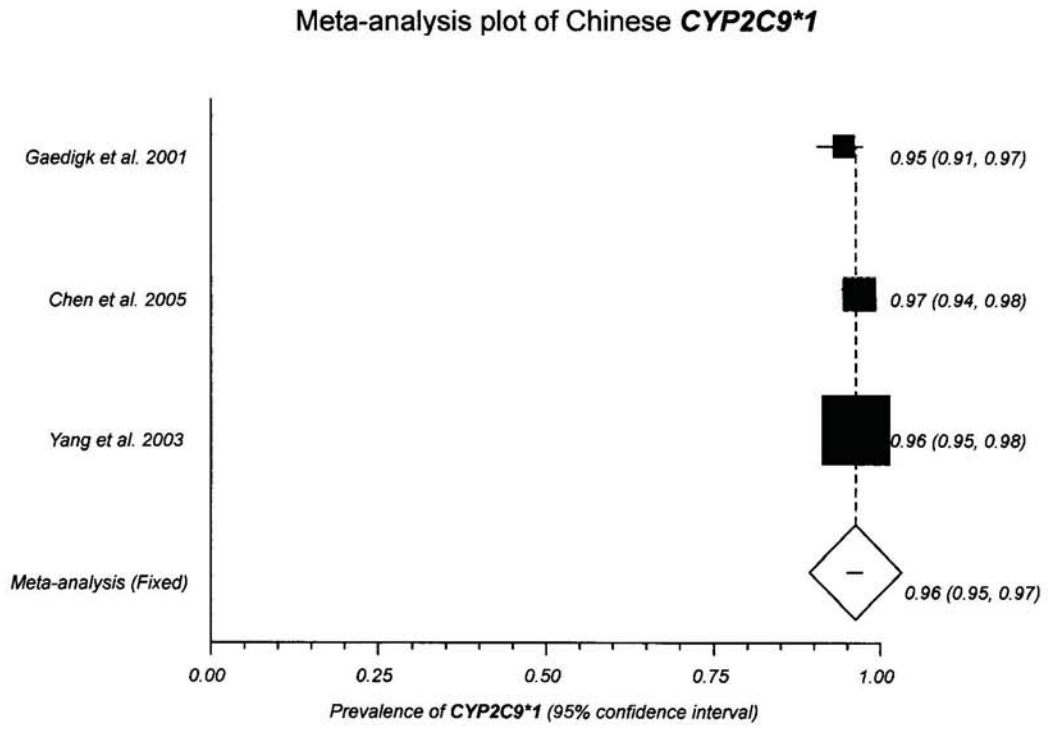


Figure 3.3.6a: Meta-analysis of ***CYP2C9*1*** prevalence in Chinese populations. (Q = 1.64, df = 2)

The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

The pooled prevalence for *CYP2C9*1* in Chinese population was 96%, with 95% confidence interval 95 to 97%.

The overall prevalence for *CYP2C9*2* in Chinese population was less than 0.13%, and the pooled prevalence for *CYP2C9*3* in Chinese population was 4%, with 95% confidence interval 3 to 5% as shown in figure 3.3.6b and c.

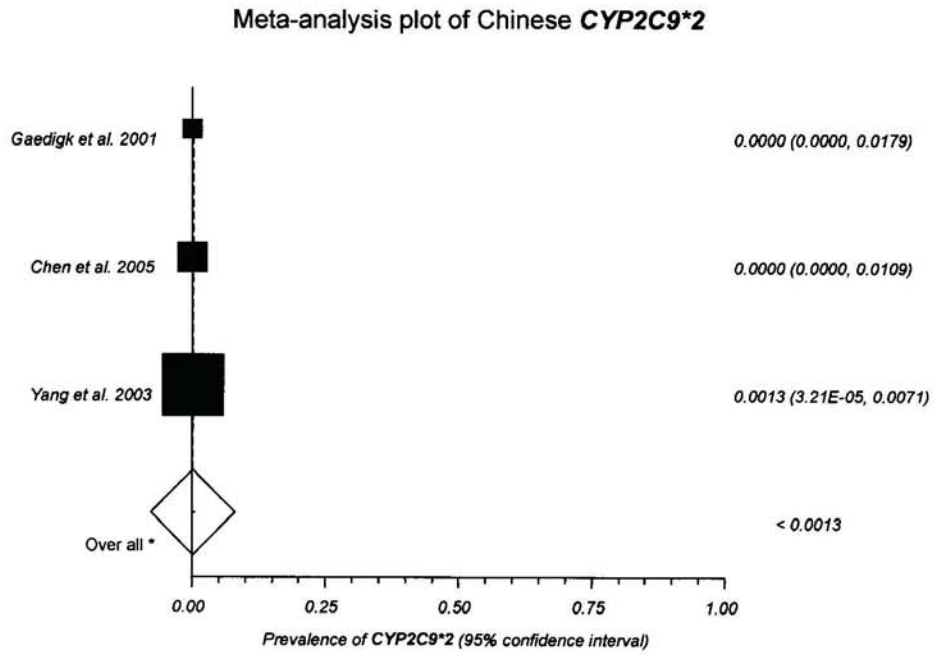


Figure 3.3.6b: Meta-analysis of ***CYP2C9*2*** prevalence in Chinese populations.
*The prevalence of *CYP2C9*2* was zero in two studies of Chinese populations. Therefore, the pooled prevalence and the Cochrane Q value cannot be calculated. The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square, which is sample size of each study.

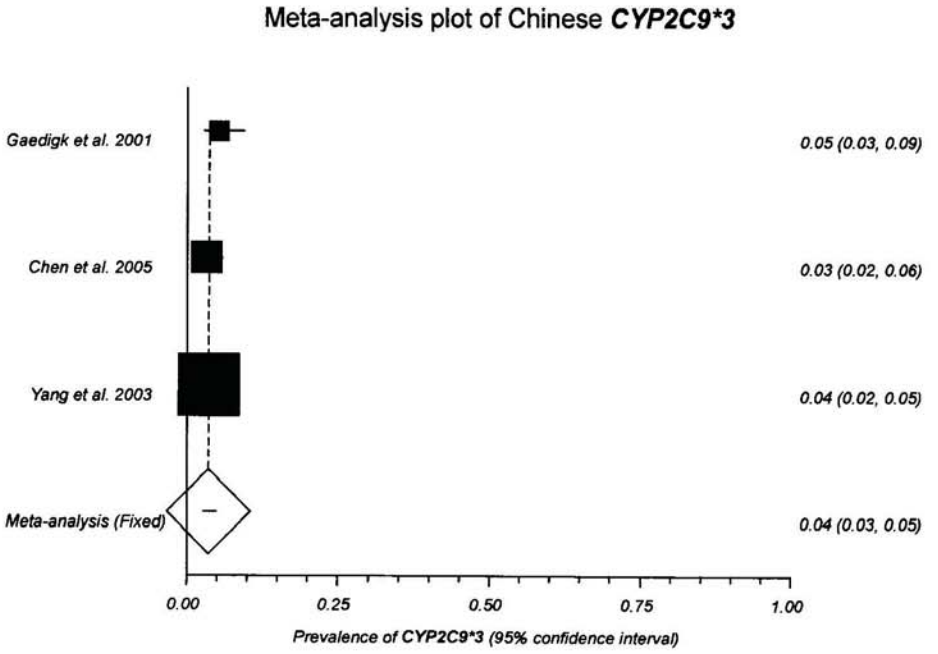


Figure 3.3.6c: Meta-analysis of ***CYP2C9*3*** prevalence in Chinese populations. (Q = 1.75, df = 2)
The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

3. ***Japanese*** population studies:

One of six studies of *CYP2C9* polymorphism in Japanese populations had recruited epilepsy patients only. Therefore this was excluded from meta-analysis for estimation of the CYP2C9 prevalence. Six studies all fitted the Hardy-Weinberg equilibrium.

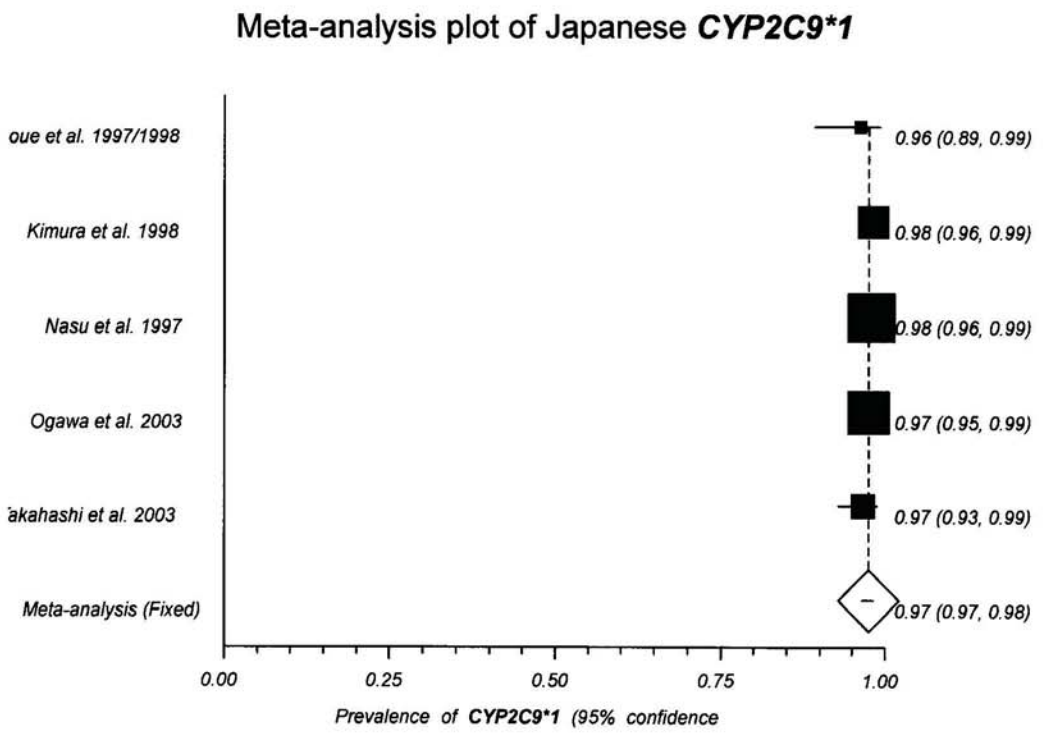


Figure 3.3.7a: Meta-analysis of ***CYP2C9*1*** prevalence in Japanese populations. (Q = 2.23, df = 4)

The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study is shown by the solid square. Pooled estimation of prevalence was presented by diamond.

The pooled prevalence for *CYP2C9*1* in Japanese population was 97%, with 95% confidence interval 97 to 98%.

The overall prevalence for *CYP2C9*2* in Japanese population was rare and the pooled prevalence for *CYP2C9*3* in Japanese population was 2.5%, with 95% confidence interval 1.8 to 3.4% as shown in figure 3.3.7b and c.

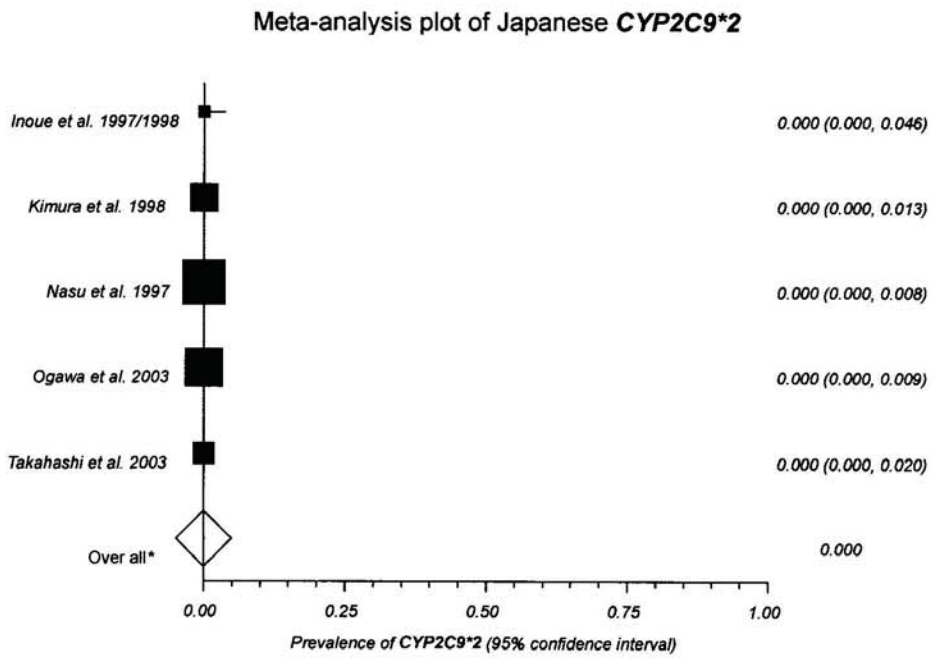


Figure 3.3.7b: Meta-analysis of ***CYP2C9*2*** prevalence in Japanese populations.
*The weight and the Cochrane Q value cannot be calculated because the prevalence of *CYP2C9*2* in all studies was zero. However, the overall estimation can be accepted as zero since there is not possibility of heterogeneity among studies. The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight is shown using the sample size the solid square of each study.

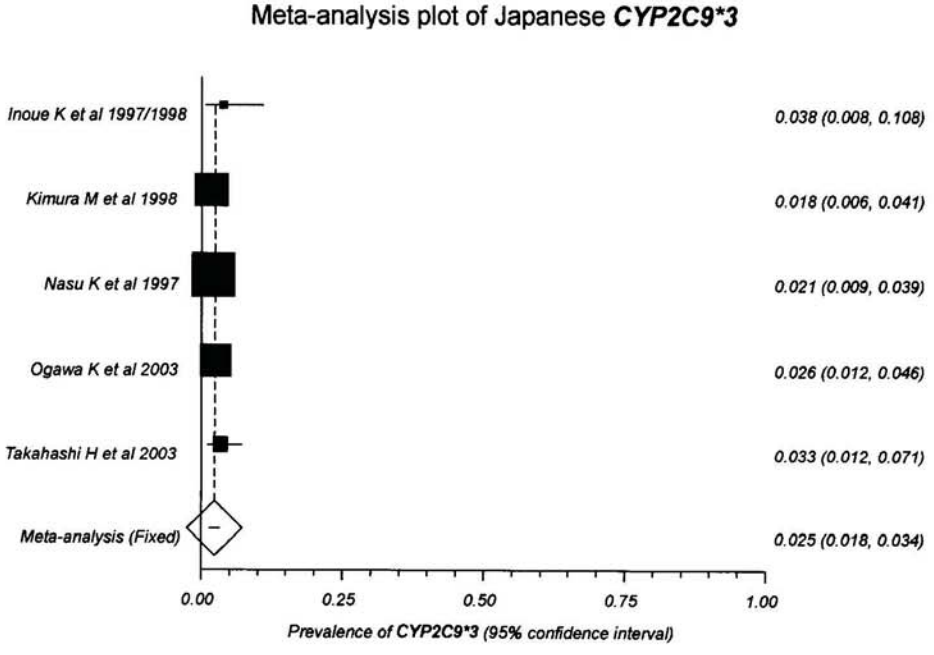


Figure 3.3.7c: Meta-analysis of ***CYP2C9*3*** prevalence in Japanese populations. (Q = 2.23, df = 4)
The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study is shown by the solid square. Pooled estimation of prevalence was presented by diamond.

4. ***African*** population studies:

Among the five studies of *CYP2C9* polymorphism in African populations, three did not provide completed genotype data, which included one study recruited patients with lung cancer only. Therefore they were excluded from meta-analysis for estimation of the CYP2C9 prevalence. Furthermore, apart from *CYP2C9*1*, **2*, and **3* alleles, the study of Allabi *et al* (2003) and Dickmann *et al* (2001) have found other *CYP2C9* variants, *CYP2C9*5*, **11* in studied African populations; therefore, more studies of African populations are required to elucidate the genetic model of *CYP2C9* and the estimated prevalence of *CYP2C9*2* and *3 here is inconclusive. The Cochrane Q value did not calculate for *CYP2C9*2* and *3 due to lack of data.

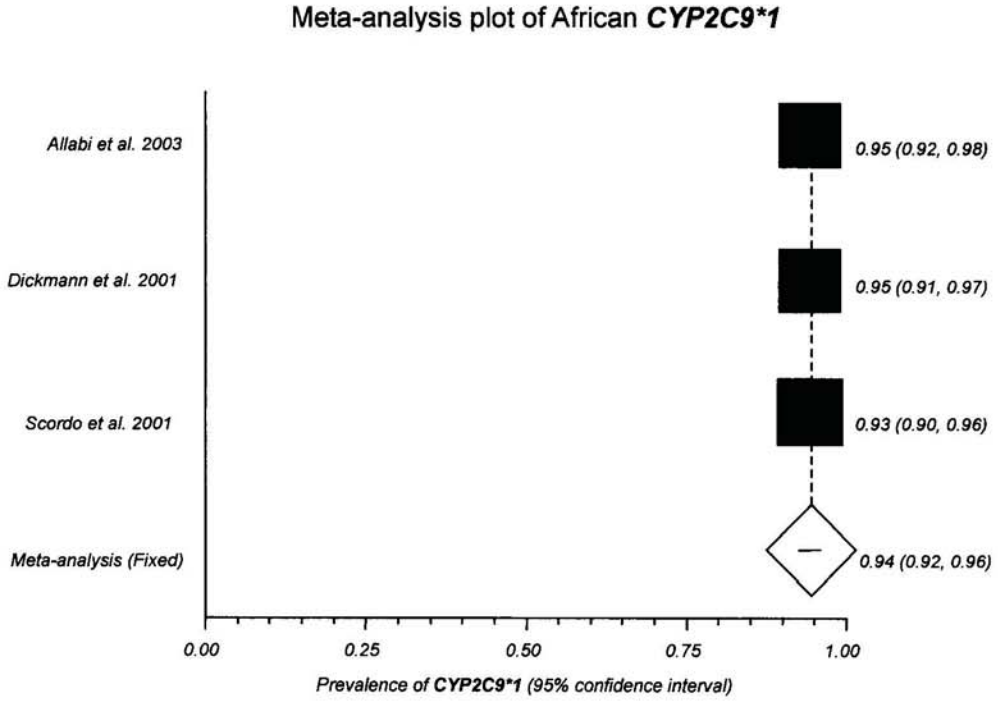


Figure 3.3.8a: Meta-analysis of ***CYP2C9*1*** prevalence in African populations. (Q = 1.08, df = 2)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

The pooled prevalence for *CYP2C9*1* in African population was 94% with 95% confidence interval 92 to 96%. Further more, the prevalence for *CYP2C9*2* in African population was less than 3.3% with 95% confidence interval 2.1 to 4.8%, and the prevalence for *CYP2C9*3* in African population was 1.7% with 95% confidence interval 0.9 to 2.9% as shown in figure 3.3.8b and c.

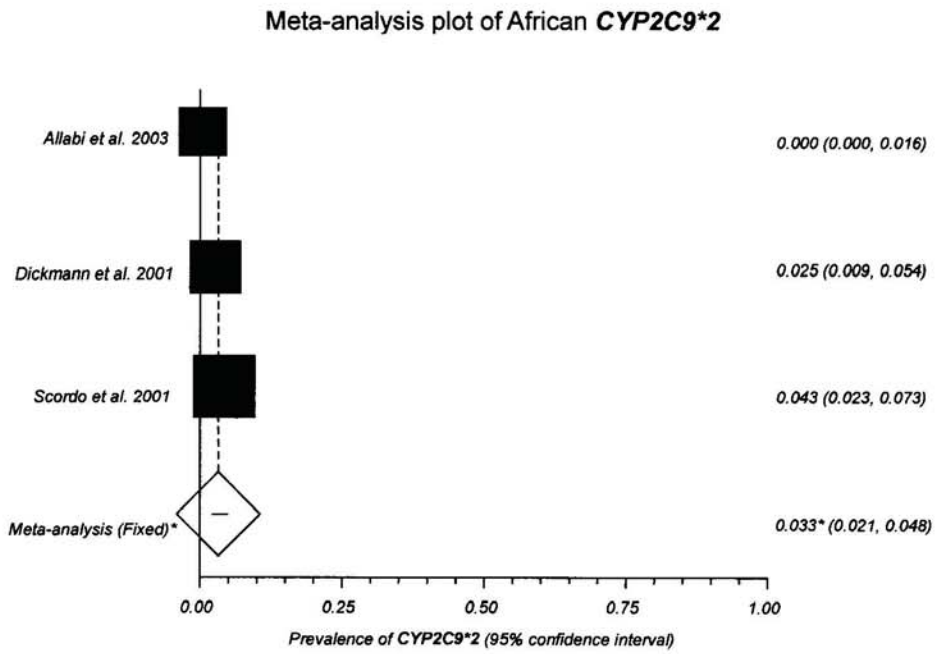


Figure 3.3.8b: Meta-analysis of ***CYP2C9*2*** prevalence in African populations.
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight shown by the solid square is the sample size of each study. *Pooled estimation of prevalence was presented by diamond, which are calculated from two studies that found *CYP2C9*2*.

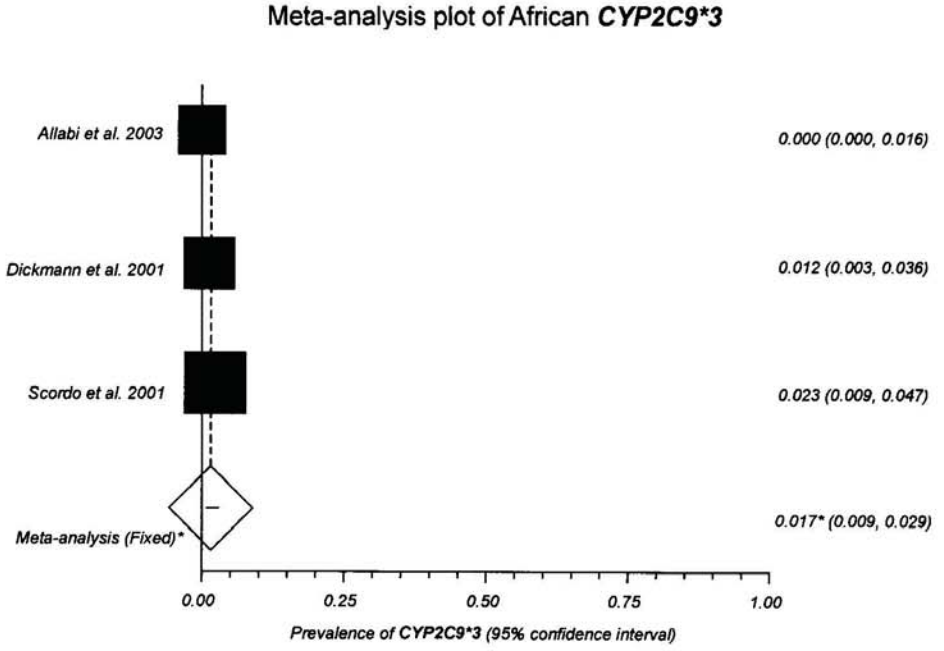


Figure 3.3.8c: Meta-analysis of ***CYP2C9*3*** prevalence in African populations.
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight shown by the solid square is the sample size of each study. *Pooled estimation of prevalence was presented by diamond, which are calculated from two studies that found *CYP2C9*3*.

*Prevalence of CYP2C9*1, *2 and *3 for Caucasian healthy subjects and patients under warfarin treatment:*

As previously mentioned, in order to estimate the population distribution of the three *CYP2C9* variants, the subjects of each study should have clearly defined ethnicity and the preferable recruitment method for subjects whose genotype was determined was randomly population-based sampling. However, among 49 Caucasian population studies of *CYP2C9*, eight studies recruited subjects in order to observe the association between CYP2C9 genotype and warfarin dose requirement in patients under anticoagulant treatment. Although some studies had provided patients pathological details, these patients cannot be considered as random sample of Caucasian populations. Furthermore, although most studies stated that the prevalence of three *CYP2C9* alleles was not different from a previous report, none of them have recruited healthy subjects as control group. Therefore, these studies were excluded for estimation of *CYP2C9*1*, **2* and **3* prevalence. Furthermore, it was interesting that when the population differentiation test was employed in GENEPOP software, the genotype data from studies of Caucasian patients appeared significantly heterogeneous; nevertheless, the genotype data from studies of healthy Caucasians appeared significantly homogeneous. Therefore, the meta-analysis has been applied to studies with healthy subjects and patients respectively, in order to find whether there is significant difference in the prevalence between them. The estimated prevalence of three *CYP2C9* alleles from healthy subjects and patients is summarized here in table 3.3.3.

Table 3.3.3: Summary of estimated *CYP2C9* allele prevalence with 95% CI in healthy and unhealthy **Caucasians**

Health status of subjects	*CYP2C9*1*	*CYP2C9*2*	*CYP2C9*3*	Number of individuals	Number of studies
Healthy	0.7880 (0.7705~0.8050)	0.1225 (0.1108~0.1348)	0.0885 (0.0784~0.0992)	2862	8
Patients	0.8156 (0.7991~0.8315)	0.1268 (0.1171~0.1368)	0.0670 (0.0491~0.0874)	6304	8

Results of healthy subjects and patients are individually presented for each allele in following figure 3.3.9-12 for *CYP2C9*1*, **2* and **3* respectively.

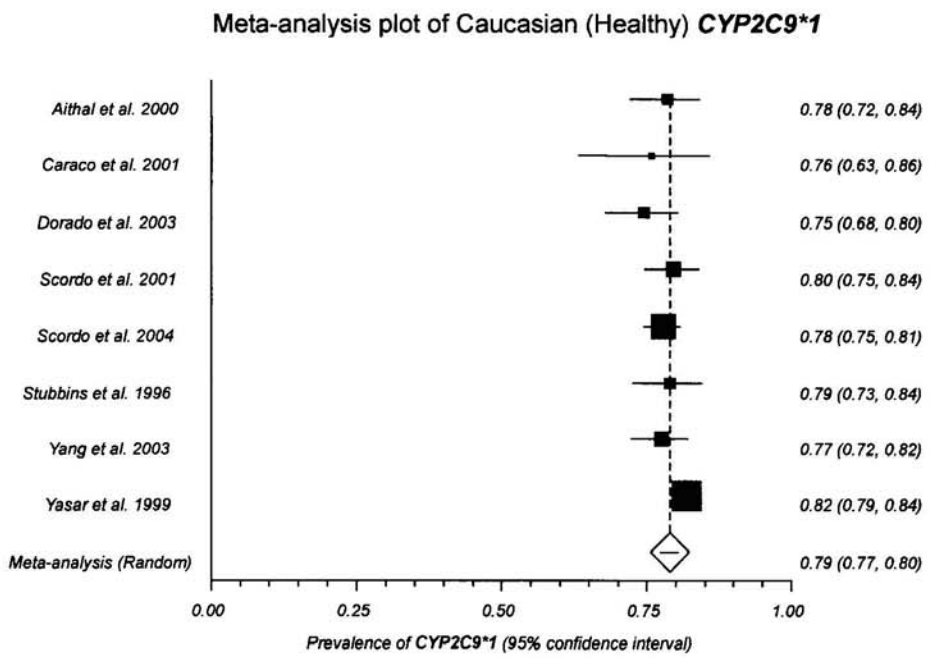


Figure 3.3.9a: Meta-analysis of ***CYP2C9*1*** prevalence in **healthy** Caucasian populations (Q = 8.51, df = 7). The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. Random model of meta-analysis was employed due to the Cochrane Q value.

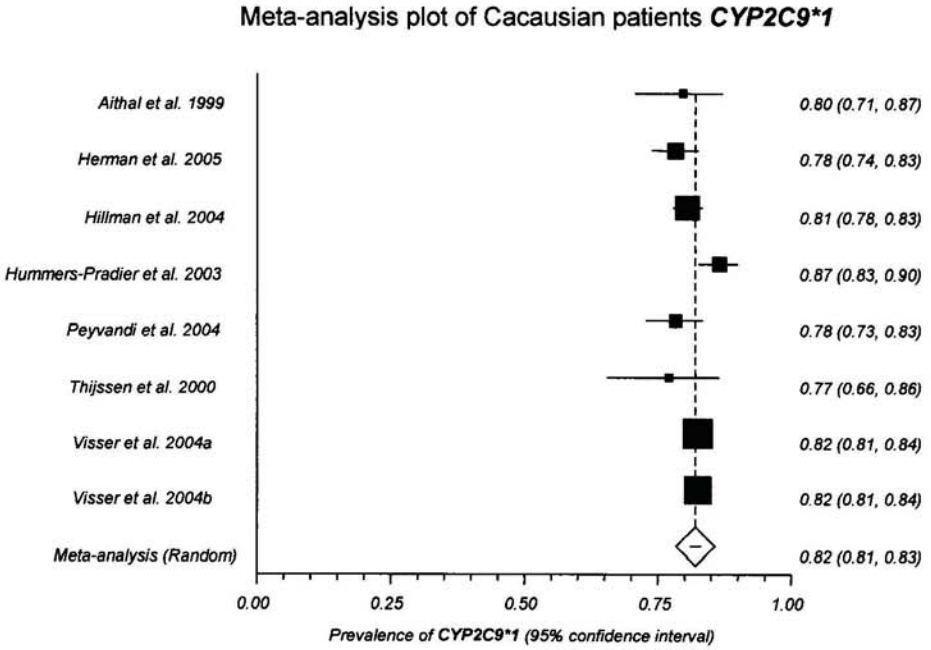


Figure 3.3.9b: Meta-analysis of ***CYP2C9*1*** prevalence in Caucasian **Patients** (Q = 14.07, df = 7)
The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. Random model of meta-analysis was employed due to the Cochrane Q value.

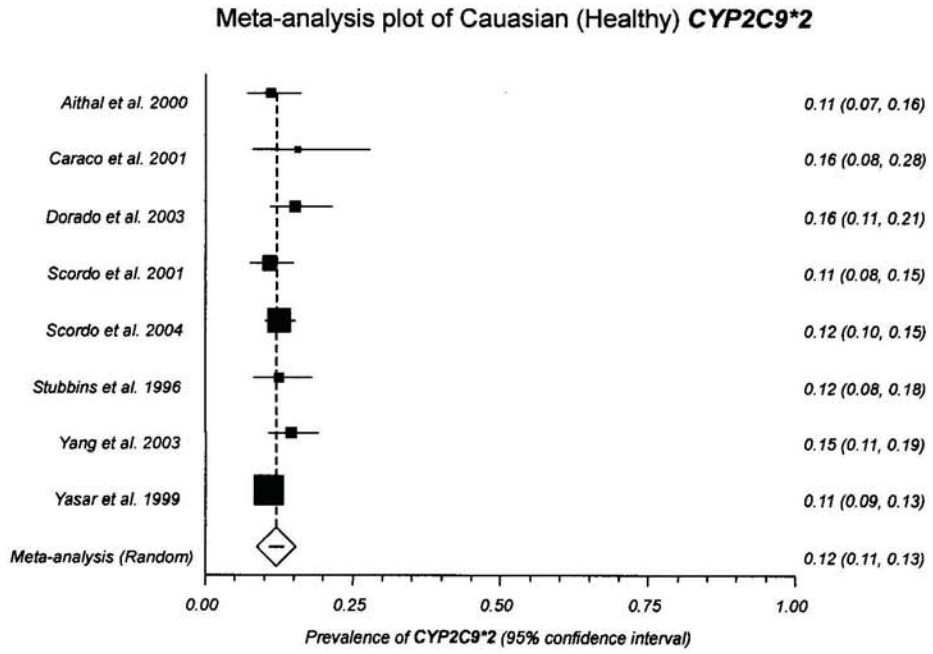


Figure 3.3.10a: Meta-analysis of ***CYP2C9*2*** prevalence in **healthy** Caucasian populations (Q =7.50, df = 7). The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. Random model of meta-analysis was employed due to the Cochrane Q value.

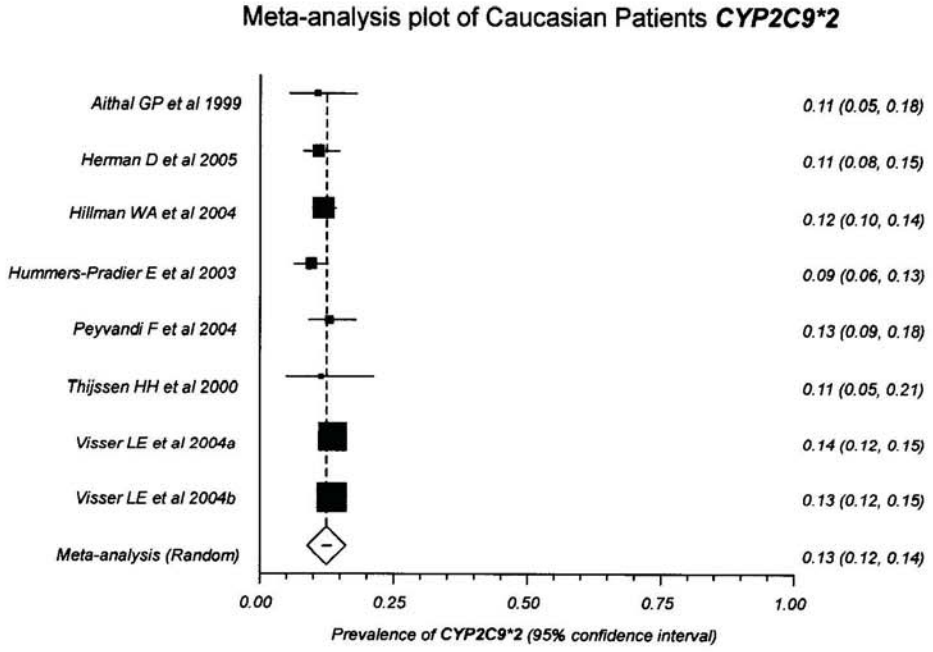


Figure 3.3.10b: Meta-analysis of ***CYP2C9*2*** prevalence in Caucasian **Patients** (Q = 8.40, df = 7)
The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. Random model of meta-analysis was employed due to the Cochrane Q value.

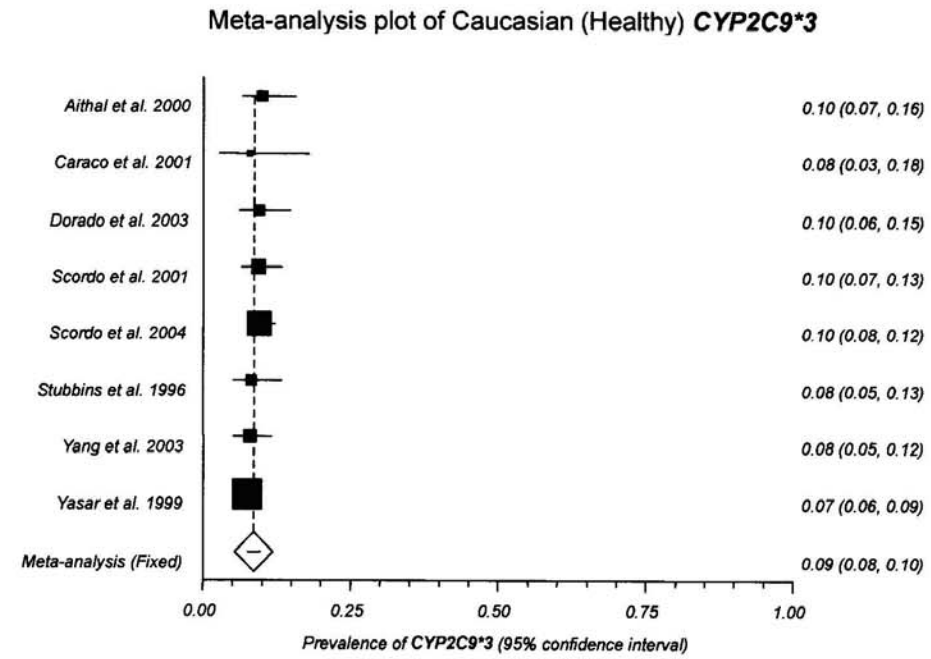


Figure 3.3.11a: Meta-analysis of ***CYP2C9*3*** prevalence in **healthy** Caucasian populations (Q = 4.45, df = 7). The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

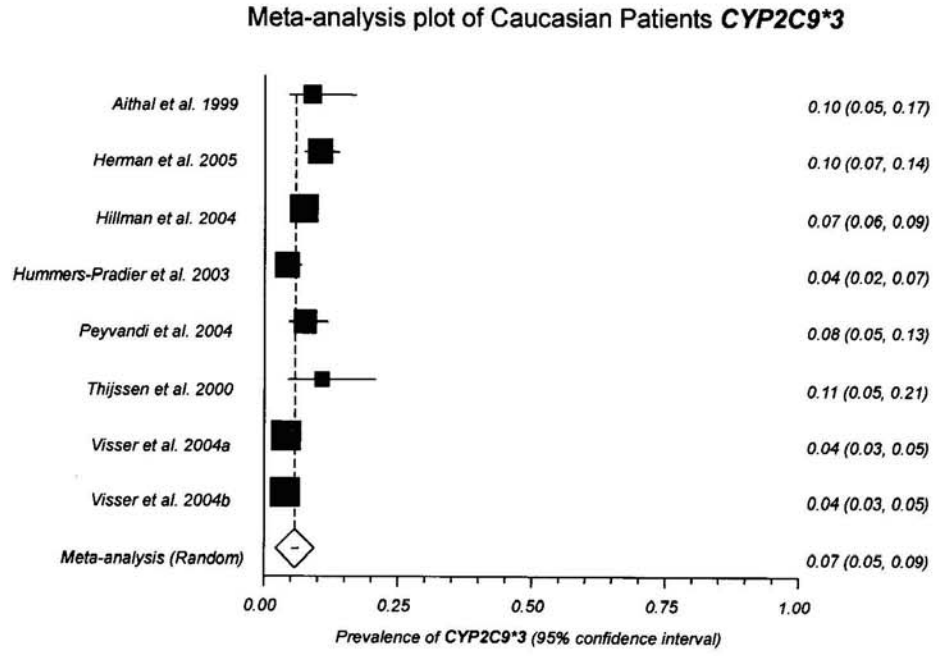


Figure 3.3.11b: Meta-analysis of ***CYP2C9*3*** prevalence in Caucasian **Patients** (Q = 49.38, df = 7)
The prevalence of each allele is presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which is indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. Random model of meta-analysis was employed due to the Cochrane Q value.

It can be seen from table 3.3.3 and figure 3.3.9-12 that apart from *CYP2C9*3*, the estimated prevalence of *CYP2C9*1* and **2* using a meta-analysis method appear relative close between studies of healthy and studies of unhealthy subjects. However, as previously mentioned, the population differentiation test in the GENEPOP program considers three *CYP2C9* alleles together, which showed significantly homogeneity between studies of healthy Caucasian subjects and exhibited significant heterogeneity between those studies of patients with Caucasian origin. Therefore, trinomial contours were employed to represent the dispersal of *CYP2C9* population distribution between studies sampling from Caucasians patients. As a comparison, the *CYP2C9* distribution in Caucasian population was represented here using the pooled prevalence of the three *CYP2C9* alleles from those homogenous studies of healthy Caucasians.

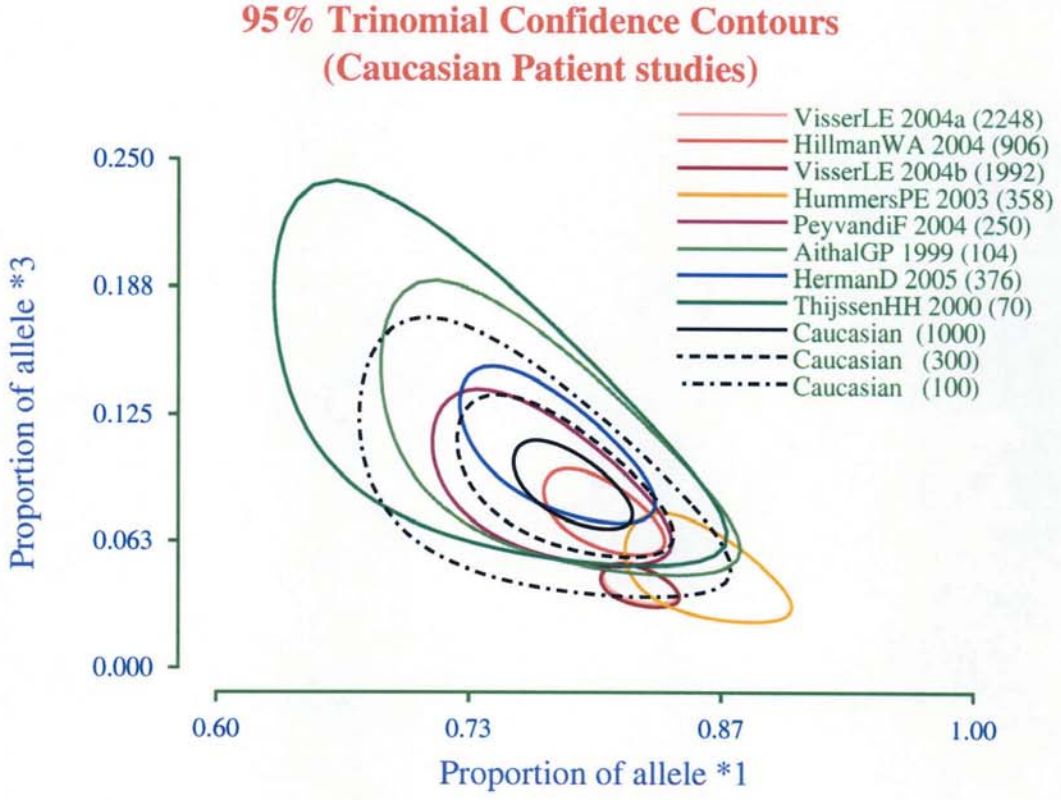


Figure 5.1.1: Trinomial contours of CYP2C9*1, *2 and *3 in Caucasian patients
Studies were presented according to name of first author with sample size in bracket. Caucasian was drawn according to the pooled prevalence of three CYP2C9 alleles from healthy Caucasian populations, which were represented in three sample sizes, 1000, 300 and 100, for visual comparison

As it can be seen in the figure, the sampling from patient populations would exhibit

different confidence contours for the three alleles. Three studies have no overlap with other studies, such as studies of Visser *et al* (2004a and b) and the study of Hummer-Pradier *et al.* (2003). Therefore, the prevalence estimation of *CYP2C9* alleles would be inappropriate if the data from these patient studies were included.

Meanwhile, the different recruitment of patients may partially explain the dispersal in these trinomial contours. According to the including and excluding criteria, these studies of Caucasian patients can be accepted as a Cohort study design, which have included patients undergoing anticoagulant treatments in clinics, and excluded subjects who have known underlying conditions that influence anticoagulant dosage or metabolism, such as cancer, renal or hepatic insufficiency and co-medications. However, referring table 3.2.1, of these Cohort studies, Visser *et al.* (2004a and b) and Hummer-Pradier *et al.* (2003) recruited patients from thrombosis clinics for a period time, which was not considered in the remaining studies. It is interesting that these three studies appeared to show dispersal from other Caucasian patient studies.

*Population distribution of CYP2C19*1, *2 and *3 in Caucasian, Chinese, Japanese and African:*

1. ***Caucasian*** population studies:

Among twenty-five studies of *CYP2C19* polymorphism in Caucasian populations, seven studies did not provide complete genotype data and were excluded from meta-analysis for estimation of the *CYP2C19* prevalence. Furthermore, three studies had a particular interest in poor-metabolizers. They were excluded as well. Another four studies were excluded due to the genotype data not fitting Hardy-Weinberg equilibrium. The study of Brosen *et al* (1995) was excluded because subjects were recruited from members of several families known as poor-metabolizers from previous studies.

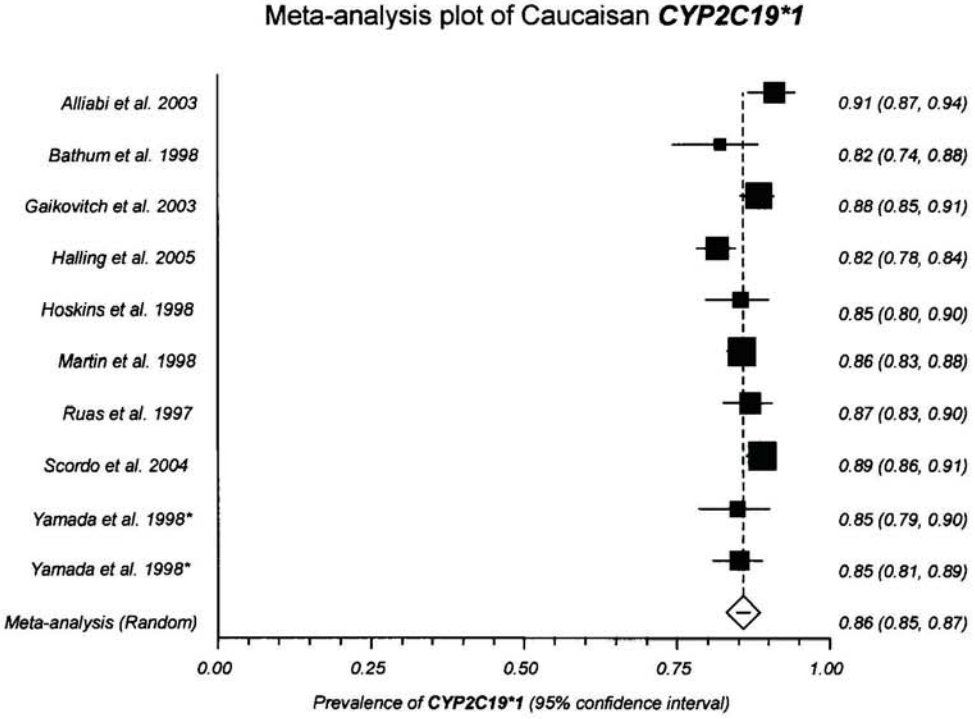


Figure 3.3.13a: Meta-analysis of ***CYP2C19*1*** prevalence in Caucasian populations. (Q = 25.12, df = 9)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for two subgroup populations of same ethnicity.

The pooled prevalence for *CYP2C19*1* in Caucasian population was 86% with 95% confidence interval 85 to 87%. Furthermore, the pooled prevalence for *CYP2C19*2* in Caucasian population was less than 14% with 95% confidence interval 13to 15%, and the prevalence for *CYP2C19*3* in Caucasian population was less 0.38% as shown in figure 3.3.13b and c.

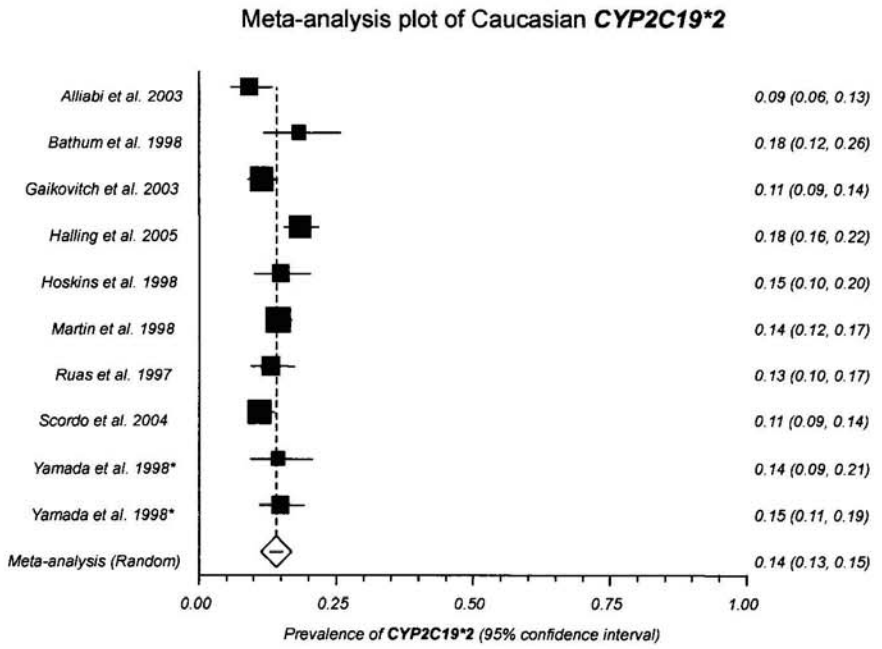


Figure 3.3.13b: Meta-analysis of ***CYP2C19*2*** prevalence in Caucasian populations. (Q = 25.74, df = 9)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for two subgroup populations of same ethnicity.

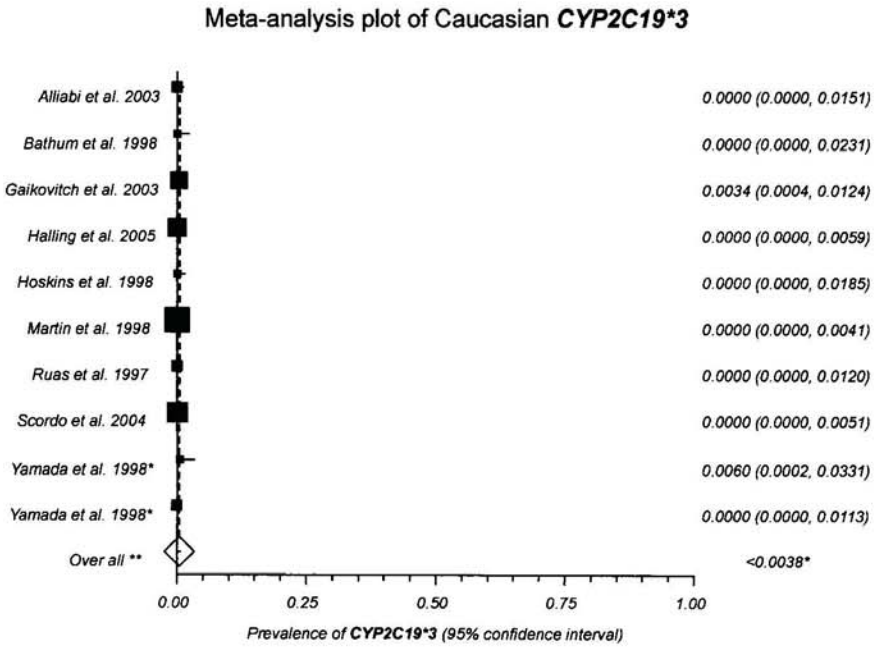


Figure 3.3.13c: Meta-analysis of ***CYP2C19*3*** prevalence in Caucasian populations.
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight shown by the solid square is the sample size of each study. The Q value in studies of *CYP2C19*3* cannot be calculated since most studies did not find the allele. *One publication provided data for two subgroup populations of same ethnicity. **The overall estimation is calculated by arithmetic average of entire studies.

2. ***Chinese*** population studies:

Four of fourteen retrieved Chinese population studies of *CYP2C19* polymorphism did not have complete genotype data for *CYP2C19*. They were excluded from the prevalence estimation. Among the remaining ten studies, the genotype data of Chinese (Bai) population (Xiao *et al.* 1997) did not fit the Hardy-Weinberg equilibrium, therefore it was excluded from meta-analysis. Furthermore, other two Chinese studies based on epilepsy patients only were also excluded.

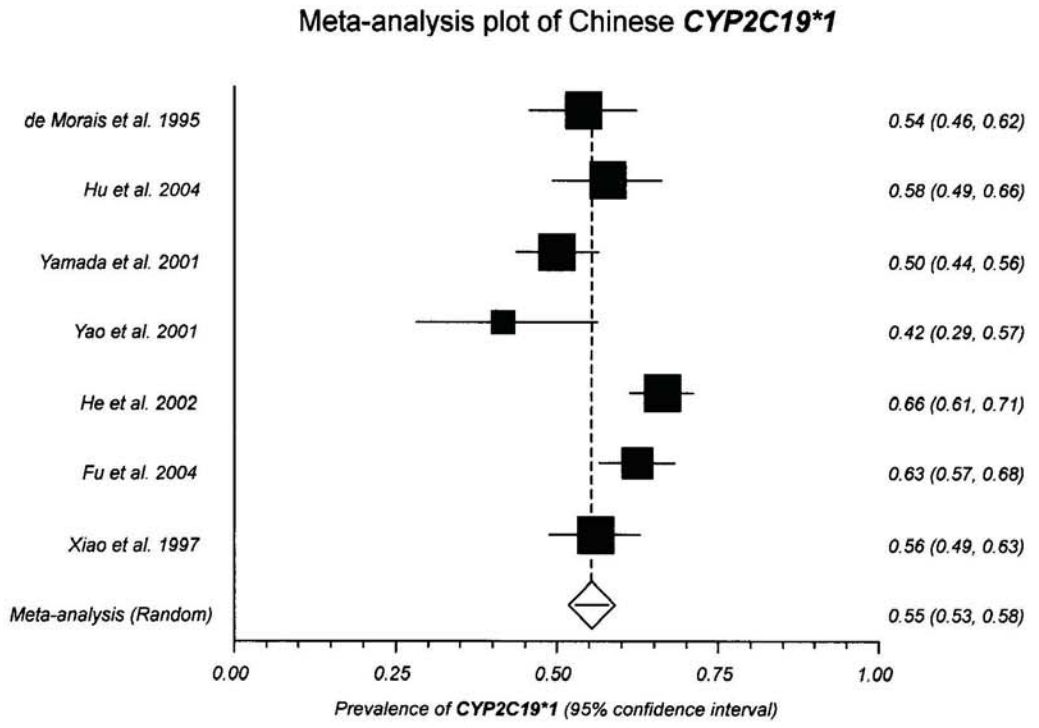


Figure 3.3.14a: Meta-analysis of ***CYP2C19*1*** prevalence in Chinese populations. (Q = 26.24, df = 6)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

The pooled prevalence for *CYP2C19*1* in Chinese population was 55% with 95% confidence interval 53 to 58%. Furthermore, the pooled prevalence for *CYP2C19*2* in Chinese population was less than 39% with 95% confidence interval 36 to 42%, and the pooled prevalence for *CYP2C19*3* in Chinese population was 6% with 95% confidence interval 5 to 7% as shown in figure 3.3.14b and c.

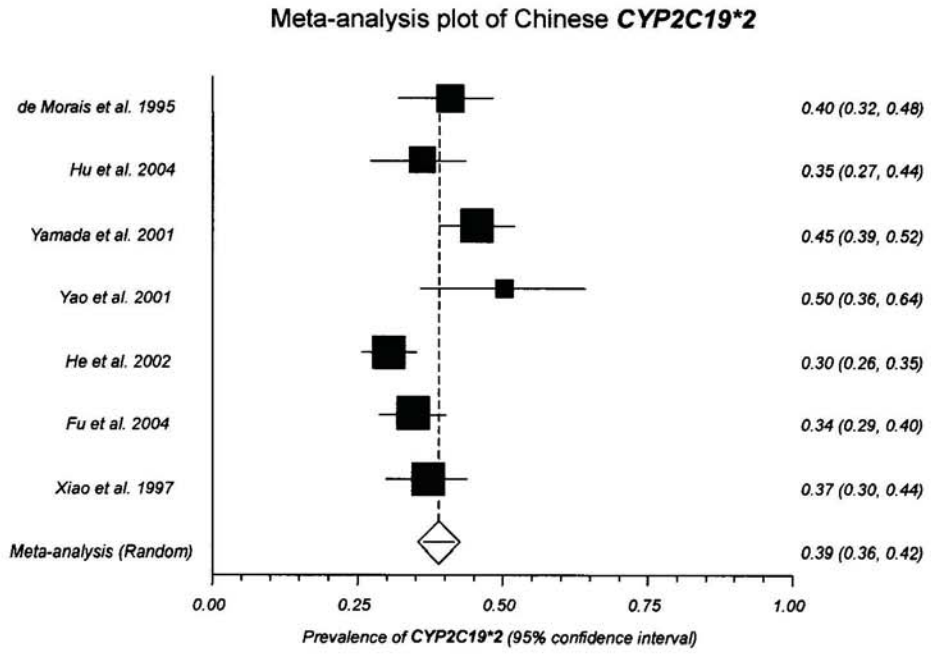


Figure 3.3.14b: Meta-analysis of ***CYP2C19*2*** prevalence in Chinese populations. (Q = 20.07, df = 6)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

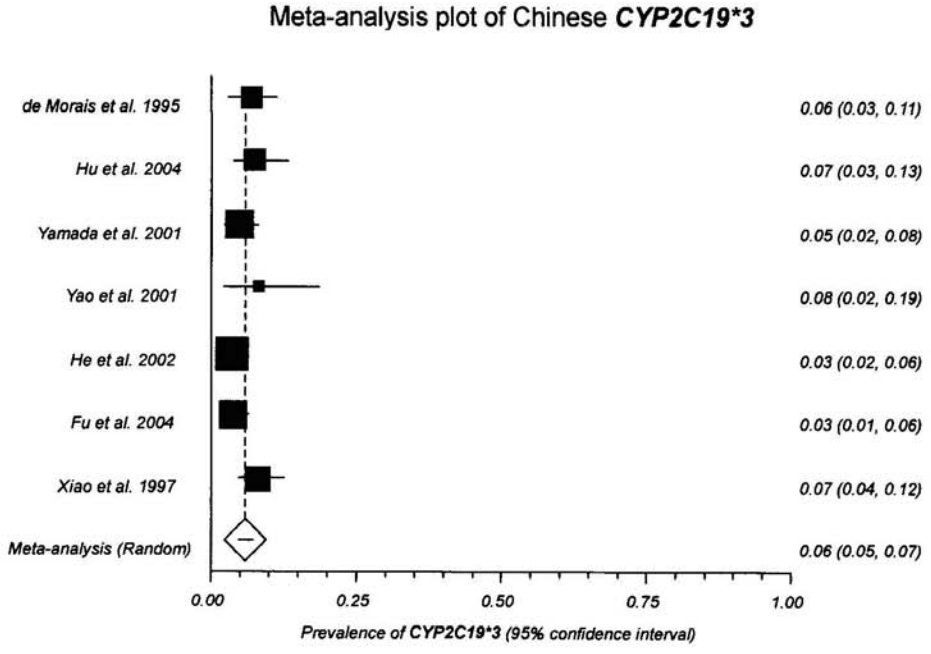


Figure 3.3.14c: Meta-analysis of ***CYP2C19*3*** prevalence in Chinese populations. (Q = 9.50, df = 6)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond.

3. ***Japanese*** population studies:

There was only one Japanese population study describing association of CYP2C19 polymorphism that did not provide complete genotype data, which was excluded from meta-analysis. Three studies that recruited subjects with same disease, such epilepsy or heart disease, were also excluded for the prevalence estimation. One study was excluded due to the small sample size (6 subjects). The remaining nine studies were analyzed using a meta-analysis method to estimate the prevalence of *CYP2C19*1*, **2* and **3* in the Japanese population, where all data was fitting Hardy-Weinberg equilibrium.

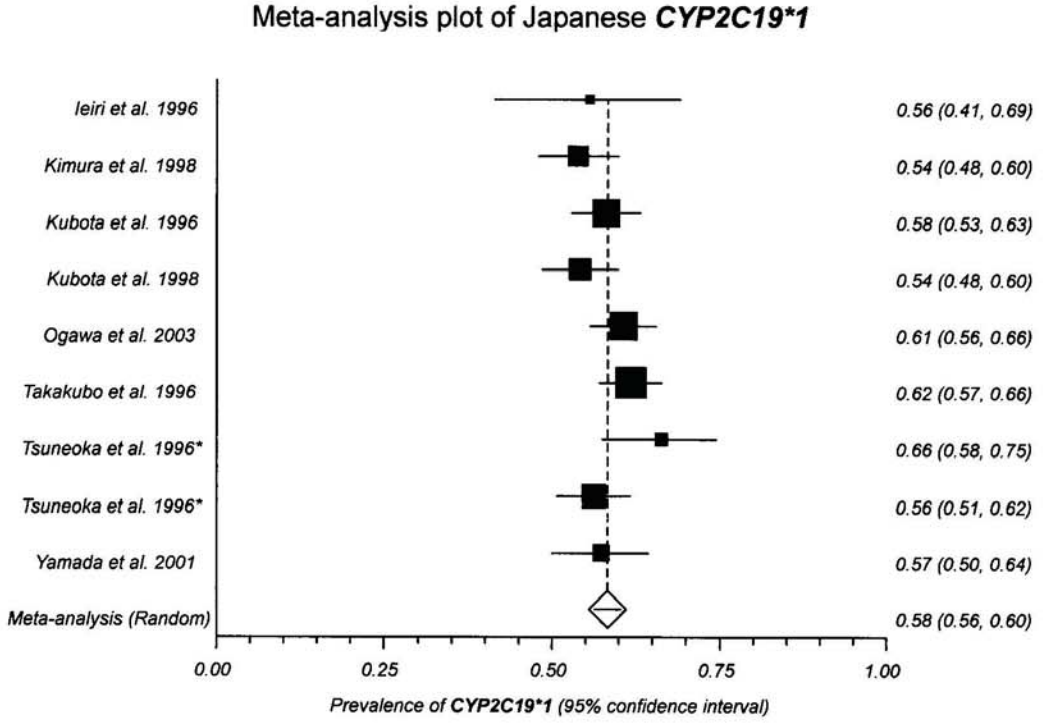


Figure 3.3.15a: Meta-analysis of ***CYP2C19*1*** prevalence in Japanese populations. (Q = 11.75, df = 8)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for two subgroup populations of same ethnicity.

The pooled prevalence for *CYP2C19*1* in Japanese population was 58% with 95% confidence interval 56 to 60%. Furthermore, the pooled prevalence for *CYP2C19*2* in Japanese population was 30% with 95% confidence interval 28 to 32%, and the pooled prevalence for *CYP2C19*3* in Japanese population was 12% with 95% confidence interval 11 to 13% as shown in figure 3.3.15b and c.

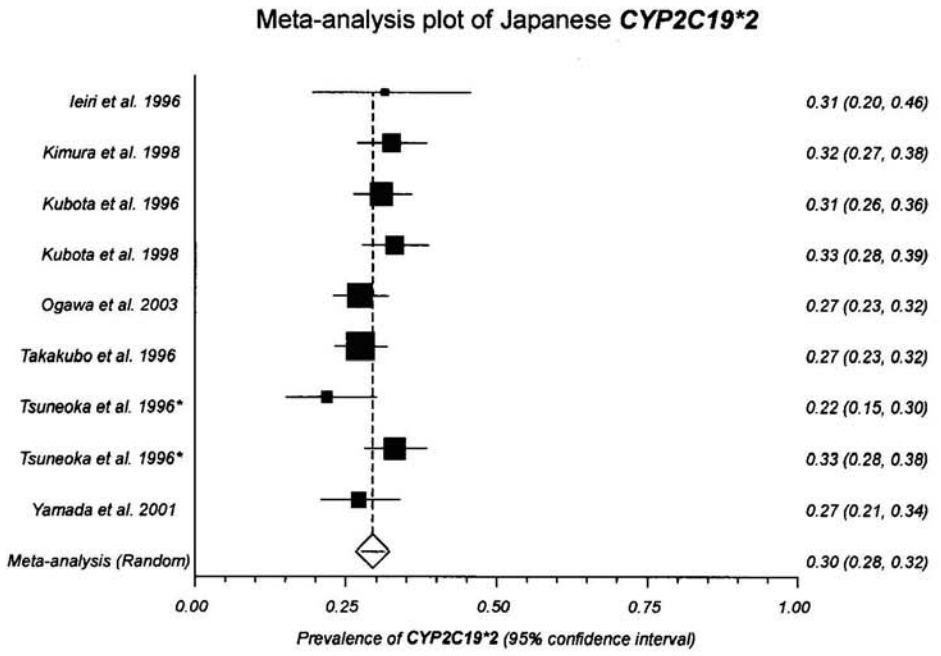


Figure 3.3.15b: Meta-analysis of ***CYP2C19*2*** prevalence in Japanese populations. (Q = 11.73, df = 8)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for two subgroup populations of same ethnicity.

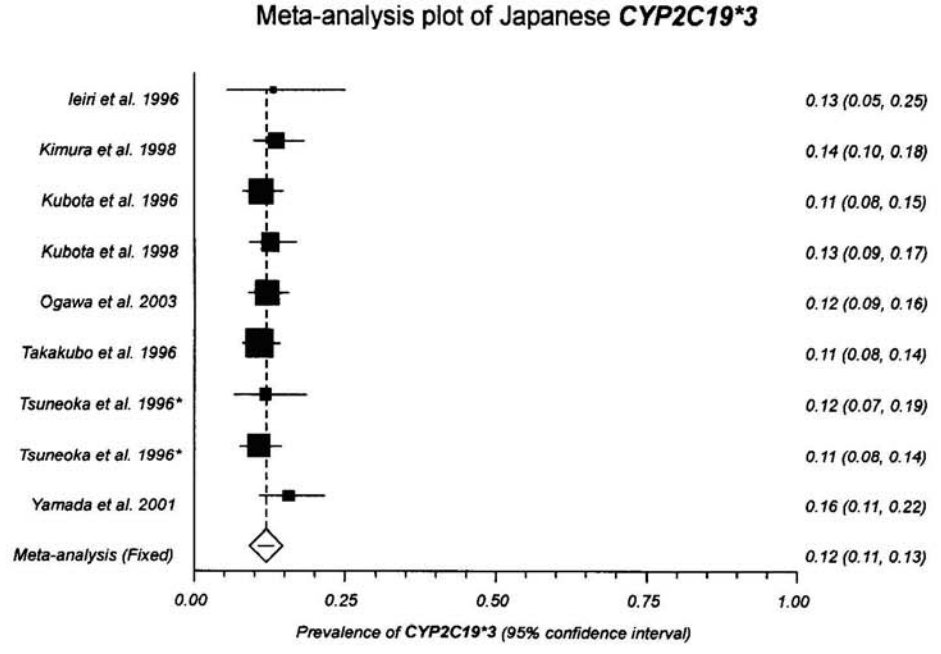


Figure 3.3.15c: Meta-analysis of ***CYP2C19*3*** prevalence in Japanese populations. (Q = 4.63, df = 8)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for two subgroup populations of same ethnicity.

4. ***African*** population studies:

After excluding one study of psychiatric patients, two more studies were excluded due to incomplete genotype data, another two studies were also excluded because the data did not fit the Hardy-Weinberg equilibrium. Therefore, the prevalence of *CYP2C19*1*, **2*, and **3* was estimated for the African population by employing meta-analysis method on the remaining ten studies.

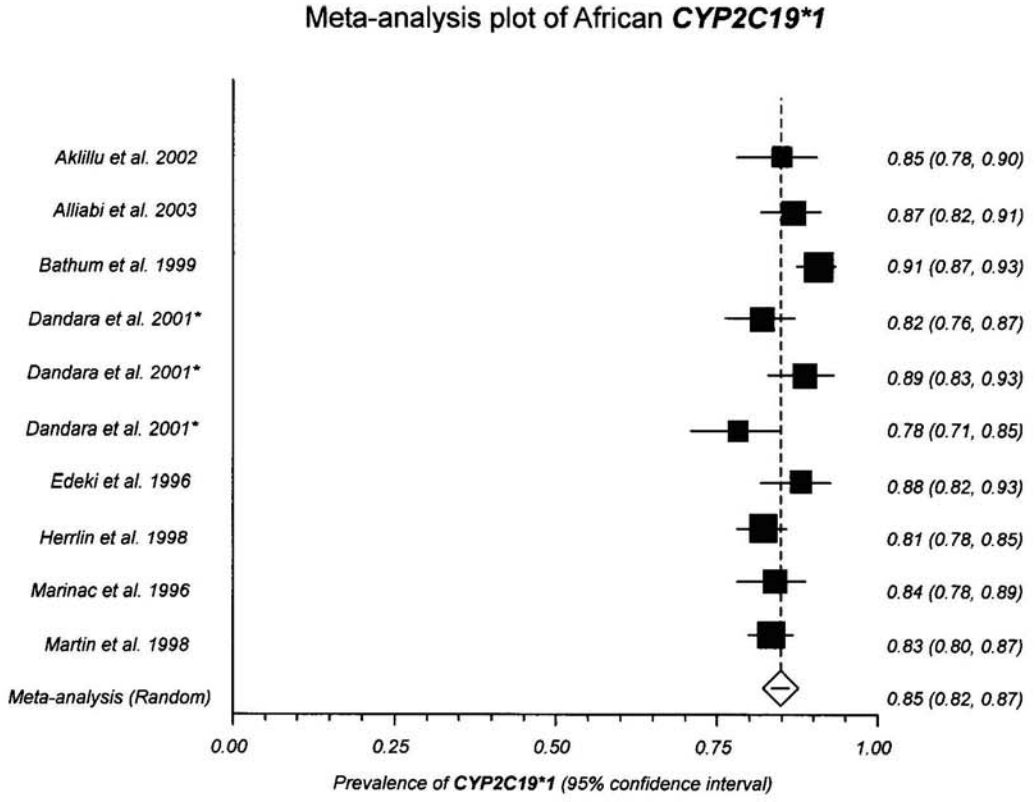


Figure 3.3.16a: Meta-analysis of ***CYP2C19*1*** prevalence in African populations. (Q = 26.95, df = 9)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for three subgroup populations of same ethnicity.

The pooled prevalence for *CYP2C19*1* in African population was 85% with 95% confidence interval 82 to 87%.

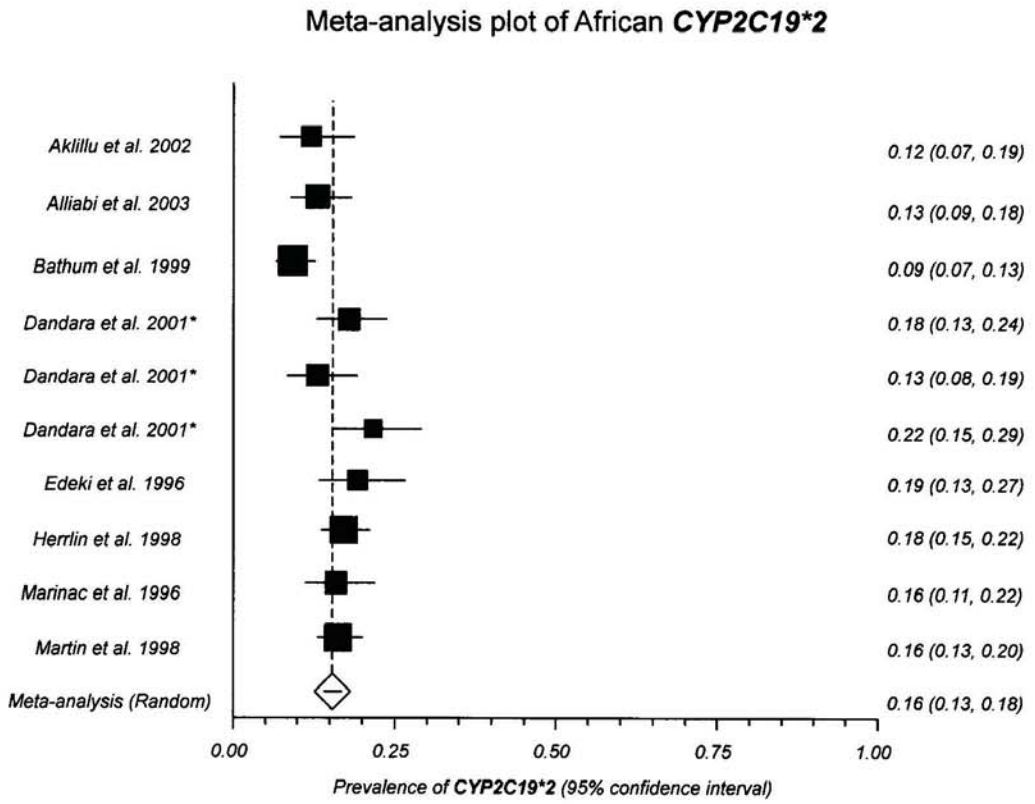


Figure 3.3.16b: Meta-analysis of ***CYP2C19*2*** prevalence in African populations. (Q = 25.71, df = 9)
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for three subgroup populations of same ethnicity.

As shown in figure 3.3.16b, the pooled prevalence for *CYP2C19*2* in African population was 16% with 95% confidence interval 13 to 18%.

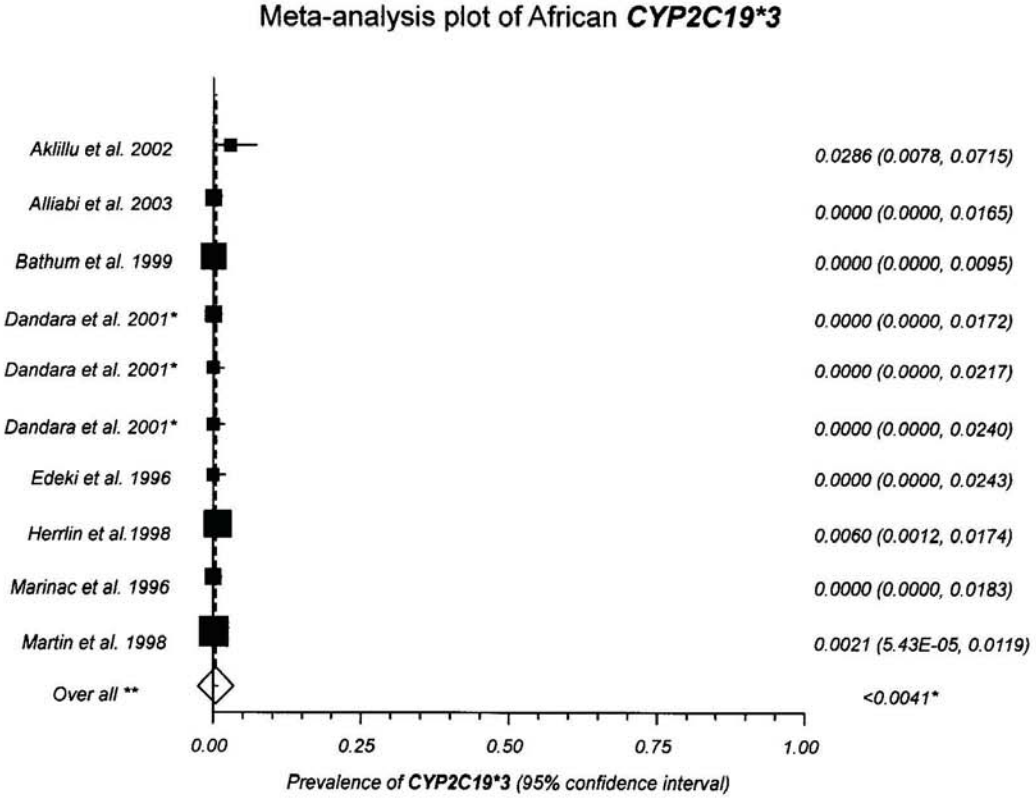


Figure 3.3.16c: Meta-analysis of ***CYP2C19*3*** prevalence in African populations.
The prevalence of each allele was presented with 95% confidence interval in the bracket on the left side of the graph according to each original publication, which was indicated by first author name and year of publication. The weight of each study was shown by the solid square. Pooled estimation of prevalence was presented by diamond. *One publication provided data for three subgroup populations of same ethnicity. **The pooled prevalence and the Cochrane Q value in African population studies of *CYP2C19*3* cannot be calculated because most studies reported the prevalence of this allele as zero. The over all estimation is calculated by arithmetic average of all studies.

The overall prevalence for *CYP2C19*3* in African population was less than 0.4%.

3.3.5 Prevalence estimation of six *CYP2C9/2C19* genotypes for different ethnic populations

As previously mentioned, for the population fitting the Hardy-Weinberg equilibrium (HWE), the prevalence of each possible genotype can be estimated using the prevalence of each allele described in the population. Therefore, when these four ethnic populations follow the HWE, the proportion of six *CYP2C9/2C19* genotypes can be predicted for each population using the estimated prevalence of the three *CYP2C9/2C19* alleles from the meta-analysis result in previous section. If the prevalence of *1, *2 and *3 was denoted as p, q, r, according to HWE, the prevalence of six genotypes (*1/*1, *1/*2, *1/*3, *2/*2, *2/*3, *3/*3) can be calculated as following:

Allele prevalence: $p + q + r = 1$

Expected prevalence of six genotypes: $p^2 + 2pq + 2pr + q^2 + 2qr + r^2 = 1$

Results are shown in table 3.3.4a and b for *CYP2C9* and *CYP2C19* respectively.

Table 3.3.4a: summary of estimated six *CYP2C9* genotypes prevalence in different ethnicity populations

Ethnicity	*CYP2C9 *1/*1*	*CYP2C9 *1/*2*	*CYP2C9 *1/*3*	*CYP2C9 *2/*2*	*CYP2C9 *2/*3*	*CYP2C9 *3/*3*
Caucasian	0.649	0.191	0.122	0.0140	0.0179	0.00576
Chinese	0.924	<0.00244 [a]	0.0736	<1.61E-06 [a]	<9.7E-05 [a]	0.00147
Japanese	0.950	0 [a]	0.0480	0 [a]	0 [a]	6.50 E-04
African	0.888	0.0616 [b]	0.0318 [b]	0.00107	0.00111 [b]	2.86 E-04 [b]

Note: a. The value is presented for data compliance only. b. The value was inconclusive, presented for data compliance.

Table 3.3.4b: summary of estimated six *CYP2C19* genotypes prevalence in different ethnicity populations

Ethnicity	*CYP2C19 *1/*1*	*CYP2C19 *1/*2*	*CYP2C19 *1/*3*	*CYP2C19 *2/*2*	*CYP2C19 *2/*3*	*CYP2C19 *3/*3*	PM [b]
Caucasian	0.741	0.238	<0.00654 [a]	0.0190	<0.00105 [a]	<1.44 E-05 [a]	<0.0210
Chinese	0.320	0.428	0.0591	0.143	0.0395	0.00272	0.185
Japanese	0.339	0.346	0.141	0.0883	0.0720	0.0147	0.175
African	0.721	0.264	<0.00951 [a]	0.0241	<0.00174 [a]	<3.14 E-05 [a]	<0.0260

Note: a. The value was obtained for data compliance b. PM stands for poor metabolizers, which are subjects who have one of three genotypes, *CYP2C19*2/*2*, **2/*3*, **3/*3*. The value is the sum of the prevalence of these three genotypes.

3.4 Trinomial distribution plot of three CYP2C9/2C19 alleles for specific population studies

Studies in the miscellaneous group provided CYP2C9/2C19 variant distribution for different individual ethnic populations. Due to the distinctive human evolution and immigration history, these populations have quite unique characteristics from each other. These studies cannot be easily categorized into ethnic group such as Caucasian or African. Furthermore, each ethnic population has only one or two studies available. Therefore, a meta-analysis method cannot be employed for the prevalence estimation of these ethnicities. In order to present the *CYP2C9/2C19* polymorphism distributions, trinomial contours were produced using the data of these ethnic populations. Meanwhile, according to the phylogenetic tree (Cavalli-Sforza *et al.* 1994), these ethnic populations have close genetic distance with one or some ethnicities, such as Caucasian, Chinese, Japanese and African, whose prevalence of the three *CYP2C9/2C19* alleles have been estimated in previous section using a meta-analysis methods. Therefore, trinomial contours of these populations in the miscellaneous group are presented together with the ethnicity or ethnicities that show close genetic distance in the phylogenetic tree, which are shown in figure 3.3.1-6 to figure 3.4.1-5 for *CYP2C9* and *CYP2C19* respectively. The *CYP2C9* and *CYP2C19* polymorphism in Caucasian, Chinese, Japanese and African population are first shown in figure 3.3.1 and figure 3.4.1 using the estimated prevalence of the three alleles in previous section.

For studies of *CYP2C9* polymorphism in the miscellaneous group, the trinomial contours are represented in five groups: 1.populations close to Caucasian; 2.population close to Chinese and Japanese; 3.population with unique evolution history; 4.population with ancestries from different ethnic origins; 5.studies of subjects from different ethnic groups.

For studies of *CYP2C19* polymorphism in the miscellaneous group, the trinomial contours are represented in four groups: 1.populations close to Caucasian; 2.population close to Chinese and Japanese; 3.population with unique evolution history; 4.population of Pacific Islander.

*Distribution of CYP2C9*1, *2 and *3 in populations of the miscellaneous groups:*

The trinomial contours were produced in SimFit software for Caucasian, Chinese, Japanese and African populations prevalence of CYP2C9*1, *2 and *3 with 95% confidence.

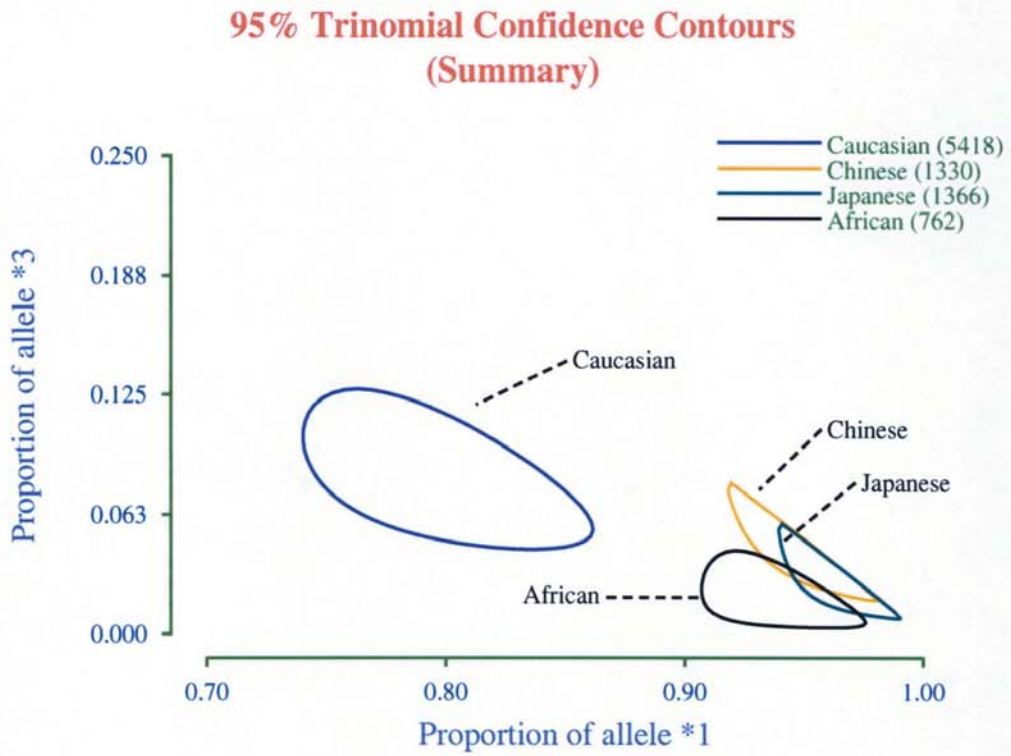


Figure 3.4.1: Prevalence of *CYP2C9*1*, **2* and **3* in different ethnic populations (meta-analysis pooling results)

The trinomial contours visually presented the different distributions of three *CYP2C9* alleles between Caucasian and other ethnic populations. As the previous description of the *CYP2C9* prevalence estimation for these populations showed, studies of Chinese, Japanese populations found that the *CYP2C9*2* variant is a rare allele in these two populations, and approximately 2% of Chinese and 3.6% of Japanese populations have the *CYP2C9*3* variant. Therefore, the trinomial contours in both Chinese and Japanese populations have one side straight line and some overlapping regions. Meanwhile, the *CYP2C9*2* and **3* variant in African populations also have a relatively lower prevalence compared with Caucasian populations. Furthermore, apart from *CYP2C9*2* and **3* mutants, two of three African population studies have found other CYP2C9 variants. Therefore, trinomial contours of African population are not conclusive results. For data completeness, it was represented in this trinomial contours section.

The trinomial contours of Caucasian, Chinese, Japanese and African populations were represented in the following figures 3.4.2-6 with consistent colour legend for visual presentation of the *CYP2C9* polymorphism distribution in different ethnic populations.

1. Populations close to Caucasian:

In the miscellaneous group, several populations are quite close to Caucasian populations according to phylogenetic tree created by DNA markers by Cavalli-Sforza *et al.* in 1994. As each of these populations has not been reported by more than one study reported, meta-analysis is not employed. Therefore, allele frequencies of *CYP2C9*1*, **2* and **3* are presented as 95% trinomial confidence contours here in figure 3.4.2 in order to show the diversity of the three alleles distributed among these populations.

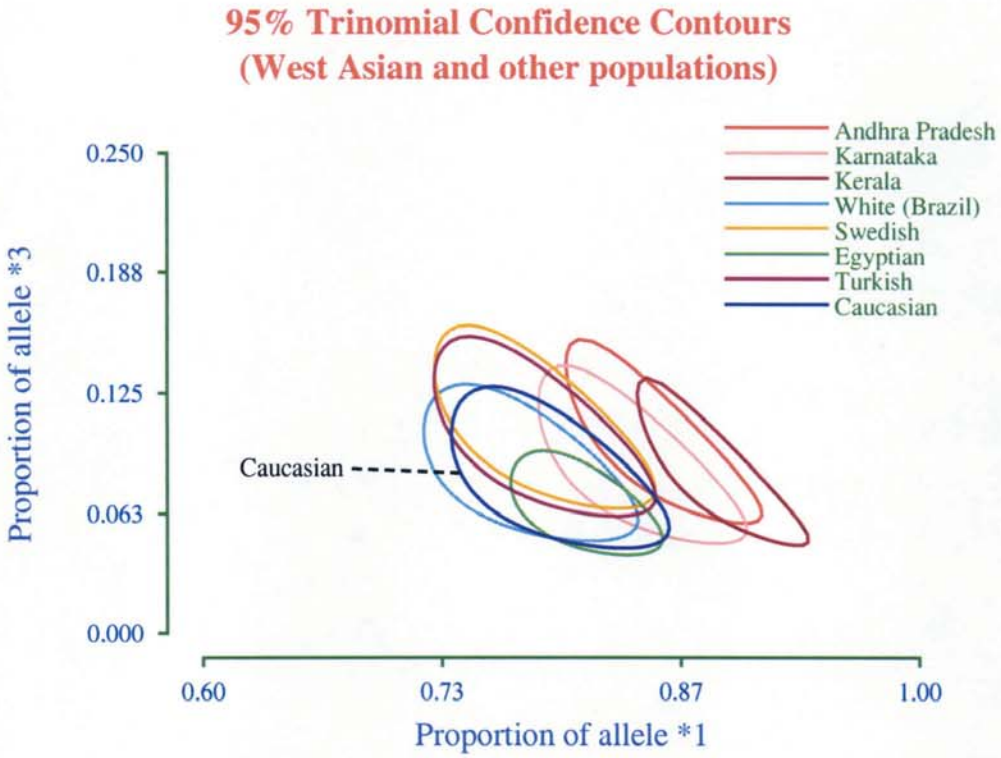


Figure 3.4.2: Trinomial confidence contours of *CYP2C9*1*, **2* and **3* in west Asian and other populations comparing with Caucasians.
First three populations are South India. One study did not provide the ethnicity directly, however the studied population is highly possible Swedish, the data is presented in the figure.

The *CYP2C9* polymorphism distribution in populations of South India, Egypt and Turkey appear as largely overlapping with the Caucasian population. This fits the phylogenetic tree derived by DNA markers (Cavalli-Sforza *et al.* 1994).

2. Populations close to Chinese:

In the miscellaneous group, three populations (one Vietnamese, two Korean) are predicted to be close to Chinese and Japanese populations according to the phylogenetic tree (Cavalli-Sforza *et al.* 1994). One study of two Korean population studies only found *CYP2C9*1*, and the data could not be used to produce trinomial contours. As the remaining populations have not been studied more than once, it is not possible to apply meta-analysis. Therefore, allele frequencies of *CYP2C9*1*, **2* and **3* in these populations are presented in 95% trinomial confidence contours. Furthermore, in order to show the diversity of the *CYP2C9* distribution among Asian populations, data from populations of South India and Turkey are also presented here. The data of south India was obtained by pooling data of three individual studies, which have no significant heterogeneity.

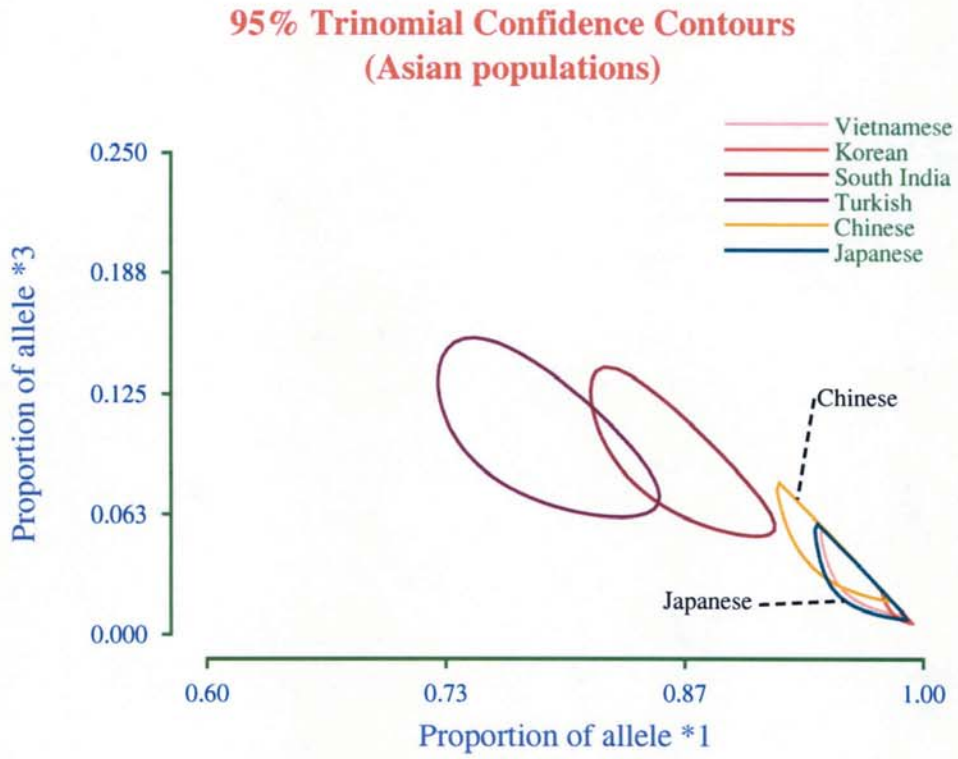


Figure 3.4.3: Trinomial confidence contours of *CYP2C9*1*, **2* and **3* in different Asian populations

The trinomial contours in figure 3.4.3 show that populations of South India and Turkish have no overlapping area with Chinese, Japanese, however, Vietnamese and Korean populations appear overlapping with Chinese and Japanese. The *CYP2C9* distribution in those Asian populations appear to fit the phylogenetic tree proposed by Cavalli-Sforza *et al.* in 1994.

3. Specific populations:

In the miscellaneous group, several populations are quite unique from other ethnic populations according to human evolutionary history. Their allele frequencies of CYP2C9*1, *2 and *3 were presented in 95% trinomial confidence contours here in order to show the diversity of the three alleles population distribution, where the trinomial contours of Caucasian, Chinese, Japanese and African populations are represented together for comparison.

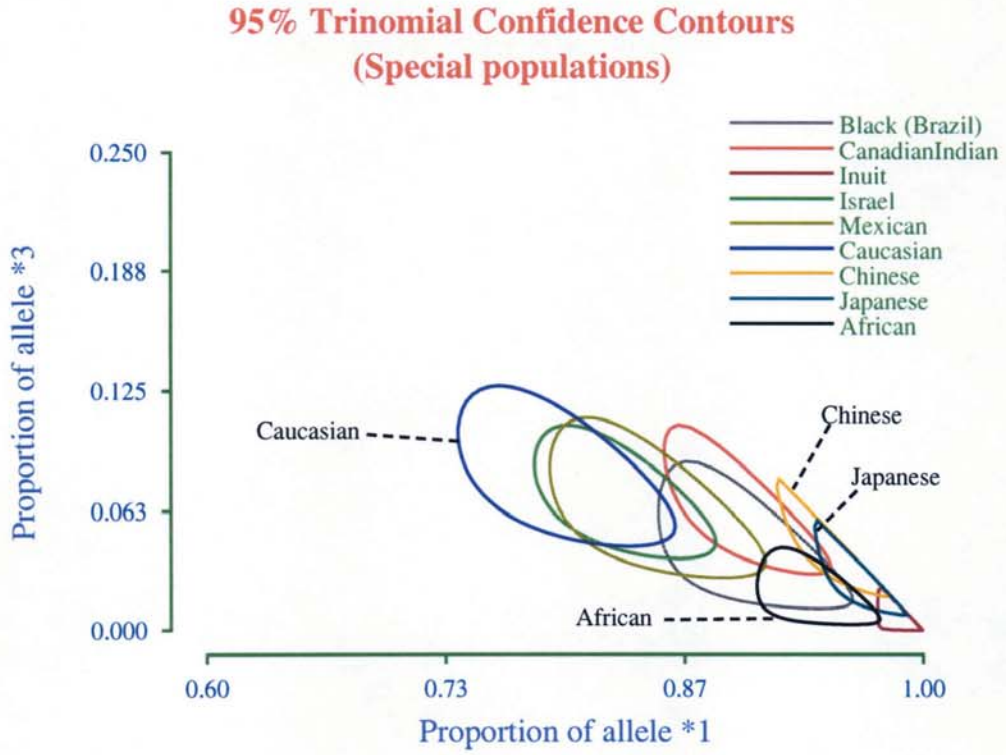


Figure 3.4.4: Trinomial confidence contours of *CYP2C9*1*, **2* and **3* in different populations comparing with Caucasian, Chinese, Japanese and African populations

There are some overlapping regions between black population in Brazil and African populations. According to the Brazil census (2000), the majority of the black population in Brazil is descended from those African people who are transported into Brazil as slavers. However, due to the complexity of immigration and intermarriage with other ethnic populations, the black population in Brazil cannot be simplified as an African population. Furthermore, the genotype data from this black population of Brazil did not fit HWE, which may indicate the possibility of unknown variants in the population. Similarly, Israeli

Mexican population are also difficult to define into any ethnic category. In the CYP2C9 trinomial contours, these two populations have partly overlapping contours with Caucasian and several other ethnic populations. Canadian native Indian and Inuit populations are two unique populations from Caucasian, Chinese, Japanese and African. Although the trinomial contours of Inuit population have overlapping with Japanese population, as one genetic polymorphism is insufficient to population diversity, any conclusion of the population differentiation between these populations will be inappropriate. Therefore, more studies with these specific populations are required to elucidate the difference of CYP2C9 distribution between these ethnic population and other ethnicities.

4. Admixed populations:

Some studies provided three CYP2C9 allele frequencies for several unique populations, which are subject with ancestors from different ethnic origins, such as subjects with 100% Canadian Native Indian origin, and subjects with 50% Canadian Native Indian origin. Meanwhile, as the authors described in the original publications, the Bolivian population was descendant of Caucasian and America Indian and the intermediate population in Brazil is descendant of Caucasian and African. Therefore, the trinomial contours of these populations were presented in the following figure 3.4.5 with Caucasian, Canadian Native Indian, and African population together. As previously mentioned, it is inappropriate to conclude the differentiation between populations by one genetic polymorphism; the trinomial contours shown here are for visual presentation of the data from those population only and emphasize the diversity of genetic polymorphism between populations. No any conclusive results can be derived from these figures.

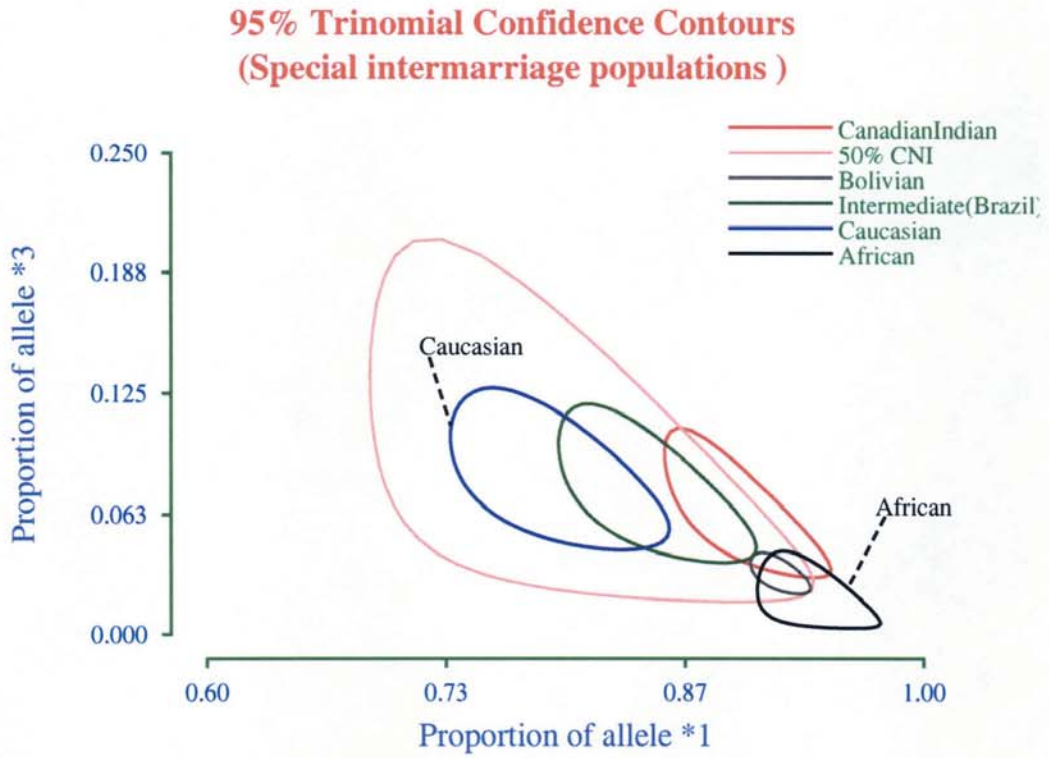


Figure 3.4.5: Frequencies of *CYP2C9*1*, **2* and **3* in some specific populations.
Note: CNI: Canadian Native Indian.

5. Studies with subjects from different ethnic populations:

From the previous figures 3.4.1 to 3.4.5, it is obvious that different ethnic populations have different CYP2C9 polymorphism distributions. However, in the thesis, several studies of CYP2C9 polymorphism have investigated subjects with different ethnic origins, and provided the genotype data for subjects with different ethnicities together. Furthermore, they evaluated the effect of genetic polymorphism on drug metabolism or drug response (Tabrizi *et al.* 2002, Freeman *et al.* 2000, Pirmohamed *et al.* 2000, Joffe *et al.* 2004, Higashi *et al.* 2002, Linder *et al.* 2002). The different distribution of polymorphism between ethnicities may assign dissimilar candidate genes in the pharmacogenetic studies for various ethnic populations. Therefore, it is essential for pharmacogenetic studies to clearly define ethnicity of subjects and provide the data for individual ethnic populations. In the miscellaneous group, several studies recruited subjects with either Caucasian and African origin or other ethnic origin together for CYP2C9 polymorphism investigations. Some studies provided the composition of different ethnic subjects in percentage; some did

not provide any detail. In the following figure 3.4.6, the trinomial contours of these studied populations exhibited irregular contours overlapping with Caucasian. This is not surprising since the allele prevalence of *CYP2C9*1*, **2* and **3* is significantly different between African and Caucasian populations.

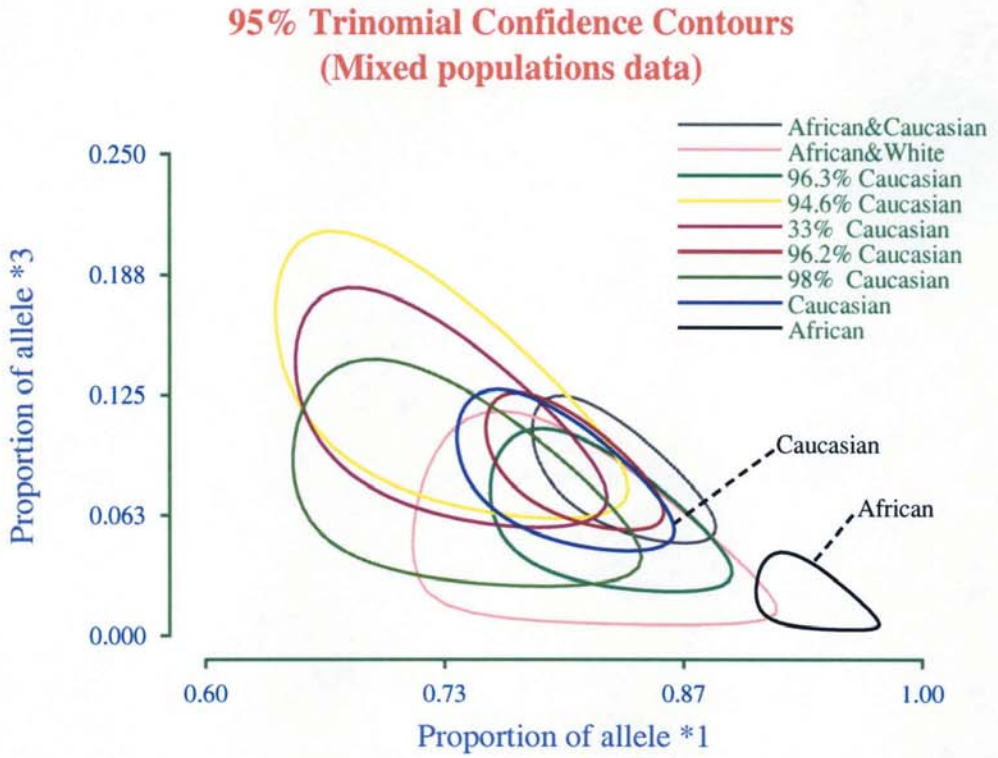


Figure 3.4.6: Trinomial confidence contours of *CYP2C9*1*, **2* and **3* in population studies with mixed ethnic peoples

*Distribution of **CYP2C19***1, *2 and *3 in populations of the miscellaneous group:*

The estimated prevalence of *CYP2C19*1*, **2* and **3* in the trinomial confidence contours have been produced for Caucasian, Chinese, Japanese and African populations according to meta-analysis results in section 3.3, which is shown in figure 3.4.7. Studies in the miscellaneous group cannot be easily categorized in any ethnic groups, and there are not more than one or two studies available for each ethnic population. Therefore, a meta-analysis cannot be applied to the data provided in these studies. Again following the phylogenetic tree (Cavalli-Sforza *et al.* 1994), contours of Caucasian, Chinese, Japanese and African population are shown with the three *CYP2C19* allele distributions of population studies in the miscellaneous group. The visual comparisons perform in four groups: 1.populations close Caucasian (including South India, North India, Egyptian and Turkish); 2.populations close to Chinese and Japanese (Filipino, Thai and Korean); 3.specific populations such as Canadian native Indian and Bolivian; 4.Pacific Islander populations. They were presented in following figure 3.4.7-11.

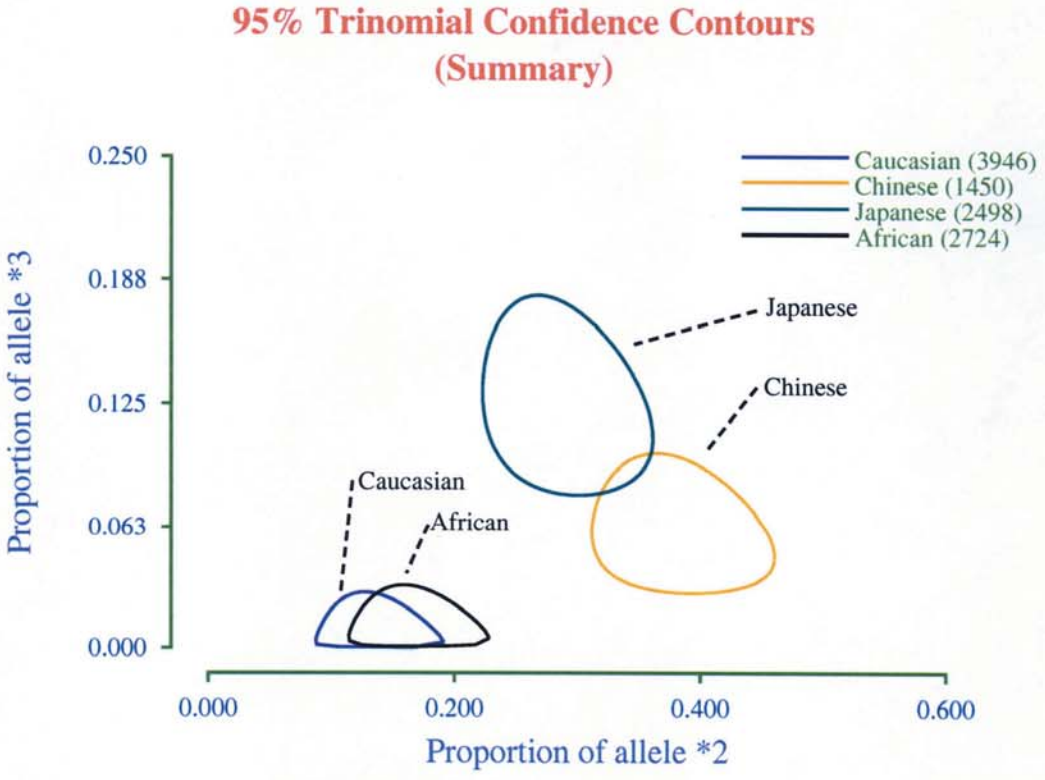


Figure 3.4.7: Prevalence of *CYP2C19*1*, **2* and **3* in different ethnic populations (meta-analysis pooling results)

Referring the population distributions of *CYP2C19* shown in the figure 3.4.7, the

Caucasian population appears have a large overlap with the African population. The Chinese population is dispersed from the Japanese population, which appears to have different diversity from the population distribution of *CYP2C9* between the other ethnic populations. It is shown that the population distributions of different ethnic populations differ from individual genetic polymorphisms. Furthermore, the low prevalence of *CYP2C19*3* in Caucasian and African population cause the straight side in the trinomial contours of these two populations.

1. Populations close to Caucasian:

According to the phylogenetic tree model followed in this thesis, West Asian and North African populations are at a relatively small genetic distance away from Caucasian populations. Therefore, allele frequencies of *CYP2C19*1*, **2* and **3* are presented in 95% trinomial confidence contours here in order to show the diversity of distributions of the three alleles among these populations.

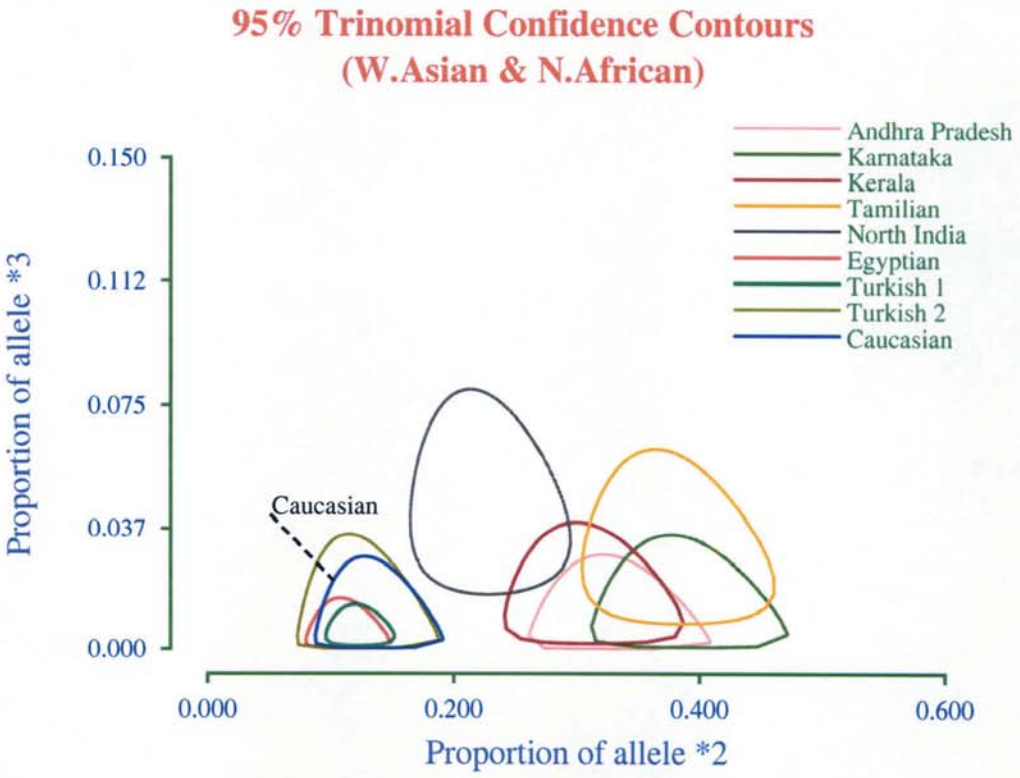


Figure 3.4.8: Prevalence of *CYP2C19*1*, **2* and **3* among ethnic populations with short genetic distance from Caucasians

Although two studies of Turkish populations have been completed by the same research

group and recruited subjects from same population, these two studies are presented in the figure exhibiting different contours of confidence interval due to significantly different sample sizes (Turkish1: 808; Turkish2: 188). Apart from populations from Egypt and Turkey, the trinomial contours of South India and North India populations have no overlapping region with the contour of Caucasian populations. Furthermore, there is no overlap between populations of South India and North India. The distribution of CYP2C19 also appears less homogenous within populations of South India.

2. Populations close to Chinese and Japanese:

Trinomial confidence contours of several populations close to Chinese and Japanese populations are presented in figure 3.4.9. Meanwhile, the prevalence of *CYP2C19*1*, **2* and **3* of North India and South India populations is also presented in the contours in order to show the diversity of three *CYP2C19* alleles distributed between Asian populations.

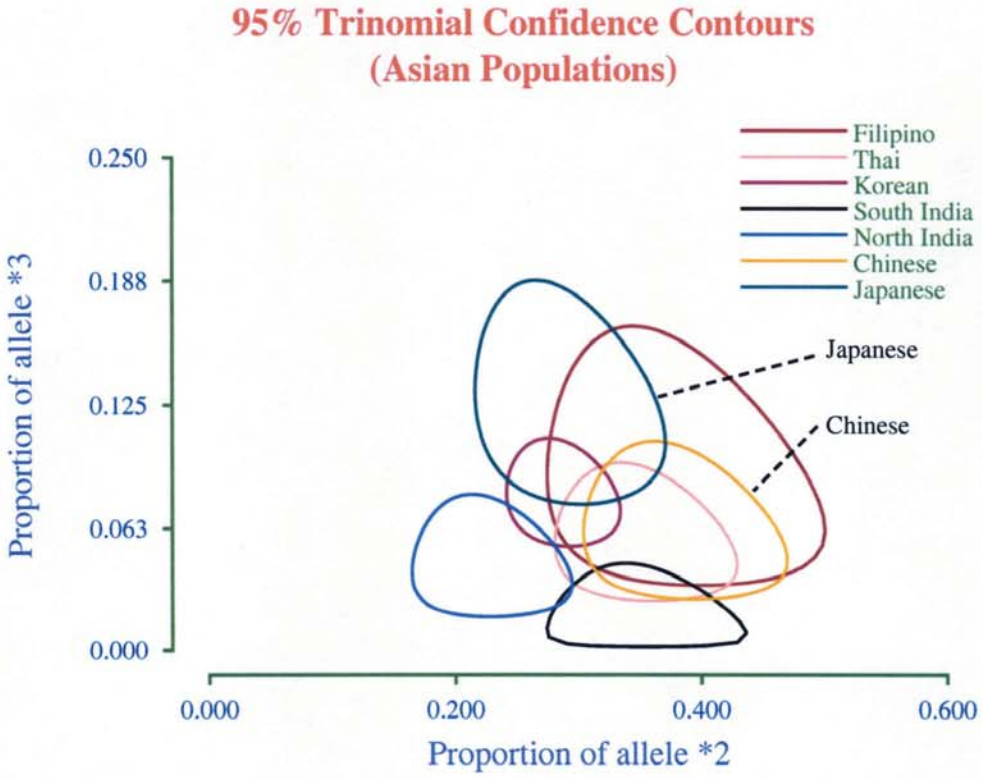


Figure 3.4.9: Trinomial confidence contours of *CYP2C19*1*, **2* and **3* in different Asian populations

Among these populations, the trinomial confidence contours of Filipino, Korean, Thai populations have overlapping with both Chinese and Japanese populations, where the

overlapping region is different between these populations. Meanwhile the *CYP2C19* distribution of South India and North India appears less or no overlapping with either Chinese or Japanese populations. As each ethnic population only have one study available and each study have investigated different number of subjects, the differentiation of population distribution between these ethnic populations cannot derive a certain result.

3. Specific populations:

Two studies provided three CYP2C19 allele frequencies for subjects with two different ancestries; the authors described that the Bolivian population was descended from Caucasian and America Indian. Additionally, Gadgik *et al.* 1998 recruited subjects with grandparent(s) of Canadian native Indian (CNI). Those subjects with two CNI grandparents are described as 50% CNI; 25% CNI refers to those subjects with one CNI grandparent. However, the ethnicity of the other grandparents was not provided. The trinomial contour of allele distributions is presented in figure 3.4.10.

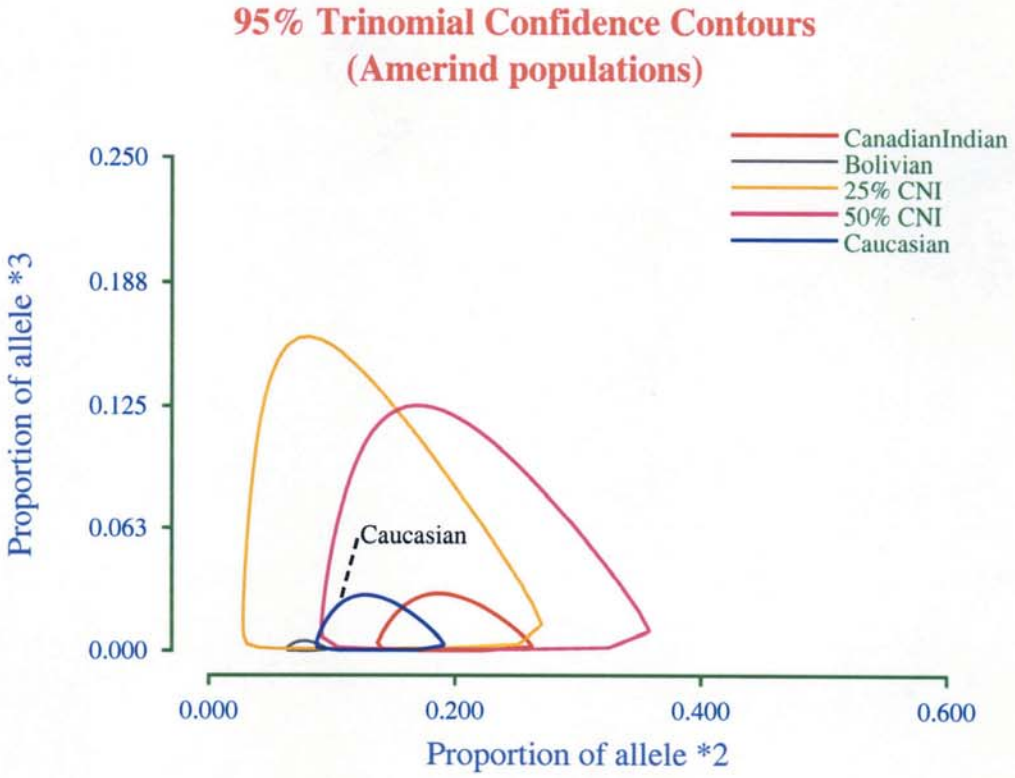


Figure 3.4.10: Trinomial confidence contours of *CYP2C19*1*, **2* and **3* in some specific populations
Note: CNI: Canadian Native Indian.

In figure 3.4.10, Bolivian population have no any overlapping with populations of

Caucasian and Canadian native Indian. Meanwhile, due to small number of 25% CNI and 50% CNI subjects (36 and 48), the *CYP2C19* allele distribution of these two population have relative larger confidence contours, which cover the contours of Caucasian population and Canadian native Indian population. The contours cannot show the diversity of CYP2C19 allele distribution between 25% or 50% CNI and Canadian native Indian populations.

4. Populations of Pacific Islanders:

Several studies have investigated the *CYP2C19* polymorphism in Pacific Islander populations. Kaneko *et al.* (1999) had studied Pacific Islander from Vanuatu islanders. According to the language and geographic location, these subjects are categorized in 24 subpopulations. As the figure 3.3.3a in section 3.3, these twenty-four subpopulations have significantly different prevalence of CYP2C19 alleles. The homogenous between these subpopulations is very low. The category of ethnicity within the pacific population is very difficult due to a lack of sufficient evidence of the historical human immigration patterns. As both geneticists and archaeologists agreed that the ancestries of Pacific Islander are from two or more different ethnicities (Gibbons *et al.* 2001), it is not surprising that the allele distribution between these subpopulations have highly diversity than other ethnic populations retrieved in the thesis. Therefore, it may be inappropriate to pool the data of these subpopulations together or apply a meta-analysis method to estimate the CYP2C19 allele prevalence. In this thesis, trinomial confidence contours are applied to show the significant diversity between these populations of Pacific Islander.

Since it is difficult to identify 24 subpopulations in one figure, the data of Northern-central subpopulations that included numerous populations from same geographic origin is chosen as an example to represent the dissimilar CYP2C19 distribution within Pacific Islander population (figure 3.4.11a). Furthermore, the study of Kaneko *et al.* (1997), the study of Masta *et al.* (2003) and the study of Griese *et al.* (2001) have examined the CYP2C19 allele distributions for other Pacific Islander populations, such as populations from New Guinea and the Australian Aborigine. The data of CYP2C19 polymorphism from these studies are represented in figure 3.4.11b. These trinomial contours are a visual presentation of the diverse distribution of CYP2C19 in populations of Pacific Islanders.

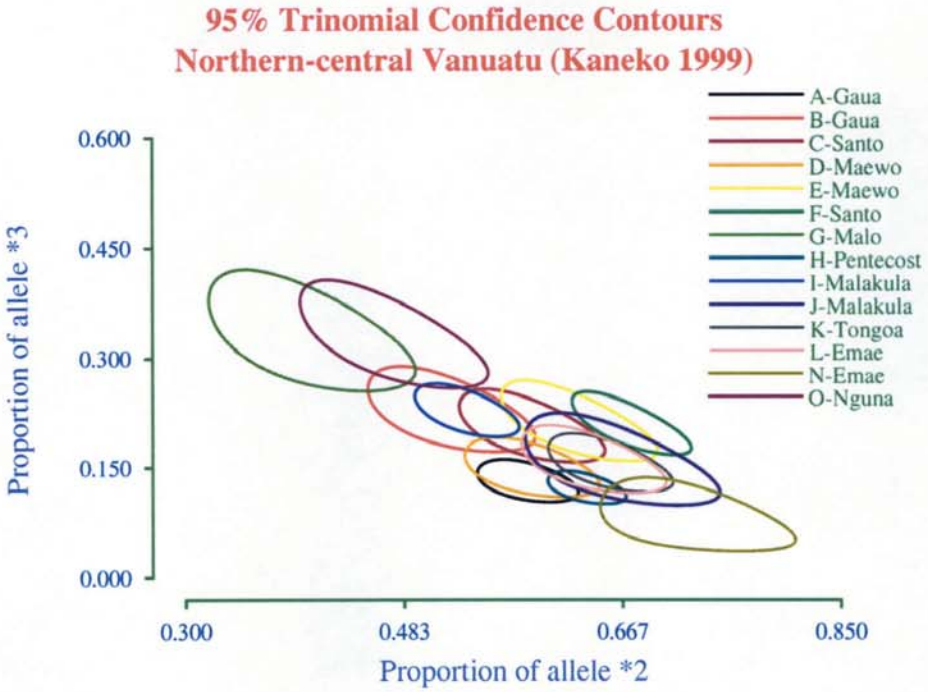


Figure 3.4.11a: Trinomial confidence contours of *CYP2C19* alleles in Northern-central Vanuatu populations. Subpopulations are named with their geographic origin following a capital letter that is defined according to the language phylogenetic tree in Kaneko *et al* (1999)

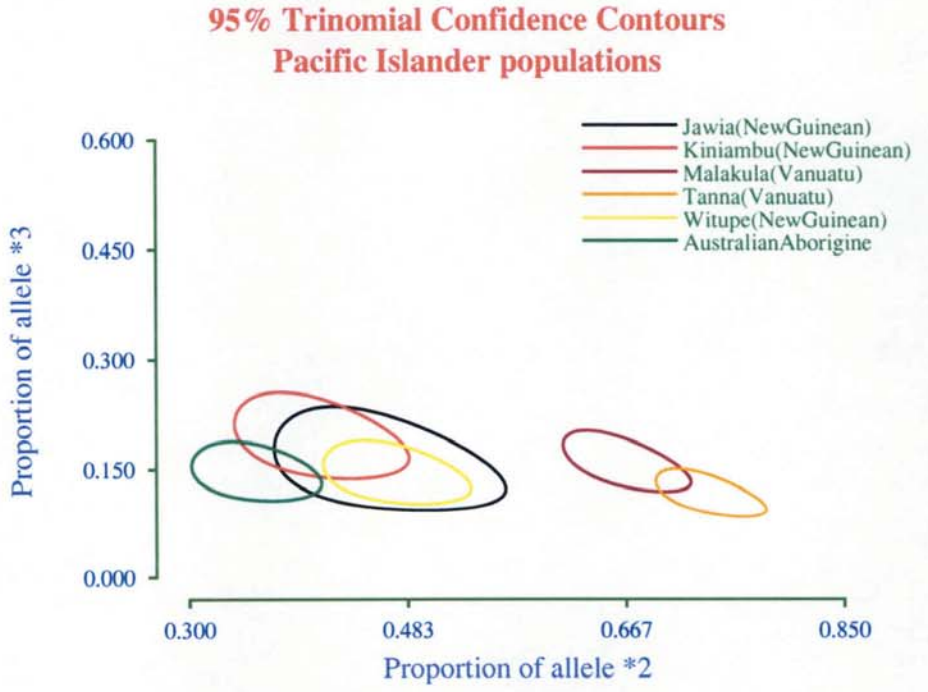


Figure 3.4.11b: Trinomial confidence contours of *CYP2C19* alleles in some Pacific Islander populations Jawia, Kiniambu and Witupe are from Papua New Guinea (Masta *et al* 2003); Malakula and Tanna are from Vanuatu (Kaneko *et al* 1997); Australian Aborigine is represented from the study of Griese *et al* 2001.

3.5 Conclusion

Numerous studies had reported the *CYP2C9/2C19* allele/genotype data for different ethnic populations. There are 74 and 76 studies retrieved from the PubMed database for *CYP2C9* and *CYP2C19* respectively. *CYP2C9* genotypes are available for 49 Caucasian populations, 7 Chinese populations, 5 Japanese populations, 6 African populations and 29 miscellaneous populations; *CYP2C19* genotypes are available for 25 Caucasian populations, 14 Chinese populations, 16 Japanese populations, 16 African populations and 28 miscellaneous populations. However, many studies did not contribute to the investigation of population distribution. Many studies were excluded due to kinds of reasons, such as, studies that did not provide the complete genotype data, studies that combined genotype data into groups and studies that provided the genotype data not fitting the Hardy-Weinberg equilibrium. Therefore, only 14 Caucasian, 3 Chinese, 5 Japanese and 3 African populations studies actually contributed to the estimation of *CYP2C9 *1, *2* and **3* prevalence, and 10 Caucasian, 7 Chinese, 9 Japanese and 11 African population studies contributed to the prevalence estimation of *CYP2C19 *1, *2* and **3*.

According to heterogeneity between studies of each ethnicity, the fixed or random meta-analysis model was applied. The prevalence of *CYP2C9* and *CYP2C19* are significant different between ethnic populations. Caucasian populations have higher prevalence of *CYP2C9*2/*3* variants than any of other ethnicities. *CYP2C9*2* allele is rare in Chinese and Japanese populations. Two studies of African populations have found other CYP variants existing in African subjects, which required more studies to elucidate, apart from the *CYP2C9*1*, **2* and **3*, other alleles are more frequent in those populations. In the population distribution of *CYP2C19* polymorphism, Chinese and Japanese populations have higher prevalence of *CYP2C19*2* variant than Caucasian and African populations. The variant of *CYP2C19*3* are rare in Caucasian and African populations, however it have over 5% and 11% prevalence in Chinese and Japanese population respectively. The polymorphism of *CYP2C9* and *CYP2C19* appeared different population distributions between same ethnic populations.

Studies in the miscellaneous groups provided *CYP2C9/2C19* genotypes for some specific populations. They either have no more than two studies to apply meta-analysis for estimation of *CYP2C9/2C19* alleles prevalence, or have investigated subjects with ancestries from different ethnic origin due to unique human evolution history. Therefore, trinomial confidence contours are employed to visually present the diverse population distribution of *CYP2C9/2C19* between these populations and previously studied four ethnic populations using a meta-analysis method, *i.e.* Caucasian, Chinese, Japanese and African populations. The population distributions as indicated by the trinomial contours are obviously different. Apart from overlap of *2C9/2C19* alleles of Egyptian and Turkish population with Caucasian population, the overall pattern is one of dissimilarity between distribution of *2C9* and *2C19* between ethnicities. Although one or two polymorphism cannot fully explain the population differentiations, the population distribution of *CYP2C9/2C19* is significant different between different ethnic populations. It emphasizes that the different prevalence of *CYP2C9/2C19* polymorphism between ethnic populations may lead to the different molecular basis of variability in drug response. Therefore, the candidate gene in the pharmacogenetics studies may need to consider differently between different ethnicities.

Chapter 4 Influence of polymorphic *CYP2C9/CYP2C19* on drug metabolism

4.1 Preface

The aim of pharmacogenetic studies is to find out the genetic factors that are responsible for variable drug responses within and between populations. The association between genetic polymorphism and varied drug response is important to evaluate the qualitative or quantitative effect of the relevant mutated genes on a given drug. In the initial search of pharmacogenetic studies related to antiepileptic drug responses, 56 articles were retrieved and 37 of those articles (over 66%) were concerned with polymorphism of cytochrome P450s (CYP450s), especially of *CYP2C9* and *CYP2C19*. This confirmed that most current available pharmacogenetic studies of antiepileptic drugs were focused on the polymorphism of mutated *CYP2C9/2C19* alleles.

According to the latest WHO survey in 155 countries across the world, phenytoin is still one of the most-often used drugs in the pharmacotherapy of epilepsy. It is primarily metabolised by CYP2C9 and CYP2C19. Therefore, phenytoin has been chosen as an example to evaluate the effect of genetic polymorphisms in *CYP2C9/2C19* on antiepileptic drug responses.

The population distribution of mutated *CYP2C9/2C19* reported in chapter three of this thesis provides the estimated variable prevalence of three major *CYP2C9/2C19* alleles for people with different ethnic origins. This may help physicians to predict the possibility of epileptic patients carrying mutated *CYP2C9/2C19* very early on and consequently select optimal pharmacotherapy with desired antiepileptic drugs using a suitable regime and monitoring protocol. Furthermore, if patients are identified as carriers of those *CYP2C9/2C19* variants, it is expected that the qualitative or quantitative effect of those variants on the metabolism of phenytoin could help the physician to adjust therapeutic regime of phenytoin according to each patient's genotype.

In general, the genotype data is easily obtained from modern DNA amplification techniques when information on variant alleles or genotypes is available from genetic and population genetic studies. Conversely, the subsequent phenotype data, such as the capability of the enzyme in metabolism, is often more difficult to determine than the genotype data, due to the combined effect of genetics, environment and endogenous factors on enzyme activity, in addition to the pharmacokinetic characteristics of the given drug. However, CYP genotyping cannot be clinically useful unless the relevant phenotype data is accessible. A desired study should provide the genotype of *CYP2C9/2C19* and the phenotype relevant to each genotype; this could be any pharmacokinetic parameter of the given drug, where the parameter could demonstrate the activity or capability of the enzymes. Although, currently there is little agreement on the best method of characterising enzyme activity, two kinds of approaches, in *vitro* or in *vivo*, have been widely applied in research or clinics to explore enzyme phenotypes, in terms of metabolic activities (Venkatakrishnam *et al.* 2001, Rodrigues and Rushmore 2002). In this thesis, *in-vivo* studies were selected for measuring the association between the genotype of *CYP2C9/2C19* and the phenotype of enzyme activity, which was the pharmacokinetic parameter related to phenytoin metabolism that is consistently correlated to genotype.

However, the retrieved articles from the search performed have failed to provide sufficient evidence to clarify the association between the genetic polymorphism of *CYP2C9/2C19* and phenytoin metabolism. For *in-vivo* studies, potential phenotyping probe drugs have been proposed for most CYP enzymes, where the recommendations of appropriate probe drugs for each CYP enzyme have been mainly based on the knowledge of substrate specificity of the individual CYP isoforms (Smith *et al.* 1998). Therefore, studies on warfarin and mephenytoin, which are two probe-drugs used broadly to study polymorphism of *CYP2C9* and *CYP2C19* respectively, were retrieved to further explore the possible impact of *CYP2C9/2C19* genetic polymorphism on drug metabolism.

4.2 Pharmacogenetic studies in AEDs -- phenytoin

4.2.1 Pharmacogenetic aspect of antiepileptic drugs

In the beginning of this project, a search was undertaken to explore genetic factors involved in antiepileptic drugs metabolism, which was expected to provide an over-view of current available pharmacogenetic studies on variable antiepileptic drug responses. This search result compelled the project to focus initially on the population distribution studies of *CYP2C9/2C19* polymorphism. The association between *CYP2C9/2C19* polymorphism and phenytoin metabolism was explored after definition of ethnicity.

Searches were performed in the PubMed database using EndNote software to compare the results. Publications were retrieved when they had the MeSH term antiepileptic or anticonvulsant drug and the general terms, polymorphism and pharmacogenetic simultaneously in the abstract. The following two general terms were searched in both singular and plural formats.

After deleting duplicated publications, 92 articles were retrieved, where 39 journal articles, 10 reviews, and 7 case reports appeared describing the polymorphism of antiepileptic drugs response. Among these selected 56 articles, 37 publications described the effect of cytochrome P450s polymorphism on antiepileptic drugs response. The detail is shown in figure 4.2.1.

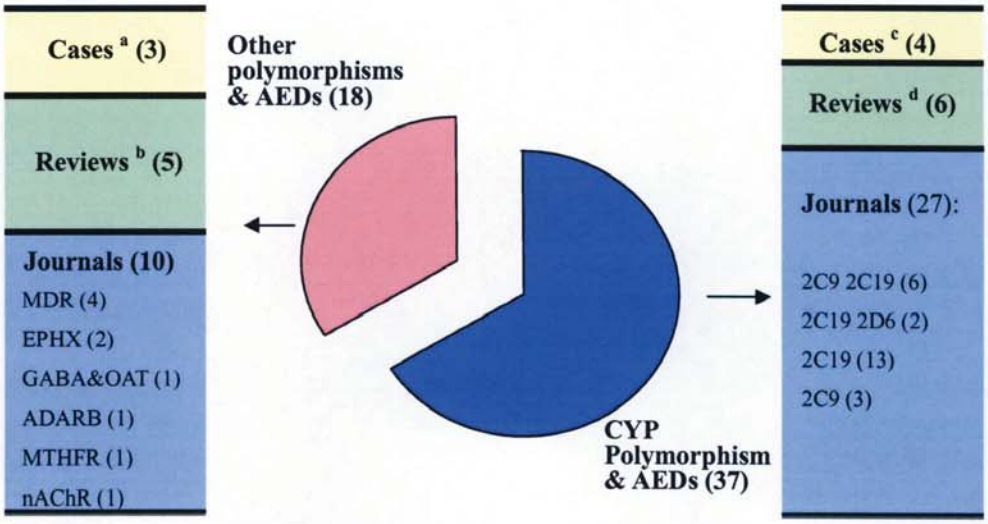


Figure 4.2.1: The distribution of information from retrieved studies on polymorphism of AEDs response
Note: MDR multiple drug resistant gene (P-glycoprotein); EPHX epoxide hydrolase gene; GABA gamma aminobutyric acid, OAT ornithine aminotransferase gene; ADARB adenosine deaminase gene; MTHFR

methyleneter-ahydrofolate reductase gene; nAChR nicotinic acetycholine receptor gene
Number of retrieved articles which were classified as cases or reviews: a. calcium channel (1), nAChR (1), others (1); b. MDR (1), epilepsy genes (2), in *vitro* study (2); c. CYP2C19 (1), CYP2C9 (2), CYP2C9/2C19 (1); d. CYP2C9 (2), CYP2C19 (1), CYP2C (1), CYP2C9/2C19 (2)

It was obvious that existing pharmacogenetic research of AEDs largely focused on the polymorphism of cytochrome P450s. However AEDs have multiple targets, and have a complicated mechanism of action; a single gene polymorphism would not be sufficient to explain the various responses to AEDs among epilepsy patients. Besides the polymorphism of CYPs, it is necessary to give attention to other genetic polymorphisms that also have a potentially important impact on the various AED responses (Spear 2001, Kaneko *et al.* 2002, Loscher and Potschka 2002a, Ferraro and Buono 2005). The potential genetic influences on response to antiepileptic drugs identified from searches undertaken in the thesis are summarised in table 4.2.1.

Table 4.2.1 Genes associated with variable antiepileptic drug (AED) responses

Pharmacological aspects of AEDs		**Polymorphic genes and relevant AEDs**
Pharmacokinetic	Metabolism	CYP3A4 (CBZ); CYP2C9, 2C19(PHT); CYP2C19 (PBB); CYP2C9, 2A6 (VPA); UGT (CBZ, PBB)
	Protein binding	α1 acidic glycoprotein
	Transporters	ABCB1 or MDR1 (VPA); SLCO2A
Pharmacodynamic	Ion Channels	SCN2A (PHT), CACNA (PHT) CACNA, CACNB (PBB)
	GABA receptor	GABAa (PBB)

CBZ: carbamazepine, PBB: phenobarbital, PHT: phenytoin, VPA: valproic acid; UGT: uridine diphosphate glucuronosyl transferase; ABCB1: ATP-binding cassette subfamily B (MDR/TAP) member 1 transporter, MDR1: multiple drug resistance 1, SLCO2A: solute carrier monocarboxylate, SCN2A: sodium channelα subunit, CACNA: calcium channel αsubunit, CACNB: calcium channelβsubunit, GABAa: γ-aminobutyric acid ionotropic type A receptor

Additionally, according to recent epilepsy associated genetic research in population-based twin studies, certain genes were suspected to be responsible for some seizure phenotypes in epilepsy, such as the sodium channel β1, α1 subunit, voltage-gated K^+ channel, nicotinic acetcholine receptor subunit andγ-aminobutyric acid receptor (GABA), which are also potential targets of many antiepileptic drugs as shown in table 4.2.2. (Berkobic and Scheffer 2001, Jones-Davis and Macdonald 2003, Kaneko *et al.*2002, Kjeldsen *et al.* 2002, Kjeldsen

et al. 2003, Treiman 2001). These research results confirmed that the studies of a polymorphic gene or genes relevant to variable drug responses must consider those genes responsible for the underlying disease condition. As previously mentioned, the objective genes in pharmacogenetic research must include both the genes relevant to pharmacological aspects of drugs and the genes involved in the underlying disease condition or intermediate phenotypes. These are obviously important for improving the drug therapy of epilepsy especially when the administered antiepileptic drugs mainly target ion-channels or GABA receptors.

Table 4.2.2. Some of epilepsy syndrome with possible genetic basis

Epilepsy syndrome	Possible genetic basis
Generalized epilepsy with febrile seizure	Mutation of voltage-gated sodium channel
Generalized epilepsy with febrile seizure	Mutation of γ-aminobutyric acid receptor
Juvenile myoclonic epilepsy; idiopathic generalized epilepsy	Mutation of high voltage-gated calcium channel

As the time of starting this thesis (October 2002), very few reports were available describing in relevant to genes other than *CYP2C9/2C19* in drug response, therefore this thesis has focused on cytochrome P4502C9 and 2C19 in relation to antiepileptic drug responses only.

4.2.2 Association between phenytoin and CYP2C9/2C19 polymorphism

4.2.2.1 Database search result

Two searches were made for association studies between phenytoin and *CYP2C9/2C19* polymorphism and were performed in PubMed database using EndNote software.

In the first search, publications were retrieved if phenytoin and two general terms, polymorphism and pharmacogenetic, were found simultaneously in any free text of the articles; the later two terms were searched in both singular and plural formats. In total, 128 unduplicated articles were collected in an EndNote file for further selection. This library included 71 journal articles, 40 reviews and 16 other types of publication, such as comments, and case reports.

In the second search, publications were retrieved when both phenytoin and cytochrome p450 were found in the keywords of articles; 204 unduplicated articles were found. They comprised of 160 journal articles, 20 reviews and 24 other types of publication.

Since the *CYP2C9/2C19* alleles were first discovered in December 1994, the selection of articles was limited to a publication date of 1995 or later. Meanwhile, another selection was carried out in those articles with 'human' as a keyword; those publications, which had studied the CYP enzyme's activity in-*vitro* with isolated liver microsome instead of individual subjects, were excluded.

Finally, 41 journal articles and 19 reviews were selected from first search result; 66 journal articles and 14 reviews were selected from the second search result. Reading the abstracts of these selected articles, 22 unduplicated articles were retrieved for further data collection, where 20 and 17 publications came from the first and second search results respectively. All searches were completed by August 2005.

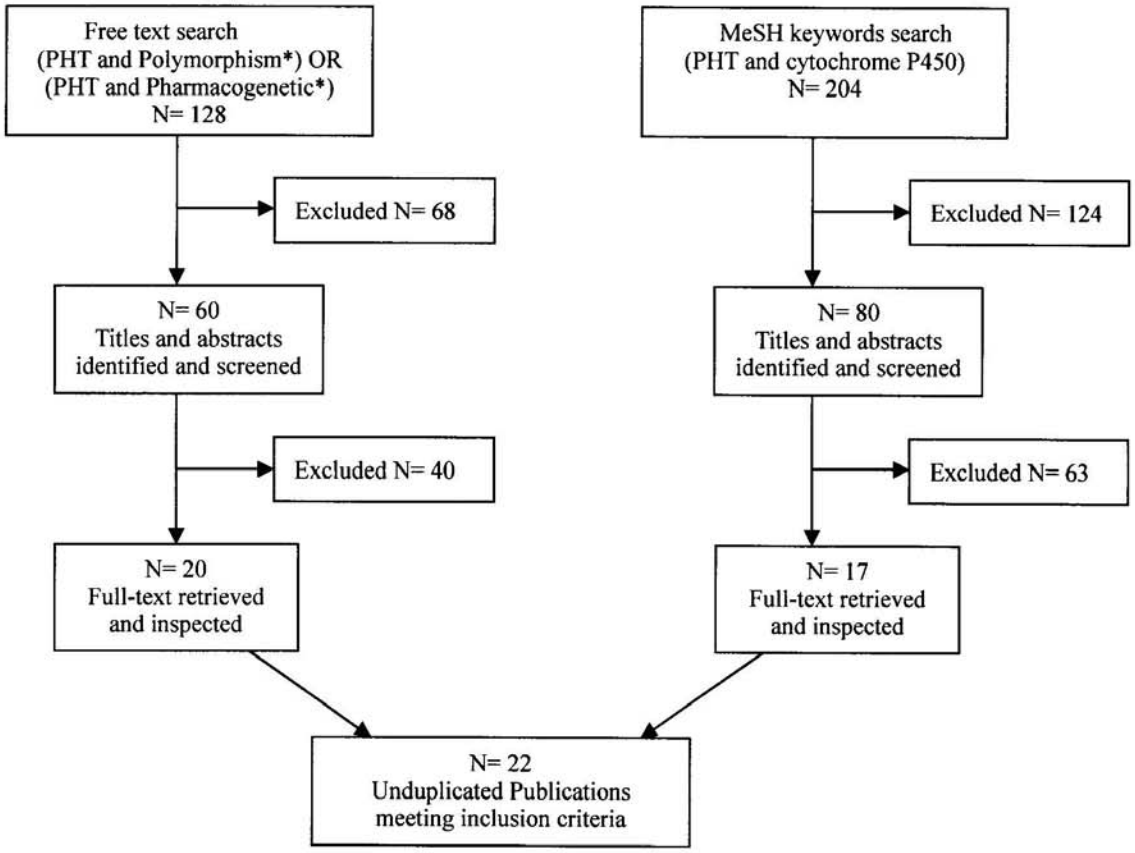


Figure 4.2.2: The search flowchart of studies related to CYP polymorphism and phenytoin metabolism

4.2.2.2 Retrieved data

Twenty-two articles were gained from the library or other resources. As mentioned previously, a desired study would provide the genotype of CYP2C9/2C19 and the phenotype relevant to each genotype. The phenotype could be described by any pharmacokinetic parameter of phenytoin, such as steady state concentration, required dosage, or clearance rate, as long as the parameter could demonstrate the activity or capability of the enzymes. Unfortunately, most of the articles mentioned the impact of *CYP2C9/2C19* polymorphism on phenytoin metabolism without providing complete original genotype data with corresponding phenotype data. Only eight of the selected 22 articles met the criteria. Even those eight studies had provided data in relatively incompatible form as shown in following table 4.2.2. Regarding the unique pharmacokinetic characters of phenytoin, the administration of one dose loading in healthy subjects contrasts with the long-term application in epilepsy patients and would require different pharmacokinetic parameters to establish the activity or capability of enzymes CYP2C9 and CYP2C19; therefore the eight studies were summarized into table 4.2.2a and table 4.2.2b according to the physiological feature of recruited subjects, *i. e.* healthy subjects and epilepsy patients.

Table 4.2.3a: General information of studies in **healthy** subjects that describe CYP polymorphisms and phenytoin pharmacokinetic parameters

Author (year)	Ethnicity	Dose (mg)	Drug Brand	Time of blood sample	Plasma Concentration	Metabolic ratio (serum)	Bioavailability	Elimination	Urine content	Genotype [a]
Aynacioglu *et al*, 1999	Turkish	300	NS	After 12 hours	PHT, p-HPPH	p-HPPH/PHT	NA	NA	NA	CYP2C9
Caraco *et al*, 2001	Caucasian	100*3	Epanutin, Parke-Davis	0~96hours	PHT, p-HPPH	NA	AUC	$t_{1/2}$	(S)-p-HPPH (R)-p-HPPH	CYP2C9/2C19
Kerb *et al*, 2001	Turkish	300	NS	After 12 hours	PHT	p-HPPH/PHT	NA	NA	NA	CYP2C9 CYP2C19

Table 4.2.3b: General information of studies in **Epilepsy** patients that describe CYP polymorphisms and phenytoin pharmacokinetic parameters

Author (year)	Ethnicity	Genotype [a]	Dose (mg/d/kg) (mean of group)	Drug Brand	PHT Treatment Duration	Plasma concentration	Kinetic parameters	Renal, liver function status	Co-medication [b]
Hashimoto *et al*, 1996	Japanese	CYP2C9/2C19	NA [c]	Aleviatin, Dainippon Co.	NA	NA	Km, Vmax	NA	Yes
Hung *et al*, 2004	Chinese (Taiwan)	CYP2C9/2C19	Yes	NS	≥ 1 Month	PHT (Css)	Km, Vmax	Normal	Yes
Mamiya *et al*, 1998	Japanese	CYP2C9/2C19	Yes	Aleviatin, Dainippon Co.	≥ 1 Month	PHT (Css) (R)-p-HPPH, (S)-p-HPPH	NA	Normal	Yes
Odani *et al*, 1997	Japanese	CYP2C9/2C19	NA	Aleviatin, Dainippon Co.	NA	NA	Km, Vmax	Normal	Yes
van der Weide *et al*, 2001	Caucasian (Dutch)	CYP2C9	NA	NS	≥ 6 Month	PHT (Css)	NA	NA	Yes

Note: a: In studies reporting both of *CYP2C9* and *CYP2C19* genotypes, patients were grouped according to the detected genotype co-occurrence; the detail is shown in table 4.2.4.; b: the co-medications reported include carbamazepine, valproic acid, zonisamide, phenobarbital, and primidone, while no further details were provided; c: the mean daily dose of phenytoin was reported for all patients taking part in the study instead of each genotype group separately. NA: not applicable; NS: not stated; PHT: phenytoin; p-HPPH: 5-(4-hydroxyphenyl)-5-phenylhydantoin; S-: S-enantiomer; R-: R-enantiomer; Vmax: maximal elimination rate; Km: Michaelis-Menten constant; Css: steady state plasma concentration; AUC: area under curve, indicate the bioavailability of correlated drug; $t_{1/2}$: half-life of elimination.

It can be seen from the table 4.2.3 that the phenotype data was described by different pharmacokinetic parameters and the correlated genotypes were also reported differently between studies. For instance, some articles studied subjects' *CYP2C9* genotype only, the others provided both of *CYP2C9* and *CYP2C19* genotypes.

In the three studies of healthy subjects, the oral dose of phenytoin was 300mg, however subjects in Caraco *et al.* (2001) took three tablets (100mg/each) instead of one, which could have different bioavailability due to different product formulation from the other two studies. The collection times of blood sample were different according to the authors' study aims. Caraco *et al.* (2001) reported metabolism of phenytoin in plasma and urine after 96 hours in order to find a suitable pharmacokinetic parameter of phenytoin as measurement of CYP2C9 activity in *vivo*. However, Aynacioglu *et al.* (1998) and Kerb *et al.* (2001) examined plasma concentration of phenytoin and its metabolite once, 12 hours after the drug was administered. Moreover, the studies of Aynacioglu *et al.* (1998) and Kerb *et al.* (2001) were done by the same research group with very similar recruitment of subjects; their data were highly likely to be duplicated. Therefore, comparison of data from these healthy subjects was not undertaken.

In those studies of epilepsy patients, most patients also had been prescribed other medicines with phenytoin, which included CYP2C9 inducers, such as phenobarbital, primidone and carbamazepine or CYP2C9 inhibitors, such as valproic acid and zonisamide. However, the detail of treatment regime for these co-medications was not provided in any publication. As the metabolism of phenytoin can be influenced by the drug interactions, the variation of Km and Vm from those epilepsy patients could not be concluded as the consequence of *CYP2C9/2C19* polymorphism alone. Hung *et al.* (2004), Haslimoto *et al.* (1996), Odani *et al.* (1997) and Mamiya *et al.* (1998) had independently employed a pharmacokinetics program, the NONMEM program, to estimate the phenytoin parameters, Km and Vmax. The co-medication was considered as a co-variable, the impact of co-medication, *i.e.* the parameter for the estimation, was individually derived from previous pharmacokinetic studies of epilepsy patients. These are unique parameters regarding the recruitment of patients and they were not provided in any of the publications.

Furthermore, epilepsy patients had already taken phenytoin for more than one month or

even longer, and achieved a steady state of phenytoin plasma concentration; however, the individual dosage, which is highly relevant to the therapeutic and toxic effects, is not provided at all (Gatti *et al.* 2001).

Moreover, the genotypes of *CYP2C9/2C19* reported in three studies of epilepsy patients were assumed to present all possible combination of patients with normal/deficient CYP2C9 enzyme activity and normal/deficient CYP2C19 activity. However, the mutated *CYP2C9/2C19* alleles found in patients were variable between studies. For instance, some patients in the study of Mamiya *et al.* (1998) were found as homozygous for both *CYP2C9*1* and *CYP2C19*3*, which was not found in any of the other studies. The detail of the genotype groups found in epilepsy patients is shown in table 4.2.4. Additionally, the weight (proportion) of mutated CYP2C9/2C19 alleles in genotype groups also varied between studies. Therefore, further analysis cannot be employed to achieve a general conclusion of the extent to which mutated *CYP2C9/2C19* allele can influence the metabolism of phenytoin.

Table 4.2.4 Genotype groups in studies relevant phenytoin and CYP polymorphisms in **Epilepsy** patients

Genotype groups	CYP2C9 *1/*1						CYP2C9 *1/*3			
	CYP2C19 *1/*1	CYP2C19 *1/*2	CYP2C19 *1/*3	CYP2C19 *2/*2	CYP2C19 *2/*3	CYP2C19 *3/*3	CYP2C19 *1/*1	CYP2C19 *1/*2	CYP2C19 *1/*3	CYP2C19 *2/*3
Hashimoto et al, 1996	**G1**	**G2**		**G3**			**G4**			
	4	8		3			2			
Mamiya et al, 1998	**G1**	**G2**		**G3**			**G4**			
	52	47	17	7	5	3	3			
Hung et al, 2004	**G1**	**G2**		**G3**			**G4**			**G5**
	47	79	9	10	6		6	9	2	1
Odani et al, 1997	**G1**	**G2**		**G3**			**G4**	**G5**		
	15	18		5			3	3		

Note: G1 stands for genotype group one; number underneath the genotype group describes the number of subjects than fell into the group.

4.2.2.3 Summary CYP2C9/2C19 polymorphism and phenytoin metabolism

Phenytoin (5,5-diphenylhydantoin) administered in humans is principally metabolised into 5-(4-hydroxyphenyl)-5-phenyhydantoin (p-HPPH), which involved both of CYP2C9 and CYP2C19 enzymes (Browne and LeDuc 2002). Evidence in vivo and in vitro suggests that CYP2C9 is responsible for the stereo-selective hydroxylation of phenytoin into formation of (S)-p-HPPH, and CYP2C19 is mainly responsible for the hydroxylation of phenytoin into formation of (R)-p-HPPH. Since (S)-p-HPPH accounts for over 90% of total p-HPPH

(Fritz *et al.* 1987, Ieiri *et al.* 1997, Caraco *et al.* 2001), the polymorphism of *CYP2C9* and *CYP2C19* will influence the formation of (S)-p-HPPH and (R)-p-HPPH respectively.

Because (R)-p-HPPH produced following metabolism by CYP2C19 only contributes a small proportion of p-HPPH, Aynacioglu *et al.* (1998) and Kerb *et al.* (2001) examined the correlation of *CYP2C9* genotypes with the metabolic ratio of p-HPPH *vs* phenytoin plasma concentration, where the S or R enantiomer were not detected separately. According to their results, the phenytoin plasma concentration was statistically significantly correlated with the *CYP2C9* genotype; this was also found in the Caraco *et al.* (2001) study. Furthermore, Aynacioglu *et al.* (1998) and Kerb *et al.* (2001) found the metabolic ratio of p-HPPH *vs* phenytoin plasma concentration in healthy subjects decreased with increasing number of defective *CYP2C9* alleles; they proposed the metabolic ratio could reflect the altered CYP2C9 activity better than the phenytoin concentration due to lower dependence on bioavailablity and volume of distribution in individuals. In Caraco's study (2001), the S-p-HPPH and R-p-HPPH in urine were examined separately for those healthy subjects taking phenytoin from 0 to 96 hours; the result confirmed that the urinary excretion of R-p-HPPH was much smaller than the excretion of S-p-HPPH, which only account for 2.26 ±1.97% of the phenytoin dose. The urinary content of R-p-HPPH in subjects with mutated *CYP2C9* alleles is higher than its level in subjects without any mutated alleles; however, the difference does not reach a statistically significant level. In contrast, the R-p-HPPH urinary content was significantly lower in subjects with mutated *CYP2C19* than in subjects without any mutated allele of *CYP2C19*. This study further confirmed that CYP2C19 is responsible for the formation of R-enantiomer of phenytoin metabolite, R-p-HPPH. Although the urinary content of S-p-HPPH exhibits a significant association with *CYP2C9* genotypes, obviating the need for tedious and time consuming enantio selective assays, Caraco *et al.* (2001) proposed phenytoin metabolic ratio, defined as the ratio of p-HPPH in urine collection to mid-interval phenytoin plasma concentration, can be used as a putative marker of CYP2C9 activity in *vivo* and can represent the altered pharmacokinetic features of phenytoin metabolism due to genetic polymorphism of *CYP2C9*.

Meanwhile, phenytoin metabolism by CYP2C9 appears saturable within the therapeutic dose range. The non-linear relationship between steady-state serum concentration and

dosage follows Michaelis-Menten kinetics as described by the equation below. In contrast, CYP2C19 metabolism is not saturable at therapeutic phenytoin concentration (Browne and LeDuc 2002).

Michaelis-Menten kinetic equation:

$$Css = (R*Km)/(Vmax-R)$$

where Css is the steady-state phenytoin concentration, R is dosing rate, Vmax is the maximum velocity of the enzyme system, Km is the Michaelis constant of the enzyme system (Km also equal to plasma phenytoin concentration at which half of the maximum velocity of the enzyme system is attained). In the metabolism of phenytoin, the values of Vmax and Km appear to be determined principally by CYP2C9 and CYP2C19.

In the studies of phenytoin metabolism in epilepsy patients, four of five publications studied the association of phenytoin metabolism with both of *CYP2C9* and *CYP2C19* genotype. Among the four studies, three articles had provided the Km and Vmax parameters of phenytoin in epilepsy patients by applying population pharmacokinetic estimation methods performed in NONMEN software. Although the genotype groups were defined differently between studies, each of them had confirmed that overall contribution of *CYP2C19* polymorphism toward phenytoin metabolism was less important than *CYP2C9* polymorphism; those patients with mutated *CYP2C9* alleles appeared to have higher Km and lower Vmax than patients without CYP2C9 mutated alleles. Mamiya *et al* (1998) did not provide the value of Km and Vmax parameters for patients with different genotype of *CYP2C9/2C19*, but they reported the impact of genetic polymorphism *CYP2C9/2C19* on phenytoin metabolism in a similar way to other publications. Furthermore, they emphasized that the plasma concentration of phenytoin in patients with *CYP2C9*3* increased dramatically even at lower daily dose compared with patients without mutated *CYP2C9*; however the difference of plasma concentration among patients with mutated *CYP2C19* alleles became larger with the daily dose increases. Therefore, at high daily doses of phenytoin, patients with mutated *CYP2C19* alleles have to be treated carefully. van der Weide *et al.* (1997) did not provide the genotype of *CYP2C19*, and focused on *CYP2C9* only. According to their result, there is a strong association between *CYP2C9* genotype and phenytoin maintenance dose requirement, however *CYP2C19* polymorphism seems to have

no effect on phenytoin dose requirement.

Overall, with current available evidence, it is agreed that *CYP2C9/2C19* polymorphisms have an effect on the pharmacokinetic parameters of phenytoin, especially for *CYP2C9*, but further research works are still required to quantify the genetic effect on phenytoin metabolism using standardized pharmacokinetic parameters.

4.2.2.4 Discussion

In addition to the characteristics of phenytoin metabolism, the observation of *CYP2C9/2C19* polymorphism in healthy subjects would require consideration of difference in pharmacokinetic parameters compared with epilepsy patients. For healthy subjects, one oral dose of phenytoin could not reach steady-state plasma concentration, therefore the estimation of pharmacokinetic parameters, Km and Vmax is not appropriate. According to Aynacioglu *et al.* (1999) and Caraco *et al.* (2001), the molar plasma ratio of p-HPPH (S/R) to phenytoin in the blood sample drawn 12 hours post phenytoin administration and the ratio of p-HPPH in urine collection to mid-interval phenytoin plasma concentration are highly correlated with phenytoin metabolic clearance. Therefore, the production of (S)-p-HPPH correlates with the number of mutated *CYP2C9* alleles; and these metabolic ratios could be applied in future association studies of phenytoin metabolism and *CYP2C9* polymorphism in healthy subjects. While in clinical practice, if these two parameters could detect the impact of *CYP2C9* genetic polymorphism on phenytoin metabolism with equal sensitivity and accuracy, the molar plasma ratio proposed by Aynacioglu *et al.* (1998) would be more convenient to obtain than the urine collection intended for the metabolic ratio suggested by Caraco *et al.* (2004). Among the current available publications, the consolidation of pharmacokinetic parameters to observe the association between genetic polymorphism of CYP2C9/2C19 and phenytoin metabolism has not been established yet. In order to quantify the genetic effect on altered phenytoin metabolism in healthy subjects, more studies in healthy subjects are required with application of comparable pharmacokinetic parameters.

Furthermore, other antiepileptic drugs, carbamazepine, valproic acid, and phenobarbital, are often co-administrated with phenytoin among epilepsy patients, which are well recognized

as inducers or inhibitors of CYP2C9/2C19. Meanwhile, the rate of phenytoin metabolism is also subject to considerable variation under the influence of physiological factors, such as age, gender, and pregnancy (Gatti *et al.* 2001), which can bring more variability into phenytoin metabolism in epilepsy patients. Additionally, the absorption of phenytoin depends on the product formulation; therefore different generic preparations of phenytoin have the potential to differ in absolute bioavailability from the brand names, which had been reported to vary by +/-14% or more (Tsai *et al.* 1992, Browne and LeDuc 2002). Thus, the quantitative genetic effect on phenytoin metabolism determined in healthy subjects could not be easily applied to epilepsy patients. Clinical application of *CYP2C9/2C19* genotypes in epilepsy pharmacotherapy with phenytoin would be more applicable and provide useful data for adjustment of phenytoin therapeutic regimes, if epilepsy patients were studied directly.

In epilepsy patients, the consequences of saturated metabolism of phenytoin are most readily observed under steady state conditions, when the drug is administered at a fixed rate until the rate of elimination and the average rate of absorption are equal. With experienced clinical monitoring of phenytoin, the estimation of pharmacokinetic parameters, such as Km and Vmax, are easy to achieve. From studies retrieved in the thesis, the association of mutated *CYP2C9/2C19* alleles with altered metabolism of phenytoin was confirmed. As CYP2C19-mediated metabolism to (R)-p-HPPH is minor metabolic pathway for phenytoin, most studies established that the deficient function of CYP2C19 enzyme would not influence phenytoin steady-state plasma concentration effectively. Therefore, the polymorphism of *CYP2C9* is a more important factor for variable phenytoin response than *CYP2C19*, particularly for those toxicity syndromes that were related with impaired metabolism of phenytoin. According to studies of epilepsy patients, *CYP2C9* polymorphism can be useful in the dosage adjustment and evaluation of phenytoin serum concentration for epilepsy pharmacotherapy. Furthermore, three articles estimated the Km and Vmax parameters by considering the level of co-administered drugs and physiological features of the patients, such as body weight. However, the value of Km and Vmax was less feasible to employ in a meta-analysis to achieve quantitative results for various *CYP2C9/2C19* genotype groups due to poorly comparable group components in each study as described

previously. In the genotype groups, G1-3, G1 described epilepsy patients without any mutated *CYP2C9* alleles, while G2 has one mutated *CYP2C19* allele and G3 has two mutated alleles; the estimated values and 95% confidence interval (CI) of Km and Vmax overlapped each other (shown in table 4.2.4), which indicates that the polymorphism of *CYP2C19* has less impact on the pharmacokinetic parameters. These results are supported by studies of healthy subjects. Mamiya *et al.* (1998) measured the plasma concentration ratio of (R)- and (S)-p-HPPH *vs* phenytoin in epilepsy patients and found higher ratio in patients with mutated CYP2C9 alleles than patients without the mutations, however the difference was not statistically significant as the results in studies of healthy subjects. None of other epilepsy patient studies had evaluated the ratio of phenytoin metabolites; therefore further comparison could not be taken.

Table 4.2.5: Estimated Km and Vmax with 95% CI for CYP2C19 genotypes in epilepsy patients without CYP2C9 mutations.

Author of studies	Genotype group with CYP2C9*1/*1					
	G1: CYP2C19*1/*1		G2: CYP2C19*1/*2 or *1/*3		G3: CYP2C19*2/*2 or *2/*3	
	Km (ug/ml)	Vmax (mg/day/kg)	Km (ug/ml)	Vmax (mg/day/kg)	Km (ug/ml)	Vmax (mg/day/kg)
Hashimoto et al, 1996	8.9 (6.8-10.2)	10.4 (9.5-11.3)	9.7 (7.2-10.6)	9.6 (8.5-11.6)	8.9 (7.7-10.0)	8.9 (8.1-9.9)
Hung et al, 2004	8.15 (7.99-8.31)	10.01 (9.83-10.19)	9.03 (8.91-9.15)	9.77 (9.65-9.89)	9.38 (9.02-9.56)	9.18 (9.03-9.33)
Odani et al, 1997	8.52 (7.66-9.38)	10.7 (9.7-11.7)	9.38 (8.27-10.49)	9.75 (8.61-10.89)	9.16 (8.24-10.08)	9.18 (8.39-9.97)

As phenytoin has a narrow therapeutic range between the minimum effective serum concentration and the minimum toxic serum concentration, individuals with deficient metabolic enzymes would be at high risk to develop drug toxicity. The quantitative or qualitative estimation of the impact of *CYP2C9/2C19* polymorphism on phenytoin metabolism can aid the physician to create an optimal therapeutic regime for each patient, either at the early stage of phenytoin treatment or after achieving seizure-control but requiring dosage adjustment. However, more studies either with a large population of healthy subjects or epilepsy patients are required to quantify the genetic effects of *CYP2C9* polymorphism on the pharmacokinetics of phenytoin.

4.3 Association studies of CYP2C19 polymorphism and mephenytoin metabolism

A few probe drugs are employed widely in the phenotyping of CYP2C9 and CYP2C19 enzyme activity (Streetman *et al.* 2000, Smith *et al.* 1998). In studies of CYP2C19 polymorphism, mephenytoin, omeprazole and proguanil have been employed, where the mephenytoin S/R ratio has long been the standard CYP2C19 phenotyping method. However, mephenytoin has a few limitations related to the convenience of analysis, such as long-term sample stability, low urinary (S)-mephenytoin concentration and the comforts of subjects who may experience side effects. Meanwhile, mephenytoin belongs to the same drug family, hydantoins, as phenytoin, although it is no longer used as an antiepileptic drug due to a greater incidence of serious toxicity in patients with CYP2C19 enzyme deficiency. Therefore, the polymorphic *CYP2C19* studies on mephenytoin have been included in the project for a better understanding of the influence of *CYP2C19* polymorphism on drug metabolism, where the metabolic ratio of S-mephenytoin and R-mephenytoin (S/R) was accepted as a measurement of phenotype data.

4.3.1 Database search result

Mephenytoin was consequently searched as a keyword combined with polymorphism and cytochrome P450 in the PubMed database using EndNote software. There were 111 and 208 publications found in each of the combined keyword searches respectively. The entire search was completed by September 2004.

Publications with CYP2C19 in the abstract were gathered for further selection; this identified 98 unduplicated articles. In those publications without an abstract available, selection was carried out by reading the title. Only one review article was accepted as an association study of CYP2C19 polymorphism and mephenytoin metabolism. Based on the content of the abstract, 40 journal articles and 5 reviews were chosen for relevant data retrieval. Another five articles provided the *CYP2C19* genotypes with relevant mephenytoin metabolic ratios. These were added into the table. Four of these articles were gained from

Key journal searches and one was found in a related review.

In these articles, genotype data was successfully retrieved from 26 articles and was available for 30 ethnic populations. However phenotype data (mephenytoin metabolic ratio) was only provided in 17 articles in 21 different ethnic populations.

Furthermore, the entire articles had already been retrieved in chapter three, so the summary of these studies can be found in the relevant table 3.2.1.

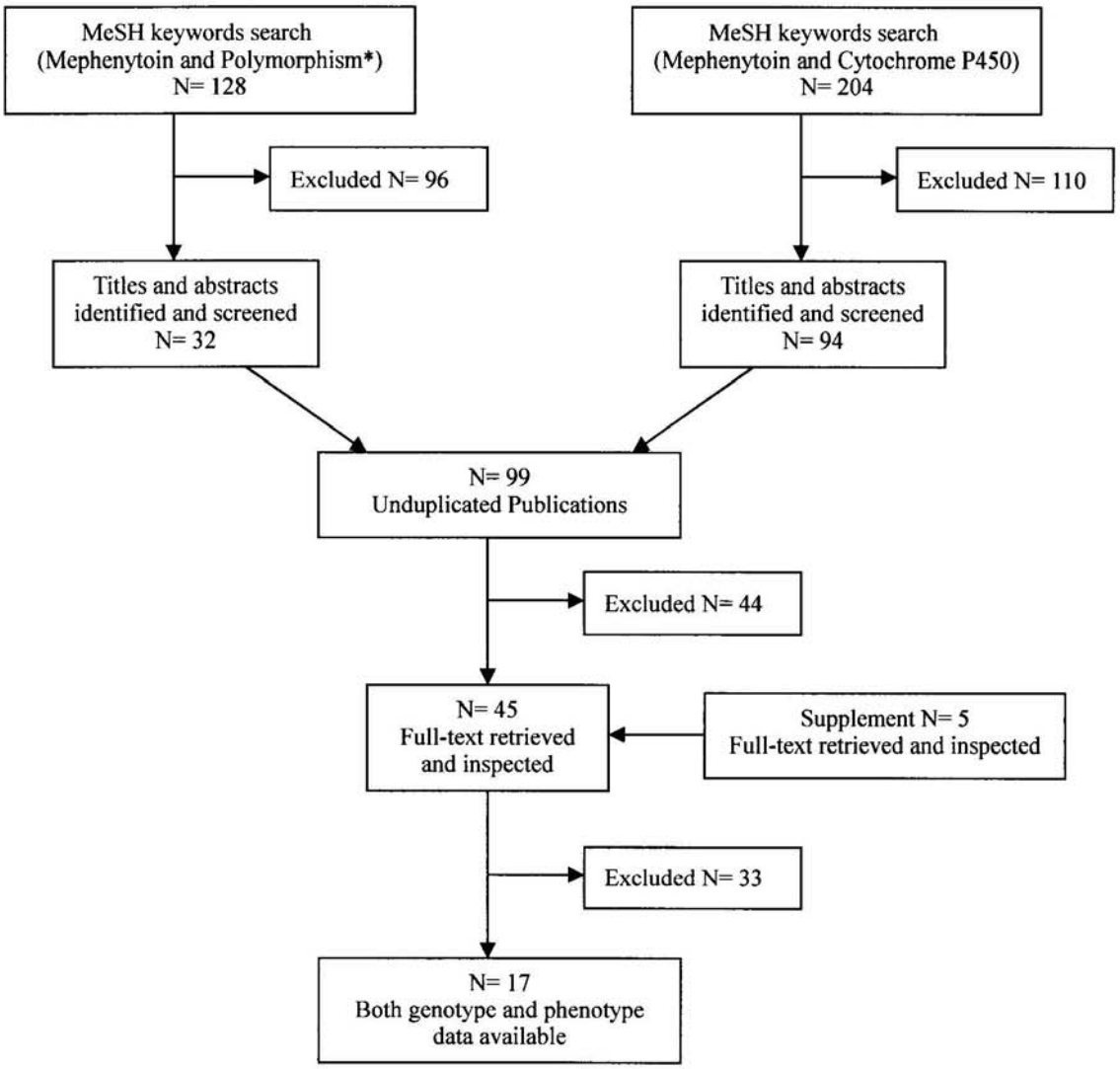


Figure 4.3.1: The search flowchart of studies related to CYP polymorphism and mephenytoin metabolism

4.3.2 Retrieved data

Kupfer *et al.* (1982) first reported the stereoselective metabolism of racemic mephenytoin; S-mephenytoin was rapidly and quantitatively converted into 4'-hydroxymephenytoin (4-OH-M), then excreted in urine. R-mephenytoin is also slowly excreted in urine as other metabolites. Later CYP2C19 was confirmed as the major enzyme responsible for the stereoselecive metabolism of mephenytoin; the S/R mephenytoin ratio in urine from 0-8 hours after taking mephenytoin is reported to be the most preferable measurement of CYPC19 metabolic activity comparing with other urinary analysis of mephenytoin metabolism, such as hydroxylation index (micromoles S-mephenytoin dosed *vs* micromoles 4-OH-M excreted in the urine), the log 10 of the hydroxylation index and the log10 of the percent of the dose of mephenytoin excreted as 4-OH-M (Wrighton *et al.* 1993, Goldstein *et al.*1994, de Morais *et al.* 1994a,b, 1995). The poor metabolizers of mephenytoin are usually subjects with genotype *CYP2C19*2/*2*, *CYP2C19*2/*3*, or *CYP2C19*3/*3*, *i.e.* carrier of two mutated *CYP2C19* alleles (Xie *et al.* 1997, 1999).

In this thesis, the urinary S/R mephenytoin ratio was accepted as a measure of phenotype that is correlated with *CYP2C19* genotypes. The information on both genotype and phenotype data is presented in table 4.3.1.

The mephenytoin phenotype data, S/R ratio, was retrieved from text, table and graph individually for each publication. In those article with graphic data only, a graph digitalizer program – Grafula is employed, which is summarized in table 4.3.1 as well. All studies recruited healthy subjects, therefore no further description of health status was shown in the table.

Table 4.3.1: Association studies of Mephenytoin (MP) metabolic ratio with CYP2C19 polymorphism

Author & year	Ethnicity	Dose	Brand of Mephenytoin	Measurement of MP metabolism	Graph digitalizer application [c]	Phenotype data with Correlated CYP2C19 genotypes
Aklillu E *et al* 2002	African (Ethiopian)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Missing *2/*2, *2/*3
Brockmoller J *et al* 1995	Caucasian	50mg	Epilan (half tablet)	S/R ratio in urine (0-5, 5-8hrs)	Yes	Not detected for *3
Chang M *et al* 1995	Caucasian (Swedish)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Un-retrievable
de Morais *et al*1995	Chinese (Dong)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Completed
Edeki TI *et al* 1996	African	100 mg	Not available	S/R ratio in urine (0-8hrs)	Yes	Completed
Ferguson RJ *et al* 1998	Caucasian (French)	100 mg	Not available	Hydroxylation Index	Yes	Completed
Goldstein JA *et al* 1997	Filipino, Saudi Arabian, Japanese	100 mg	Not available	S/R ratio in urine (0-8hrs)	Yes	Completed [d]
He N *et al* 2002	Chinese (Dai)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Not necessary	Completed [d]
Herrlin K *et al* 1998	African	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Completed [d]
Jurima-Romet M *et al* 1996	Inuit	100 mg	Mesantoin	S/R ratio in urine (11.9 +/-1.6 hrs)	NA*	Un-retrievable
Kubota T *et al* 1996	Japanese	100 mg	Mesantoin	Log_{10}(4-OH-M) in urine	Yes	Completed [d]
Masimirembwa C *et al* 1995	African (Zimbabweans)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Not detect for *3
Persson I *et al* 1996	African (Ethiopian)	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Completed
Sviri S *et al* 1999	Jewish	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Completed
Xiao ZS *et al* 1997	Chinese	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Not necessary	Completed [d]
Xie HG *et al* 1997	Chinese	100 mg	Mesantoin	S/R ratio in urine (0-8hrs)	Yes	Completed [d]
Yao TW *et al* 2001	Chinese	100 mg	Not available	S/R ratio in urine (0-8hrs or other [b])	no	SD missing

Note: a. 50mg tablet was given to Filipino subjects due to side effect reports of somnolence; b. Urine collection was done at 0,2,4,6,8,10,12,14,18,24, and 32hrs after taking MP; c. the graph digitalizer program, Grafula, was applied when articles provided the phenotype data only in graph; d. the phenotype data of mephenytoin was reported according to detected genotype of CYP2C19, in some of cases, CYP2C19 mutant genotypes were grouped together; NA*. the phenotype data was reported in graphically but impossible to retrieve by any method due to highly overlapping data presentation markers.

Among the seventeen articles gained from libraries, two provided other pharmacokinetic parameters such as mephenytoin phenotype measurement instead of S/R mephenytoin ratio in urine and are excluded from further analysis. Another four studies were excluded due to incomplete genotype data or phenotype data. One article provided phenotype data in a graphical form, which is too difficult to retrieve by any methods; therefore it was excluded from further analysis as well. The data from Yao *et al.* (2001) was also excluded because of failure to retrieve the standard deviation (SD) for each phenotype data. Nine of the original seventeen articles reported complete phenotype data, S/R-mephenytoin ratio, correlated with *CYP2C19* genotypes. However, six of them had provided S/R-mephenytoin ratio according to groups of subjects with one or two *CYP2C19* mutation, *i.e.* subjects with genotype *CYP2C19*1/*2* or *CYP2C19*1/*3* in one group, or subjects with genotype *CYP2C19*2/*2*, *CYP2C19*2/*3*, or *CYP2C19*3/*3* in one group. This was correlated with the clinically categorized terms, extensive metabolizers and poor metabolizers, respectively. Furthermore, the prevalence of CYP2C19*3 in most studies was low; most studies did not find subjects who were homozygous for CYP2C19*3 alleles, and several studies found less than five subjects who were heterozygous of *CYP2C19*2* and **3*. Therefore, the meta-analysis was undertaken for three groups according to subjects genotype, which are: 1.*CYP2C19*1/*1*; 2.*CYP2C19*1/*2* or *CYP2C19*1/*3*; 3.*CYP2C19*2/*2* or *CYP2C19*2/*3* or *CYP2C19*3/*3*. The data from these nine studies is presented in table 4.3.2.

Furthermore, quite a few S/R mephenytoin ratios were extracted by employing Grafula software. Therefore, articles with phenotype data, S/R-mephenytoin ratio, both in table and graphs were chosen to check the accuracy of data transferring from graphs into exact numbers. The table 4.3.3 presents both the data retrieved from table and the correlated data digitalized from graph in the same publication. Based on the paired T-test, there is no statistically significant difference between them ($\alpha > 0.05$; $t = 1.42$).

Table 4.3.2: Data derived from studies S/R-mephenytoin ratio and correlated CYP2C19 genotypes

Data source		Phenotype category Extensive Metabolizers (EM)						Phenotype category Poor Metabolizers (PM)			Difference of S/R ratio between subjects with one and two mutated CYP2C19 allele/s		
AUTHOR	Ethnicity	CYP2C19 *1/*1	S/R ratio Mean	SD	CYP2C19 *1/*2 *1/*3	S/R ratio Mean	SD	CYP2C19 *2/*2 *2/*3 *3/*3	S/R ratio Mean	SD	CYP2C19*1/*1 *vs* CYP2C19*1/*2&*1/*3	CYP2C19*2/*2 &*2/*3&*3/*3 *vs* CYP2C19*1/*1	CYP2C19*2/*2 &*2/*3&*3/*3 *vs* CYP2C19*1/*2 &*1/*3
de Morais *et al* 1995	Chinese Dong	25	0.240	0.030	30	0.330	0.045	19	0.979	0.020	-0.090	0.739	0.649
Xie HG *et al* 1997	Chinese	23	0.258	0.120	30	0.321	0.145	19	0.993	0.073	-0.063	0.735	0.672
Xiao ZS *et al* 1997	Chinese Bai	102	0.177	0.119	73	0.260	0.020	26	0.990	0.020	-0.083	0.813	0.730
He N *et al* 2002	Chinese Dai	81	0.215	0.150	93	0.330	0.181	18	0.962	0.040	-0.115	0.747	0.632
Xiao ZS et al 1997	Chinese Han	32	0.185	0.162	49	0.270	0.020	20	1.000	0.020	-0.085	0.815	0.730
Goldstein JA *et al* 1997	Filipinos	16	0.255	0.187	24	0.279	0.189	12	1.107	0.057	-0.024	0.852	0.828
Sviri S *et al* 1999	Jewish	99	0.244	0.173	37	0.350	0.202	4	0.962	0.085	-0.107	0.719	0.612
Edeki TI *et al* 1996	Negroid (African-American)	48	0.245	0.173	27	0.390	0.211	1	1.099	NA	-0.144	0.854	0.710
Persson I *et al* 1996	Negroid (Ethiopians)	85	0.207	0.182	23	0.327	0.219	6	1.053	0.056	-0.121	0.846	0.726
Herrlin K *et al* 1998	Negroid (Tanzanian)	71	0.329	0.153	30	0.639	0.266	5	1.070	0.047	-0.310	0.741	0.430
Goldstein JA et al 1997	Saudi Arabian	70	0.193	0.158	25	0.220	0.155	2	1.088	0.055	-0.027	0.895	0.868

Note: SD. Standard deviation; NA. not applicable.

Table 4.3.3: Accuracy checking of digitalizer program: comparing retrieved data provided in table/text and in graph

Article	He *et al* 2002		Xiao *et al* 1997				Xie *et al* 1997							
Ethnicity	Chinese Dai		Chinese Bai		Chinese Han		Chinese							
Subjects	CYP2C19*1/*1		CYP2C19 *1/*1	PM	CYP2C19 *1/*1	PM	CYP2C19*1/*1 (Female)		CYP2C19*1/*1 (male)		EM (female)		EM (male)	
Parameters	S/R ratio	SD	S/R ratio	S/R ratio	S/R ratio	S/R ratio	S/R ratio	SD	S/R ratio	SD	S/R ratio	SD	S/R ratio	SD
Original Value	0.2287	0.1455	0.16	0.99	0.17	1.00	0.22	0.14	0.33	0.09	0.24	0.15	0.28	0.17
Value read from graph	0.2153	0.1503	0.1768	1.069	0.1854	1.0531	0.2211	0.1343	0.32793	0.08597	0.2427	0.142367	0.278867	0.167567

Note: paired test results. t value is 1.42, p-value is 17.9% at alpha error 0.05 level (two tails).

4.3.3 Statistical analysis

The Hardy-Weinberg equilibrium test has been applied to all the studies included in chapter three. Of the nine articles that reported phenotype and genotype data in a suitable format, the article of Persson *et al.* (1998) is the only study that does not fit the HWE. Meta-analysis was employed to nine studies for the difference of S/R-mephenytoin ratio according to the group of subjects' *CYP2C19* genotypes, where subjects with one mutated CYP2C19 allele, either *CYP2C19*2* or *CYP2C19*3*, are grouped together and subjects with two mutated CYP2C19 alleles, either homozygous or heterozygous for *CYP2C19*2* and *CYP2C19*3*, are put into one group. All studies recruited healthy subjects, who were administered racemic mephenytoin (Mesantoin) 100mg. Each study measured S/R-mephenytoin ratio in the urine 0 to 8 hours after taking mephenytoin.

The S/R-mephenytoin ratios are presented for three subject groups in table 4.3.2; the first is for subjects homozygous for *CYP2C19*1*, the second is for subjects with genotype *CYP2C19*1/*2* or *CYP2C19*1/*3*, the third is for subjects without *CYP2C19*1* allele. As mentioned previously, subjects in the first and second group are normally considered as extensive metabolizers of mephenytoin according to their pharmacokinetic characters in clinical observation. Subjects in the third group are usually poor metabolizers, who are at high risk of side effects due to deficient metabolism of mephenytoin. Therefore, the three groups are named as no mutation group, one mutation group and two mutations group. The comparisons undertaken between these three groups were for no mutation with one mutation, no mutation with two mutations, and one mutation with two mutations respectively. This provides a defined level of deficiency in terms of mephenytoin metabolic ratio due to the effect of mutated CYP2C19 allele/s. The difference between S/R-mephenytoin ratio among three group subjects without or with mutated allele/s is shown in table 4.3.2. The heterogeneity of the nine studies was analysed by employing the Cochrane Q test as proposed by Whitehead *et al.* (1991). The S/R-mephenytoin ratio was treated as a continuous variable with approximately a normal distribution.

The Q value of the heterogeneity test showed no statistically significant heterogeneity

apparent among these nine studies. Therefore, meta-analysis in a fixed model was applied for the estimation of S/R-mephenytoin ratio reduction due to genetic polymorphism of CYP2C19. The results are presented in figure 4.3.2, figure 4.3.3 and figure 4.3.4. for the three comparisons of metabolic efficiency respectively.

The difference of S/R mephenytoin is presented on the left side according to each individual publication, which is shown by first author name and publication year on the right side of the figure. The 95% confidence interval of each S/R ratio is presented on the left side of the figure in bracket following each S/R mephenytoin ratio. The dotted line in the graph represents the over-all difference of S/R ratio between subjects falling into the different genotype groups. The solid line is zero, which indicates no difference between the S/R mephenytoin of two groups in the comparison. Each comparison is described in the title of graph, where the CYP2C9 variant is emphasized in the legend with bold lettering.

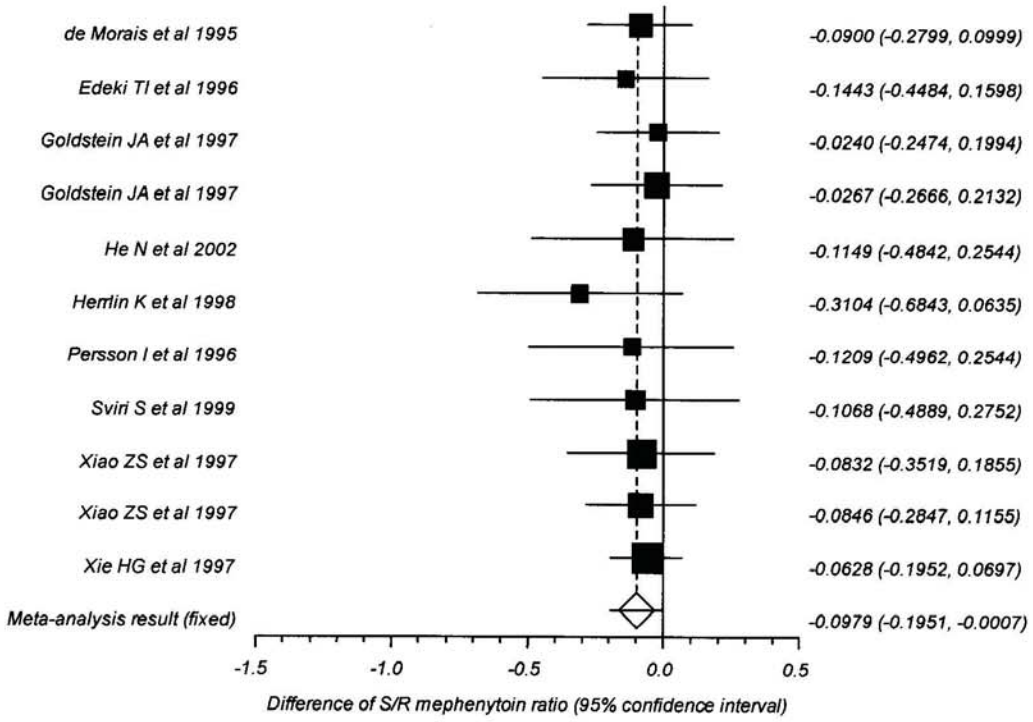


Figure 4.3.2: Meta-analysis (fixed model) of the difference of S/R mephenytoin ratio in subjects carrying **one** *CYP2C19* mutation comparing with subjects without any mutated *CYP2C19* alleles. (Q= 1.765, df = 8)

This analysis shows that healthy subjects carrying one mutated *CYP2C19* allele have a marginally higher S/R mephenytoin ratio compared to healthy subjects without any

CYP2C19 mutant, which indicates a deficient metabolism of mephenytoin. In the absence of any mutated CYP2C19 alleles, the median S/R ratio from healthy subjects within the nine selected studies is 0.24. However, in subjects carrying one mutated CYP2C19 allele, which are clinically categorized as extensive metabolizers in the same group with subjects without any mutation; a difference of S/R mephenytoin ratio between these two group subjects is 0.0979 using a meta-analysis in the fixed model (95% CI: -0.1951~ -0.0007; heterogeneity chi-square Q = 1.765 at df = 8). This corresponds to approximately 37.5 percent of the median S/R ratio of subjects without mutation.

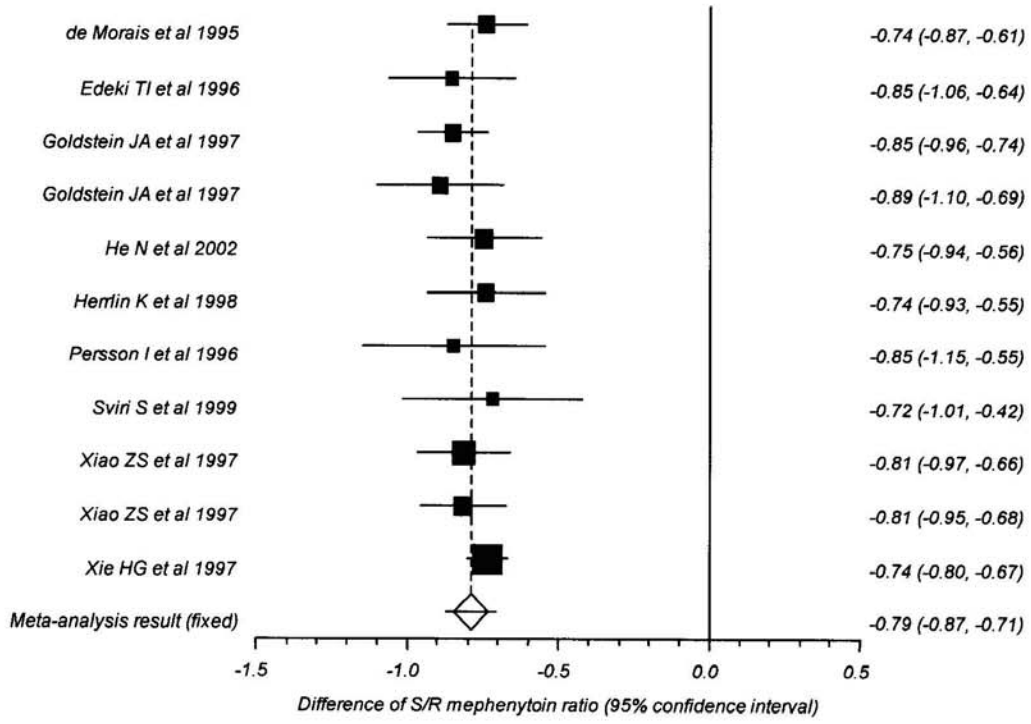


Figure 4.3.3: Meta-analysis (fixed model) of the difference of S/R mephenytoin ratio in subjects carrying **two** *CYP2C19* mutation compared with subjects without mutated *CYP2C19* alleles. (Q= 2.402, df = 8)

In figure 4.3.3, the difference of S/R mephenytoin ratio between subjects with two mutations and subjects without any mutation was 0.79 using a meta-analysis in the fixed model (95% CI: -0.87~ -0.71; heterogeneity chi-square Q = 2.402 at df = 8). This is more than three times of the median S/R ratio (0.24 referring table 4.3.2) of subjects without any mutated CYP2C19 alleles and presents significant deficiency of enzyme activity due to

CYP2C19 mutations.

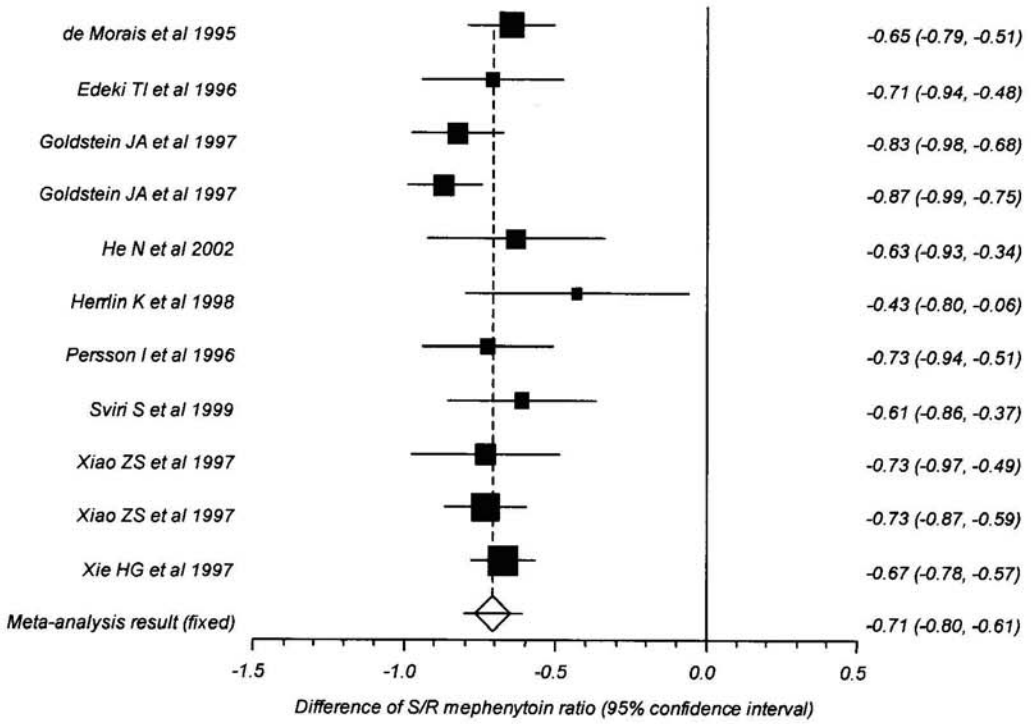


Figure 4.3.4: Meta-analysis (fixed model) of the difference of S/R mephenytoin ratio in subjects carrying **two** *CYP2C19* mutations comparing with subjects with **one** mutated *CYP2C19* allele. (Q= 3.736, df = 8)

In figure 4.3.4, a difference of S/R ratio between healthy subjects with one or two mutated CYP2C19 allele/s was 0.71 using a meta-analysis in the fixed model (95% CI: -0.80~ -0.61; heterogeneity chi-square Q = 3.736 at df = 8). This is only about 10 percent less than the difference of S/R mephenytoin ratio between subjects with no mutation and one mutation (0.79).

4.3.4 Conclusion

These results show subjects carrying more than one mutated *CYP2C19* alleles have higher S/R mephenytoin ratio than subjects without or with only on mutated *CYP2C19* alleles. The deficient metabolism of mephenytoin due to mutation *CYP2C19* can increase the S/R mephenytoin ratio by more than three times. However, the genetic effect of mutated *CYP2C19* on mephenytoin metabolism does not show a direct quantitative relationship between subjects with one mutant and two mutants. Therefore, the additive effect between subjects carrying one and two mutated *CYP2C19* alleles is not found.

Overall, the genetic polymorphism of *CYP2C19* is responsible for the altered metabolism of mephenytoin. Subjects with *CYP2C19*2* or **3* have deficient CYP2C19 enzyme activity. This effect is particularly prevalent in those subjects who are homozygous or heterozygous for *CYP2C19*2* and **3*, as the capability of this enzyme in metabolism of relevant substrates is reduced significantly. According to the prevalence estimation of *CYP2C19* alleles in chapter 3, 2% of Caucasian, 17-20% of Asian such as Chinese or Japanese, and 2.5% of African may have risk of experiencing adverse effect resulting from deficient CYP2C19 enzyme activity for any administered drug that is metabolized by this enzyme. Although CYP2C19 is responsible for only proportion of metabolism of the antiepileptic drug, phenytoin, caution should be applied in the prescription of phenytoin for epilepsy patients with two mutated CYP2C19 alleles.

4.4 Association studies of CYP2C9 polymorphism and warfarin dose requirement

Tolbutamide, phenytoin and warfarin have been used as probe drugs for CYP2C9 studies. However, warfarin is considered as the desired probe drug when considering the extremely narrow therapeutic index and the possible need for steady-state sampling (Streetman *et al.* 2000, Smith *et al.* 1998), which is a major issue in the pharmacotherapy of epilepsy. The studies in the section 4.2 of this chapter have failed to provide sufficient evidence to clarify any association of *CYP2C9* polymorphism with phenytoin intake. Studies on warfarin and polymorphic *CYP2C9* are expected to provide some indication of the impact of *CYP2C9* genetic polymorphism on phenytoin metabolism, where the warfarin dose requirement is considered as a phenotypic marker.

4.4.1 Database search result

A search in the PubMed database was performed through EndNote software using a similar strategy to that applied to studies in the section 4.3, except the keyword mephenytoin was replaced by warfarin. The search identified 63 and 217 publications with the keyword warfarin combinations with polymorphism and cytochrome P450 respectively. The entire search was completed by August 2005.

Selection of those articles was based on the appearance of CYP2C9 in the abstract or in any free text when no abstract available and 46 articles were finally chosen for data retrieval. Following individual review, 32 articles were recovered which described *CYP2C9* genotype data and 13 of them provided phenotype data (warfarin dose) as well. They are presented in table 4.3.1.

Among the 13 articles, data from seven articles was excluded for further analysis; in four of them, the author did not describe the ethnicity of recruited subjects directly or recruited subjects with different ethnic origins, in another two articles, the warfarin dose was not reported for subjects with mutation in *CYP2C9* genotype and one presented warfarin dose per kilogram of subject weight. Six of the studies have completely described warfarin dose

with genotype.

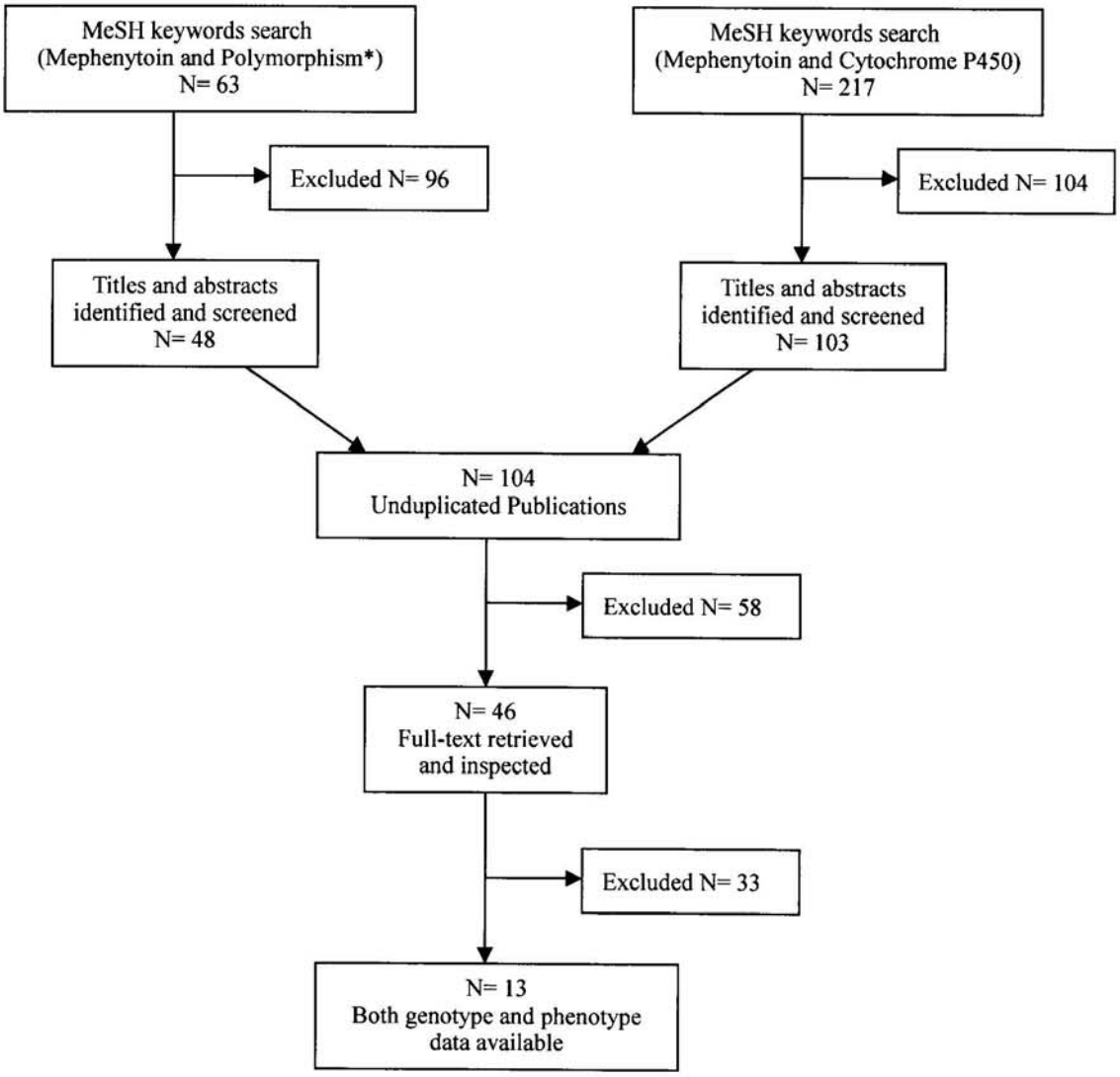


Figure 4.4.1: The search flowchart of studies related to CYP polymorphism and warfarin metabolism

4.4.2 Retrieved data

The thirteen articles were gathered from libraries. The details extracted from each study are presented in table 4.4.1.

Most studies focus on genetic polymorphism of *CYP2C9* only, however Scordo *et al.* (2002) genotyped both of CYP2C9 and CYP2C19, and concluded that the genetic polymorphism of CYP2C19 is unlikely to represent major factor of variability in warfarin dose requirement or metabolism supporting the earlier work of Takahashi *et al.* (1998).

The subjects in these studies were patients receiving warfarin treatment for various disease or pathological reasons, which were not provided in detail by all authors. Although all studies provided maintenance doses of warfarin, not all recruited patients were initially receiving fixed warfarin doses and restricted INR ranges (international normalized ratio of prothrombin time). Besides genetic factors, some authors considered whether warfarin dose requirement is also influenced by a number of pathologic, physiologic and environmental factors of patients, such as age, body weight, co-existing disease, and nutrition in daily diet. However, those factors were not detailed in all studies.

Among those studies with completed phenotype data, *i.e.* warfarin dose, related to each *CYP2C9* genotype, there were only five studies that recruited patients who were initially receiving stable warfarin dosage and with defined INR range. Furthermore, within these five studies, Herman *et al.* (2005) and Linder *et al.* (2002) concluded that patients age contributed significantly to the inter-individual variability in the warfarin dose requirement, however, Scordo *et al.* (2002) failed to find a statistically significant association between patients' age and warfarin dose requirement. Whereas Linder *et al.* (2002) provided the daily maintenance dose of warfarin adjusted by mean body weight of patients. Moreover, in these five studies, Linder *et al.* (2002) only included patients with atrial fibrillation; Aithal *et al.* (1999) and Taube *et al.* (2000) did not provide any patients pathology information in their studies; Herman *et al.* (2005) and Scordo *et al.* (2002) studied patients with a combination of different diseases under the indication of warfarin treatment. Due to difference in study design, meta-analysis has not been undertaken for these association studies between polymorphism of *CYP2C9* genotype and warfarin dose requirement.

Table 4.4.1. Studies of association polymorphism *CYP2C9* and warfarin dose requirement

Author	Ethnicity	Target INR	Stable dose	Disease (%) of recruited patients	Excluded factors	Other factors considered	Warfarin relevant phenotype parameters
Aithal GP *et al* 1999	Caucasian	2.0-3.0	Yes	Not Reported	H; D	None	Dose (mg/day)
Herman D *et al* 2005	Caucasian (Slovenian)	2.0-3.0 (2.5-3.5)	Yes	AF (83.5); DVT (5.9); PHV (1.6); others (9.0)	C; H; R	Age; D; LBW; V (vitK)	Dose(mg/day); warfarin (S/R) plasma concentration(mg/l) & clearance(ml/min/kg)
Higashi MK *et al* 2002	Mixture (96.2%Caucasian, 3.85 Hispanic)	None	None	AF (51.4); DCM (20.5), DVT & PE (21.6); PHV (6.5)	Ethnicity [b]	None	Bleeding event, over-anticoagulant by INR
Hillman WA *et al* 2004	Caucasian	None	None	AF (53.4); DVT (17.7); PHV (22.3); others (6.6)	C; H; R [c]	Age; gender; BSA; PHV; diabetes	Dose (mg/week)
Joffe HV *et al* 2004	Mixture (89%Caucasian, 33%African-American, 3other races)	None	Yes	AF (45.2); VTE (27.4); PHV (11.0); others (16.4)	H; A; V; underweight, age over 80	Age; BM; Ethnicity; acetaminophen	Dose (mg/day); out-of range INR event, bleeding & clotting complications
Kamali F *et al* 2004	Not reported	2.0-3.0	Yes	AF (62.8); VTE (23.1); others (PHV included) (14.1)	H; R; D [d]	Age; gender; BW; BSA; plasma vitamin K concentration	Dose (mg/day), warfarin (S/R) clearance (ml/min) [e]
Linder MW *et al* 2002	Mixture (98%Caucasians with 1 Hispanic, 2% Africa-American)	2.0-3.0	Yes	AF (100)	Not Reported	Age; gender	Daily dose (mg/kg/day); warfarin clearance; S/R plasma concentration (mg/ml) half life
Loebstein R *et al* 2001	Not reported	None	Yes	Not Reported	Not Reported	Age; gender; creatinine & albumin concentration; BSA; A; plasma vitamin K; D (esp. Amiodarone)	Dose (mg/day, regarding each allele); warfarin plasma concentration (ug/l) & clearance (ml/min) [e]
Pchelina SN *et al* 2005	Caucasian (Slavic)	2.0-3.0 (2.5-3.5)	None	AF, VTE	H; R; D	None	Dose (mg/kg/week); time to reach INR [e]
Scordo MG *et al* 2002	Caucasians (Italian)	2.0-3.0	Yes	AF; VTE; VHT; others	D	Age; gender	Dose (mg/week); warfarin (S/R) plasma concentration (mg/ml) & clearance (ml/min)
Tabrizi AR *et al* 2002	Mix (African-American and Caucasians)	1.8-3.5	Yes	VHT (32.7); DVT (17.0); CAD (15.0); AF (14.4); PVD (5.2), others [a]	D; H; ethnicity (two)	Age; BW	Dose (mg/week regarding each allele)
Taube J *et al* 2000	Not reported	2.5	Yes	Not Reported	Not Reported	None	Dose (mg)
Topic E *et al* 2004	Caucasian (Croatians)	None	Yes	AF (16.7); DVT (27.1); PHV (14.5); CAD (13.1); different fractures (11.3); PE (8.5); PVD (2.8); others (6.0)	D	None	Dose (mg/day)

Abbreviation: AF. atrial fibrillation; **DCM**. dilated cardiomyopathy; **DVT**. deep venous thrombosis; **PE**. pulmonary embolism; **PHV**. prosthetic heart valve; **VHD**. valvular heart disease; **VTE**. venous thromboembolism; **Others**. arterial, coronary or heart disease, also included undefined disease. **A**. alcohol consumption; **H**. hepatic dysfunction; **R**. renal dysfunction; **D**. drug interaction with warfarin; **C**. cancer; **V**. diet with rich vegetable, possible high vitamin K content. **BSA**. body surface area; **BM**. body mass; **LBW**. lean body weight; **BW**. body weight; **INR**. International normalized ratio of prothrombin time.

Note: a. few various diseases; b. index date of first warfarin treatment; c. congestive heart failure, taking long acting barbiturate; d. rifampin or herbal use, oral acetaminophen, aminodarone caused and changed warfarin dose, and drug influence vitamin K dispositions; e. phenotype data uncompleted for subjects carrying mutated CYP2C9 allele.

4.4.3 Discussion

Studies describing warfarin dose requirement association with CYP2C9 polymorphism:

Of the five studies focusing on warfarin dose requirement, Athial *et al.* (1999) and Topic *et al.* (2004) stated that age and gender of patients in their studies had no statistically significant association with warfarin maintenance dose without providing detail. Pchelina *et al.* (2005) and Taube *et al.* (2000) did not provide any information about the consideration of age or gender influence on warfarin dosage. Patients recruited in the Higashi *et al.* (2002) study have similar mean ages in both groups with or without variant *CYP2C9* alleles. However, despite these differences, the association between *CYP2C9* variant alleles and reduced warfarin dose requirement was confirmed in all five studies.

Furthermore, Higashi *et al.* (2002) found that patients with at least one variant allele required more time to achieve stable dosing. According to patients records of INR (international normalized ratio of prothrombin time) value, patients with mutated alleles had an increased risk of elevated warfarin levels over the therapeutic INR range and experienced a first bleeding event sooner than patients without mutations, *i.e.* the warfarin sensitivity in carriers of *CYP2C9*2 or *3* was higher than those homozygous for *CYP2C9*1*. This is consistent with the result in the Pchelina *et al.* (2005) study. Patients carrying *CYP2C9*2* and *CYP2C9*3* alleles appeared to require less time to achieve target INR. Therefore, the authors of these studies proposed that determination of *CYP2C9* genotypes could help prediction of individual warfarin dosage requirements prior to the initiation of warfarin therapy, and decrease the risk of serious bleeding events or over-anticoagulation.

Studies observed other variables together with warfarin dose:

In the remaining seven studies, the association of *CYP2C9* polymorphism with warfarin dose requirement was also confirmed, and the contribution was further evaluated by combining metabolic effects with other factors, such as age, body weight, and gender. Both Scordo *et al.* (2002) and Tabrizi *et al.* (2000) found the prevalence of variant *CYP2C9* to be

higher among patients in lower-dosage group than in the median or high dosage groups. Although all seven studies found both age and *CYP2C9* genotype were correlated with warfarin dose requirement, the contribution of these two factors was reported quite differently from each other. Tabrizi *et al.* (2000) stated that 26% of the variability of warfarin dose is determined by *CYP2C9*, age and weight when these three independent variables were considered in the linear regression model together, while age and weight alone will account for 14% of the variation. However, Kamali *et al.* (2004) declared that 20% of inter-patient variability of warfarin dose requirement was determined by *CYP2C9* and age together. In the study of Hillman *et al.* (2004), the contribution of *CYP2C9* and age to the variability of warfarin maintenance dose are clearly quantified as 19.8% and 14.6% respectively, which is significantly different from other studies.

As described earlier, patients recruited into these studies have different combinations of pathological states, which might explain the varied results. Higashi *et al.* (2002) proposed that the data from patients with prosthetic heart valve would potentially skew a generalized evaluation of *CYP2C9* contribution to warfarin dose in a small sample due to their unusual high dosage requirement, but did not provide clear explanation. Consistently, Hillman *et al.* (2004) found the cardiac valve replacement was as an indication for higher warfarin dose requirement and accounted for 5.4% of the variability.

Furthermore, Loebstein *et al.* (2001) proposed that among ambulatory patients, the warfarin sensitivity at optimized steady state dosage was mainly determined by *CYP2C9* and age, without significant association with plasma concentration of vitamin K, creatinine and albumin. This is confirmed by results in studies of Linder *et al.* (2002) and Kamali *et al.* (2004) as well. Additionally, both studies of Scordo *et al.* (2002) and Herman *et al.* (2005) confirmed that *CYP2C9* genetic polymorphism is markedly related to the clearance of the S-warfarin enantiomer, which is a major active anticoagulant component of warfarin. Linder *et al.* (2002) found that during the stable anticoagulation therapy, the plasma concentration S-warfarin did not differ between subjects with different *CYP2C9* genotypes. This study clearly demonstrated that subjects carrying variant *CYP2C9* allele have

decreased metabolism and S-warfarin clearance that can be compensated for by reducing the maintenance dose to achieve a therapeutic concentration of S-warfarin.

Additionally, the Hardy-Weinberg equilibrium (HWE) test was applied to studies that provided both *CYP2C9* allele and genotype data. Apart from Higashi *et al.* (2002) and Topic *et al.* (2004), the rest of the studies fit HWE. The lack of fit to the HWE could be the result of a non-random sample selection.

4.4.4 Conclusion

Overall, most studies reported an association between polymorphism in *CYP2C9* and warfarin dose requirement, but poor comparability of these studies did not allow an actual quantitative estimation of this genetic effect. However, the results in the previous chapter showed that less than 1% of Chinese and Japanese carry the **2* variant, and 2.2-3.6% of them have the **3* variant, in contrast, over 11% and 7% Caucasian people appear have *CYP2C9*2* and **3* respectively. If polymorphism of *CYP2C9* is the major genetic factor relevant to warfarin dose requirment, the potential risk due to warfarin over-dose is higher for Caucasian populations.

The most recent studies (Shikata *et al.* 2004, Wadelinus *et al.* 2004, Taguchi *et al.* 2005) have proposed that other candidate genes or proteins that might play a role in the inter-individual variability of warfarin therapy, and polymorphism of *CYP2C9* might only partially contribute to the variability. Therefore, more studies that consider those co-factors would be able to provide more evidence to quantify the *CYP2C9* genetic effect on warfarin metabolism and may consequently provide information to improve the therapeutic safety and efficiency of other CYP2C9 substrates, such as phenytoin.

4.5 Summary and Conclusion

Topic *et al.* (2004) proposed that differences in the variant allele frequency of CYP in Caucasian subgroups could be responsible for the variability in percentage of individuals with altered drug response among different Caucasian populations. According to the estimated prevalence of *CYP2C9/2C19* that were reported in chapter three, it is believed that patients with different ethnic origin would have different metabolic ratios or dose requirements due to various prevalence of *CYP2C9/2C19* variants. Therefore, the consequent risk of adverse effect, such as sedation because of high plasma concentration results from deficient metabolism of mephenytoin, and serious bleeding event or over-anticoagulation related to warfarin over-dosage, will be different among patients from different ethnic populations.

Overall, a generalized quantification of CYP2C9/2C19 genetic polymorphism cannot be achieved from the currently available studies. As phenytoin exhibits non-linear metabolism, altered enzyme activity could change the concentration of free phenytoin significantly. The genetic polymorphism of *CYP2C9/2C19* is important to help clinical therapeutic monitoring. However, polymorphism of metabolizing enzymes cannot fully explain the variability of therapeutic response either for phenytoin or warfarin. Undetermined co-variables remain to be identified. Present pharmacogenetic studies lack established procedures to minimise confounding variables and deliver reproducible results. Monogenetic studies are inappropriate for explaining the variability in complex drug action; besides drug metabolism enzymes, other genetic polymorphisms associated with drug metabolic enzymes have to be emphasized and investigated. Non-genetic factors and genetic factors combine to contribute to polymorphism of drug response, therefore a multiple variable model is required to establish the quantitative combined effect of genotype and environmental factors.

It is still difficult for pharmacogenetic studies to reach clinical pharmacotherapy. Even though the testing-kit for CYP polymorphism is commercially available, many inconclusive genetic polymorphisms remain to be identified achieving before personalized medicine.

Chapter 5 Discussion and Conclusions

5.1 Preface

Pharmacogenetics holds the promise that drugs might one day be tailor-made for individuals and adapted to each person's own genetic makeup. Much pharmacogenetic research has been done since the human genome was published in 2001. In this thesis, the pharmacogenetic factors relevant to the antiepileptic drug, phenytoin, were assessed, and have focussed on the polymorphisms of *CYP2C9/2C19* using available evidence. Initially, the population distribution of *CYP2C9/2C19* was investigated for populations with different ethnic origins. Where several datasets were provided in studies of Caucasian, Chinese, Japanese and African populations, meta-analysis methods were applied to estimate the prevalence of the three major *CYP2C9/2C19* alleles. The results confirmed that a difference in *CYP2C9/2C19* prevalence exists between those populations. This provides a molecular basis for the prevalence of altered metabolism, *i.e.* poor metabolizer or extensive metabolizer between people with different ethnic origins. In addition, studies of phenytoin metabolism and *CYP2C9/2C19* polymorphism were examined. The association between *CYP2C9/2C19* variants and altered phenytoin metabolism was retrieved from previously reported studies. However, insufficient evidence precluded a generalized quantification of this genetic effect on phenytoin metabolism for clinical application. Therefore, the association studies of CYP2C9/2C19 polymorphism and an appropriate probe drug were further explored. Mephenytoin studies investigated the enzyme CYP2C19 activity according to *CYP2C19* genotypes. The altered metabolism exhibits an association with *CYP2C19* allelic variants; metabolism was significantly reduced in healthy subjects who were homozygous or heterozygous of *CYP2C19*2* and *CYP2C19*3*. Warfarin dose requirement also appeared to associate with *CYP2C9* genotypes in Caucasian patients under anticoagulant treatments. However, the data were heterogeneous and no further analysis was undertaken. Further studies are required to establish replicable association results for different ethnic populations.

5.2 Discussion

5.2.1 Evidence search and selection:

Although pharmacogenetics was introduced since 1950's, the standard MeSH keywords related to pharmacogenetic studies are limited. As found in the previous search of this thesis, the most commonly used keywords in pharmacogenetic publications were 'variant', 'variation' and 'polymorphism', which are actually frequently used subject headings in genetic studies as well. Therefore, many references retrieved using those terms were associated with genetic sequencing or variation studies instead of drug response studies. This complicated the selection of relevant publications. However, in the future, with the construction of specified pharmacogenetic databases, such as PharmGKB, (a recently developed database for genetic variation correlated with drug responses by Stanford University funded by the National Institutes of Health), a more integrated resource related to pharmacogenetic aspects will be systematically organized and easily accessed (PharmGKB database website).

Meanwhile, the nomenclature of genetic variations or mutations was variable between individual publications. For instance, among the studies of *CYP2C9*2* polymorphisms, many authors used 430C>T, the nucleotide change in cDNA, or Arg144Cys, the amino acid change, instead of the variant name introduced by Human Cytochrome P450 Nomenclature system, despite the introduction of the nomenclature rule of Cytochrome P-450 (CYP) in 1987. Furthermore, among studies of *CYP2C19* polymorphism, some authors followed the earliest publications by naming *CYP2C19*1* as 'wt' allele, *CYP2C19*2* as 'm1', and *CYP2C19*3* as 'm2', which indicated wild type allele, mutation one and two respectively. From the prevalence estimation undertaken in this thesis, the ethnic differences in the CYP2C19 genetic polymorphism unequivocally exist; the most common allele in one ethnic population is not always the most common allele in another ethnicity. For instance, *CYP2C19*1* can be found in 85.8% of Caucasian people; however only 22.2% of Vanuatu people would have the allele (data presented in Pacific population studies section 3.2.4). Therefore, the term 'wild-type' is not suitable for the description of this allele in human

populations any more. It is necessary to stop using ambiguous terms and apply the standardized human genetic nomenclature rules consistently in future pharmacogenetic research.

5.2.2 Ethnicity and populations:

The inter-ethnic variation in drug response has been observed since the early nineteenth century (Evans *et al.* 2001). Before molecular or genetic markers were developed, the identification of ethnicity was commonly made using the physiological characteristics of human populations, such as skin colour, skull shape and face features. The term used interchangeable with ethnicity, *i.e.* race, is frequently used and linked with inequality of health care or discrimination of a 'super population' concept. However, in this thesis, the ethnicity classification and grouping aims to allow determination of the difference in allele frequency among population groups, as it can provide the molecular basis for population differences in drug-metabolism enzymes and contribute to the discovery of differential drug responses.

The metabolism of phenytoin predominantly involves two Cytochrome P450s; CYP2C9 and CYP2C19. The current available evidence of a polymorphic phenytoin response is from the association studies between three CYP2C9/2C19 alleles and phenytoin metabolism. As previously mentioned, the different prevalence of poor metabolizer (PM) phenotypes, such as for mephenytoin, was reported over fairly wide range. For instance 14-22% of Asian population and 2-6% Caucasian were detected as mephenytoin poor metabolizers. Meanwhile, the CYP2C9 variants were found in 33-51% of Caucasian populations, 2-10% of Asian populations and 3-13% of African populations. Both phenotype observations and genotype studies have given us a relatively imprecise estimation of PM prevalence in different ethnic populations. Therefore, the thesis initially employed a meta-analysis method to study CYP2C9/2C19 population distribution, and intended to provide a more accurate estimated prevalence of deficient metabolizing enzyme activity among patients with different ethnic origins. As the substrate list for CYP2C9/2C19 enzymes has been

developed comprehensively in pharmacological studies, it is believed that the population distribution of CYP2C9/2C19 variants could contribute to selection of the clinical therapeutic strategy at very early on stage, *i.e.* when an administered drug was metabolized principally by CYP2C9, patients with Caucasian ethnic origin would require more attention or clinical monitoring than Chinese patients. On the other hand, prescription of drugs primarily metabolized by CYP2C19 in Chinese patients would require more caution.

According to the results in the thesis, *CYP2C9*2/*3* variants are found at relatively lower prevalence in Chinese, Japanese and African populations (around 1-3%); in contrast, the prevalence of *CYP2C19*2/*3* variants, especially *CYP2C19*2,* in those populations is higher, *i.e.* over 38% Chinese and 29% Japanese have the mutated allele *CYP2C19*2.* However, most population studies found that the polymorphism of *CYP2C9* is responsible for the polymorphism of phenytoin metabolism, whereas the polymorphism of *CYP2C19* seems to be un-important in the metabolism of phenytoin. Such lower prevalence of *CYP2C9*2/*3* variants in Chinese and Japanese population is hard to explain in the clinical context of polymorphism in phenytoin metabolism or maintenance dose requirement within populations. Therefore, other genetic or non-genetic polymorphisms have to be considered for these populations, as Taguchi *et al.* (2005) found recently that polymorphism of *CYP2C9/2C19* failed to identify the large variability of phenytoin dose requirement in routinely treated Japanese epilepsy patients. This highlighted an important aspect of the allele prevalence studies among different ethnic populations. The interethnic variability of drug response might have a slightly different molecular basis between different populations, which would be partially presented in the different allele frequencies of functional variants among populations (McLeod and Ameyaw 2002). Therefore, the identification of ethnicity is very important in studies of genetic polymorphism.

However, two major demographic trends, migration and intermarriage, have created difficulties for measurement and identification of the population according to either race or ethnicity. Consequently, the pool of potential multiple-ethnicity respondents has become higher and higher, and is common in some developed countries, such as the United States.

For instance, by the 1990s, 17% of immigrants to the US came from Europe or Canada, 30% from Asia, and around 50% from Latin America. However, in 1997, 61% of Asians, 38% of Hispanic, 8% of whites, 6% of Blacks and 6% of American Indians had migrated from other countries. Correspondingly, the pattern of intermarriage changed tremendously with the composition of migrants as well. According to the 1990 census in US, 5% of the population reported an ancestry that differed from their primary ethnicity. However, in 1995, 7% of the population were reported as having multiple ancestries (Waters 2000). Therefore, the categories of ethnicity have to be localized according to the composition of the populations in different regions or countries. Unfortunately, among the retrieved publications in the thesis, many authors had recruited subjects from the local community or residences without further detail of either the method used for ethnicity classification or the local population composition. Furthermore, some studies observed the prevalence of *CYP2C9* polymorphism in European countries. These articles did not define the ethnicity of subjects in the study directly, nevertheless they had compared the allele prevalence of the studied population with the allele frequencies published for Caucasian and even further concluded there was no statistically significant difference between the uncertain ethnic population and the Caucasian populations (Sandberg *et al.* 2004). Meanwhile, some studies grouped into miscellaneous group declared that the inter-ethnic crosses are normally difficult to detect in many populations. For instance, the study with subjects living in Brazil categorized their populations into white, black and intermediate according to the Brazilian Census (Vianna-Jorge *et al.* 2004). They found *CYP2C9*5* in a self-identified white individual, which is a variant previously only detected in African populations (Kirchheiner *et al.* 2005). In order to elucidate the apparent discordance, the authors accomplished a further genealogical study with the subject's parents and two brothers by using ancestry-informative DNA markers developed for the trihybrid Brazilian population, where the mother and one brother identified themselves as intermediate and the father and the other brother identified themselves as white. The results revealed the relative contribution of European, African and Amerindian roots to the proband's genetic pool with 92.0%, 7.5% and 0.5% respectively. Furthermore, DNA markers for matrilineal inheritance indicated the presence of Western African population roots in the family. Therefore, the researchers

concluded that the *CYP2C9*5* allele in the proband and one brother were inherited from the matrilineal, African ancestry. Their results clearly stressed the dilemma of global extrapolation of pharmacogenetic research.

Currently the reporting of ethnicity data by self-identification is recognized as the more accurate method of data collection (Bierman *et al.* 2002). Daar and Singer proposed that for the association studies between genetic polymorphism and observed phenotypic differences, self-identified ethnicity combined with information about continental ancestry would provide a relatively higher predictive power than self-identified ethnicity only (Daar and Singer 2002). However, in this thesis, the articles of only a few authors have traced three generations of ancestry of subjects, and this only accounts for approximately 7% of the entire retrieved publications (10/129).

Additionally, some frequently used terms in pharmacogenetic publications, such as Caucasian, Chinese, Japanese and African, actually represent different ranks of populations; where Chinese and Japanese have more specific meaning than Caucasian and African. Therefore, studies of Chinese and Japanese populations would be predicted to have less heterogeneity compared with Caucasian and African population studies, which were confirmed with population differentiation tests and heterogeneity tests. Moreover, as previously mentioned, the category of ethnicities has not been consistently established in biological science yet. Regarding the validity of ethnic categorisation, there is debate about the relationship between human genetic variation and ethnicity in biomedical science. Population genetic theories are acknowledged as a valuable methodology to construct population structures by implying human evolutionary history and genetic variation (Risch *et al.* 2002). These theories propose that the human population is subdivided into major old world continents of African, Asia and Europe, and whilst around 85-90% of genetic variation exists within these subdivisions, only 10-15% of variation exists between them. This estimate is highly consistent for many types of autosomal systems, such as protein polymorphisms (Jorde *et al.* 2001). It is exhibited in the trinomial contours of studies in miscellaneous group, where the *CYP2C19* distribution is highly diverse among Asian

populations, such as Chinese, Japanese, Thai and Korean peoples. The phylogenetic tree of populations followed in the thesis was created using a number of DNA molecular markers. Hey (1998) proposed that any single measurement of genetic distance confounds divergence over time since population subdivision and divergence may due to limited gene flow (genetic migration). For instance, two populations with limited gene flow, which separated a long time ago, will reveal lower numerical distance value than two population separated recently but with zero gene flow. Therefore, the different population histories will lead to an important distinction between genetic distance tree and population trees (Jody 1998). In the thesis, Caucasian, Chinese, Japanese and African populations showed different allelic distributions when comparing with the population diversity between *CYP2C9* and *CYP2C19*, which illustrates the divergence of distance value by a single genetic locus. At the same time, the overlap or distance between other Asian population and Chinese/Japanese also exhibited dissimilarity between the trinomial contours of *CYP2C9* and *CYP2C19*. Similar results were shown among those populations close to Caucasian or African as well. Therefore, this supports that proposal that constructing a human haplotype map from multiple loci is expected to give a more conclusive population structure in the future (Tabor *et al.* 2002, Daar and Singer 2005).

Apart from the complexity of inter-ethnic origin and population structures, populations with same ethnic origin may present with different phenotypes due to interaction between genetic factors and environmental factors. Although Aklillu *et al.* (2002) did not find a significantly different contribution of environmental factor to the CYP2C19 activity between Ethiopians living in Sweden and Ethiopians living in Ethiopia, a significant environmental contribution on CYP2D6 activity was found between the two groups of Ethiopians. However, this is the only one study reporting such evidence that was available at the time of writing the thesis. Further research is required.

5.2.3 Study design: random sampling and polygenetic model

An absolute random sample is hard to achieve in a practical experiment. For gene

prevalence estimation, subject recruitment is acceptable as long as the sample can represent the objective population. Meanwhile, most studies tend to combine the convenience and cost of sampling within a random sample. The results show that the majority of studies retrieved in the thesis sampled unrelated subjects from the local community or residences. However, some studies recruited subjects according to restricted criteria, such as epilepsy patients or patients receiving the same clinical treatment, which were excluded from any prevalence estimation in this thesis as they may be less generalizable.

As warfarin is frequently used as a therapeutic drug, the phenotypic studies of CYP polymorphisms were predominantly carried out by recruitment of patients undergoing anticoagulant treatment. Aside from the debate whether warfarin is suitable as a probe-drug for observation of CYP2C9 activity as it is an active therapeutic medicine; the clarification of association between warfarin maintenance dose and *CYP2C9* polymorphism may contribute effectively to clinical management for those patients under anticoagulant treatment. However, these studies offered highly inconsistent conclusions without compatibility in study design. As most studies acknowledged that non-genetic factors of patients, such as demographic characteristics, co-existing diseases and environmental exposures (diet) (Hillman *et al.* 2004, Joffe *et al.* 2004, Loebstein *et al.* 2001), could be confounding variables and affect warfarin dose requirement, a lack of replicated results among these studies was not a surprise. Therefore, for association studies between genetic polymorphism and drug response, the selection of subjects with fewer confounding variables is essential to empower the generalize ability of any conclusion.

Furthermore, the detection of an association between genetic polymorphism and drug response requires a scrupulous selection of candidate genes. Since Meyer proposed that genetic polymorphisms are named for those mutant or variant genes that exist at a prevalence of more than 1% in the normal population (Meyer 1991), it is conventionally accepted that the genetic variation occurring in more than 1% of a population would be considered a useful polymorphism for pharmacogenetic studies (see definition of polymorphism in National Institutes of Health). As the prevalence of *CYP2C9*2* and **3*

occurs at a relatively low frequency in Chinese and Japanese populations, apart from those non-genetic factors, it is possible that among Chinese and Japanese patients, other undiscovered *CYP2C9* variants may contribute to the polymorphism of warfarin dose requirement, or unknown molecular mechanisms may exist, which are responsible for the phenotypic polymorphism.

As a drug response involves many biological activities that are comprised of various protein or enzymes, a multiple-genes model would be more appropriate than monogenetic observations (Hattersley and McCarthy 2005). Furthermore, the monogenetic model is not practical for complex disease treatment, such as epilepsy. More appropriate methodologies are required to discover the association between polygenetics and response to drug treatments.

5.2.4 Measurement of association between genotype and phenotype:

Although genotypic data is easy to obtain with currently available biological techniques, the primer design and development has to be carefully checked to reduce poor or non-specific amplification of polymerase chain reaction products, which is particularly important for those genes closely located in the same region of same chromosome, *i.e.* one gene cluster. Polymorphic genes of *CYP2C9/2C19* that are the focus of this thesis belong to the same gene cluster, CYP2C cluster, where four members of cytochrome P450 subfamily C (*CYP2C8*, *CYP2C9*, *CYP2C18*, and *CYP2C19*) are located adjacent to one another on chromosome 10. In a study of Hong Kong Chinese patients, Leung *et al.* (2001) reported the discovery of four new single nucleotide polymorphism (SNPs) in the coding region of CYP2C9. They further declared that patients with one of the new variants exhibited a 26-34% reduction in warfarin dose requirement. Since the lower frequency of CYP2C9 *2 and *3 in Chinese populations cannot explain the evidence that Asian patients require only about 50% of warfarin dose comparing with Caucasian patients, the study had been considered significant. Therefore, a few more studies have undertaken according to their design, but they all failed to find any of the variants in other Asian populations (Zarza *et al.*

2002, Goldstein 2002, Chang 2003). However, in 2003, Rettie *et al.* found that the primers designed in Leung's study were not adequate for the discovery of *CYP2C9* SNPs. Apart from the first ten bases, the forward primer actually exhibited 100% homology to the exon 4 sequences within *CYP2C8*, *CYP2C9*, *CYP2C18*, and *CYP2C19*. Similarly, after the first eight bases the reverse primer exhibits 100% match to exon 4/intron 4 sequences within *CYP2C9* and *CYP2C19*. Therefore, the genotype methods applied in these studies were referred back to the original publications in order to identify whether those genotype data had introduced a similar error.

The phenotypic data related to a particular polymorphic gene is normally hard to detect, due to multiple factors involved in drug administration from drug absorption, metabolism, disposition, and elimination. Besides consideration of these co-existing factors or co-variables, the appropriate measurement has to consider the nature of individual drug action, and defined according to pharmacokinetic feature of the drugs, such as S/R mephenytoin ratio for mephenytoin studies, phenytoin plasma concentration for phenytoin studies, warfarin maintenance dose for warfarin studies. Furthermore, the phenotype measurement may need different parameters or observation methods for each metabolism feature and varied study designs for healthy subjects and patients.

5.2.5 Statistical significance:

Statistical principles have been frequently applied in pharmacogenetics, however, the application of the statistical method should be undertaken with caution. Consideration of the genetic assumption or particular hypothesis underlying is essential. For instance, the Chi-square goodness of fit test is a basic principle applied to detect population deviation from the Hardy-Weinberg equilibrium (HWE). However, the calculation to test whether a population fits Hardy-Weinberg equilibrium requires the observed genotype frequencies in order to calculate the allele frequencies. Allele frequencies are then used to predict the expected genotype frequencies. Therefore the degree of freedom in the statistical test has to be reduced by the number of allele frequencies estimated from the observed genotype

frequencies. In the earlier statistical HWE test of section 3.2.2, the Chi-square goodness fit was performed using the Statistica (6.0) software. As a default setting in most programs, the number of degrees of freedom is equal to the number of observed variables minus one. Therefore, the probability of goodness of fit was calculated in an inappropriate way. However, in the GENEPOP program, the test was designed for verifying the deviation from HWE. Therefore, probability was computed according to the appropriate degrees of freedom for each study.

Furthermore, some articles summarized the polymorphic prevalence of CYP2C9/2C19 variants among different ethnic populations (Kirchheiner *et al.* 2005, Lee *et al.* 2001, Xie *et al.* 2000), however, none of them have applied a meta-analysis method to estimate the prevalence of variants for different ethnic populations.

Xie *et al.* (2000) studied the polymorphism of CYP2C19 variants in Chinese populations, where they applied the Chi-square test to the number of CYP2C19 genotypes for the detection of heterogeneity between studies. The frequencies of CYP2C19 genotypes were then estimated using the arithmetic average of four homogenous Chinese population studies. In this thesis, the population genetic principle, the Hardy-Weinberg equilibrium (HWE), was initially tested each study by Chi-square goodness of fit test performed in the GENEPOP program. Those studies not fitting HWE were excluded for further prevalence estimation. The result showed that the possibility of one Chinese population (Bai, Xiao *et al.* 1997) fitting HWE was less than 5%, therefore it was excluded from the prevalence estimation. However, this study was included in the genotype estimation of Xie *et al.* (2000).

The meta-analysis approach always requires heterogeneity test before pooling results from individual studies. Then different pooling methods will be used according to whether heterogeneity exists among studies; fixed and random models will apply to meta-analysis of studies without or with heterogeneity respectively. The Cochrane Q test is often employed for the heterogeneity test. In the thesis, besides applying the Q-test for each allele individually, the heterogeneity among populations was also assessed by considering the

three CYP2C9/2C19 alleles together; this population differentiation test was performed in the GENEPOP. Furthermore, because some studies had found less than five subjects with mutated alleles, and some studies had not found certain mutated allele, the power of the Q test must be questioned due to variable skew from assumption of normality (Hardy and Thompson 1998). The population differentiation test in the GENEPOP program has employed the Fisher exact test using the Markov Chain principle, which would give acceptable estimation of heterogeneity among populations and between paired populations in these circumstances. Therefore, population differentiations between those studies were identified in the GENEPOP program as an addition to the Cochrane Q test.

For the prevalence estimations, conclusive results cannot be given to the *CYP2C9* variants in the African and Chinese populations due to insufficient studies. There are only three studies currently available for each population after selecting the random recruitment of subjects and testing the derivation from the HWE among those studies provided complete *CYP2C9* genotypes. Furthermore, apart from *CYP2C9*2* and **3*, two of the African population studies had found other variants. The frequency of these variants was greater than 1 in 100 subjects, therefore the *CYP2C9*1*, **2*, and**3* may not the major alleles in the African population. Before the prevalence estimation of *CYP2C9* alleles can be undertaken, more studies are required to examine the genetic variants existing in the African populations.

The effect of CYP2C9/2C19 polymorphism on metabolic activity towards phenytoin, mephenytoin and warfarin metabolism was confirmed in the research. However, a quantitative effect of each *CYP2C9/2C19* variant was not achieved due to poorly comparable studies.

5.2.6 Other issues:

As in other medical evidence searches, this thesis retrieved quite a few publications with duplicated data. This is particularly common when the recruitment for genotype is difficult

and expensive in some countries; some research groups studied the same subjects with different objectives and consequently published similar association studies more than once. The data from these studies was included once according to the completeness and most recent publishing date. Meanwhile, the publications retrieved in this thesis are primarily in English language, which limited the journal selection to most west or developed countries. Therefore, studies of other ethnic populations were insufficient for any further analysis comparing with the studies of Caucasian or Japanese populations. The studies of African populations were principally completed with American-Africans instead of indigenous Africans. As mentioned previously, the environmental factors and admixed ethnic composition would have an impact on the prevalence estimation of genetic variants. The estimated prevalence of geographic variation in Africans may be biased by these factors.

In the thesis, the exclusion criteria have included: 1.studies with no ethnicity definition, 2.studies which do not provide complete data, 3.studies of subjects sampled from restricted populations except restricted ethnicity, 4.studies failing to fit to the Hardy-Weinberg equilibrium, 5.studies performed in-*vitro*. Among these excluded studies, in-*vitro* pharmacogenetic studies can be considered as pre-experimental of studies performing on individual subjects. They are important for studies of the molecular basis of a relevant drug response. Meanwhile, two articles investigated the association between warfarin dose requirement and the single nucleotide polymorphisms (SNPs) of *CYP2C9* by constructing a population haplotype structure (Takahasi *et al.* 2004, Veenstra *et al.* 2005). One study identified thirteen SNPs, and found some of the SNPs were in linkage disequilibrium with functionally defective coding region variants in the Japanese population. Another study proposed that the *CYP2C9* genetic variations in exons rather than in the promoter or other regulatory region are responsible for warfarin sensitivity in European American (Caucasian) patients. Their results proved the possibility that different molecular bases of genetic variants may contribute to the variability of warfarin treatment between Caucasian and Japanese. However, more studies are needed in the future to obtain any conclusive results. This further affirmed there is possibility that the different molecular bases contribute to the variability of a given drug response when variant prevalence between populations was

significantly different. Also the variant studies of the encoding region cannot always explain the variability of drug responses. The constructions of haplotyope maps between different populations can provide more valuable evidence of genetic polymorphism that is responsible for the variability of drug responses in pharmacogenetic studies.

Future of pharmacogenetic research:

Pharmacogenetics has not really been applied to clinical therapy yet. Most pharmacogenetic studies have provided a qualitative association between genetic polymorphism and drug response. However, a large amount of association studies are related to drug metabolism enzymes. The genetic polymorphism of drug receptors and transport proteins still remain to be investigated. Therefore, the quantitative estimation of a genetic effect on drug response remains uncertain. Meanwhile, many methodological issues exist in genetic association studies. Regarding the pharmacogenetic studies investigated in this thesis, there is still a lack of sufficient evidence to achieve any generalized conclusion and contribute to advice on personalized medicine.

Monogenetic polymorphism often does not explain the variability of drug response. Identification of the potential candidate genes remains questionable. Multiple-gene models need to be established with improved knowledge of potential candidate genes. Quantitative analysis of genetic effect on drug response will become achievable with the aid of both the advanced knowledge of molecular biology in drug actions and the development of statistical methodologies suitable for multiple-gene models and complex diseases.

5.3 Conclusion

The prevalence of *CYP2C9*1*, **2* and **3* is different among Caucasian, Chinese, Japanese and African populations. The difference is significant between Caucasian and other populations. Besides the Turkish and south India populations, the population of other Asian peoples have similar allele prevalence to the Chinese and Japanese populations, where the *CYP2C9*2* allele is rare for Chinese and Japanese. Populations from West Asia, such as from Turkey and Southern India do not have overlapping trinomial contours for alleles of CYP2C9*1, *2 and *3, where the contours of the Southern India population is not overlapping with Caucasian populations either. However, some populations from north of Africa, such as Egyptian, and West Asia, such as Turkish, appear to show overlapping trinomial contours with Caucasians for the three alleles. Those populations, which defined themselves as descendants of two different continental ancestries, have partially overlapped trinomial contours with both of the continental ancestries. Other populations that have a distinct identification of ethnicity exhibit discrete distributions from each other. The trinomial contours of those studies that studied mixture of subjects from different ethnic origin exhibited an unconventional distribution from any of ethnic populations, which represents a composite of the mixed populations.

Considering the population distributions of *CYP2C19* alleles, *CYP2C19*2* variants are rare alleles for Caucasian and African populations. Chinese and Japanese population have little overlapping contours of *CYP2C9*1*, **2*, and **3* from each other. Populations of other Asian peoples have overlapping trinomial contours of allelic variants with Chinese and Japanese populations except the populations of south and north India. Similarly, in studying the *CYP2C9* population distributions, those populations defined themselves as descendants of two different continental ancestries and have partially overlapping trinomial contours with both of the continental ancestries. Among those unique ethnicities, except Inuit, most populations have highly overlapped trinomial contours with Caucasian and African populations, and *CYP2C19*2* appear rarely. In contrast, trinomial contour of the Inuit population is located separated from any other populations. Furthermore, trinomial contours

of *CYP2C19* alleles of Turkish and Egyptians markedly overlap with Caucasian populations, which is also true for trinomial contours of *CYP2C9* alleles. In the contrast, some India populations did not overlap region with Caucasians.

An association between *CYP2C9/2C19* allelic variants and phenytoin metabolism was found in healthy subjects. Epilepsy patients exhibited a similar association between CYP2C9/2C19 variants and two phenytoin pharmacokinetic parameters, the maximum velocity of the enzyme system (Vmax) and the Michaelis constant of the enzyme system (Km). However, consolidation of data and identification of preferable phenotype measurements remain to be confirmed in large population studies. This may identify different parameters for healthy subjects compared with epilepsy patients due to their different pharmacokinetic activities. The metabolic ratio of the probe-drug mephenytoin reduced significantly when people carried two mutated CYP2C19 alleles. When comparing people with one mutation and two mutations of *CYP2C19*, the allelic variants did not have an additive effect. As an active therapeutic drug, numerous association studies between warfarin dose requirement and CYP2C9 polymorphism have been performed using Caucasian patients under clinical anticoagulant treatments. However, patients recruited into such studies do not necessarily have similar pathological circumstances. Meanwhile, among these studies, some recruited patients immediately on starting of warfarin treatment or whereas others had already achieved a stable warfarin maintain dosage with or without a restricted INR range. Therefore, the data derived from those patients have less comparability for undertaking further analysis. Additionally the influence of other genetic factors or non-genetic factors precludes any conclusive results from those studies. Further studies are required to elucidate these confounding variables and to evaluate the effect of *CYP2C9* polymorphism on warfarin dose requirement.

Reference list:

Aiken S. P. and Brown W. M., Treatment of epilepsy: existing therapies and future developments. Front Biosci, 2000. 5: p. E124-52.

Aklillu E., Herrlin K., Gustafsson L. L., Bertilsson L. and Ingelman-Sundberg M., Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. Pharmacogenetics, 2002. 12(5): p. 375-83.

Alain Li Wan P, Evidence-based Pharmacotherapy. The pharmaceutical Journal, 1996 256: 308-312

Ambudkar S. V., Dey S., Hrycyna C. A., Ramachandra M., Pastan I. and Gottesman M. M., Biochemical, cellular, and pharmacological aspects of the multidrug transporter. Annu Rev Pharmacol Toxicol, 1999. 39: p. 361-98.

Anantharam A., Markowitz S. M. and Abbott G. W., Pharmacogenetic considerations in diseases of cardiac ion channels. J Pharmacol Exp Ther, 2003. 307(3): p. 831-8.

Anderson G. D., A mechanistic approach to antiepileptic drug interactions. Ann Pharmacother, 1998. 32(5): p. 554-63.

Anderson G. D., Pharmacogenetics and enzyme induction/inhibition properties of antiepileptic drugs. Neurology, 2004. 63(10 Suppl 4): p. S3-8.

Asconape J. J., Some common issues in the use of antiepileptic drugs. Semin Neurol, 2002. 22(1): p. 27-39.

Bailey D. S., Bondar A. and Furness L. M., Pharmacogenomics - it's not just pharmacogenetics. Curr Opin Biotechnol, 1998. 9(6): p. 595-601.

Beghi E., Efficacy and tolerability of the new antiepileptic drugs: comparison of two recent guidelines. Lancet Neurol, 2004. 3(10): p. 618-21.

Berg A. T., Testa F. M., Levy S. R. and Shinnar S., The epidemiology of epilepsy. Past, present, and future. Neurol Clin, 1996. 14(2): p. 383-98.

Berkovic S. F. and Scheffer I. E., Genetics of the epilepsies. Epilepsia, 2001. 42 Suppl 5: p. 16-23.

Bertilsson L. and Kalow W. (2001), Interethnic differences in drug disposition and effects. In Pacifici G.M. and Pelkonen O. Interindividual Variability in Human Drug Metabolism, 15-74.

Bertz R. J. and Granneman G. R., Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. Clin Pharmacokinet, 1997. 32(3): p. 210-58.

Bierman A. S., Lurie N., Collins K. S. and Eisenberg J. M., Addressing racial and ethnic barriers to effective health care: the need for better data. Health Aff (Millwood), 2002. 21(3): p. 91-102.

Birbeck G. L. and Kalichi E. M., Epilepsy prevalence in rural Zambia: a door-to-door survey. Trop Med Int Health, 2004. 9(1): p. 92-5.

Birbeck G. L., Seizures in rural Zambia. Epilepsia, 2000. 41(3): p. 277-81.

Bowker T. J., Clayton T. C., Ingham J., McLennan N. R., Hobson H. L., Pyke S. D., Schofield B. and Wood D. A., A British Cardiac Society survey of the potential for the secondary prevention of coronary disease: ASPIRE (Action on Secondary Prevention through Intervention to Reduce Events). Heart, 1996. 75(4): p. 334-42.

Brockmoller J., Cascorbi I., Henning S., Meisel C. and Roots I., Molecular genetics of cancer susceptibility. Pharmacology, 2000. 61(3): p. 212-27.

Brodie B. B., Axelrod J., Cooper J. R., Gaudette L., La Du B. N., Mitoma C. and Udenfriend S., Detoxication of drugs and other foreign compounds by liver microsomes. Science, 1955. 121(3147): p. 603-4.

Brodie M. J. and Dichter M. A., Antiepileptic drugs. N Engl J Med, 1996. 334(3): p. 168-75.

Browne T. R. and Leduc B. (2002). Phenytoin and other hydantoins: chemistry and biotransformation. In: Antiepileptic drugs. 5th Edition. Philadelphia, USA: Lippincott Williams & Wilkins, 565-580.

Bruce M., Griffiths C., Brock A. and Majeed A., Trends in mortality and hospital admissions associated with epilepsy in England and Wales during the 1990s. Health Stat Q, 2004(21): p. 23-9.

Bruni Joseph. (2002). Phenytoin and other hydantoins: adverse effects. In: Antiepileptic drugs. 5th Edition.

Philadelphia, USA: Lippincott Williams & Wilkins, 605-610.

Burchard E. G., Ziv E., Coyle N., Gomez S. L., Tang H., Karter A. J., Mountain J. L., Perez-Stable E. J., Sheppard D. and Risch N., The importance of race and ethnic background in biomedical research and clinical practice. N Engl J Med, 2003. 348(12): p. 1170-5.

Cascino G. D., Surgical Treatment for Extratemporal Epilepsy. Curr Treat Options Neurol, 2004. 6(3): p. 257-262.

Cavalli-Sforza L. L., Menozzi Paolo and Piazza Alberto (1996). Genetic history of world populations. In The history and Geography of Human Genes. Abridged Paperback Edition. New Jersey, USA. Princeton University Press, 60-154.

Cavalli-Sforza L. L., The DNA revolution in population genetics. Trends Genet, 1998. 14(2): p. 60-5.

Chaabani H., GM polymorphism and the evolutionary history of modern humans. Ann Genet, 2002. 45(4): p. 197-206.

Chang J. G., No exon 4 polymorphism of cytochrome P450 CYP2C9 in Taiwanese. Blood, 2003. 101(12): p. 5086-7.

Chang J. T. and Altman R. B., Extracting and characterizing gene-drug relationships from the literature. Pharmacogenetics, 2004. 14(9): p. 577-86.

Claverie J. M., Fewer genes, more noncoding RNA. Science, 2005. 309(5740): p. 1529-30.

Claverie J. M., Gene number. What if there are only 30,000 human genes? Science, 2001. 291(5507): p. 1255-7.

Crow J. F., Eighty years ago: the beginnings of population genetics. Genetics, 1988. 119(3): p. 473-6.

Daar A. S. and Singer P. A., Pharmacogenetics and geographical ancestry: implications for drug development and global health. Nat Rev Genet, 2005. 6(3): p. 241-6.

Davis A. and Long R., Pharmacogenetics research network and knowledge base: 1st annual scientific meeting. Pharmacogenomics, 2001. 2(3): p. 285-9.

Davis A. F. and Long R. M., Pharmacogenetics research network and knowledge base second scientific meeting. Pharmacogenomics J, 2002. 2(5): p. 293-6.

Davis A. F. and Long R. M., Pharmacogenetics Research Network and Knowledge Base third scientific meeting. Pharmacogenetics, 2003. 13(7): p. 437-40.

Davis A. F. and Long R. M., Pharmacogenetics research network and knowledge base fourth scientific meeting. Pharmacogenomics J, 2004. 4(6): p. 359-61.

de Morais S. M., Goldstein J. A., Xie H. G., Huang S. L., Lu Y. Q., Xia H., Xiao Z. S., Ile N. and Zhou H. H., Genetic analysis of the S-mephenytoin polymorphism in a Chinese population. Clin Pharmacol Ther, 1995. 58(4): p. 404-11. a

De Morais S. M., Wilkinson G. R., Blaisdell J., Meyer U. A., Nakamura K. and Goldstein J. A., Identification of a new genetic defect responsible for the polymorphism of (S)-mephenytoin metabolism in Japanese. Mol Pharmacol, 1994. 46(4): p. 594-8. b

de Morais S. M., Wilkinson G. R., Blaisdell J., Nakamura K., Meyer U. A. and Goldstein J. A., The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. J Biol Chem, 1994. 269(22): p. 15419-22.

Dichter M. A. and Brodie M. J., New antiepileptic drugs. N Engl J Med, 1996. 334(24): p. 1583-90.

Du B. N., Jr., Pharmacogenetics. Med Clin North Am, 1969. 53(4): p. 839-55.

Durbin R. (1998). Biological sequence analysis: probabilistic models of proteins and nucletic acids. Cambridge: Cambridge University press, 46-51.

Essious L., Destenaves B., Jais P. and Thomas F. (2002), Association studies in Pharmacogenomics. In Licinio J and Wang Ma-Li. Pharmacogenomics. Germany. WILEY-VCH Verlag GmbH Weinheim, 57-83.

Estabrook R. W., A passion for P450s (rememberances of the early history of research on cytochrome P450). Drug Metab Dispos, 2003. 31(12): p. 1461-73.

Evans D. A., Manley K. A. and Mc K. V., Genetic control of isoniazid metabolism in man. Br Med J, 1960. 5197: p. 485-91.

Evans D. A., McLeod H. L., Pritchard S., Tariq M. and Mobarek A., Interethnic variability in human drug responses. Drug Metab Dispos, 2001. 29(4 Pt 2): p. 606-10.

Evans W. E. and McLeod H. L., Pharmacogenomics--drug disposition, drug targets, and side effects. N Engl J Med, 2003. 348(6): p. 538-49.

Evans W. E. and Relling M. V., Pharmacogenomics: translating functional genomics into rational therapeutics. Science, 1999. 286(5439): p. 487-91.

Ewens W. J. (1979). Mathematical population genetics. In: Biomathematics. Vol. 9. Berlin, Germany: Spinger-Werlag Berlin Heidelberg.

Ferraro T. N. and Buono R. J., The relationship between the pharmacology of antiepileptic drugs and human gene variation: an overview. Epilepsy Behav, 2005. 7(1): p. 18-36.

Fritz S., Lindner W., Roots I., Frey B. M. and Kupfer A., Stereochemistry of aromatic phenytoin hydroxylation in various drug hydroxylation phenotypes in humans. J Pharmacol Exp Ther, 1987. 241(2): p. 615-22.

Gaitatzis A. and Sander J. W., The mortality of epilepsy revisited. Epileptic Disord, 2004. 6(1): p. 3-13.

Gaitatzis A., Purcell B., Carroll K., Sander J. W. and Majeed A., Differences in the use of health services among people with and without epilepsy in the United Kingdom: socio-economic and disease-specific determinants. Epilepsy Res, 2002. 50(3): p. 233-41.

Garfinkel D., Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. Arch Biochem Biophys, 1958. 77(2): p. 493-509.

Gervasini G., Carrillo J. A. and Benitez J., Potential role of cerebral cytochrome P450 in clinical pharmacokinetics: modulation by endogenous compounds. Clin Pharmacokinet, 2004. 43(11): p. 693-706.

Goldstein D. B., Ahmadi K. R., Weale M. E. and Wood N. W., Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. Trends Genet, 2003. 19(11): p. 615-22. a

Goldstein D. B., Pharmacogenetics in the laboratory and the clinic. N Engl J Med, 2003. 348(6): p. 553-6. b

Goldstein D. B., Tate S. K. and Sisodiya S. M., Pharmacogenetics goes genomic. Nat Rev Genet, 2003. 4(12): p. 937-47. c

Goldstein J. A. and Blaisdell J., Genetic tests which identify the principal defects in CYP2C19 responsible for the polymorphism in mephenytoin metabolism. Methods Enzymol, 1996. 272: p. 210-8.

Goldstein J. A. and de Morais S. M., Biochemistry and molecular biology of the human CYP2C subfamily. Pharmacogenetics, 1994. 4(6): p. 285-99. a

Goldstein J. A., CYP2C9 polymorphisms and CYP2C9*2 genotyping primers. Br J Clin Pharmacol, 2002. 53(4): p. 409-10.

Goldstein J. A., Faletto M. B., Romkes-Sparks M., Sullivan T., Kitareewan S., Raucy J. L., Lasker J. M. and Ghanayem B. I., Evidence that CYP2C19 is the major (S)-mephenytoin 4'-hydroxylase in humans. Biochemistry, 1994. 33(7): p. 1743-52. b

Gray I. C., Nobile C., Muresu R., Ford S. and Spurr N. K., A 2.4-megabase physical map spanning the CYP2C gene cluster on chromosome 10q24. Genomics, 1995. 28(2): p. 328-32.

Guengerich F. P., Cytochromes P450, drugs, and diseases. Mol Interv, 2003. 3(4): p. 194-204.

Guengerich F. P., Wu Z. L. and Bartleson C. J., Function of human cytochrome P450s: characterization of the orphans. Biochem Biophys Res Commun, 2005. 338(1): p. 465-9.

Guengerich Peter F. (2004). Human cytochrome P450 enzymes. In Cytochrome P450 Structure, Mechanism, and Biochemistry. 3rd Edition. New York, USA. Kluwer Academic/Plenum Publishers, 377-530.

Hartl D. L. (2000). A primer of population genetics. 3rd Edition. Sunderland, USA: Sinauer association, Inc. 29-32 (114-117).

Hartl D. L. and Clark A. G. Principles of population genetics. 3rd Edition. Sunderland, USA: Sinauer association, Inc. 20-36. (71-108, chapter 3). (111-124, chapter 4)

Hasler J. A., Pharmacogenetics of cytochromes P450. Mol Aspects Med, 1999. 20(1-2): p. 12-24, 25-137.

Hattersley A. T. and McCarthy M. I., What makes a good genetic association study? Lancet, 2005. 366(9493): p. 1315-23.

Hey J., Population genetics and human origins--haplotypes are key! Trends Genet, 1998. 14(8): p. 303-5.

Ieiri I. and Higuchi S., Pharmacogenetics of CYP2C subfamily in a Japanese population. J Toxicol Sci, 1998. 23 Suppl 2: p. 129-31.

Ieiri I., Mamiya K., Urae A., Wada Y., Kimura M., Irie S., Amamoto T., Kubota T., Yoshioka S., Nakamura K., Nakano S., Tashiro N. and Higuchi S., Stereoselective 4'-hydroxylation of phenytoin: relationship to (S)-mephenytoin polymorphism in Japanese. Br J Clin Pharmacol, 1997. 43(4): p. 441-5.

Ingelman-Sundberg M., Duplication, multiduplication, and amplification of genes encoding drug-metabolizing enzymes: evolutionary, toxicological, and clinical pharmacological aspects. Drug Metab Rev, 1999. 31(2): p. 449-59. a

Ingelman-Sundberg M., Oscarson M. and McLellan R. A., Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. Trends Pharmacol Sci, 1999. 20(8): p. 342-9. b

Ingelman-Sundberg M., Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. Trends Pharmacol Sci, 2004. 25(4): p. 193-200.

Jiang Z. P., Shu Y., Chen X. P., Huang S. L., Zhu R. H., Wang W., He N. and Zhou H. H., The role of CYP2C19 in amitriptyline N-demethylation in Chinese subjects. Eur J Clin Pharmacol, 2002. 58(2): p. 109-13.

Johnson J. A. and Evans W. E., Molecular diagnostics as a predictive tool: genetics of drug efficacy and toxicity. Trends Mol Med, 2002. 8(6): p. 300-5.

Johnson J. A., Pharmacogenetics: potential for individualized drug therapy through genetics. Trends Genet, 2003. 19(11): p. 660-6.

Johnson N.L. and Kotz S. (1969). Discrete Distributions, John Wiley & Sons, pp.58-61.

Jones-Davis D. M. and Macdonald R. L., GABA(A) receptor function and pharmacology in epilepsy and status epilepticus. Curr Opin Pharmacol, 2003. 3(1): p. 12-8.

Jorde L. B. and Wooding S. P., Genetic variation, classification and 'race'. Nat Genet, 2004. 36(11 Suppl): p. S28-33.

Jorde L. B., Watkins W. S. and Bamshad M. J., Population genomics: a bridge from evolutionary history to genetic medicine. Hum Mol Genet, 2001. 10(20): p. 2199-207.

Jody H., Population genetics and Human origins-haplotypes are key. Trend in Genetics, 1998. 14(8): p. 330-334

Kalow W., Ethnic differences in drug metabolism. Clin Pharmacokinet, 1982. 7(5): p. 373-400.

Kaneko S., Iwasa H. and Okada M., Genetic identifiers of epilepsy. Epilepsia, 2002. 43 Suppl 9: p. 16-20.

Kemeny A. A., Surgery for epilepsy. Seizure, 2001. 10(6): p. 461-5.

Kim K., Johnson J. A. and Derendorf H., Differences in drug pharmacokinetics between East Asians and Caucasians and the role of genetic polymorphisms. J Clin Pharmacol, 2004. 44(10): p. 1083-105.

Kim R. B., Leake B. F., Choo E. F., Dresser G. K., Kubba S. V., Schwarz U. I., Taylor A., Xie H. G., McKinsey J., Zhou S., Lan L. B., Schuetz J. D., Schuetz E. G. and Wilkinson G. R., Identification of functionally variant MDR1 alleles among European Americans and African Americans. Clin Pharmacol Ther, 2001. 70(2): p. 189-99.

Kim R. B., Pharmacogenetics of CYP enzymes and drug transporters: remarkable recent advances. Adv Drug Deliv Rev, 2002. 54(10): p. 1241-2.

Kirchheiner J. and Brockmoller J., Clinical consequences of cytochrome P450 2C9 polymorphisms. Clin Pharmacol Ther, 2005. 77(1): p. 1-16.

Kjeldsen M. J., Corey L. A., Christensen K. and Friis M. L., Epileptic seizures and syndromes in twins: the importance of genetic factors. Epilepsy Res, 2003. 55(1-2): p. 137-46.

Kjeldsen M. J., Kyvik K. O., Christensen K. and Friis M. L., Genetic and environmental factors in epilepsy: a population-based study of 11900 Danish twin pairs. Epilepsy Res, 2001. 44(2-3): p. 167-78.

Kjeldsen M. J., Kyvik K. O., Friis M. L. and Christensen K., Genetic and environmental factors in febrile seizures: a Danish population-based twin study. Epilepsy Res, 2002. 51(1-2): p. 167-77.

Klassen B. D. and Sadler R. M., Induction of hypersensitivity to a previously tolerated antiepileptic drug by a second antiepileptic drug. Epilepsia, 2001. 42(3): p. 433-5.

Kohno M., Yokokawa K., Minami M., Kano H., Yasunari K., Hanehira T. and Yoshikawa J., Association between angiotensin-converting enzyme gene polymorphisms and regression of left ventricular hypertrophy in patients treated with angiotensin-converting enzyme inhibitors. Am J Med, 1999. 106(5): p. 544-9.

Kupfer A. and Preisig R., Pharmacogenetics of mephenytoin: a new drug hydroxylation polymorphism in man. Eur J Clin Pharmacol, 1984. 26(6): p. 753-9.

Kupfer A., Desmond P. V., Schenker S. and Branch R. A., Stereoselective metabolism and disposition of the enantiomers of mephenytoin during chronic oral administration of the racemic drug in man. J Pharmacol Exp Ther, 1982. 221(3): p. 590-7.

La Du B. N., Pharmacogenetics: defective enzymes in relation to reactions to drugs. Annu Rev Med, 1972. 23: p. 453-68.

LaRoche S. M. and Helmers S. L., The new antiepileptic drugs: clinical applications. Jama, 2004. 291(5): p. 615-20.

Lee C. R., Goldstein J. A. and Pieper J. A., Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. Pharmacogenetics, 2002. 12(3): p. 251-63.

Leung A. Y., Chow H. C., Kwong Y. L., Lie A. K., Fung A. T., Chow W. H., Yip A. S. and Liang R., Genetic polymorphism in exon 4 of cytochrome P450 CYP2C9 may be associated with warfarin sensitivity in Chinese patients. Blood, 2001. 98(8): p. 2584-7.

Levy R. H., Cytochrome P450 isozymes and antiepileptic drug interactions. Epilepsia, 1995. 36 Suppl 5: p. S8-13.

Lewis C. M., Genetic association studies: design, analysis and interpretation. Brief Bioinform, 2002. 3(2): p. 146-53.

Lewis D. F., Watson E. and Lake B. G., Evolution of the cytochrome P450 superfamily: sequence alignments and pharmacogenetics. Mutat Res, 1998. 410(3): p. 245-70.

Lima J. M., The new drugs and the strategies to manage epilepsy. Curr Pharm Des, 2000. 6(8): p. 873-8.

Lindhout D. and Omtzigt J. G., Pregnancy and the risk of teratogenicity. Epilepsia, 1992. 33 Suppl 4: p. S41-8.

Little J., Bradley L., Bray M. S., Clyne M., Dorman J., Ellsworth D. L., Hanson J., Khoury M., Lau J., O'Brien T. R., Rothman N., Stroup D., Taioli E., Thomas D., Vainio H., Wacholder S. and Weinberg C., Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol, 2002. 156(4): p. 300-10.

Loscher W. and Potschka H., Role of multidrug transporters in pharmacoresistance to antiepileptic drugs. J Pharmacol Exp Ther, 2002. 301(1): p. 7-14. a

Loscher W., Current status and future directions in the pharmacotherapy of epilepsy. Trends Pharmacol Sci, 2002. 23(3): p. 113-8. b

Louis E.J. and Dempster E.R. An exact test for Hardy-Weinberg and multiple alleles. Biometrics, 1987, 43: 805-911.

Ma J. D., Nafziger A. N. and Bertino J. S., Jr., Genetic polymorphisms of cytochrome P450 enzymes and the effect on interindividual, pharmacokinetic variability in extensive metabolizers. J Clin Pharmacol, 2004. 44(5): p. 447-56.

Ma M. K., Woo M. H. and McLeod H. L., Genetic basis of drug metabolism. Am J Health Syst Pharm, 2002. 59(21): p. 2061-9.

Macdonald R. L. and Kelly K. M., Antiepileptic drug mechanisms of action. Epilepsia, 1995. 36 Suppl 2: p. S2-12.

Mahgoub A., Idle J. R., Dring L. G., Lancaster R. and Smith R. L., Polymorphic hydroxylation of Debrisoquine in man. Lancet, 1977. 2(8038): p. 584-6.

Mattick J. S., The functional genomics of noncoding RNA. Science, 2005. 309(5740): p. 1527-8.

McFayden M. C., Melvin W. T. and Murray G. I., Regional distribution of individual forms of cytochrome P450 mRNA in normal adult human brain. Biochem Pharmacol, 1998. 55(6): p. 825-30.

McLeod H. L. and Evans W. E., Pharmacogenomics: unlocking the human genome for better drug therapy. Annu Rev Pharmacol Toxicol, 2001. 41: p. 101-21.

Meier U. T. and Meyer U. A., Genetic polymorphism of human cytochrome P-450 (S)-mephenytoin 4-hydroxylase. Studies with human autoantibodies suggest a functionally altered cytochrome P-450 isozyme

as cause of the genetic deficiency. Biochemistry, 1987. 26(25): p. 8466-74.

Meyer U. A. and Zanger U. M., Molecular mechanisms of genetic polymorphisms of drug metabolism. Annu Rev Pharmacol Toxicol, 1997. 37: p. 269-96.

Meyer U. A., Genotype or phenotype: the definition of a pharmacogenetic polymorphism. Pharmacogenetics, 1991. 1(2): p. 66-7.

Meyer U. A., Molecular genetics and the future of pharmacogenetics. Pharmacol Ther, 1990. 46(3): p. 349-55.

Meyer U. A., Pharmacogenetics and adverse drug reactions. Lancet, 2000. 356(9242): p. 1667-71.

Meyer U. A., Pharmacogenetics: the slow, the rapid, and the ultrarapid. Proc Natl Acad Sci U S A, 1994. 91(6): p. 1983-4. a

Meyer U. A., The molecular basis of genetic polymorphisms of drug metabolism. J Pharm Pharmacol, 1994. 46 Suppl 1: p. 409-15. b

Miksys S. and Tyndale R. F., The unique regulation of brain cytochrome P450 2 (CYP2) family enzymes by drugs and genetics. Drug Metab Rev, 2004. 36(2): p. 313-33.

Miners J. O. and Birkett D. J., Cytochrome P4502C9: an enzyme of major importance in human drug metabolism. Br J Clin Pharmacol, 1998. 45(6): p. 525-38.

Moore M. J., From birth to death: the complex lives of eukaryotic mRNAs. Science, 2005. 309(5740): p. 1514-8.

Motulsky A. G., Drug reactions enzymes, and biochemical genetics. J Am Med Assoc, 1957. 165(7): p. 835-7.

Mountain J. L. and Risch N., Assessing genetic contributions to phenotypic differences among 'racial' and 'ethnic' groups. Nat Genet, 2004. 36(11 Suppl): p. S48-53.

Nebert D. W. and Gonzalez F. J., P450 genes: structure, evolution, and regulation. Annu Rev Biochem, 1987. 56: p. 945-93. a

Nebert D. W. and Russell D. W., Clinical importance of the cytochromes P450. Lancet, 2002. 360(9340): p. 1155-62.

Nebert D. W., Adesnik M., Coon M. J., Estabrook R. W., Gonzalez F. J., Guengerich F. P., Gunsalus I. C., Johnson E. F., Kemper B., Levin W. and et al., The P450 gene superfamily: recommended nomenclature. DNA, 1987. 6(1): p. 1-11. b

Nebert D. W., Ingelman-Sundberg M. and Daly A. K., Genetic epidemiology of environmental toxicity and cancer susceptibility: human allelic polymorphisms in drug-metabolizing enzyme genes, their functional importance, and nomenclature issues. Drug Metab Rev, 1999. 31(2): p. 467-87.

Nebert D. W., McKinnon R. A. and Puga A., Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. DNA Cell Biol, 1996. 15(4): p. 273-80.

Nebert D. W., Nelson D. R., Adesnik M., Coon M. J., Estabrook R. W., Gonzalez F. J., Guengerich F. P., Gunsalus I. C., Johnson E. F., Kemper B. and et al., The P450 superfamily: updated listing of all genes and recommended nomenclature for the chromosomal loci. DNA, 1989. 8(1): p. 1-13.

Nebert D. W., Nelson D. R., Coon M. J., Estabrook R. W., Feyereisen R., Fujii-Kuriyama Y., Gonzalez F. J., Guengerich F. P., Gunsalus I. C., Johnson E. F. and et al., The P450 superfamily: update on new sequences, gene mapping, and recommended nomenclature. DNA Cell Biol, 1991. 10(1): p. 1-14.

Nelson D. R., Kamataki T., Waxman D. J., Guengerich F. P., Estabrook R. W., Feyereisen R., Gonzalez F. J., Coon M. J., Gunsalus I. C., Gotoh O. and et al., The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol, 1993. 12(1): p. 1-51.

Omura T. and Sato R., A new cytochrome in liver microsomes. J Biol Chem, 1962. 237: p. 1375-6.

Ortiz de Montellano P.R. and De Voss J.J. (2004), Substrate oxidation by cytochrome P450 enzymes. In Cytochrome P450 Structure, Mechanism, and Biochemistry. 3rd Edition. New York, USA. Kluwer Academic/Plenum publishers, 183-246.

Patsalos P. N., Froscher W., Pisani F. and van Rijn C. M., The importance of drug interactions in epilepsy therapy. Epilepsia, 2002. 43(4): p. 365-85.

Perucca E., Beghi E., Dulac O., Shorvon S. and Tomson T., Assessing risk to benefit ratio in antiepileptic drug therapy. Epilepsy Res, 2000. 41(2): p. 107-39.

Pfeffer N., Theories of race, ethnicity and culture. Bmj, 1998. 317(7169): p. 1381-4.

Phillips K. A., Veenstra D. L., Oren E., Lee J. K. and Sadee W., Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. Jama, 2001. 286(18): p. 2270-9.

Pirmohamed M. and Park B. K., Cytochrome P450 enzyme polymorphisms and adverse drug reactions. Toxicology, 2003. 192(1): p. 23-32.

Pirmohamed M. and Park B. K., Genetic susceptibility to adverse drug reactions. Trends Pharmacol Sci, 2001. 22(6): p. 298-305.

Preux P. M. and Druet-Cabanac M., Epidemiology and aetiology of epilepsy in sub-Saharan Africa. Lancet Neurol, 2005. 4(1): p. 21-31.

Purcell B., Gaitatzis A., Sander J.W. and Majeed A., Epilepsy prevalence and prescribing patterns in England and Wales. Health Statistics Quaterly, 2002, 15: 23-30.

Raymond M and Rousset F. An exact test for population differentiation. Evolution, 1995, 49: 1280-1283

Regesta G. and Tanganelli P., Clinical aspects and biological bases of drug-resistant epilepsies. Epilepsy Res, 1999. 34(2-3): p. 109-22.

Rendic S., Summary of information on human CYP enzymes: human P450 metabolism data. Drug Metab Rev, 2002. 34(1-2): p. 83-448.

Rettie A. E., Tai G., Veenstra D. L., Farin F. M., Srinouanprachan S., Lin Y. S., Thummel K. E. and Hines R. N., CYP2C9 exon 4 mutations and warfarin dose phenotype in Asians. Blood, 2003. 101(7): p. 2896-7.

Risch N., Burchard E., Ziv E. and Tang H., Categorization of humans in biomedical research: genes, race and disease. Genome Biol, 2002. 3(7): p. comment2007.

Rodrigues A. D. and Rushmore T. H., Cytochrome P450 pharmacogenetics in drug development: in vitro studies and clinical consequences. Curr Drug Metab, 2002. 3(3): p. 289-309.

Rogawski M. A. (2002). General principles: principles of antiepileptic drug action. In: Antiepileptic drugs. 5th Edition. Philadelphia, USA: Lippincott Williams & Wilkins, 3-22.

Rosenberg N. A., Pritchard J. K., Weber J. L., Cann H. M., Kidd K. K., Zhivotovsky L. A. and Feldman M. W., Genetic structure of human populations. Science, 2002. 298(5602): p. 2381-5.

Rousset F. and Raymond M., Testing heterozygote excess and deficiency. Genetics, 1995. 140(4): p. 1413-9.

Sachidanandam R., Weissman D., Schmidt S. C., Kakol J. M., Stein L. D., Marth G., Sherry S., Mullikin J. C., Mortimore B. J., Willey D. L., Hunt S. E., Cole C. G., Coggill P. C., Rice C. M., Ning Z., Rogers J., Bentley D. R., Kwok P. Y., Mardis E. R., Yeh R. T., Schultz B., Cook L., Davenport R., Dante M., Fulton L., Hillier L., Waterston R. H., McPherson J. D., Gilman B., Schaffner S., Van Etten W. J., Reich D., Higgins J., Daly M. J., Blumenstiel B., Baldwin J., Stange-Thomann N., Zody M. C., Linton L., Lander E. S. and Altshuler D., A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature, 2001. 409(6822): p. 928-33.

Scheffer I. E. and Berkovic S. F., The genetics of human epilepsy. Trends Pharmacol Sci, 2003. 24(8): p. 428-33.

Schmidt D. and Loscher W., How effective is surgery to cure seizures in drug-resistant temporal lobe epilepsy? Epilepsy Res, 2003. 56(2-3): p. 85-91.

Shikata E., Ieiri I., Ishiguro S., Aono H., Inoue K., Koide T., Ohgi S. and Otsubo K., Association of pharmacokinetic (CYP2C9) and pharmacodynamic (factors II, VII, IX, and X; proteins S and C; and gamma-glutamyl carboxylase) gene variants with warfarin sensitivity. Blood, 2004. 103(7): p. 2630-5.

Shimada T., Yamazaki H., Mimura M., Inui Y. and Guengerich F. P., Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther, 1994. 270(1): p. 414-23.

Shimizu W., The long QT syndrome: therapeutic implications of a genetic diagnosis. Cardiovasc Res, 2005. 67(3): p. 347-56.

Smith D. A., Abel S. M., Hyland R. and Jones B. C., Human cytochrome P450s: selectivity and measurement in vivo. Xenobiotica, 1998. 28(12): p. 1095-128.

Solus J. F., Arietta B. J., Harris J. R., Sexton D. P., Steward J. Q., McMunn C., Ihrie P., Mehall J. M., Edwards T. L. and Dawson E. P., Genetic variation in eleven phase I drug metabolism genes in an ethnically diverse population. Pharmacogenomics, 2004. 5(7): p. 895-931.

Spear B. B., Pharmacogenetics and antiepileptic drugs. Epilepsia, 2001. 42 Suppl 5: p. 31-4.

Stephen L. J. and Brodie M. J., Seizure freedom with more than one antiepileptic drug. Seizure, 2002. 11(6): p. 349-51.

Streetman D. S., Bertino J. S., Jr. and Nafziger A. N., Phenotyping of drug-metabolizing enzymes in adults: a review of in-vivo cytochrome P450 phenotyping probes. Pharmacogenetics, 2000. 10(3): p. 187-216.

Tabor H. K., Risch N. J. and Myers R. M., Opinion: Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet, 2002. 3(5): p. 391-7.

Taguchi M., Hongou K., Yagi S., Miyawaki T., Takizawa M., Aiba T. and Hashimoto Y., Evaluation of phenytoin dosage regimens based on genotyping of CYP2C subfamily in routinely treated Japanese patients. Drug Metab Pharmacokinet, 2005. 20(2): p. 107-12.

Takahashi H., Ieiri I., Wilkinson G. R., Mayo G., Kashima T., Kimura S., Otsubo K. and Echizen H., 5'-Flanking region polymorphisms of CYP2C9 and their relationship to S-warfarin metabolism in white and Japanese patients. Blood, 2004. 103(8): p. 3055-7.

Tanaka E., Clinical importance of non-genetic and genetic cytochrome P450 function tests in liver disease. J Clin Pharm Ther, 1998. 23(3): p. 161-70. b

Tanaka E., Clinically important pharmacokinetic drug-drug interactions: role of cytochrome P450 enzymes. J Clin Pharm Ther, 1998. 23(6): p. 403-16. a

Tanaka E., Clinically significant pharmacokinetic drug interactions between antiepileptic drugs. J Clin Pharm Ther, 1999. 24(2): p. 87-92. a

Tanaka E., Update: genetic polymorphism of drug metabolizing enzymes in humans. J Clin Pharm Ther, 1999. 24(5): p. 323-9. b

Tate S. K. and Goldstein D. B., Will tomorrow's medicines work for everyone? Nat Genet, 2004. 36(11 Suppl): p. S34-42.

Treiman D. M., GABAergic mechanisms in epilepsy. Epilepsia, 2001. 42 Suppl 3: p. 8-12.

Tsai J. J., Lai M. L., Yang Y. H. and Huang J. D., Comparison on bioequivalence of four phenytoin preparations in patients with multiple-dose treatment. J Clin Pharmacol, 1992. 32(3): p. 272-6.

van der Weide J. and Steijns L. S., Cytochrome P450 enzyme system: genetic polymorphisms and impact on clinical pharmacology. Ann Clin Biochem, 1999. 36 (Pt 6): p. 722-9.

Veenstra D. L., Blough D. K., Higashi M. K., Farin F. M., Srinouanprachan S., Rieder M. J. and Rettie A. E., CYP2C9 haplotype structure in European American warfarin patients and association with clinical outcomes. Clin Pharmacol Ther, 2005. 77(5): p. 353-64.

Venkatakrishnan K., Von Moltke L. L. and Greenblatt D. J., Human drug metabolism and the cytochromes P450: application and relevance of in vitro models. J Clin Pharmacol, 2001. 41(11): p. 1149-79.

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., Amanatides P., Ballew R. M., Huson D. H., Wortman J. R., Zhang Q., Kodira C. D., Zheng X. H., Chen L., Skupski M., Subramanian G., Thomas P. D., Zhang J., Gabor Miklos G. L., Nelson C., Broder S., Clark A. G., Nadeau J., McKusick V. A., Zinder N., Levine A. J., Roberts R. J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A. E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T. J., Higgins M. E., Ji R. R., Ke Z., Ketchum K. A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G. V., Milshina N., Moore H. M., Naik A. K., Narayan V. A., Neelam B., Nusskern D., Rusch D. B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M. L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y. H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N. N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J. F., Guigo R., Campbell M. J., Sjolander K. V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha

M., Carnes-Stine J., Caulk P., Chiang Y. H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M., Pan S., Peck J., Peterson M., Rowe W., Sanders R., Scott J., Simpson M., Smith T., Sprague A., Stockwell T., Turner R., Venter E., Wang M., Wen M., Wu D., Wu M., Xia A., Zandieh A. and Zhu X., The sequence of the human genome. Science, 2001. 291(5507): p. 1304-51.

Vesell E. S., Genetic and environmental factors affecting drug response in man. Fed Proc, 1972. 31(4): p. 1253-69.

Vianna-Jorge R., Perini J. A., Rondinelli E. and Suarez-Kurtz G., CYP2C9 genotypes and the pharmacokinetics of tenoxicam in Brazilians. Clin Pharmacol Ther, 2004. 76(1): p. 18-26.

Wadelius M., Sorlin K., Wallerman O., Karlsson J., Yue Q. Y., Magnusson P. K., Wadelius C. and Melhus H., Warfarin sensitivity related to CYP2C9, CYP3A5, ABCB1 (MDR1) and other factors. Pharmacogenomics J, 2004. 4(1): p. 40-8.

Waters M. C., Immigration, intermarriage, and the challenges of measuring racial/ethnic identities. Am J Public Health, 2000. 90(11): p. 1735-7.

Weber W. W. and Hein D. W., Clinical pharmacokinetics of isoniazid. Clin Pharmacokinet, 1979. 4(6): p. 401-22.

Wedlund P. J., The CYP2C19 enzyme polymorphism. Pharmacology, 2000. 61(3): p. 174-83.

Weinshilboum R., Inheritance and drug response. N Engl J Med, 2003. 348(6): p. 529-37.

Weiss K. M. (1993). Genetic variation and human disease: principles and evolutionary approaches. Cambridge: Cambridge University press, 52-65.

Werner E., Genome semantics, in silico multicellular systems and the Central Dogma. FEBS Lett, 2005. 579(8): p. 1779-82.

Whitehead A. and Whitehead J., A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med, 1991. 10(11): p. 1665-77.

Whitehead J. Measurement of treatment difference. In The Design and Analysis of sequential clinical trials. 2nd Edition. London. Elli Horwood publisher, 39-73.

Wilby J., Kainth A., Hawkins N., Epstein D., McIntosh H., McDaid C., Mason A., Golder S., O'Meara S., Sculpher M., Drummond M. and Forbes C., Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation. Health Technol Assess, 2005. 9(15): p. 1-157, iii-iv.

Willingham A. T., Orth A. P., Batalov S., Peters E. C., Wen B. G., Aza-Blanc P., Hogenesch J. B. and Schultz P. G., A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science, 2005. 309(5740): p. 1570-3.

Wilson J. F., Weale M. E., Smith A. C., Gratrix F., Fletcher B., Thomas M. G., Bradman N. and Goldstein D. B., Population genetic structure of variable drug response. Nat Genet, 2001. 29(3): p. 265-9.

Winter P.C., Hickey G.I. and Fletcher H.L. (1998), Population genetics and evolution. In Instant notes in Genetics. Oxford. BIOS Scientific publisher, 1999-251.

World Healthy Organization technical report series No. 524, Scientific Group on Pharmacogenetics (1973). Pharmacogenetics, report of a WHO Scientific group. Geneva, W.H.O.

Wright J., Pickard N., Whitfield A. and Hakin N., A population-based study of the prevalence, clinical characteristics and effect of ethnicity in epilepsy. Seizure, 2000. 9(5): p. 309-13.

Wrighton S. A., Stevens J. C., Becker G. W. and VandenBranden M., Isolation and characterization of human liver cytochrome P450 2C19: correlation between 2C19 and S-mephenytoin 4'-hydroxylation. Arch Biochem Biophys, 1993. 306(1): p. 240-5.

Xiao Z. S., Goldstein J. A., Xie H. G., Blaisdell J., Wang W., Jiang C. H., Yan F. X., He N., Huang S. L., Xu Z. H. and Zhou H. H., Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther, 1997. 281(1): p. 604-9.

Xie H. G., Genetic variations of S-mephenytoin 4'-hydroxylase (CYP2C19) in the Chinese population. Life Sci, 2000. 66(14): p. PL175-81.

Xie H. G., Huang S. L., Xu Z. H., Xiao Z. S., He N. and Zhou H. H., Evidence for the effect of gender on activity of (S)-mephenytoin 4'-hydroxylase (CYP2C19) in a Chinese population. Pharmacogenetics, 1997. 7(2): p. 115-9.

Xie H. G., Prasad H. C., Kim R. B. and Stein C. M., CYP2C9 allelic variants: ethnic distribution and functional significance. Adv Drug Deliv Rev, 2002. 54(10): p. 1257-70.

Xie H. G., Stein C. M., Kim R. B., Wilkinson G. R., Flockhart D. A. and Wood A. J., Allelic, genotypic and phenotypic distributions of S-mephenytoin 4'-hydroxylase (CYP2C19) in healthy Caucasian populations of European descent throughout the world. Pharmacogenetics, 1999. 9(5): p. 539-49.

Xu J., Turner A., Little J., Bleecker E. R. and Meyers D. A., Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? Hum Genet, 2002. 111(6): p. 573-4.

Yasar U., Lundgren S., Eliasson E., Bennet A., Wiman B., de Faire U. and Rane A., Linkage between the CYP2C8 and CYP2C9 genetic polymorphisms. Biochem Biophys Res Commun, 2002. 299(1): p. 25-8.

Zarza J., Hermida J., Montes R., Alberca I., Lopez M. L. and Rocha E., Leu208Val and Ile181Leu variants of cytochrome P450 CYP2C9 are not related to the acenocoumarol dose requirement in a Spanish population. Blood, 2002. 100(2): p. 734.

Zhang Y. and Benet L. Z., The gut as a barrier to drug absorption: combined role of cytochrome P450 3A and P-glycoprotein. Clin Pharmacokinet, 2001. 40(3): p. 159-68.

http://www.nigms.nih.gov/pharmacogenetics (accessed on 30th Apr. 2004)

http://www.hapmap.org/ (accessed on 10th Mar. 2006)

http://www.pharmgkb.org/index.jsp (accessed on 10th Mar. 2006).

http://www.ilae-epilepsy.org/ (accessed on 30th Apr. 2004).

http://www.who.int/mental_health/management/globaepilepsycampaign/en/ accessed on 30th Apr. 2004).

http://www.fda.gov/cber/gdlns/racethclin.htm (accessed on 30th Apr. 2004).

http://www.imm.ki.se/CYPalleles/ accessed on 30th Apr. 2004).

http://drnelson.utmem.edu/Nomenclature.html (accessed on 30th Apr. 2004).

http://www-hto.usc.edu/~cbmp/2001/SNP/index.html (accessed on 30th Apr. 2004)

http://wbiomed.curtin.edu.au/genepop/ (accessed on 30th Apr. 2004)

http://www.simfit.man.ac.uk/default.htm (accessed on 30th Apr. 2004)

Appendix 1a: List of studies included in the thesis for the population distribution of *CYP2C9* and *CYP2C19* polymorphism

A L. L., Dorado P., O'Kirwan F., Jepson R., Licinio J. and Wong M. L., Lower frequency of CYP2C9*2 in Mexican-Americans compared to Spaniards. Pharmacogenomics J, 2004. 4(6): p. 403-6.

Adithan C., Gerard N., Vasu S., Rosemary J., Shashindran C. H. and Krishnamoorthy R., Allele and genotype frequency of CYP2C19 in a Tamilian population. Br J Clin Pharmacol, 2003. 56(3): p. 331-3.

Aithal G. P., Day C. P., Kesteven P. J. and Daly A. K., Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet, 1999. 353(9154): p. 717-9.

Aithal G. P., Day C. P., Leathart J. B. and Daly A. K., Relationship of polymorphism in CYP2C9 to genetic susceptibility to diclofenac-induced hepatitis. Pharmacogenetics, 2000. 10(6): p. 511-8.

Aklillu E., Herrlin K., Gustafsson L. L., Bertilsson L. and Ingelman-Sundberg M., Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. Pharmacogenetics, 2002. 12(5): p. 375-83.

Allabi A. C., Gala J. L., Desager J. P., Heusterspreute M. and Horsmans Y., Genetic polymorphisms of CYP2C9 and CYP2C19 in the Beninese and Belgian populations. Br J Clin Pharmacol, 2003. 56(6): p. 653-7.

Aynacioglu A. S., Brockmoller J., Bauer S., Sachse C., Guzelbey P., Ongen Z., Nacak M. and Roots I., Frequency of cytochrome P450 CYP2C9 variants in a Turkish population and functional relevance for phenytoin. Br J Clin Pharmacol, 1999. 48(3): p. 409-15.

Aynacioglu A. S., Sachse C., Bozkurt A., Kortunay S., Nacak M., Schroder T., Kayaalp S. O., Roots I. and Brockmoller J., Low frequency of defective alleles of cytochrome P450 enzymes 2C19 and 2D6 in the Turkish population. Clin Pharmacol Ther, 1999. 66(2): p. 185-92.

Babaoglu M. O., Yasar U., Sandberg M., Eliasson E., Dahl M. L., Kayaalp S. O. and Bozkurt A., CYP2C9 genetic variants and losartan oxidation in a Turkish population. Eur J Clin Pharmacol, 2004.

Bathum L., Andersen-Ranberg K., Boldsen J., Brosen K. and Jeune B., Genotypes for the cytochrome P450 enzymes CYP2D6 and CYP2C19 in human longevitY. Role of CYP2D6 and CYP2C19 in longevity. Eur J Clin Pharmacol, 1998. 54(5): p. 427-30.

Bathum L., Skjelbo E., Mutabingwa T. K., Madsen H., Horder M. and Brosen K., Phenotypes and genotypes for CYP2D6 and CYP2C19 in a black Tanzanian population. Br J Clin Pharmacol, 1999. 48(3): p. 395-401.

Bozina N., Granic P., Lalic Z., Tramisak I., Lovric M. and Stavljenic-Rukavina A., Genetic polymorphisms of cytochromes P450: CYP2C9, CYP2C19, and CYP2D6 in Croatian population. Croat Med J, 2003. 44(4): p. 425-8.

Bramness J. G., Skurtveit S., Fauske L., Grung M., Molven A., Morland J. and Steen V. M., Association between blood carisoprodol:meprobamate concentration ratios and CYP2C19 genotype in carisoprodol-drugged drivers: decreased metabolic capacity in heterozygous CYP2C19*1/CYP2C19*2 subjects? Pharmacogenetics, 2003. 13(7): p. 383-8.

Bravo-Villalta H. V., Yamamoto K., Nakamura K., Baya A., Okada Y. and Horiuchi R., Genetic polymorphism of CYP2C9 and CYP2C19 in a Bolivian population: an investigative and comparative study. Eur J Clin Pharmacol, 2005. 61(3): p. 179-84.

Brockmoller J., Rost K. L., Gross D., Schenkel A. and Roots I., Phenotyping of CYP2C19 with enantiospecific HPLC-quantification of R- and S-mephenytoin and comparison with the intron4/exon5 G-->A-splice site mutation. Pharmacogenetics, 1995. 5(2): p. 80-8.

Brosen K., de Morais S. M., Meyer U. A. and Goldstein J. A., A multifamily study on the relationship between CYP2C19 genotype and s-mephenytoin oxidation phenotype. Pharmacogenetics, 1995. 5(5): p. 312-7.

Burian M., Grosch S., Tegeder I. and Geisslinger G., Validation of a new fluorogenic real-time PCR assay for detection of CYP2C9 allelic variants and CYP2C9 allelic distribution in a German population. Br J Clin Pharmacol, 2002. 54(5): p. 518-21.

Caraco Y., Muszkat M. and Wood A. J., Phenytoin metabolic ratio: a putative marker of CYP2C9 activity in vivo. Pharmacogenetics, 2001. 11(7): p. 587-96.

Chang M., Dahl M. L., Tybring G., Gotharson E. and Bertilsson L., Use of omeprazole as a probe drug for

CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and CYP2C19 genotype. Pharmacogenetics, 1995. 5(6): p. 358-63.

Chen K., Wang R., Wen S. Y., Li J. and Wang S. Q., Relationship of P450 2C9 genetic polymorphisms in Chinese and the pharmacokinetics of tolbutamide. J Clin Pharm Ther, 2005. 30(3): p. 241-9.

Dandara C., Masimirembwa C. M., Magimba A., Sayi J., Kaaya S., Sommers D. K., Snyman J. R. and Hasler J. A., Genetic polymorphism of CYP2D6 and CYP2C19 in east- and southern African populations including psychiatric patients. Eur J Clin Pharmacol, 2001. 57(1): p. 11-7.

de Morais S. M., Wilkinson G. R., Blaisdell J., Meyer U. A., Nakamura K. and Goldstein J. A., Identification of a new genetic defect responsible for the polymorphism of (S)-mephenytoin metabolism in Japanese. Mol Pharmacol, 1994. 46(4): p. 594-8.

de Morais S. M., Wilkinson G. R., Blaisdell J., Nakamura K., Meyer U. A. and Goldstein J. A., The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. J Biol Chem, 1994. 269(22): p. 15419-22.

de Morais S. M., Goldstein J. A., Xie H. G., Huang S. L., Lu Y. Q., Xia H., Xiao Z. S., Ile N. and Zhou H. H., Genetic analysis of the S-mephenytoin polymorphism in a Chinese population. Clin Pharmacol Ther, 1995. 58(4): p. 404-11.

Dickmann L. J., Rettie A. E., Kneller M. B., Kim R. B., Wood A. J., Stein C. M., Wilkinson G. R. and Schwarz U. I., Identification and functional characterization of a new CYP2C9 variant (CYP2C9*5) expressed among African Americans. Mol Pharmacol, 2001. 60(2): p. 382-7.

Dorado P., Berecz R., Norberto M. J., Yasar U., Dahl M. L. and A L. L., CYP2C9 genotypes and diclofenac metabolism in Spanish healthy volunteers. Eur J Clin Pharmacol, 2003. 59(3): p. 221-5.

Edeki T. I., Goldstein J. A., de Morais S. M., Hajiloo L., Butler M., Chapdelaine P. and Wilkinson G. R., Genetic polymorphism of S-mephenytoin 4'-hydroxylation in African-Americans. Pharmacogenetics, 1996. 6(4): p. 357-60.

Edeki T. I., Goldstein J. A., de Morais S. M., Hajiloo L., Butler M., Chapdelaine P. and Wilkinson G. R., Genetic polymorphism of S-mephenytoin 4'-hydroxylation in African-Americans. Pharmacogenetics, 1996. 6(4): p. 357-60.

Ferguson R. J., De Morais S. M., Benhamou S., Bouchardy C., Blaisdell J., Ibeanu G., Wilkinson G. R., Sarich T. C., Wright J. M., Dayer P. and Goldstein J. A., A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. J Pharmacol Exp Ther, 1998. 284(1): p. 356-61.

Freeman B. D., Zehnbauer B. A., McGrath S., Borecki I. and Buchman T. G., Cytochrome P450 polymorphisms are associated with reduced warfarin dose. Surgery, 2000. 128(2): p. 281-5.

Fu L. Q., Huang F., Wu D. Z. and Guo J. H., [Comparison of genetic polymorphism of cytochrome CYP2C19 between men and women in Chinese population]. Yao Xue Xue Bao, 2004. 39(3): p. 161-3.

Furuta T., Shirai N., Takashima M., Xiao F., Hanai H., Nakagawa K., Sugimura H., Ohashi K. and Ishizaki T., Effects of genotypic differences in CYP2C19 status on cure rates for Helicobacter pylori infection by dual therapy with rabeprazole plus amoxicillin. Pharmacogenetics, 2001. 11(4): p. 341-8.

Gaedigk A., Interethnic differences of drug-metabolizing enzymes. Int J Clin Pharmacol Ther, 2000. 38(2): p. 61-8.

Gaedigk A., Casley W. L., Tyndale R. F., Sellers E. M., Jurima-Romet M. and Leeder J. S., Cytochrome P4502C9 (CYP2C9) allele frequencies in Canadian Native Indian and Inuit populations. Can J Physiol Pharmacol, 2001. 79(10): p. 841-7.

Gaikovitch E. A., Cascorbi I., Mrozikiewicz P. M., Brockmoller J., Frotschl R., Kopke K., Gerloff T., Chernov J. N. and Roots I., Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. Eur J Clin Pharmacol, 2003. 59(4): p. 303-12.

Garcia-Barcelo M., Chow L. Y., Kum Chiu H. F., Wing Y. K., Shing Lee D. T., Lam K. L. and Waye M. M., Frequencies of defective CYP2C19 alleles in a Hong Kong Chinese population: detection of the rare allele CYP2C19*4. Clin Chem, 1999. 45(12): p. 2273-4.

Garcia-Martin E., Martinez C., Ladero J. M., Gamito F. J. and Agundez J. A., High frequency of mutations related to impaired CYP2C9 metabolism in a Caucasian population. Eur J Clin Pharmacol, 2001. 57(1): p.

47-9.

Garcia-Martin E., Martinez C., Tabares B., Frias J. and Agundez J. A., Interindividual variability in ibuprofen pharmacokinetics is related to interaction of cytochrome P450 2C8 and 2C9 amino acid polymorphisms. Clin Pharmacol Ther, 2004. 76(2): p. 119-27.

Goldstein J. A., Ishizaki T., Chiba K., de Morais S. M., Bell D., Krahn P. M. and Evans D. A., Frequencies of the defective CYP2C19 alleles responsible for the mephenytoin poor metabolizer phenotype in various Oriental, Caucasian, Saudi Arabian and American black populations. Pharmacogenetics, 1997. 7(1): p. 59-64.

Griese E. U., Ilett K. F., Kitteringham N. R., Eichelbaum M., Powell H., Spargo R. M., LeSouef P. N., Musk A. W. and Minchin R. F., Allele and genotype frequencies of polymorphic cytochromes P4502D6, 2C19 and 2E1 in aborigines from western Australia. Pharmacogenetics, 2001. 11(1): p. 69-76.

Hallberg P., Karlsson J., Kurland L., Lind L., Kahan T., Malmqvist K., Ohman K. P., Nystrom F. and Melhus H., The CYP2C9 genotype predicts the blood pressure response to irbesartan: results from the Swedish Irbesartan Left Ventricular Hypertrophy Investigation vs Atenolol (SILVHIA) trial. J Hypertens, 2002. 20(10): p. 2089-93.

Halling J., Petersen M. S., Damkier P., Nielsen F., Grandjean P., Weihe P., Lundgren S., Lundblad M. S. and Brosen K., Polymorphism of CYP2D6, CYP2C19, CYP2C9 and CYP2C8 in the Faroese population. Eur J Clin Pharmacol, 2005. 61(7): p. 491-7.

Hamdy S. I., Hiratsuka M., Narahara K., El-Enany M., Moursi N., Ahmed M. S. and Mizugaki M., Allele and genotype frequencies of polymorphic cytochromes P450 (CYP2C9, CYP2C19, CYP2E1) and dihydropyrimidine dehydrogenase (DPYD) in the Egyptian population. Br J Clin Pharmacol, 2002. 53(6): p. 596-603.

Hashimoto Y., Otsuki Y., Odani A., Takano M., Hattori H., Furusho K. and Iui K., Effect of CYP2C polymorphisms on the pharmacokinetics of phenytoin in Japanese patients with epilepsy. Biol Pharm Bull, 1996. 19(8): p. 1103-5.

He N., Yan F. X., Huang S. L., Wang W., Xiao Z. S., Liu Z. Q. and Zhou H. H., CYP2C19 genotype and S-mephenytoin 4'-hydroxylation phenotype in a Chinese Dai population. Eur J Clin Pharmacol, 2002. 58(1): p. 15-8.

Herman D., Locatelli I., Grabnar I., Peternel P., Stegnar M., Mrhar A., Breskvar K. and Dolzan V., Influence of CYP2C9 polymorphisms, demographic factors and concomitant drug therapy on warfarin metabolism and maintenance dose. Pharmacogenomics J, 2005. 5(3): p. 193-202.

Herrlin K., Massele A. Y., Jande M., Alm C., Tybring G., Abdi Y. A., Wennerholm A., Johansson I., Dahl M. L., Bertilsson L. and Gustafsson L. L., Bantu Tanzanians have a decreased capacity to metabolize omeprazole and mephenytoin in relation to their CYP2C19 genotype. Clin Pharmacol Ther, 1998. 64(4): p. 391-401.

Higashi M. K., Veenstra D. L., Kondo L. M., Wittkowsky A. K., Srinouanprachanh S. L., Farin F. M. and Rettie A. E., Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. Jama, 2002. 287(13): p. 1690-8.

Hillman M. A., Wilke R. A., Caldwell M. D., Berg R. L., Glurich I. and Burmester J. K., Relative impact of covariates in prescribing warfarin according to CYP2C9 genotype. Pharmacogenetics, 2004. 14(8): p. 539-47.

Ho G. T., Moodie F. M. and Satsangi J., Multidrug resistance 1 gene (P-glycoprotein 170): an important determinant in gastrointestinal disease? Gut, 2003. 52(5): p. 759-66.

Hoskins J. M., Shenfield G. M. and Gross A. S., Relationship between proguanil metabolic ratio and CYP2C19 genotype in a Caucasian population. Br J Clin Pharmacol, 1998. 46(5): p. 499-504.

Hu Y. R., Qiao H. L. and Kan Q. C., Pharmacokinetics of lansoprazole in Chinese healthy subjects in relation to CYP2C19 genotypes. Acta Pharmacol Sin, 2004. 25(8): p. 986-90.

Huang Y., Yang J. F., Qi X. L., Wang Y. Q., Wang W. Z. and Chen B., [Association between genetic polymorphisms of CYP2C19 and CYP2C9 and phenytoin serum concentration]. Zhonghua Yi Xue Za Zhi, 2004. 84(20): p. 1686-9.

Hummers-Pradier E., Hess S., Adham I. M., Papke T., Pieske B. and Kochen M. M., Determination of bleeding risk using genetic markers in patients taking phenprocoumon. Eur J Clin Pharmacol, 2003. 59(3): p. 213-9.

Hung C. C., Lin C. J., Chen C. C., Chang C. J. and Liou H. H., Dosage recommendation of phenytoin for

patients with epilepsy with different CYP2C9/CYP2C19 polymorphisms. Ther Drug Monit, 2004. 26(5): p. 534-40.

Ieiri I., Kubota T., Urae A., Kimura M., Wada Y., Mamiya K., Yoshioka S., Irie S., Amamoto T., Nakamura K., Nakano S. and Higuchi S., Pharmacokinetics of omeprazole (a substrate of CYP2C19) and comparison with two mutant alleles, C gamma P2C19m1 in exon 5 and C gamma P2C19m2 in exon 4, in Japanese subjects. Clin Pharmacol Ther, 1996. 59(6): p. 647-53.

Ieiri I., Mamiya K., Urae A., Wada Y., Kimura M., Irie S., Amamoto T., Kubota T., Yoshioka S., Nakamura K., Nakano S., Tashiro N. and Higuchi S., Stereoselective 4'-hydroxylation of phenytoin: relationship to (S)-mephenytoin polymorphism in Japanese. Br J Clin Pharmacol, 1997. 43(4): p. 441-5.

Inoue K., Yamazaki H., Imiya K., Akasaka S., Guengerich F. P. and Shimada T., Relationship between CYP2C9 and 2C19 genotypes and tolbutamide methyl hydroxylation and S-mephenytoin 4'-hydroxylation activities in livers of Japanese and Caucasian populations. Pharmacogenetics, 1997. 7(2): p. 103-13.

Inoue K., Yamazaki H. and Shimada T., Linkage between the distribution of mutations in the CYP2C18 and CYP2C19 genes in the Japanese and Caucasian. Xenobiotica, 1998. 28(4): p. 403-11.

Joffe H. V., Xu R., Johnson F. B., Longtine J., Kucher N. and Goldhaber S. Z., Warfarin dosing and cytochrome P450 2C9 polymorphisms. Thromb Haemost, 2004. 91(6): p. 1123-8.

Jose R., Chandrasekaran A., Sam S. S., Gerard N., Chanolean S., Abraham B. K., Satyanarayanamoorthy K., Peter A. and Rajagopal K., CYP2C9 and CYP2C19 genetic polymorphisms: frequencies in the south Indian population. Fundam Clin Pharmacol, 2005. 19(1): p. 101-5.

Jurima-Romet M., Goldstein J. A., LeBelle M., Aubin R. A., Foster B. C., Walop W. and Rode A., CYP2C19 genotyping and associated mephenytoin hydroxylation polymorphism in a Canadian Inuit population. Pharmacogenetics, 1996. 6(4): p. 329-39.

Kaneko A., Kaneko O., Taleo G., Bjorkman A. and Kobayakawa T., High frequencies of CYP2C19 mutations and poor metabolism of proguanil in Vanuatu. Lancet, 1997. 349(9056): p. 921-2.

Kaneko A., Lum J. K., Yaviong L., Takahashi N., Ishizaki T., Bertilsson L., Kobayakawa T. and Bjorkman A., High and variable frequencies of CYP2C19 mutations: medical consequences of poor drug metabolism in Vanuatu and other Pacific islands. Pharmacogenetics, 1999. 9(5): p. 581-90.

Kerb R., Aynacioglu A. S., Brockmoller J., Schlagenhaufer R., Bauer S., Szekeres T., Hamwi A., Fritzer-Szekeres M., Baumgartner C., Ongen H. Z., Guzelbey P., Roots I. and Brinkmann U., The predictive value of MDR1, CYP2C9, and CYP2C19 polymorphisms for phenytoin plasma levels. Pharmacogenomics J, 2001. 1(3): p. 204-10.

Kimura M., Ieiri I., Mamiya K., Urae A. and Higuchi S., Genetic polymorphism of cytochrome P450s, CYP2C19, and CYP2C9 in a Japanese population. Ther Drug Monit, 1998. 20(3): p. 243-7.

Kubota T., Chiba K. and Ishizaki T., Genotyping of S-mephenytoin 4'-hydroxylation in an extended Japanese population. Clin Pharmacol Ther, 1996. 60(6): p. 661-6.

Kubota T., Hibi N. and Chiba K., Linkage of mutant alleles of CYP2C18 and CYP2C19 in a Japanese population. Biochem Pharmacol, 1998. 55(12): p. 2039-42.

Lamba J. K., Dhiman R. K. and Kohli K. K., CYP2C19 genetic mutations in North Indians. Clin Pharmacol Ther, 2000. 68(3): p. 328-35.

Lee C. R., Goldstein J. A. and Pieper J. A., Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. Pharmacogenetics, 2002. 12(3): p. 251-63.

Lee S., Kim J. M., Chung C. S., Cho K. J. and Kim J. H., Polymorphism in CYP2C9 as a non-critical factor of warfarin dosage adjustment in Korean patients. Arch Pharm Res, 2003. 26(11): p. 967-72.

Lee S. S., Kim K. M., Thi-Le H., Yea S. S., Cha I. J. and Shin J. G., Genetic polymorphism of CYP2C9 in a Vietnamese Kinh population. Ther Drug Monit, 2005. 27(2): p. 208-10.

Linder M. W., Looney S., Adams J. E., 3rd, Johnson N., Antonino-Green D., Lacefield N., Bukaveckas B. L. and Valdes R., Jr., Warfarin dose adjustments based on CYP2C9 genetic polymorphisms. J Thromb Thrombolysis, 2002. 14(3): p. 227-32.

Loebstein R., Yonath H., Peleg D., Almog S., Rotenberg M., Lubetsky A., Roitelman J., Harats D., Halkin H. and Ezra D., Interindividual variability in sensitivity to warfarin--Nature or nurture? Clin Pharmacol Ther,

2001. 70(2): p. 159-64.

London S. J., Daly A. K., Leathart J. B., Navidi W. C. and Idle J. R., Lung cancer risk in relation to the CYP2C9*1/CYP2C9*2 genetic polymorphism among African-Americans and Caucasians in Los Angeles County, California. Pharmacogenetics, 1996. 6(6): p. 527-33.

Mamiya K., Ieiri I., Shimamoto J., Yukawa E., Imai J., Ninomiya H., Yamada H., Otsubo K., Higuchi S. and Tashiro N., The effects of genetic polymorphisms of CYP2C9 and CYP2C19 on phenytoin metabolism in Japanese adult patients with epilepsy: studies in stereoselective hydroxylation and population pharmacokinetics. Epilepsia, 1998. 39(12): p. 1317-23.

Marandi T., Dahl M. L., Rago L., Kiivet R. and Sjoqvist F., Debrisoquine and S-mephenytoin hydroxylation polymorphisms in a Russian population living in Estonia. Eur J Clin Pharmacol, 1997. 53(3-4): p. 257-60.

Marinac J. S., Balian J. D., Foxworth J. W., Willsie S. K., Daus J. C., Owen R. and Flockhart D. A., Determination of CYP2C19 phenotype in black Americans with omeprazole: correlation with genotype. Clin Pharmacol Ther, 1996. 60(2): p. 138-44.

Martin J. H., Begg E. J., Kennedy M. A., Roberts R. and Barclay M. L., Is cytochrome P450 2C9 genotype associated with NSAID gastric ulceration? Br J Clin Pharmacol, 2001. 51(6): p. 627-30.

Martinez C., Garcia-Martin E., Ladero J. M., Sastre J., Garcia-Gamito F., Diaz-Rubio M. and Agundez J. A., Association of CYP2C9 genotypes leading to high enzyme activity and colorectal cancer risk. Carcinogenesis, 2001. 22(8): p. 1323-6.

Masimirembwa C., Bertilsson L., Johansson I., Hasler J. A. and Ingelman-Sundberg M., Phenotyping and genotyping of S-mephenytoin hydroxylase (cytochrome P450 2C19) in a Shona population of Zimbabwe. Clin Pharmacol Ther, 1995. 57(6): p. 656-61.

Masta A., Lum J. K., Tsukahara T., Hwaihwanje I., Kaneko A., Paniu M. M., Sapuri M., Takahashi N., Ishizaki T., Kobayakawa T. and Hombhanje F. W., Analysis of Sepik populations of Papua New Guinea suggests an increase of CYP2C19 null allele frequencies during the colonization of Melanesia. Pharmacogenetics, 2003. 13(11): p. 697-700.

Nasu K., Kubota T. and Ishizaki T., Genetic analysis of CYP2C9 polymorphism in a Japanese population. Pharmacogenetics, 1997. 7(5): p. 405-9.

Nowak M. P., Sellers E. M. and Tyndale R. F., Canadian Native Indians exhibit unique CYP2A6 and CYP2C19 mutant allele frequencies. Clin Pharmacol Ther, 1998. 64(4): p. 378-83.

Odani A., Hashimoto Y., Otsuki Y., Uwai Y., Hattori H., Furusho K. and Inui K., Genetic polymorphism of the CYP2C subfamily and its effect on the pharmacokinetics of phenytoin in Japanese patients with epilepsy. Clin Pharmacol Ther, 1997. 62(3): p. 287-92.

Ogawa K., Suno M., Shimizu K., Yoshida M., Awaya T., Matsubara K. and Shiono H., Genotyping of cytochrome p450 isoform genes is useful for forensic identification of cadaver. Leg Med (Tokyo), 2003. 5(3): p. 132-8.

Ozawa S., Schoket B., McDaniel L. P., Tang Y. M., Ambrosone C. B., Kostic S., Vincze I. and Kadlubar F. F., Analyses of bronchial bulky DNA adduct levels and CYP2C9, GSTP1 and NQO1 genotypes in a Hungarian study population with pulmonary diseases. Carcinogenesis, 1999. 20(6): p. 991-5.

Pchelina S. N., Sirotkina O. V., Taraskina A. E., Vavilova T. V., Shwarzman A. L. and Schwartz E. I., The frequency of cytochrome P450 2C9 genetic variants in the Russian population and their associations with individual sensitivity to warfarin therapy. Thromb Res, 2005. 115(3): p. 199-203.

Pederson R. S., Verstuyft C., Becquemont L., Jaillon P. and Brosen K., Cytochrome P4502C9 (CYP2C9) genotypes in a Nordic population in Denmark. Basic Clin Pharmacol Toxicol, 2004. 94(3): p. 151-2.

Persson I., Aklillu E., Rodrigues F., Bertilsson L. and Ingelman-Sundberg M., S-mephenytoin hydroxylation phenotype and CYP2C19 genotype among Ethiopians. Pharmacogenetics, 1996. 6(6): p. 521-6.

Peyvandi F., Spreafico M., Siboni S. M., Moia M. and Mannucci P. M., CYP2C9 genotypes and dose requirements during the induction phase of oral anticoagulant therapy. Clin Pharmacol Ther, 2004. 75(3): p. 198-203.

Pirmohamed M., Alfirevic A., Vilar J., Stalford A., Wilkins E. G., Sim E. and Park B. K., Association analysis of drug metabolizing enzyme gene polymorphisms in HIV-positive patients with co-trimoxazole hypersensitivity. Pharmacogenetics, 2000. 10(8): p. 705-13.

Roddam P. L., Rollinson S., Kane E., Roman E., Moorman A., Cartwright R. and Morgan G. J., Poor metabolizers at the cytochrome P450 2D6 and 2C19 loci are at increased risk of developing adult acute leukaemia. Pharmacogenetics, 2000. 10(7): p. 605-15.

Roh H. K., Dahl M. L., Johansson I., Ingelman-Sundberg M., Cha Y. N. and Bertilsson L., Debrisoquine and S-mephenytoin hydroxylation phenotypes and genotypes in a Korean population. Pharmacogenetics, 1996. 6(5): p. 441-7.

Roh H. K., Dahl M. L., Tybring G., Yamada H., Cha Y. N. and Bertilsson L., CYP2C19 genotype and phenotype determined by omeprazole in a Korean population. Pharmacogenetics, 1996. 6(6): p. 547-51.

Ruas J. L. and Lechner M. C., Allele frequency of CYP2C19 in a Portuguese population. Pharmacogenetics, 1997. 7(4): p. 333-5.

Sandberg M., Johansson I., Christensen M., Rane A. and Eliasson E., The impact of CYP2C9 genetics and oral contraceptives on cytochrome P450 2C9 phenotype. Drug Metab Dispos, 2004. 32(5): p. 484-9.

Scordo M. G., Aklillu E., Yasar U., Dahl M. L., Spina E. and Ingelman-Sundberg M., Genetic polymorphism of cytochrome P450 2C9 in a Caucasian and a black African population. Br J Clin Pharmacol, 2001. 52(4): p. 447-50.

Scordo M. G., Pengo V., Spina E., Dahl M. L., Gusella M. and Padrini R., Influence of CYP2C9 and CYP2C19 genetic polymorphisms on warfarin maintenance dose and metabolic clearance. Clin Pharmacol Ther, 2002. 72(6): p. 702-10.

Scordo M. G., Caputi A. P., D'Arrigo C., Fava G. and Spina E., Allele and genotype frequencies of CYP2C9, CYP2C19 and CYP2D6 in an Italian population. Pharmacol Res, 2004. 50(2): p. 195-200.

Si D., Guo Y., Zhang Y., Yang L., Zhou H. and Zhong D., Identification of a novel variant CYP2C9 allele in Chinese. Pharmacogenetics, 2004. 14(7): p. 465-9.

Stubbins M. J., Harries L. W., Smith G., Tarbit M. H. and Wolf C. R., Genetic analysis of the human cytochrome P450 CYP2C9 locus. Pharmacogenetics, 1996. 6(5): p. 429-39.

Sullivan-Klose T. H., Ghanayem B. I., Bell D. A., Zhang Z. Y., Kaminsky L. S., Shenfield G. M., Miners J. O., Birkett D. J. and Goldstein J. A., The role of the CYP2C9-Leu359 allelic variant in the tolbutamide polymorphism. Pharmacogenetics, 1996. 6(4): p. 341-9.

Sviri S., Shpizen S., Leitersdorf E., Levy M. and Caraco Y., Phenotypic-genotypic analysis of CYP2C19 in the Jewish Israeli population. Clin Pharmacol Ther, 1999. 65(3): p. 275-82.

Tabrizi A. R., Zehnbauer B. A., Borecki I. B., McGrath S. D., Buchman T. G. and Freeman B. D., The frequency and effects of cytochrome P450 (CYP) 2C9 polymorphisms in patients receiving warfarin. J Am Coll Surg, 2002. 194(3): p. 267-73.

Takahashi H., Kashima T., Nomizo Y., Muramoto N., Shimizu T., Nasu K., Kubota T., Kimura S. and Echizen H., Metabolism of warfarin enantiomers in Japanese patients with heart disease having different CYP2C9 and CYP2C19 genotypes. Clin Pharmacol Ther, 1998. 63(5): p. 519-28.

Takahashi H., Wilkinson G. R., Caraco Y., Muszkat M., Kim R. B., Kashima T., Kimura S. and Echizen H., Population differences in S-warfarin metabolism between CYP2C9 genotype-matched Caucasian and Japanese patients. Clin Pharmacol Ther, 2003. 73(3): p. 253-63.

Takakubo F., Kuwano A. and Kondo I., Evidence that poor metabolizers of (S)-mephenytoin could be identified by haplotypes of CYP2C19 in Japanese. Pharmacogenetics, 1996. 6(3): p. 265-7.

Tamminga W. J., Wemer J., Oosterhuis B., de Zeeuw R. A., de Leij L. F. and Jonkman J. H., The prevalence of CYP2D6 and CYP2C19 genotypes in a population of healthy Dutch volunteers. Eur J Clin Pharmacol, 2001. 57(10): p. 717-22.

Tassaneeyakul W., Tawalee A., Kukongviriyapan V., Blaisdell J., Goldstein J. A. and Gaysornsiri D., Analysis of the CYP2C19 polymorphism in a North-eastern Thai population. Pharmacogenetics, 2002. 12(3): p. 221-5.

Thijssen H. H., Verkooijen I. W. and Frank H. L., The possession of the CYP2C9*3 allele is associated with low dose requirement of acenocoumarol. Pharmacogenetics, 2000. 10(8): p. 757-60.

Topic E., Stefanovic M. and Samardzija M., Association between the CYP2C9 polymorphism and the drug metabolism phenotype. Clin Chem Lab Med, 2004. 42(1): p. 72-8.

Tsuneoka Y., Fukushima K., Matsuo Y., Ichikawa Y. and Watanabe Y., Genotype analysis of the CYP2C19 gene in the Japanese population. Life Sci, 1996. 59(20): p. 1711-5.

van der Weide J., Steijns L. S., van Weelden M. J. and de Haan K., The effect of genetic polymorphism of cytochrome P450 CYP2C9 on phenytoin dose requirement. Pharmacogenetics, 2001. 11(4): p. 287-91.

Vianna-Jorge R., Perini J. A., Rondinelli E. and Suarez-Kurtz G., CYP2C9 genotypes and the pharmacokinetics of tenoxicam in Brazilians. Clin Pharmacol Ther, 2004. 76(1): p. 18-26.

Visser L. E., Schaik R. H., Vliet Mv M., Trienekens P. H., De Smet P. A., Vulto A. G., Hofman A., Van Duijn C. M. and Stricker B. H., The risk of bleeding complications in patients with cytochrome P450 CYP2C9*2 or CYP2C9*3 alleles on acenocoumarol or phenprocoumon. Thromb Haemost, 2004. 92(1): p. 61-6. a

Visser L. E., van Vliet M., van Schaik R. H., Kasbergen A. A., De Smet P. A., Vulto A. G., Hofman A., van Duijn C. M. and Stricker B. H., The risk of overanticoagulation in patients with cytochrome P450 CYP2C9*2 or CYP2C9*3 alleles on acenocoumarol or phenprocoumon. Pharmacogenetics, 2004. 14(1): p. 27-33. b

Wang S. L., Huang J., Lai M. D. and Tsai J. J., Detection of CYP2C9 polymorphism based on the polymerase chain reaction in Chinese. Pharmacogenetics, 1995. 5(1): p. 37-42.

Xiao Z. S., Goldstein J. A., Xie H. G., Blaisdell J., Wang W., Jiang C. H., Yan F. X., He N., Huang S. L., Xu Z. H. and Zhou H. H., Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther, 1997. 281(1): p. 604-9.

Xie H. G., Huang S. L., Xu Z. H., Xiao Z. S., He N. and Zhou H. H., Evidence for the effect of gender on activity of (S)-mephenytoin 4'-hydroxylase (CYP2C19) in a Chinese population. Pharmacogenetics, 1997. 7(2): p. 115-9.

Yamada H., Dahl M. L., Lannfelt L., Viitanen M., Winblad B. and Sjoqvist F., CYP2D6 and CYP2C19 genotypes in an elderly Swedish population. Eur J Clin Pharmacol, 1998. 54(6): p. 479-81.

Yamada S., Onda M., Kato S., Matsuda N., Matsuhisa T., Yamada N., Miki M. and Matsukura N., Genetic differences in CYP2C19 single nucleotide polymorphisms among four Asian populations. J Gastroenterol, 2001. 36(10): p. 669-72.

Yang J. Q., Morin S., Verstuyft C., Fan L. A., Zhang Y., Xu C. D., Barbu V., Jaillon P. and Becquemont L., Frequency of cytochrome P450 2C9 allelic variants in the Chinese and French populations. Fundam Clin Pharmacol, 2003. 17(3): p. 373-6.

Yao T. W., Zeng S., Wang T. W. and Chen S. Q., Phenotype analysis of cytochrome P450 2C19 in Chinese subjects with mephenytoin S/R enantiomeric ratio in urine measured by chiral GC. Biomed Chromatogr, 2001. 15(1): p. 9-13.

Yasar U., Eliasson E., Dahl M. L., Johansson I., Ingelman-Sundberg M. and Sjoqvist F., Validation of methods for CYP2C9 genotyping: frequencies of mutant alleles in a Swedish population. Biochem Biophys Res Commun, 1999. 254(3): p. 628-31.

Yasar U., Aklillu E., Canaparo R., Sandberg M., Sayi J., Roh H. K. and Wennerholm A., Analysis of CYP2C9*5 in Caucasian, Oriental and black-African populations. Eur J Clin Pharmacol, 2002. 58(8): p. 555-8.

Yasar U., Bennet A. M., Eliasson E., Lundgren S., Wiman B., De Faire U. and Rane A., Allelic variants of cytochromes P450 2C modify the risk for acute myocardial infarction. Pharmacogenetics, 2003. 13(12): p. 715-20.

Yilmaz N., Erbagci A. B. and Aynacioglu A. S., Cytochrome P4502C9 genotype in Southeast Anatolia and possible relation with some serum tumour markers and cytokines. Acta Biochim Pol, 2001. 48(3): p. 775-82.

Yoon Y. R., Shon J. H., Kim M. K., Lim Y. C., Lee H. R., Park J. Y., Cha I. J. and Shin J. G., Frequency of cytochrome P450 2C9 mutant alleles in a Korean population. Br J Clin Pharmacol, 2001. 51(3): p. 277-80.

Appendix 1b: List of original studies for the genotyping method of *CYP2C9/2C19* (corresponding to those notes in section 3.2 of chapter 3)

1. Sullivan-Klose, T.H., et al., The role of the CYP2C9-Leu359 allelic variant in the tolbutamide polymorphism. Pharmacogenetics, 1996. 6(4): p. 341-9.

2. Nasu, K., T. Kubota, and T. Ishizaki, Genetic analysis of CYP2C9 polymorphism in a Japanese population. Pharmacogenetics, 1997. 7(5): p. 405-9.

3. Yasar, U., et al., Validation of methods for CYP2C9 genotyping: frequencies of mutant alleles in a Swedish population. Biochem Biophys Res Commun, 1999. 254(3): p. 628-31.

4. Sviri, S., et al., Phenotypic-genotypic analysis of CYP2C19 in the Jewish Israeli population. Clin Pharmacol Ther, 1999. 65(3): p. 275-82.

5. Wang, S.L., et al., Detection of CYP2C9 polymorphism based on the polymerase chain reaction in Chinese. Pharmacogenetics, 1995. 5(1): p. 37-42.

6. Stubbins, M.J., et al., Genetic analysis of the human cytochrome P450 CYP2C9 locus. Pharmacogenetics, 1996. 6(5): p. 429-39.

7. Coller, J.K., A.A. Somogyi, and F. Bochner, Comparison of (S)-mephenytoin and proguanil oxidation in vitro: contribution of several CYP isoforms. Br J Clin Pharmacol, 1999. 48(2): p. 158-67.

8. Kristensen, V.N., et al., High-throughput methods for detection of genetic variation. Biotechniques, 2001. 30(2): p. 318-22, 324, 326 passim.

9. Ninomiya, H., et al., Genetic polymorphism of the CYP2C subfamily and excessive serum phenytoin concentration with central nervous system intoxication. Ther Drug Monit, 2000. 22(2): p. 230-2.

10. Veronese, M.E., et al., Site-directed mutation studies of human liver cytochrome P-450 isoenzymes in the CYP2C subfamily. Biochem J, 1993. 289 (Pt 2): p. 533-8.

11. London, S.J., et al., Genetic polymorphism of CYP2D6 and lung cancer risk in African-Americans and Caucasians in Los Angeles County. Carcinogenesis, 1997. 18(6): p. 1203-14.

12. Suno M, Awaya T, Ogawa K, Ohtaki K, Chiba K, Hayase N,Matsubara K. A wax-mediated hot-start polymerase chain reaction improves the genotyping of CYP2C9*1 and *3 alleles. Jpn J Clin Pharmacol Ther 2002;33:215–8.

13. Furuya, H., et al., Genetic polymorphism of CYP2C9 and its effect on warfarin maintenance dose requirement in patients undergoing anticoagulation therapy. Pharmacogenetics, 1995. 5(6): p. 389-92.

14. Gill, H.J., et al., The effect of genetic polymorphisms in CYP2C9 on sulphamethoxazole N-hydroxylation. Pharmacogenetics, 1999. 9(1): p. 43-53.

15. Steward, D.J., et al., Genetic association between sensitivity to warfarin and expression of CYP2C9*3. Pharmacogenetics, 1997. 7(5): p. 361-7.

A. De Morais, S.M., et al., Identification of a new genetic defect responsible for the polymorphism of (S)-mephenytoin metabolism in Japanese. Mol Pharmacol, 1994. 46(4): p. 594-8.

A'. de Morais, S.M., et al., The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. J Biol Chem, 1994. 269(22): p. 15419-22.

B. Persson, I., et al., S-mephenytoin hydroxylation phenotype and CYP2C19 genotype among Ethiopians. Pharmacogenetics, 1996. 6(6): p. 521-6.

C. Ferguson, R.J., et al., A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. J Pharmacol Exp Ther, 1998. 284(1): p. 356-61.

D. Ibeanu, G.C., et al., Identification of new human CYP2C19 alleles (CYP2C19*6 and CYP2C19*2B) in a Caucasian poor metabolizer of mephenytoin. J Pharmacol Exp Ther, 1998. 286(3): p. 1490-5.

E. Hersberger, M., et al., Two single-tube tetra-primer assays to detect the CYP2C19*2 and *3 alleles of S-mephenytoin hydroxylase. Clin Chem, 2001. 47(4): p. 772-4.

F. Xiao, Z.S., et al., Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther, 1997. 281(1): p. 604-9.

G. Brockmoller, J., et al., Phenotyping of CYP2C19 with enantiospecific HPLC-quantification of R- and S-mephenytoin and comparison with the intron4/exon5 G-->A-splice site mutation. Pharmacogenetics, 1995. 5(2): p. 80-8.

H. Bathum, L., et al., A dual label oligonucleotide ligation assay for detection of the CYP2C19*1, CYP2C19*2, and CYP2C19*3 alleles involving time-resolved fluorometry. Ther Drug Monit, 1998. 20(1): p. 1-6.

I. Goldstein, J.A. and J. Blaisdell, Genetic tests which identify the principal defects in CYP2C19 responsible for the polymorphism in mephenytoin metabolism. Methods Enzymol, 1996. 272: p. 210-8.

J. Kimura, M., et al., Genetic polymorphism of cytochrome P450s, CYP2C19, and CYP2C9 in a Japanese population. Ther Drug Monit, 1998. 20(3): p. 243-7.

K. Borlak, J. and T. Thum, Identification of major CYP2C9 and CYP2C19 polymorphisms by fluorescence resonance energy transfer analysis. Clin Chem, 2002. 48(9): p. 1592-4.

L. Ibeanu, G.C., et al., An additional defective allele, CYP2C19*5, contributes to the S-mephenytoin poor metabolizer phenotype in Caucasians. Pharmacogenetics, 1998. 8(2): p. 129-35.

M. Griese, E.U., F. Lapple, and M. Eichelbaum, Detection of CYP2C19 alleles *1, *2 and *3 by multiplex polymerase chain reaction. Pharmacogenetics, 1999. 9(3): p. 389-91.

N. Hiratsuka, M., et al., High throughput detection of drug-metabolizing enzyme polymorphisms by allele-specific fluorogenic 5' nuclease chain reaction assay. Biol Pharm Bull, 2000. 23(10): p. 1131-5.

O. Chang, M., et al., Use of omeprazole as a probe drug for CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and CYP2C19 genotype. Pharmacogenetics, 1995. 5(6): p. 358-63.

P. Goldstein, J.A., et al., Evidence that CYP2C19 is the major (S)-mephenytoin 4'-hydroxylase in humans. Biochemistry, 1994. 33(7): p. 1743-52.

Appendix 1c: List of studies included in the thesis for the association between phenytoin metabolism and *CYP2C9/2C19* polymorphism

Aynacioglu A. S., Brockmoller J., Bauer S., Sachse C., Guzelbey P., Ongen Z., Nacak M. and Roots I., Frequency of cytochrome P450 CYP2C9 variants in a Turkish population and functional relevance for phenytoin. Br J Clin Pharmacol, 1999. 48(3): p. 409-15.

Caraco Y., Muszkat M. and Wood A. J., Phenytoin metabolic ratio: a putative marker of CYP2C9 activity in vivo. Pharmacogenetics, 2001. 11(7): p. 587-96.

Hashimoto Y., Otsuki Y., Odani A., Takano M., Hattori H., Furusho K. and Iui K., Effect of CYP2C polymorphisms on the pharmacokinetics of phenytoin in Japanese patients with epilepsy. Biol Pharm Bull, 1996. 19(8): p. 1103-5.

Hung C. C., Lin C. J., Chen C. C., Chang C. J. and Liou H. H., Dosage recommendation of phenytoin for patients with epilepsy with different CYP2C9/CYP2C19 polymorphisms. Ther Drug Monit, 2004. 26(5): p. 534-40.

Kerb R., Aynacioglu A. S., Brockmoller J., Schlagenhaufer R., Bauer S., Szekeres T., Hamwi A., Fritzer-Szekeres M., Baumgartner C., Ongen H. Z., Guzelbey P., Roots I. and Brinkmann U., The predictive value of MDR1, CYP2C9, and CYP2C19 polymorphisms for phenytoin plasma levels. Pharmacogenomics J, 2001. 1(3): p. 204-10.

Mamiya K., Ieiri I., Shimamoto J., Yukawa E., Imai J., Ninomiya H., Yamada H., Otsubo K., Higuchi S. and Tashiro N., The effects of genetic polymorphisms of CYP2C9 and CYP2C19 on phenytoin metabolism in Japanese adult patients with epilepsy: studies in stereoselective hydroxylation and population pharmacokinetics. Epilepsia, 1998. 39(12): p. 1317-23.

Odani A., Hashimoto Y., Otsuki Y., Uwai Y., Hattori H., Furusho K. and Inui K., Genetic polymorphism of the CYP2C subfamily and its effect on the pharmacokinetics of phenytoin in Japanese patients with epilepsy. Clin Pharmacol Ther, 1997. 62(3): p. 287-92.

van der Weide J., Steijns L. S., van Weelden M. J. and de Haan K., The effect of genetic polymorphism of cytochrome P450 CYP2C9 on phenytoin dose requirement. Pharmacogenetics, 2001. 11(4): p. 287-91.

Appendix 1d: List of studies included in the thesis for the association between mephenytoin metabolic ratio and *CYP2C19* polymorphism

Aklillu E., Herrlin K., Gustafsson L. L., Bertilsson L. and Ingelman-Sundberg M., Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. Pharmacogenetics, 2002. 12(5): p. 375-83.

Brockmoller J., Rost K. L., Gross D., Schenkel A. and Roots I., Phenotyping of CYP2C19 with enantiospecific HPLC-quantification of R- and S-mephenytoin and comparison with the intron4/exon5 G-->A-splice site mutation. Pharmacogenetics, 1995. 5(2): p. 80-8.

Chang M., Dahl M. L., Tybring G., Gotharson E. and Bertilsson L., Use of omeprazole as a probe drug for CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and CYP2C19 genotype. Pharmacogenetics, 1995. 5(6): p. 358-63.

de Morais S. M., Goldstein J. A., Xie H. G., Huang S. L., Lu Y. Q., Xia H., Xiao Z. S., Ile N. and Zhou H. H., Genetic analysis of the S-mephenytoin polymorphism in a Chinese population. Clin Pharmacol Ther, 1995. 58(4): p. 404-11.

Edeki T. I., Goldstein J. A., de Morais S. M., Hajiloo L., Butler M., Chapdelaine P. and Wilkinson G. R., Genetic polymorphism of S-mephenytoin 4'-hydroxylation in African-Americans. Pharmacogenetics, 1996. 6(4): p. 357-60.

Ferguson R. J., De Morais S. M., Benhamou S., Bouchardy C., Blaisdell J., Ibeanu G., Wilkinson G. R., Sarich T. C., Wright J. M., Dayer P. and Goldstein J. A., A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. J Pharmacol Exp Ther, 1998. 284(1): p. 356-61.

Goldstein J. A., Ishizaki T., Chiba K., de Morais S. M., Bell D., Krahn P. M. and Evans D. A., Frequencies of the defective CYP2C19 alleles responsible for the mephenytoin poor metabolizer phenotype in various Oriental, Caucasian, Saudi Arabian and American black populations. Pharmacogenetics, 1997. 7(1): p. 59-64.

He N., Yan F. X., Huang S. L., Wang W., Xiao Z. S., Liu Z. Q. and Zhou H. H., CYP2C19 genotype and S-mephenytoin 4'-hydroxylation phenotype in a Chinese Dai population. Eur J Clin Pharmacol, 2002. 58(1): p. 15-8.

Herrlin K., Massele A. Y., Jande M., Alm C., Tybring G., Abdi Y. A., Wennerholm A., Johansson I., Dahl M. L., Bertilsson L. and Gustafsson L. L., Bantu Tanzanians have a decreased capacity to metabolize omeprazole and mephenytoin in relation to their CYP2C19 genotype. Clin Pharmacol Ther, 1998. 64(4): p. 391-401.

Jurima-Romet M., Goldstein J. A., LeBelle M., Aubin R. A., Foster B. C., Walop W. and Rode A., CYP2C19 genotyping and associated mephenytoin hydroxylation polymorphism in a Canadian Inuit population. Pharmacogenetics, 1996. 6(4): p. 329-39.

Kubota T., Chiba K. and Ishizaki T., Genotyping of S-mephenytoin 4'-hydroxylation in an extended Japanese population. Clin Pharmacol Ther, 1996. 60(6): p. 661-6.

Masimirembwa C., Bertilsson L., Johansson I., Hasler J. A. and Ingelman-Sundberg M., Phenotyping and genotyping of S-mephenytoin hydroxylase (cytochrome P450 2C19) in a Shona population of Zimbabwe. Clin Pharmacol Ther, 1995. 57(6): p. 656-61.

Persson I., Aklillu E., Rodrigues F., Bertilsson L. and Ingelman-Sundberg M., S-mephenytoin hydroxylation phenotype and CYP2C19 genotype among Ethiopians. Pharmacogenetics, 1996. 6(6): p. 521-6.

Sviri S., Shpizen S., Leitersdorf E., Levy M. and Caraco Y., Phenotypic-genotypic analysis of CYP2C19 in the Jewish Israeli population. Clin Pharmacol Ther, 1999. 65(3): p. 275-82.

Xiao Z. S., Goldstein J. A., Xie H. G., Blaisdell J., Wang W., Jiang C. H., Yan F. X., He N., Huang S. L., Xu Z. H. and Zhou H. H., Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare CYP2C19 mutant allele. J Pharmacol Exp Ther, 1997. 281(1): p. 604-9.

Xie H. G., Huang S. L., Xu Z. H., Xiao Z. S., He N. and Zhou H. H., Evidence for the effect of gender on activity of (S)-mephenytoin 4'-hydroxylase (CYP2C19) in a Chinese population. Pharmacogenetics, 1997. 7(2): p. 115-9.

Yao T. W., Zeng S., Wang T. W. and Chen S. Q., Phenotype analysis of cytochrome P450 2C19 in Chinese subjects with mephenytoin S/R enantiomeric ratio in urine measured by chiral GC. Biomed Chromatogr, 2001. 15(1): p. 9-13.

Appendix 1e: List of studies included in the thesis for the association between warfarin dose requirement and *CYP2C9* polymorphism

Aithal G. P., Day C. P., Kesteven P. J. and Daly A. K., Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet, 1999. 353(9154): p. 717-9.

Herman D., Locatelli I., Grabnar I., Peternel P., Stegnar M., Mrhar A., Breskvar K. and Dolzan V., Influence of CYP2C9 polymorphisms, demographic factors and concomitant drug therapy on warfarin metabolism and maintenance dose. Pharmacogenomics J, 2005. 5(3): p. 193-202.

Higashi M. K., Veenstra D. L., Kondo L. M., Wittkowsky A. K., Srinouanprachanh S. L., Farin F. M. and Rettie A. E., Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. Jama, 2002. 287(13): p. 1690-8.

Hillman M. A., Wilke R. A., Caldwell M. D., Berg R. L., Glurich I. and Burmester J. K., Relative impact of covariates in prescribing warfarin according to CYP2C9 genotype. Pharmacogenetics, 2004. 14(8): p. 539-47.

Joffe H. V., Xu R., Johnson F. B., Longtine J., Kucher N. and Goldhaber S. Z., Warfarin dosing and cytochrome P450 2C9 polymorphisms. Thromb Haemost, 2004. 91(6): p. 1123-8.

Kamali F., Khan T. I., King B. P., Frearson R., Kesteven P., Wood P., Daly A. K. and Wynne H., Contribution of age, body size, and CYP2C9 genotype to anticoagulant response to warfarin. Clin Pharmacol Ther, 2004. 75(3): p. 204-12.

Linder M. W., Looney S., Adams J. E., 3rd, Johnson N., Antonino-Green D., Lacefield N., Bukaveckas B. L. and Valdes R., Jr., Warfarin dose adjustments based on CYP2C9 genetic polymorphisms. J Thromb Thrombolysis, 2002. 14(3): p. 227-32.

Loebstein R., Yonath H., Peleg D., Almog S., Rotenberg M., Lubetsky A., Roitelman J., Harats D., Halkin H. and Ezra D., Interindividual variability in sensitivity to warfarin--Nature or nurture? Clin Pharmacol Ther, 2001. 70(2): p. 159-64.

Pchelina S. N., Sirotkina O. V., Taraskina A. E., Vavilova T. V., Shwarzman A. L. and Schwartz E. I., The frequency of cytochrome P450 2C9 genetic variants in the Russian population and their associations with individual sensitivity to warfarin therapy. Thromb Res, 2005. 115(3): p. 199-203.

Scordo M. G., Pengo V., Spina E., Dahl M. L., Gusella M. and Padrini R., Influence of CYP2C9 and CYP2C19 genetic polymorphisms on warfarin maintenance dose and metabolic clearance. Clin Pharmacol Ther, 2002. 72(6): p. 702-10.

Tabrizi A. R., Zehnbauer B. A., Borecki I. B., McGrath S. D., Buchman T. G. and Freeman B. D., The frequency and effects of cytochrome P450 (CYP) 2C9 polymorphisms in patients receiving warfarin. J Am Coll Surg, 2002. 194(3): p. 267-73.

Taube J., Halsall D. and Baglin T., Influence of cytochrome P-450 CYP2C9 polymorphisms on warfarin sensitivity and risk of over-anticoagulation in patients on long-term treatment. Blood, 2000. 96(5): p. 1816-9.

Topic E., Stefanovic M. and Samardzija M., Association between the CYP2C9 polymorphism and the drug metabolism phenotype. Clin Chem Lab Med, 2004. 42(1): p. 72-8.

Appendix 2: Results of Hardy-Weinberg equilibrium test using the GENEPOP program

Studies of CYP2C9 polymorphism of Caucasian, Chinese, Japanese and African populations:

Author	Ethnicity	Health status	Drug	HWE test result P-value	*CYP2C9* 1*	*CYP2C9* 2*	*CYP2C9* 3*	Allele numbers	Subject numbers
Aithal GP, 1999	Caucasian?	patients from anticoagulant clinic	warfarin	0.5272	0.798	0.106	0.096	104	52
Aithal GP, 2000	Caucasian	patient with diclofenac hepatotoxicity	diclofenac	1	0.854	0.104	0.042	48	24
Aithal GP, 2000	Caucasian	healthy	N/A	0.8144	0.785	0.11	0.105	200	100
Allabi A, 2003	Caucasian (Belgian White)	no certain	N/A	0.6085	0.825	0.1	0.075	240	121
Brockmoller J, 1995	Caucasian	healthy	mephenotype S-	1	0.866	0.134	0	254	127
Burian M, 2003	Caucasian (German)	no certain	N/A	0.703	0.809	0.14	0.051	236	118
Caraco, 2001	Caucasian	healthy	phenytoin	1	0.758	0.161	0.081	62	31
Chen Kun, 2005	Chinese (Han?)	healthy	tolbutamide	0.1546	0.967	0	0.033	338	169
Dickmann L, 2001	Caucasian (European American)	no certain	warfarin, diclofenac, Lauric Acid	0.5364	0.825	0.132	0.043	280	140
Dorado P, 2003	Caucasian (Spanish white)	healthy	diclofenac	0.4071	0.745	0.157	0.098	204	102
Gaedigk A, 2001	Caucasian (White)	patients	N/A	0.4194	0.78	0.149	0.071	650	325
Gaikovitch, 2003	Caucasian (Russian)	mixture: patients, healthy	N/A	0.9624	0.828	0.105	0.067	580	290
Garcia ME, 2001	Caucasian (White Spanish)	healthy	N/A	**0.0058**	0.694	0.143	0.162	314	157
Garcia ME, 2004	Caucasian (Spanish white)	healthy	ibuprofen	**0.0068**	0.704	0.188	0.108	260	130
Hallberg, 2002	Caucasian (White Swedish)	patient with hypertension and LVH	irbesartan atenolol	0.7772	0.814	0.127	0.059	204	102
Halling J, 2005	Caucasian (Nordic, Faroese)	no certain	Mephenytoin (Racemic)	0.1939	0.847	0.093	0.059	622	311
Herman D, 2005	Caucasian?	patients on warfarin maintenance therapy	warfarin	0.5571	0.785	0.112	0.104	376	188
Hillman WA, 2004	Caucasian	patient under warfarin treatment	warfarin	0.9851	0.807	0.119	0.074	906	453
Ho PC, 2003	Caucasians	no certain	valproic acid	0.8298	0.538	0.256	0.205	78	39
Hummers-P E, 2003	Caucasian (German)?	patient with anticoagulant treatment	phenprocoumon	0.6991	0.866	0.095	0.039	358	179
Inoue K, 1997 1998	Caucasian	no certain	tobutamide	1	0.856	0.111	0.033	90	45
Kamali F, 2004	Caucasian?	patient with stable warfarin dose requirement	warfarin	0.2224	0.798	0.128	0.074	242	121
London SJ, 1996	Caucasian (non-hispanic)	lung cancer	N/A	0.7702	0.85	0.15	0	354	177
London SJ, 1996	Caucasian (non-hispanic)	no certain	N/A	0.6031	0.9	0.1	0	922	461
Martin J H, 2001	Caucasian	Rheumatology out-patient	NSAIDs	1	0.848	0.087	0.065	46	23
Martin J H, 2001	Caucasian	patient with gastric ulcers	N/A	1	0.79	0.145	0.065	62	31

Author	Ethnicity	Health status	Drug	HWE test result P-value	*CYP2C9* 1*	*CYP2C9* 2*	*CYP2C9* 3*	Allele numbers	Subject numbers
Continuous...									
Martinez C, 2001	Caucasian (White Spanish)	healthy	N/A	**0.0097**	0.697	0.143	0.16	300	150
Martinez C, 2001	Caucasian (White Spanish)	patients	N/A	**0.0315**	0.767	0.132	0.101	258	129
Pchelina SN, 2005	Caucasian (Slavic, Russian)	healthy male, patient	warfarin	0.9819	0.824	0.111	0.065	596	298
Pchelina SN, 2005	Caucasian (Slavic, Russian)	patient	warfarin	0.7192	0.863	0.089	0.048	124	62
Pedersen RS, 2004	Caucasian (Danish)	healthy	N/A	**0.0002**	0.826	0.111	0.063	552	276
Peyvandi F, 2004	Caucasian (Italian)	patients with oral anticoagulant treatment	Warfarin	0.3051	0.784	0.132	0.084	250	125
Schalekamp T, 2004	Caucasian?	patient with anticoagulant treatment	phenprocoumon	0.6419	0.805	0.127	0.069	568	284
Schalekamp T, 2004	Caucasian (Duntch)	patient	acenocoumarol	0.2938	0.801	0.095	0.104	462	231
Scordo MG, 2001	Caucasians (Italian)	healthy	N/A	0.1122	0.796	0.108	0.096	314	157
Scordo MG, 2002	Caucasians (Italian)	patient with cardiovascular disease	warfarin	0.563	0.747	0.124	0.129	186	93
Scordo MG, 2004	Caucasians (Italian)	healthy	N/A	0.1334	0.778	0.125	0.097	720	360
Stubbins MJ, 1996	Caucasian	healthy	N/A	0.1845	0.79	0.125	0.085	200	100
Takahashi H, 2003	Caucasians	patients	warfarin	0.3833	0.723	0.223	0.053	94	47
Taube J, 2000	Caucasian?	patient with anticoagulant treatment	warfarin	0.3352	0.841	0.106	0.053	1122	561
Thijssen HH, 2000	Caucasian?	anticoagulated patients	acenocoumarol	1	0.771	0.114	0.114	70	35
Topic E, 2004	Caucasian (Croatians)	patient with thromboembolism	warfarin	**0.0016**	0.724	0.246	0.03	362	181
Topic E, 2004	Caucasian (Croatians)	healthy	warfarin	**0.0071**	0.808	0.186	0.006	354	177
van der Weide, 2001	Caucasian (Dutch)	patient: epilepsy	phenytoin	0.1508	0.767	0.142	0.092	120	60
van Dijk KN, 2004	Caucasian?	patient on stable acenocoumarol therapy receive one of the NSAIDs	acenocoumarol NSAID	0.6406	0.831	0.125	0.044	160	80
Verstuyft C, 2003	Caucasian?	case	acenocoumarol fluindione warfarin	0.0098	0.773	0.147	0.08	150	75
Verstuyft C, 2003	Caucasian?	control	acenocoumarol fluindione warfarin	0.7032	0.82	0.113	0.067	150	75
Visser LE, 2004	Caucasian	patient with anticoagulant treatment	acenocoumarol phenprocoumon	0.1616	0.825	0.135	0.04	2248	1124
Yang JQ, 2003	Caucasian (French)	healthy	N/A	0.2715	0.775	0.146	0.079	302	151
Yasar U, 1999	Caucasian (Swedish)	healthy	N/A	0.4482	0.819	0.107	0.074	860	430
Yasar U, 2003	Caucasian (Swedish)	control	N/A	0.3342	0.826	0.106	0.067	3006	1503
Yasar U, 2003	Caucasian (Swedish)	AMI	N/A	0.0606	0.8	0.125	0.075	2344	1172
Gaedigk A, 2001	Chinese	no certain	N/A	1	0.946	0	0.054	204	102
Huang Y, 2004	Chinese	epilepsy patient	phenytoin	1	0.938	0	0.062	64	32
Wang SL, 1995	Chinese Han	healthy	N/A	1	0.983	0.017	0	230	115
Yang JQ, 2003	Chinese Han	mixture: child patient, healthy child, healthy adult	N/A	0.4141	0.963	0.001	0.036	788	394
Inoue K, 1997 1998	Japanese	no certain	tobutamide	1	0.962	0.038	0	78	39
Mamiya K, 1998	Japanese	patient with epilepsy	phenytoin	1	0.989	0.011	0	268	134

Author	Ethnicity	Health status	Drug	HWE test result P-value	*CYP2C9* 1*	*CYP2C9* 2*	*CYP2C9* 3*	Allele numbers	Subject numbers
Continuous...									
Nasu K, 1997	Japanese	healthy	N/A	1	0.979	0.021	0	436	218
Ogawa K, 2003	Japanese	healthy	N/A	1	0.974	0.026	0	392	196
Takahashi H, 2003	Japanese	patients	warfarin	1	0.972	0.028	0	180	90
Kimura M, 1998	Japanese	N/A	N/A	1	0.982	0.018	0	280	140
Allabi A, 2003	African (Beninese)	healthy	N/A	N/A	1	0	0	202	111
Dickmann L, 2001	African (African American)	no certain	warfarin, diclofenac, lauric acid	1	0.961	0.026	0.013	232	120
London SJ, 1996	African (African American)	lung cancer	N/A	1	0.97	0.03	0	304	152
London SJ, 1996	African (African American)	no certain	N/A	1	0.964	0.036	0	478	239
Scordo MG, 2001	African (Ethiopian)	healthy	N/A	1	0.933	0.043	0.023	300	150

Studies of CYP2C9 polymorphism in the miscellaneous group

Close to	Phylogenetic tree	Author	Ethnicity	Healthy status	Drug	HWE test result P-value	*CYP2C9 *1*	*CYP2C9 *2*	*CYP2C9 *3*	Allele numbers	Subject numbers
Caucasian	West Asian	Jose R, 2004	Andhra Pradesh	healthy	N/A	0.7306	0.866	0.039	0.095	232	116
Caucasian	West Asian	Jose R, 2004	Karnataka	healthy	N/A	**0.0082**	0.855	0.064	0.082	220	110
Caucasian	West Asian	Jose R, 2004	Kerala	healthy	N/A	0.7212	0.9	0.021	0.079	240	120
Caucasian	European	Vianna JR, 2004	White (Brazil)	healthy	tenoxicam*	0.103	0.798	0.121	0.081	272	136
Caucasian	European	Sandberg M, 2004	no defined (Swedish?)	healthy	losartan	0.6364	0.798	0.099	0.103	252	126
Caucasian	West Asian	Hamdy, 2002	Egyptian	no certain	N/A	0.0512	0.818	0.119	0.063	494	247
Caucasian	West Asian	Aynacioglu AS, 1999	Turkish	outpatients, healthy	phenytoin	0.3725	0.794	0.106	0.1	998	499
Caucasian	West Asian	Babaoglu MO, 2004	Turkish	healthy	losartan	0.0643	0.812	0.1	0.088	170	85
Caucasian	West Asian	Kerb R, 2001	Turkish	healthy	phenytoin	0.071	0.812	0.099	0.089	192	96
Caucasian	West Asian	Yilmaz N, 2001	Turkish	healthy	serum tumour markers, cytokine	0.0594	0.789	0.102	0.109	128	64
Chinese	Southeast Asian	Lee SS, 2005	Vietnamese	healthy	N/A	1	0.978	0	0.022	314	157
Japanese	Northeast Asian	Lee S, 2003	Korean	patient	warfarin	N/A	1	0	0	180	90
Japanese	Northeast Asian	Yoon, 2001	Korean	epilepsy patients, healthy	N/A	1	0.989	0	0.011	1148	574
Negroid	African	Vianna JR, 2004	black (Brazil)	healthy	tenoxicam*	**0.0379**	0.922	0.045	0.032	154	77
Amerind	Amerind	Gaedigk A, 2001	Canadian Native Indian	healthy	N/A	1	0.912	0.031	0.057	228	114
Special	Arctic	Gaedigk A, 2001	Inuit	no certain	N/A	N/A	1	0	0	302	151
Special		Loebstein R, 2001	Israel?	patients with maintenance warfarin dose	warfarin	0.697	0.84	0.096	0.064	312	156

Close to	Phylogenetic tree	Author	Ethnicity	Healthy status	Drug	HWE test result P-value	*CYP2C9* *1	*CYP2C9* *2	*CYP2C9* *3	Allele numbers	Subject numbers
Continuous...											
Special		Llerena A, 2004	Mexican	healthy	N/A	1	0.862	0.082	0.056	196	98
Mixture	50% CNI	Gaedigk A, 2001	50% Canadian Native Indian	healthy	N/A	0.3983	0.854	0.083	0.062	48	24
Mixture	75% CNI	Gaedigk A, 2001	75% Canadian Native Indian	healthy	N/A	**0.0425**	0.767	0.133	0.1	30	15
Mixture	European & Amerind	Bravo-Villalta, 2005	Bolivian	healthy	N/A	0.5488	0.922	0.048	0.03	1556	778
Mixture	African & European	Vianna JR, 2004	Intermediate (Brazil)	healthy	tenoxicam*	1	0.86	0.072	0.068	236	118
subjects with different origin		Tabrizi AR, 2002	Mix (Arican-American and caucasians)	patients with stable warfarin dose	warfarin	0.9525	0.84	0.082	0.078	306	153
subjects with different origin		Freeman BD, 2000	Mix (White/African American)	cardiovascular inpatients under warfarin therapy	warfarin	0.6425	0.842	0.132	0.026	76	38
subjects with different origin		Pirmohamed M, 2000	Caucasian (93.2%)	HIV patients (without sulphamethoxazole hypersensitive)	co-trimoxazole	0.7663	0.837	0.112	0.051	178	89
subjects with different origin		Pirmohamed M, 2000	Caucasian (94.6%)	HIV patient (sulphamethoxazole hypersensitive)	co-trimoxazole	0.0741	0.75	0.134	0.116	112	56
subjects with different origin		Joffe HV, 2004	Mixture (89%Caucasian, 33%African-American, 3other races)	patients under warfarin therapy	warfarin	0.3418	0.747	0.151	0.103	146	73
subjects with different origin		Higashi MK, 2002	Mixture (96.2%Caucasian, 3.85 Hispanic)	patients under warfarin therapy	warfarin	**0.0038**	0.811	0.105	0.084	370	185
subjects with different origin		Linder MW, 2002	Mixture (98%Caucasians with 1 Hispanic, 2% Africa-American)	patients under stable anticoagulant therapy	warfarin	0.3096	0.759	0.179	0.062	112	56

Studies of CYP2C19 polymorphism of Caucasian, Chinese, Japanese and African populations:

Author	Ethnicity	Healthy status	Drug	HWE test result P-value	CYP2C19*1	*CYP2C19*2*	*CYP2C19*3*	Allele numbers	Subject numbers
Alliabi A, 2003	Caucasian (Belgian White)	no certain	N/A	0.2453	0.909	0.091	0	242	121
Aynacioglu S, 1999	Caucasian (German)	healthy	N/A	**0.041**	0.84	0.159	0.002	656	328
Bathum, 1998	Caucasian (Danish)	healthy	N/A	1	0.82	0.18	0	128	64
Bathum, 1998	Caucasian (Danish)	no certain	N/A	**0**	0.839	0.142	0.019	478	239
Bramness JG, 2003	Caucasian (Norwegian)	no certain	N/A	**0.0388**	0.846	0.154	0	188	94
Brockmoller J, 1995	Caucasian	healthy	mephenytoin (racemic)	0.5136	0.85	0.15	0	280	140
Brosen K, 1995	Caucasian (Danish)	healthy	mephenytoin	0.8932	0.463	0.519	0.019	108	54
Chang M, 1995	Cauasian (Swedish)	healthy	mephenytoin	0.1264	0.869	0.131	0	268	134
Chang M*, 1995	Caucasian (Swedish)	healthy	Omeprazole	0.3869	0.834	0.166	0	320	160
Ferguson RJ, 1998	Caucasian (French)	smoker	mephenytoin	**0.0049**	0.857	0.14	0.004	258	130
Gaikovitch, 2003	Caucasian (Russian)	patients, healthy	N/A	0.119	0.883	0.114	0.003	580	290
Halling J, 2005	Caucasian (Nordic, Faroese)	no certain	mephenytoin (Racemic)	0.7062	0.815	0.185	0	622	311
Hoskins, 1998	Caucasian	healthy	Proguanil	0.4283	0.854	0.146	0	198	99
Martin DE, 1998	Caucasian	healthy	N/A	0.2528	0.856	0.144	0	902	451
Roddam PL, 2000	Caucasian	healthy	N/A	**0.0153**	0.86	0.14	0	1904	952
Ruas JL, 1997	Caucasians (Portuguese)	healthy	N/A	1	0.869	0.131	0	306	153
Scordo M, 2002	Caucasian (Italian)	cariovascular outpatients	warfarin	0.6227	0.876	0.124	0	186	93
Scordo M, 2004	Caucasians (Italian)	healthy	N/A	0.4186	0.889	0.111	0	720	360
Yamada H, 1998	Caucasian (Swedish)	healthy elder	N/A	1	0.849	0.145	0.006	166	83
Yamada H, 1998	Caucasian (Swedish)	healthy	N/A	1	0.852	0.148	0	324	162
de Morais, 1994	Caucasian (American)	no certain	mephenytoin S-	0.0899	0.75	0.25	0	24	12
de Morais, 1994	Caucasian (Swiss)	no certain	mephenytoin S-	0.2979	0.5	0.5	0	30	15
de Morais, 1995	Chinese	healthy	mephenytoin S-	0.3182	0.541	0.399	0.061	148	74
Fu LQ, 2004	Chinese Han	healthy	N/A	0.8361	0.625	0.343	0.032	280	140
Garcia-Barcelo M, 1999	Chinese	depression patients, healthy	N/A	1	0.653	0.309	0.038	236	119
He N, 2002	Chinese Dai	healthy	mephenytoin	**0.006**	0.64	0.35	0.01	386	193
Huang Y, 2004	Chinese	epilepsy patient	phenytoin	0.902	0.609	0.312	0.078	64	32
Hung CC, 2004	Chinese (Taiwan)	epilepsy patients	phenytoin	**0.0085**	0.607	0.34	0.053	338	169
Xiao ZS, 1997	Chinese Bai*	healthy	mephenytoin	**0.0334**	0.689	0.259	0.052	402	202
Xiao ZS, 1997	Chinese Han*	healthy	mephenytoin	0.7343	0.559	0.366	0.074	202	101
Yamada S, 2001	Chinese	healthy	no certain	0.0751	0.512	0.442	0.045	242	121
Yao TW, 2001	Chinese	healthy	mephenytoin	0.1806	0.423	0.5	0.077	52	26
de Morais, 1994	Japanese	no certain	mephenytoin S-	0.0519	0.293	0.5	0.207	58	29

Author	Ethnicity	Healthy status	Drug	HWE test result P-value	CYP2C19*1	*CYP2C19*2*	*CYP2C19*3*	Allele numbers	Subject numbers
Continuous...									
Furuta T, 2001	Japanese	patient with *H.pylori*-positive gastritis	rabeprazole	0.898	0.589	0.282	0.129	202	101
Ieiri I, 1996	Japanese	healthy	omeprazole	0.1562	0.556	0.315	0.13	54	27
Ieiri I, 1997	Japanese	healthy	mephenytoin	0.0765	0.417	0.417	0.167	12	6
Kimura M, 1998	Japanese	no certain	N/A	0.1302	0.539	0.325	0.136	280	140
Kubota T, 1996	Japanese	healthy	mephenytoin (racemic)	0.7445	0.581	0.309	0.11	372	186
Kubota T, 1998	Japanese	healthy	N/A	0.8886	0.542	0.331	0.127	308	154
Mamiya K, 1998	Japanese	epilepsy patients	phenytoin	0.3867	0.649	0.246	0.104	268	134
Ogawa K, 2003	Japanese	healthy	N/A	0.5646	0.607	0.273	0.12	392	196
Takahashi H, 1998	Japanese	patient with heart disease	warfarin	0.6868	0.598	0.258	0.144	132	66
Takakubo F	Japanese	healthy	N/A	0.4725	0.618	0.267	0.115	434	217
Tsuneoka, 1996	Japanese	healthy	N/A	0.3634	0.664	0.219	0.117	128	64
Tsuneoka*, 1996	Japanese	patient with different disease	N/A	0.5705	0.562	0.331	0.107	338	169
Tsuneoka*, 1996	Japanese	healthy patients	N/A	0.1924	0.59	0.3	0.109	466	233
Yamada S	Japanese	healthy	no certain	0.7801	0.594	0.26	0.146	192	96
Edeki TI, 1996	African (African-American)	healthy	mephenytoin (racemic)	0.2806	0.809	0.191	0	152	76
Martin DE, 1998	African (African-American)	healthy	N/A	0.4537	0.835	0.163	0.002	466	233
Alliabi A, 2003	African (Beninese)	healthy	N/A	0.2082	0.869	0.131	0	222	111
Marinac JS, 1996	African (Black American)	healthy	omeprazole	1	0.84	0.16	0	200	100
Aklillu E, 2002	African (Ethiopian (Ethiopia))	healthy	mephenytoin (racemic)	**0.0082**	0.846	0.136	0.018	228	114
Aklillu E, 2002	African (Ethiopian (Sweden))	healthy	mephenytoin (racemic)	0.7674	0.85	0.121	0.029	140	70
Persson I, 1996	African (Ethiopians)	healthy	mephenytoin (racemic)	**0.0082**	0.846	0.136	0.018	228	114
Bathum, 1999	African (Tanzania)	healthy	mephenytoin (Racemic)	0.6693	0.907	0.093	0	388	194
Dandara, 2001	African (Tanzanian)	healthy	N/A	0.7397	0.821	0.179	0	212	106
Dandara, 2001	African (Tanzanian)	psychiatric patients	N/A	0.1893	0.808	0.186	0.006	172	86
Herrlin K, 1998	African (Tanzanian)	healthy	mephenytoin	0.3771	0.811	0.179	0.009	212	106
Herrlin K, 1998	African (Tanzanian)	healthy	omeprazole	0.4928	0.821	0.171	0.007	414	207
Dandara, 2001	African (Zimbabwean)	no certain	N/A	0.1383	0.869	0.131	0	168	84
Masimirembwa C, 1995	African (Zimbabweans)	healthy	mephenytoin	0.1383	0.869	0.131	0	168	84
Dandara, 2001	Venda	no certain	N/A	0.7402	0.783	0.217	0	152	76

Studies of CYP2C19 polymorphism in the miscellaneous group:

Close to	Phylogenetic tree	AUTHOR	Ethnicity	Healthy status	Drug	HWE test result P-value	CYP2C19 *1	CYP2C19 *2	CYP2C19 *3	Allele numbers	Subject numbers
Caucasian	West Asian	Jose R, 2004	Andhra Pradesh	healthy	N/A	0.5319	0.67	0.33	0	230	115
Caucasian	West Asian	Jose R, 2004	Karnataka	healthy	N/A	0.301	0.606	0.389	0.005	216	108
Caucasian	West Asian	Jose R, 2004	Kerala	healthy	N/A	0.8434	0.682	0.309	0.008	236	118
Caucasian	West Asian	Adithan, 2003	Tamilians	healthy	N/A	**0.021**	0.598	0.379	0.022	224	112
Caucasian	West Asian	Lamba, 1998 2000	north Indian	healthy	omeprazole	**0.0005**	0.74	0.223	0.037	242	121
Caucasian	West Asian	Hamdy, 2002	Egyptian	no certain	N/A	0.7759	0.889	0.109	0.002	494	247
Caucasian	West Asian	Aynacioglu S, 1999	Turkish	outpatients, healthy	N/A	0.6531	0.875	0.121	0.004	808	404
Caucasian	West Asian	Kerb R, 2001	Turkish	healthy	phenytoin	0.105	0.883	0.117	0	188	94
Chinese	Southeast Asian	Tassaneeyakul W, 2002	Thai	healthy	omeprazole	**0.0264**	0.71	0.266	0.023	214	107
Chinese	Southeast Asian	Yamada S, 2001	Thai	healthy	no certain	0.8039	0.607	0.347	0.045	242	121
Chinese	Southeast Asian	Yamada S, 2001	Vietnamese	healthy	no certain	**0.0491**	0.6	0.267	0.133	180	90
Japanese	Northeast Asian	Lee J, 2004	Korean	no certain	omeprazole	0.2273	0.638	0.285	0.076	564	282
Special	Arctic	Jurima-Romet M, 1999	Inuit	healthy	mephenytoin	0.4008	0.888	0.112	0	304	152
Special		Sviri S, 1999	Jewish	healthy	mephenytoin	0.8176	0.839	0.154	0.007	280	140
Special	Papua New Guinea	Masta A, 2003	Jawia	no certain	proguanil	0.1512	0.387	0.467	0.147	150	75
Special	Papua New Guinea	Masta A, 2003	Kiniambu	no certain	proguanil	**0.0221**	0.401	0.416	0.183	262	132
Special	Pacific Islander	Kaneko A, 1997	Malakula	no certain	N/A	0.1658	0.174	0.67	0.156	454	227
Special	Pacific Islander	Kaneko A, 1997	Tanna	no certain	N/A	0.7446	0.147	0.741	0.113	532	266
Special	Papua New Guinea	Masta A, 2003	Witupe	no certain	proguanil	0.3935	0.38	0.472	0.148	392	196
Special	Australian	Griese EU, 2001	Australian Aborigine	no certain	N/A	0.5143	0.502	0.355	0.143	454	227
Mixture	European & Amerind	Bravo-Villalta, 2005	Bolivian	healthy	N/A	0.1936	0.922	0.078	0.001	1556	778

Studies of CYP2C19 polymorphism of Pacific Islander in the study of Kaneko et al (1999):

Ethnicity	Healthy Status	HWE test result (P-value)	CYP2C19*1	CYP2C19*2	CYP2C19*3	Allele numbers	Subject numbers
A-Gaua	no certain	**0.0166**	0.283	0.588	0.129	798	399
B-Gaua	no certain	0.0839	0.252	0.523	0.225	298	149
C-Santo	no certain	0.7692	0.205	0.592	0.203	380	190
D-Maewo	no certain	**0.028**	0.262	0.592	0.146	446	223
E-Maewo	no certain	**0.012**	0.158	0.633	0.209	316	158
F-Santo	no certain	0.6511	0.116	0.676	0.207	516	258
G-Malo	no certain	0.9022	0.263	0.403	0.333	186	93
H-Pentecost	no certain	0.5465	0.242	0.637	0.121	1274	637
I-Malakula	no certain	0.5549	0.237	0.536	0.227	772	386
J-Malakula	no certain	0.274	0.177	0.672	0.151	192	96
K-Tongoa	no certain	0.6948	0.189	0.658	0.152	486	243
L-Emae	no certain	0.6155	0.199	0.645	0.156	366	183
M-Emae*	no certain	0.3229	0.213	0.624	0.163	178	89
N-Emae	no certain	0.485	0.19	0.738	0.071	168	84
O-Nguna	no certain	0.0551	0.197	0.474	0.329	234	117
P-Efate*	no certain	0.6189	0.255	0.605	0.14	550	275
Q-Ifira*	no certain	0.2794	0.218	0.597	0.185	330	165
R-Erromango	no certain	0.3151	0.115	0.768	0.118	908	454
S-Aniwa*	no certain	0.3205	0.286	0.633	0.081	692	346
T-Futuna*	no certain	0.7354	0.387	0.522	0.091	584	292
U-Tanna	no certain	0.6615	0.128	0.756	0.116	516	258
V-Tanna	no certain	0.3539	0.272	0.632	0.096	114	57
W-Aneityum*	no certain	0.2018	0.275	0.667	0.058	138	69
X-Aneityum	no certain	0.3852	0.196	0.784	0.021	634	317

Appendix 3: Results of population differentiation test using the GENEPOP program

Ethnicity	Study numbers	*CYP2C9*1*	*CYP2C9*2*	*CYP2C9*3*	Allele numbers	P-value *	SE	Other allele
Caucasian	14	4386	628	408	5416	0.0927	0.01666	
Chinese	3	1279	1	50	1330	0.612	0.00888	
Japanese	5	1334	0	32	1366	0.757	0.00546	
African	3	705	19	10	734	0.00148	0.00039	14

Ethnicity	Study numbers	*CYP2C19*1*	*CYP2C19*2*	*CYP2C9*3*	Allele numbers	P-value *	SE	Other allele
Caucasian	10	3608	577	3	4188	0.00036	0.00028	
Chinese	7	851	528	71	1450	0.00086	0.00041	
Japanese	9	1459	738	301	2498	0.393	0.02909	
African	11	2295	421	10	2728	0.009	0.00025	1

Note: * P-value: it is the probability of hypothesis that no differentiation exists between populations true. SE. standard error

Test was using Markov Chain method. Three Markov Chain parameters are demeorization (1000), batches (100) and iterations per batch (100).