The impact of the cytochrome CYP2C9*2 and *3 polymorphisms in the South African Caucasian population on warfarin therapy protocols

Pieter-Hendrik Green

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in the

Faculty of Health Sciences

Department of Chemical Pathology
University of Pretoria

2003

Table of contents

		Page
	ACKNOWLEDGEMENTS	I
	ABBREVIATIONS	II
	LIST OF TABLES	VI
	LIST OF FIGURES	IX
	ABSTRACT/ ABSTRAK	XIII
	Chapter 1:	
	Literature review.	
1.1	Introduction.	1
1.2	The Cytochrome P450 family of enzymes.	2
1.3	Pharmacogenetics and cytochrome P450.	5
1.4	Warfarin metabolism.	6
1.5	The CYP2C9 genes, enzymes and polymorphisms.	8
1.6	Prevalence of the CYP2C9*2 and *3 polymorphisms in different populations.	11
1.7	Detection of the CYP2C9*2 and CYP2C9*3 polymorphisms.	15
1.8	Nucleotide sequencing.	18
1.9	Screening of a population for a specific mutation.	20
1.10	The Founder effect and its impact on population screening in the South African population.	20
1.11	The aims of this study.	22
	Chapter 2:	
	Materials and Methods.	
2.1	Materials.	23
2.2	Subjects.	23
2.3	Isolation of DNA from frozen whole blood samples.	24

2.3.1	Determination of DNA concentration and quality.	25
2.3.1.1	Electrophoretic analysis of isolated DNA.	25
2.3.1.2	Spectrometric analysis of isolated DNA.	25
2.4	PCR strategy for the detection of the CYP2C9*2 polymorphism.	26
2.4.1	Optimization of the PCR methodology for the detection of the CYP2C9*2 polymorphism.	27
2.4.1.1	Optimization of $MgCl_2$ concentration.	27
2.4.1.2	Optimization of DNA concentration.	27
2.4.1.3	Optimization of primer concentration.	28
2.4.1.4	The effect of BSA, DMSO and glycerol on PCR amplification.	28
2.4.2	Amplification of the 190bp region containing the CYP2C9*2 polymorphism.	28
2.4.2.1	Ava II digestion of the CYP2C9*2, 190 bp product.	30
2.4.2.2	Polyacrylamide Gel Electrophoresis (PAGE) of PCR amplification products and restriction enzyme digests.	30
2.5	PCR strategy for the detection of the CYP2C9*3 polymorphism.	31
2.5.1	Optimization of PCR methodology for the detection of the CYP2C9*3 polymorphism.	32
2.5.1.1	Optimization of $MgCl_2$ concentration.	32
2.5.1.2	Optimization of DNA concentration.	33
2.5.1.3	Optimization of primer concentration.	33
2.5.1.4	The effect of BSA, DMSO and glycerol on PCR amplification.	33
2.5.2	Amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism.	33
2.5.2.1	Nsi I, restriction enzyme digestion and PAGE of the CYP2C9*3 digests.	34
2.6	The stability of a CYP2C9*2 and *3 PCR master mix after storage at -70°C.	35
2.7	Sequencing of CYP2C9*2 and CYP2C9*3 positive and negative controls.	35
2.7.1	Stability of the primers and DNA fragments used for sequencing.	37
2.7.2	The sequences of the 190 and the 165 bp fragments containing the CYP2C9*2 and *3 polymorphisms.	37
2.8	Population screening for the CYP2C9*2 and *3 polymorphisms in the	37

Control and the Anticoagulation Clinic population group.

2.9 Correlation between CYP2C9*2 and *3 genotypes and warfarin maintenance dosages. 38

Chapter 3:

Results.

3.1 Subjects. 39

3.2 Determination of DNA concentration and quality. 39

3.2.1 Electrophoretic analysis of isolated DNA. 39

3.2.2 Spectrometric analysis of isolated DNA. 40

3.3 Optimization of the PCR methodology for the detection of the CYP2C9*2 polymorphism. 41

3.3.1 Optimization of $MgCl_2$ concentration. 41

3.3.2 Optimization of DNA concentration. 42

3.3.3 Optimization of primer concentration. 43

3.3.4 The effect of BSA, DMSO and glycerol on PCR amplification. 44

3.4 Optimized parameters for the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism. 44

3.4.1 Restriction Enzyme Digestion, Polyacrylamide Gel Electrophoresis and Genotype Analysis for the CYP2C9*2 polymorphism. 45

3.5 Stability of the PCR reaction mix for the detection of the CYP2C9*2 polymorphism. 46

3.6 Reduction of PCR reaction volume for the detection of the CYP2C9*2 polymorphism. 47

3.7 Method development for screening CYP2C9*3 polymorphisms. 48

3.7.1 Optimization of PCR methodology for the detection of the CYP2C9*3 polymorphism. 48

3.7.1.1 Optimization of $MgCl_2$ concentration. 48

3.7.1.2 Optimization of DNA concentration. 49

3.7.1.3	Optimization of the primer concentration.	50
3.7.1.4	The effect of BSA, DMSO and glycerol on PCR amplification.	51
3.8	Optimized parameters for the amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism.	52
3.9	Restriction Enzyme Digestion, Polyacrylamide Gel Electrophoresis and Genotype Analysis for the CYP2C9*3 polymorphism.	52
3.10	Stability of the PCR reaction mix for the detection of the CYP2C9*3 polymorphism.	54
3.11	Reduction of PCR reaction volume for the detection of the CYP2C9*3 polymorphism.	55
3.12	Sequencing of CYP2C9*2 and CYP2C9*3 negative controls.	55
3.12.1	Preparation of samples for sequencing.	55
3.12.2	Stability of PCR product used for sequencing.	57
3.12.3	The sequences of the CYP2C9*2 and CYP2C9*3 polymorphisms.	57
3.13	Genotype distribution, gene and allele frequencies of the CYP2C9*2 and *3 alleles in a Control population.	59
3.13.1	Genotype distribution, gene and allele frequencies of the CYP2C9*2 allele in the Control population.	59
3.13.2	Genotype distribution, gene and allele frequencies of the CYP2C9*3 allele in the Control population.	60
3.13.3	Distribution of both CYP2C9*2 and *3 polymorphism in the Control population.	61
3.14	Genotype distribution, gene and allele frequencies of the CYP2C9*2 and *3 alleles in the Anticoagulation Clinic population.	62
3.14.1	Genotype distribution, gene and allele frequencies of the CYP2C9*2 allele in the Anticoagulation Clinic population.	62
3.14.2	Genotype distribution, gene and allele frequencies of the CYP2C9*3 allele in the Anticoagulation Clinic population.	63
3.14.3	Distribution of both CYP2C9*2 and *3 polymorphism in the Anticoagulation Clinic population.	64
3.15	Evaluation of the sensitivity and specificity of the CYP2C9*2 and *3	66

polymorphism screening methods.

3.16 Correlation between CYP2C9*2 and *3 genotypes and warfarin maintenance dosages. 66

Chapter 4:

Discussion.

4.1 Introduction. 68

4.2 Optimization of methods used for the detection of the CYP2C9*2 and *3 polymorphisms. 68

4.3 Population study. 75

4.3.1 The allele frequencies of the CYP2C9*2 and *3 polymorphisms in different normal populations. 76

4.3.2 Distribution of the various CYP2C9*2 and *3 genotypes in the British and South African Control and Anticoagulation clinic populations. 77

4.4 Recommendations for warfarin therapy. 80

4.5 Conclusion. 83

5. **References.** 84

Acknowledgements:

The Author wishes to thank:

Dr. Megan Bester, Department of Anatomy, University of Pretoria, supervisor of this project, for her guidance, patience and encouragement.

Prof. JB Ubbink, Department of Chemical Pathology, University of Pretoria, co-supervisor of this project, for his interest and advice.

Prof. WJH Vermaak, Head of the Department of Chemical Pathology, University of Pretoria, for making the laboratory available to carry out this study.

Dr. de Wet, Department of Haematology, University of Pretoria, for her clinical guidance and practical help.

His Family, for their love and encouragement, without their support this study would never have been completed.

ABOVE ALL, I WOULD LIKE TO THANK MY LORD AND SAVIOUR, JESUS CHRIST, FOR HIS GRACE THAT WAS SUFFICIENT TO COMPLETE THIS STUDY.

List of Abbreviations:

A,a	Adenine (in DNA sequence, indicating exon sequence)
Arg	Arginine
ARMS	Amplification Refractory Mutation System
BSA	Bovine serum albumin
Ava II	Restriction endonuclease isolated from *Anabaena variabilis*
bp	Base pair
BPB	Bromophenol Blue
BLAST	Basic Local Alignment Search Tool
C	Cytosine (in DNA sequence, indicating exon sequence)
C	Cysteine
°C	Degree centigrade
C	Carbon
%C	Percentage crosslinkage monomer
CYP	Cytochrome P450
Cys	Cysteine
ddH_2O	Double distilled water
dATP	2'-deoxyadenosine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
DNA	Deoxyribonucleic acid
DTT	Dithothreitol: threo-1,4-dimercapto-2,3-butanediol: $C_4H_{10}O_2S_2$
Da	Dalton
DMSO	Dimethyl Sulphoxide
DC	Direct current
EDTA	Ethylenediamine tetra-acetic acid: $C_{10}H_{16}N_2O_8$
EtBr	Ethidium bromide: $C_{21}H_{20}BrN_3$
ER	Endoplasmic Reticulum

G	Guanine (in DNA sequence, indicating exon sequence)
g	Gram
xg	Gravitational acceleration
HCl	Hydrochloric acid
HLA	Human leukocyte antigen
INR	International Normalized Ratio
I	Isoleucine
Ile	Isoleucine
KCl	Potassium chloride
kb	Kilo base
K_m	Michaelis constant
Ltd.	Limited
L	Leucine
Leu	Leucine
µ	Micro: 10^{-6}
µg	Microgram
mg	Milligram
m	Milli: 10^{-3}
M	Molar
ml	Milliliter
$MgCl_2$	Magnesium chloride
Mg^{2+}	Magnesium ions
M	Molar
Mmol/µl	Millimol per micro liter
mA	Milli Ampère
NaCl	Sodium chloride
nm	Nano meter
N	Nitrogen
n	Population size
Nsi I	Restriction endonuclease isolated from *Neisseria sicca*
NADPH	Nicotinamide Adenine Dinucleotide Phosphate (reduced)

O_2	Oxygen
Φ	phi
ρ	pico: 10^{-12}
ρmole	pico mole
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
rpm	Revolutions per minute
RFLP	Restriction fragment length polymorphism
R	Arginine
sec	Seconds
ssDNA	Single stranded DNA
S	Sulphur
SRS	Substrate recognition site
Taq DNA Polymerase	Deoxynucleosidetriphosphate: DNA deoxynucleotidyltransferase from Thermus aquaticus
T	Temperature
Thr	Threonine
T	Thymine (in DNA sequence, indicating exon sequence)
T_m	Melting temperature
TAG	Termination codon
↓	Indicates restriction enzyme cutting site
T_a	Annealing temperature
T_c	Calculated annealing temperature
TATA	Conserving non-coding DNA sequence in the 5' region of most eukaryotic gene
TBE	89.15 mM Tris [pH 8.0], 88.95 mM boric acid, 2.498 mM Na_2EDTA
Tris-HCl	2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride: $C_4H_{11}NO_3.H_2O$
Triton X-100	Triton X-100: octylphenolpoly(ethyleneglycolether)$_n$: $C_{34}H_{62}O_{11}$, for n=10
TEMED	N,N,N',N'-tetramethylethylenediamine: $C_6H_{16}N_2$

Tris	Tris: tris (hydroxymethyl) aminomethan: 2-Amino-2-(hydroxymethyl)-1,3-propanediol: $C_4H_{11}NO_3$
TBE	Tris/ Boric acid/ EDTA
TAE	Tris/ Acetate/ EDTA
%T	Total concentration polyacrylamide
U	Units
UK	United Kingdom
USA	United States of America
UV	Ultra Violet
V	Voltage
V_{max}	Enzymatic turnover number (function of time)
W	Watt
X^2	Chi-Square goodness of fit

List of tables:

Page

Chapter 1:

Literature review.

1.1	Classification of CYP2C9 Genotypes (41).	8
1.2	The effect of CYP2C9 genotypes on mean warfarin maintenance dose as mg per day (percentage reduction in comparison to prescribed doses for wild type genotype *1/*1) (43). N=size of the population studied.	10
1.3	Prevalence of CYP2C9 polymorphisms among different ethnic groups (43).	11
1.4a	The prevalence of the CYP2C9*2 and *3 polymorphisms in the community control group of North East England (4).	12
1.4b	The prevalence of the CYP2C9*2 and *3 polymorphisms in the random clinic control group of North East England (4).	13
1.4c	The prevalence of the CYP2C9*2 and *3 polymorphisms in the low dose clinic group of North East England (4).	13
1.5	The prevalence of the CYP2C9*2 and *3 polymorphisms in a clinic population of 561 patients with a target INR of 2.5 (6).	14

Chapter 2:

Materials and Methods.

2.1	Primer sequences for the detection of the CYP2C9*2 polymorphism (51).	28
2.2	Primer sequences for the detection of the CYP2C9*3 polymorphisms (41).	34

Chapter 3:

Results.

3.1	Nucleotide sequence for the 190bp DNA fragment containing the CYP2C9*2 polymorphism.	58
3.2	Nucleotide sequences for the 165bp DNA fragment containing the CYP2C9*3 polymorphism.	58
3.3a	Distribution of CYP2C9*2 polymorphism in the Control population (n=200).	59
3.3b	Chi-square for goodness of fit for the CYP2C9*2 in the Control population (n=200).	60
3.4a	Distribution of the CYP2C9*3 polymorphism in the Control population (n=200).	60
3.4b	Chi-square for goodness of fit for the CYP2C9*3 polymorphism in the Control population (n=200).	61
3.5a	Distribution of both the CYP2C9*2 and CYP2C9*3 polymorphisms in the Control population (n=200).	61
3.5b	Chi-square for goodness of fit for both the CYP2C9*2 and *3 polymorphisms in the Control population (n=200).	62
3.6a	Distribution of the CYP2C9*2 polymorphism in the Anticoagulation Clinic population (n=144).	63
3.6b	Chi-square for goodness of fit for the CYP2C9*2 polymorphism in the Anticoagulation Clinic population (n=144).	63
3.7a	Distribution of the CYP2C9*3 polymorphism in the Anticoagulation Clinic population (n=144).	64
3.7b	Chi-square for goodness of fit for the CYP2C9*3 polymorphism in the	64

Anticoagulation Clinic population (n=144).

3.8a Distribution of both the CYP2C9*2 and CYP2C9*3 polymorphisms in the Anticoagulation Clinic population (n= 144). 65

3.8b Chi-square for goodness of fit for both the CYP2C9*2 and *3 polymorphisms in the Anticoagulation Clinic population (n=144). 65

Chapter 4:

Discussion

4.1 The allele frequency of the CYP2C9*2 and *3 polymorphisms in different populations. 76

List of Figures:

Page

Chapter 1:

Literature review.

1.1 Schematic presentation of the nomenclature of the CYP450 family. 4

1.2 Model of the structure of the S-mephenytoin 4-hydroxylase enzyme (30) obtained from http://www.biozentrum.unibas.ch/schwede. 5

1.3 Indicating the molecular sites of action of Warfarin (32). 6

1.4 Sites of hydroxylation of R- and S-warfarin catalyzed by human cytochrome P450. Bold = major metabolic pathway (35). 7

1.5 Regions of CYP2C9 showing the positions of amino acid substitutions of the CYP2C9*2 and *3 polymorphisms. 8

Chapter 2:

Materials and Methods.

2.1 The predicted banding patterns of the digested PCR products for detection of individuals that are (i) Homozygous, normal (CYP2C9*2, -/-) (ii) Heterozygous (CYP2C9*2, +/-); and (iii) Homozygous (CYP2C9*2 +/+) for the CYP2C9*2 polymorphism. 26

2.2 The predicted banding patterns of the digested PCR product (165bp) for detection of individuals that are (i) Homozygous, normal (-/-) (ii) Heterozygous (+/-); and (iii) Homozygous (+/+) for the CYP2C9*3 polymorphism. 32

Chapter 3:

Results.

3.1 Electrophoretic analysis of DNA from the Control population stored for eighteen months at 4 °C and recently isolated Anticoagulation Clinic samples. Lane 1-5 40

DNA from the control population, Lane 6, No DNA and Lane 7-9 isolated from patients attending an Anticoagulation Clinic.

3.2 DNA standard curve used to determine the concentration of the isolated DNA samples. 41

3.3 The effect of different $MgCl_2$ concentrations on the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism. Lane 1: DNA size marker, Lane 2 and 3: 1 mM, Lane 4 and 5: 1.5 mM, Lane 6 and 7: 2 mM, Lane 8 and 9: 2.5 mM and Lane 10: 3.0 mM $MgCl_2$. 42

3.4 The effect of different DNA concentrations on the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism. Lane 1 and 2: 1.56 µg/µl DNA, Lane 3: DNA size marker, Lane 4 and 5: 3.12 µg/µl: Lane 6 and 7: 4.68 µg/µl, Lane 8 and 9: 6.25 µg/µl and Lane 10: 7.8 µg/µl DNA. 43

3.5 The effect of primer concentration on the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism. Lane 1: DNA size marker, Lane 2 and 3: 0.4 pmol/µl, Lane 4 and 5: 0.8 pmol/µl, Lane 6 and 7: 1.2 pmol/µl, Lane 8 and 9: 1.6 pmol/µl and Lane 2.0 pmol/µl of each of the primers. 44

3.6 CYP2C9*2 genotype analysis. Lane 1: DNA size marker, Lane 2 and 3: CYP2C9*2 heterozygous (+/-), Lane 4: CYP2C9*2 homozygous (+/+) and Lane 5: homozygous normal (-/-). 46

3.7 Effect of 6 days storage of premixed PCR reagents on the amplification of the CYP2C9*2 190 bp fragment. Lane 1: DNA size marker and Lane 2 –9: Position of the CYP2C9*2, 190 bp PCR product, indicated by arrow. 47

3.8 Effect of 11 days storage of premixed PCR reagents on the amplification of the CYP2C9*2 190 bp fragment. Lane 1: DNA size marker and Lane 2 –6: CYP2C9*2, 190 bp PCR product, indicated by arrow. 47

3.9 The effect of different $MgCl_2$ concentrations on the amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism. Lane 1 and 2 : 1 mM, Lane 3 DNA size marker, Lane 4 and 5: 1.5 mM, Lane 6 and 7: 2 mM, Lane 8 and 9: 2.5 mM and Lane 10: 3.0 mM $MgCl_2$. 49

3.10 The effect of different DNA concentrations on the amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism. Lane 1: 7 µl (11.4 µg/µl), Lane 2: 6 µl (9.75 µg/µl), Lane 3: 5 µl (8.1 µg/µl),, Lane 4: 4 µl (6.5 µg/µl), Lane 5: 3 µl (4.8 µg/µl), Lane 6: 2.5 µl (4.0 µg/µl), Lane 7: 2 µl (3.25 µg/µl), Lane 8: 1 µl (1.62 µg/µl) and Lane 9: DNA size marker. 50

3.11 The effect of primer concentration on the amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism. Lane 1: DNA size marker, Lane 2 and 3: 10 pmol (0.42 pmol/µl), Lane 4 and 5: 20 pmol (0.83 pmol/µl), Lane 6 and 7: 30 pmol (1.25 pmol/µl), Lane 8 and 9: 40 pmol (1.60 pmol/µl) and Lane 10, 50pmol (2.1 pmol/µl) of each primer in a volume of 24 µl. 51

3.12 The effect of the aduvant BSA on the amplification of the 165bp CYP2C9*3 fragment. Lane 1: DNA size marker, Lane 2-9: 0.2 µg/µl BSA. 52

3.13 CYP2C9*3 genotype analysis. Lane 1: DNA size marker; Lane 2: heterozygous (+/-); Lane 3-5: homozygous normal (-/-) and Lane 6: homozygous (+/+). 53

3.14 Stability of the CYP2C9*3, PCR mix. Lane 1-4 after 5 days storage at –70 ^{0}C, Lane 5 and 6 using freshly prepared PCR mix and Lane 7: DNA size marker. 54

3.15 Agarose gel electrophoresis of the 190bp PCR fragments containing the CYP2C9*2 polymorphism. Lane 1: No DNA, Lanes 2-5 and 7-11: Amplified PCR product and Lanes 6 and 12: DNA size marker. The same separation procedure was used to purify the 165bp fragments of the CYP2C9*3 polymorphism (not shown). 56

3.16 The PCR product formed using the CYP2C9*2, 190 bp fragment as substrate in a nested PCR. Lane 1: DNA size marker, Lane 2 and 3: Positive (CYP2C9*2 +/+) and Lane 4 and 5: Negative (CYP2C9*2 -/-) control. 56

3.17 PAGE of the PCR product formed with primers and fragments of the CYP2C9*3 polymorphism after 4 days of storage at room temperature. Lane 1: DNA size marker and Lane 2-10: PCR product of 165 bp for the CYP2C9*3 polymorphism. 57

3.18 The effect of the polymorphisms on warfarin maintenance dosage (mg/day) of 67

patients from the Anticoagulation Clinic population was determined (67). A mean daily warfarin dose of 4.85 mg (SD=3.71) was calculated for patients where both polymorphisms are absent compared to 3.71 (1.67) for patients where the CYP2C9 polymorphisms are present.

Abstract

Pharmacogenetics is a scientific field of study that focuses on biochemical variations in drug response due to genetic factors. The human cytochrome P450 (CYP) supergene family represents an important group of enzymes involved in the metabolism of exogenous substrates and xenobiotics. Genetic polymorphisms of the cytochrome P450 enzymes result in distinct subpopulations that differ in their ability to perform particular drug biotransformation reactions. One of these drugs is the anticoagulant drug warfarin, which is used for the prevention of thromboembolic disorders. Treatment is not without complications as inter-individual differences in the response to a given dose of warfarin are observed, the International Normalized Ratio (INR) is calculated to monitor treatment.

Warfarin (Coumadin, 4-hydroxy-3-(3-oxo-1-phenylbutyl)2H-1-benzopyran-2-one) is a racemic drug, and the metabolic inactivation of the more potent S-enantomere is regulated mainly by CYP2C9 that stereospecifically converts S-warfarin to metabolically inactive S-7-hydroxywarfarin. Point mutations in the CYP2C9 gene result in the expression of two allelic variants CYP2C9*2 and CYP2C9*3. Individuals where either allele is present tend to have decreased clearance of S-warfarin and heightened therapeutic response to normal dosages.

No information is available regarding the prevalence of these polymorphisms in the South African Caucasian population. Determination of the prevalence of these polymorphisms in the South African population is indicated. As a founder effect associated with population and non-population specific mutations has been observed for a number of genetic disorders within the South African Caucasian population. For the detection of the CYP2C9*2 and CYP2C9*3 polymorphisms, Restriction Fragment Length Polymorphism (RFLP) screening methodologies were established. These methods were used to screen two populations: a random clinic Control and an Anticoagulation Clinic population. In the Anticoagulation Clinic population group were divided into two sub-groups, those with a CYP2C9 polymorphism and those without. A correlation was

made between the presence of a CYP2C9 polymorphism and mean warfarin dosage. In the Control population (n=200) the prevalence of the CYP2C9*2 polymorphism was 20% heterozygous and 5% homozygous. This corresponds with a genotype frequency of 0.023 and a carrier frequency of 0.26. For the CYP2C9*3 polymorphism 19% were heterozygous and 0.5% homozygous. The prevalence of compound heterozygosity was 5.5%. For the CYP2C9*3 polymorphism the genotype and allele frequency was 0.01 and 0.18 respectively.

In the Anticoagulation Clinic population (n=144) 20.1% were CYP2C9*2 heterozygous and 0.7% were homozygous, reflecting a genotype frequency of 0.016 and a carrier frequency of 0.19. The prevalence of compound heterozygosity was 1.4%. The CYP2C9*3 genotype was found in 10.4% of the Anticoagulation Clinic patients. No homozygous individuals were found. This correlates with a genotype frequency of 0.003 and a carrier frequency of 0.1.

The South African Caucasian population is of Dutch, French and English descent. No information is available regarding the prevalence of these polymorphisms in the Dutch and French populations. The prevalence of the CYP2C9*2 and *3 polymorphisms is similar to that reported by Aithal *et al.* In the Anticoagulation Clinic population the presence of at least one CYP2C9 polymorphic allele was associated with a significant lower mean warfarin dosage (4.85 vs 3.71 mg/day).

In conclusion the CYP2C9*2 and *3 polymorphisms are found in the South African Caucasian population at a prevalence similar to the British population. The presence of one CYP2C9 polymorphic allele is associated with altered warfarin metabolism. Therefore in the South African Caucasian population warfarin treatment strategies should be adjusted and include patient screening before warfarin treatment strategies are implemented.

Abstrak

Farmakogenetika is 'n studieveld wat fokus op genetiese verklarings vir biochemiese variasies in die metabolisering van medikasie. Die Sitochroom P450 (CYP) supergeen familie speel 'n belangrike rol in die metabolisering van eksogene substrate en sekere farmaseutiese middels. Genetiese polimorfismes in die CYP familie lei tot variasie in die vermoë van 'n individu om bepaalde medikasie te metaboliseer. 'n Voorbeeld van so 'n produk is die antistol middel warfarin. Bloeding is die bekendste newe-effek van die middel en pasiënte se INR moet gereeld gekontroleer word om bloeding te verhoed.

Warfarin bestaan as twee enantiomere bekend as R- en S-warfarin. Die S-vorm het 'n baie groter effek op die stollingstyd en word gemetaboliseer deur ensieme van die CYP superfamilie na die onaktiewe vorm S-7-hidroksiwarfarin. Die CYP2C9 geenproduk is die belangrikste ensiem betrokke by die omskakeling. Puntmutasies in die geen gee aanleiding tot twee polimorfismes bekend as CYP2C9*2 en CYP2C9*3. Individue waar die polimorfismes in een of beide alele voorkom het 'n laer tempo van metabolisering van warfarin en dienooreenkomstig 'n verhoogde reaksie op normale dosisse.

Die Suid-Afrikaanse blanke populasie is van Europese herkoms. Van die mutasies wat in die Europese bevolkings beskryf is, kom ook in die Suid-Afrikaanse bevolking voor. Die voorkoms van die CYP2C9 polimorfismes is nog nie vasgestel vir die Suid-Afrikaanse blanke bevolking nie. Polimerase Ketting Reaksie (PKR) metodes is ontwikkel om die voorkoms daarvan in die Suid-Afrikaanse bevolking te bepaal. Die voorkoms is vasgestel in twee afsonderlike groepe. Die eerste groep het bestaan uit 200 blanke mediese studente wat die normale Kontrole groep vir die Suid-Afrikaanse blanke bevolking verteenwoordig. Die tweede groep is saamgestel uit die Antistol kliniek by die Pretoria Akademiese Hospitaal (144 individue). Die Antistol kliniek pasiënte is verder verdeel in twee groepe, 'n groep waar een of beide polimorfismes teenwoordig is en 'n groep waar beide afwesig is. Die gemiddelde daaglikse warfarin dosis is bereken vir beide groepe, om vas te stel of die polimorfismes dit beïnvloed.

Van die 200 individue in die bevolkings kontrole groep was 20% heterosigoties en 5% homosigoties vir die CYP2C9*2 polimorfisme. 'n Geenfrekwensie van 0.023 en 'n draerfrekwensie van 0.26 is bereken. Die CYP2C9*3 polimorfisme het heterosigoties voorgekom in 19% en homosigoties in 0.5% van die individue. Die frekensies kan uitgedruk word as 'n aleelfrekwensie van 0.18 en 'n geenfrekwensie van 0.01. Van die groep het 5.5% van die individue 'n kopie van beide polimorfismes bevat.

In die Antistolkliniek was 20.1% heterosigoties en 0.7% homosigoties vir die CYP2C9*2 polimorfisme, dit dui op 'n geenfrekwensie van 0.016 en 'n draerfrekwensie van 0.19. Slegs 1.4% van die individue het 'n kopie van beide polimorfismes bevat. Die CYP2C9*3 polimofisme is aangetref in 10.4% (heterosigoties) geen homosigote is gevind nie. Uit die data is 'n geenfrekwensie van 0.003 en 'n draerfrekwensie van 0.1 bereken.

Die Suid-Afrikaanse bevolking is van Nederlandse-, Franse- en Britse afkoms. Die voorkoms van die polimorfismes is onbekend vir die Nederlandse en Franse bevolkings; en soorgelyk aan die bevindings van Aithal vir die Britse bevolking.

In die Antistol kliniek is die teenwoordigheid van een of meer van die polimorfismes geassosieer met verlaagde daaglikse warfarin dosis. In die groep waar die polimorfismes afwesig was, was die gemiddelde daaglikse dosis 4.85mg/dag teenoor die 3.71mg/dag in die teenwoordigheid van een of meer kopieë van die polimorfismes.

Ter samevatting, die voorkoms van die CYP2C9*2 en CYP2C9*3 polimorfismes vergelyk goed met diè van die Britse bevolking. Dit is duidelik dat die teenwoordigheid van een of beide polimorfismes verband hou met verlaagde dosisse van warfarin. Die bevindings van hierdie studie dui daarop dat dit sinvol sal wees om pasiënte te toets vir hierdie polimorfismes voordat daar met warfarin behandeling begin word.

Chapter 1:

Literature review.

1.1 Introduction.

The human cytochrome P450 (CYP) supergene family represents an important group of enzymes involved in the metabolism of exogenous substrates and xenobiotics (1,2). Genetic polymorphisms of the P450 enzymes result in distinct patient sub-populations that differ in their ability to perform particular drug biotransformation reactions (3,4,5,6). The anticoagulant drug warfarin, is widely used to control and prevent thromboembolic disorders. Inter-individual differences in the response to a given dose of warfarin requires careful monitoring of the International Normalized Ratio (INR) to prevent bleeding complications (7,8).

Warfarin (Coumadin, 4-hydroxy-3-(3-oxo-1-phenylbutyl)2H-1-benzopyran-2-one (9)) is a racemic drug (10), and the metabolic inactivation of the more potent S-enantomere is controlled by CYP2C9 isoenzymes that stereospecifically converts S-warfarin to metabolically inactive S-7-hydroxywarfarin and S-6-hydroxywarfarin (4,5,6,10). Point mutations in the CYP2C9 gene results in the expression of a number of allelic variants including CYP2C9*2 and CYP2C9*3 (4,5,6). Individuals that are homozygous for either polymorphisms or compound heterozygous for both alleles, present with decreased clearance of S-warfarin and heightened therapeutic response to normal dosages (1,4,5,6,11,12).

The prevalence of the CYP2C9*2 and the CYP2C9*3 polymorphisms in white European populations varies from 8-38% and 7-19% respectively (1,13). The frequency of homozygous for CYP2C9*2 in the Caucasian population is about 20% (14,15) and only 1% in the African-American population (16). The frequency of homozygosity for CYP2C9*3 in the Caucasian population is 8.5% and is rare in the Chinese and Japanese populations (15,16).

An increased prevalence of homozygotes or compound heterozygotes for CYP2C9*2 and CYP2C9*3 was observed for patients on a low warfarin dose when compared to a random clinic control group (4,16,17). Furthermore a significant proportion of these patients

experienced difficulties at the time of induction of therapy, and had increased bleeding complications when compared to clinic controls (4). Freeman *et al* (5) tested 38 patients receiving chronic warfarin therapy, or undergoing the initiation of warfarin therapy. The CYP2C9*2 and CYP2C9*3 alleles were found in 34% of the patients tested. In addition the presence of CYP2C9*2, CYP2C9*3 or CYP2A6*2 alleles were associated with a significant reduction in the weekly warfarin dose (1,6,18,19,13). Taube *et al* (6) also observed that the maintenance dose of warfarin was significantly related to genotype where mean maintenance dosages for patients with the CYP2C9*2 and CYP2C9*3 polymorphisms were reduced when compared to wild-type (CYP2C9*1) patients.

The South African Caucasian population is of European descent and a founder effect has been observed within this population for a number of genetic mutations. This includes an increased prevalence of known mutations (e.g. Hereditary Haemochromatosis (20)) or of mutations specific (e.g Porphyria Variegate (21)) for the South African Caucasian population. For this reason it is important to determine the prevalence of the CYP2C9*2 and the CYP2C9*3 polymorphisms in the South African population before a diagnostic service can be implemented to test for these polymorphisms.

The objectives of this study is to establish and optimise screening methodologies for both the CYP2C9*2 and CYP2C9*3 polymorphisms. These procedures will be performed in a South African Caucasian random Control and Anticoagulation clinic population. The presence of a CYP2C9 polymorphism will be correlated with mean warfarin dosage. The gene and allele frequency as well as the prevalence of compound heterozygosity will be determined. This study will provide important information that will be used in the implementation of testing strategies for patients that do not conform to the standard protocols of warfarin therapy.

1. 2 The Cytochrome P450 family of enzymes.

The cytochrome P450 (CYP) super family is a highly diversified set of heme containing proteins. These proteins were discovered in 1958 by their unusual reduced carbon monoxide difference spectrum that has an absorbance at 450 nm, thus Pigment at 450 nm or P450 (2). This increase in absorbance is due to a thiolate anion acting as the 5th ligand to the heme. The most common reaction catalyzed by P450 is hydroxylation, often of a lipophilic

substrate. Consequently, this family of proteins are frequently called hydroxylases, but P450 proteins can perform a wide spectrum of reactions including N-oxidation, sulfoxidation, epoxidation, N-, S-, and O-dealkylation, peroxidation, deamination, desulfuration and dehalogenation (22). In bacteria these proteins are soluble and consist of approximately 400 amino acids while in eukaryotics the P450s are larger, about 500 amino acids and are usually membrane bound through an N-terminal hydrophobic peptide (23). The eukaryotic P450 proteins are found in the endoplasmic reticulum (ER) membrane and the mitochondrial inner membrane (24).

Cytochrome P450s are sometimes called mixed function oxidases or monooxygenases due to the manner in which molecular oxygen is incorporated into product. In the usual hydroxylation, one atom of oxygen is added to the substrate and the other contributes to forming a water molecule (22).

$$\mathbf{R}H + NADPH + H^+ + O_2 \longrightarrow \mathbf{R}OH + NADP^+ + H_2O \qquad \mathbf{R}= \textit{Substrate}$$

This process is complex and requires the donation of two electrons sequentially from an electron donor. The type of donor depends on the location of the P450 in the cell, or whether it is a bacterial protein or an eukaryotic protein. In the ER membrane, NADPH cytochrome P450 reductase is the usual electron donor, though cytochrome b5 can also participate. In the mitochondria, ferredoxin (adrenodoxin) and ferredoxin reductase (adrenodoxin reductase) form a short electron transfer chain to supply electrons. The bacterial donors are of both types. The Bacillus megaterium P450 CYP102 actually has the NADPH cytochrome P450 reductase fused to the P450 in a single gene (25).

Functionally the CYP enzymes can roughly be divided into two groups. The first group is involved in biosynthesis of low-molecular weight regulators of various biological functions. The formation of steroid hormones from cholesterol is a typical example involving several very specific CYP enzymes (25). The second class of mammalian CYP enzymes takes part in metabolism of various compounds including drugs, food preservatives, many toxicants and carcinogens (22,26).

There are more than 1500 known P450 sequences and therefore to aid in communication, a standardized nomenclature system has been established (27). This nomenclature is based on

evolution of the protein sequences, with similar sequences being clustered into families and subfamilies. The root of the cytochrome P450 names is CYP, this is the super family name. The primary structure of the families in the super family should not differ by more than 60% (40% homology) and is labelled with numerical numbers e.g. CYP1 and CYP2. The subfamilies are labelled with alphabetical symbols e.g. CYP1A and CYP1B and the primary structure should be 55% identical (i.e. 55% homology) (Figure 1.1) (2,22).

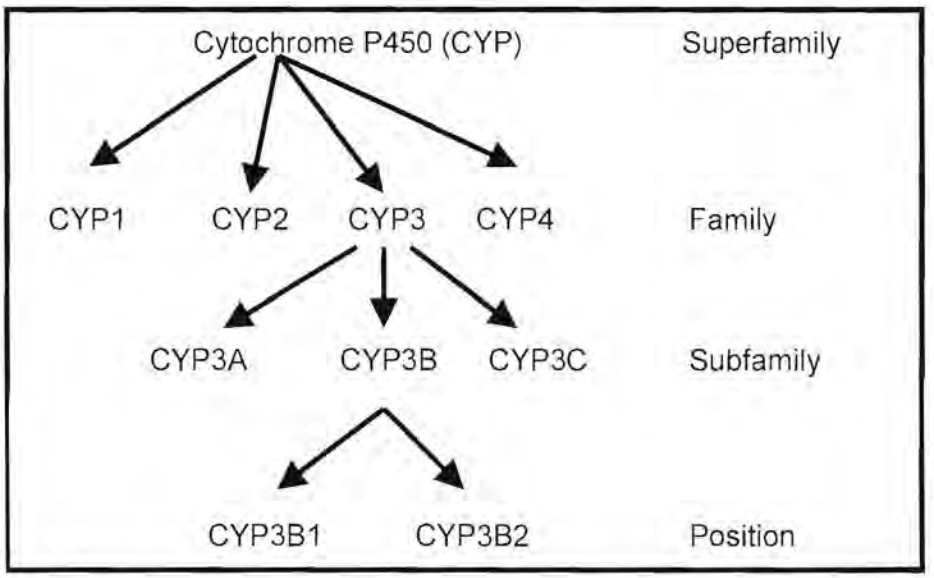


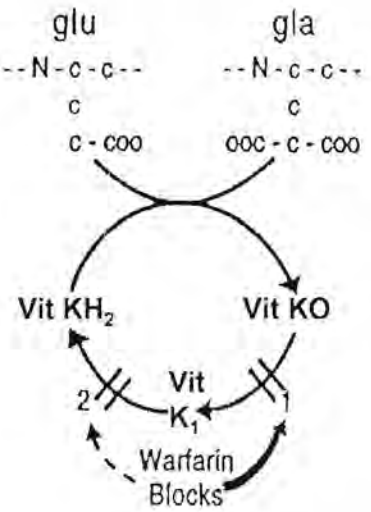
Figure 1.1: Schematic presentation of the nomenclature of the CYP450 family.

This study focuses on the CYP2C9 gene. Thus, Cytochrome P450 (CYP), the second family of enzymes (CYP2), subfamily C (CYP2C) enzyme number 9 (CYP2C9). The three polymorphisms that will be studied in this gene is known as the CYP2C9*1 (wild type), CYP2C9*2 and CYP2C9*3. The CYP2C9 gene transcribes for S-mephenytoin 4-hydroxylase (28). This enzyme consists of 490 amino acids with a molecular weight of 55 627 Da (28).

The eukaryotic P450s are membrane bound and are therefore difficult to isolate and only recently a crystal structure of a representative enzyme has been obtained (Figure 1.2). The general structure is globular, almost triangular, with the C-terminal half being helix rich and the N-terminal half being more beta sheet rich. The C-terminal half is more highly conserved. The P450 signature motif includes the heme ligand Cys and is usually represented as FXXGXXXCXG (X representing any unspecific amino acid), though there are exceptions at all three non-cys positions. This heme-binding region is about 50 amino acids from the C-terminal of the protein. The helix rich half of the protein starts with the I-helix.

1.4 Warfarin metabolism.

The anticoagulant, warfarin, produces its anticoagulant effect by preventing the activation of vitamin K-dependent coagulation factors II (prothrombin), VII, IX and X. The reduced form of Vitamin K, vitamin KH_2 is a cofactor in the carboxylation of these coagulation factors and warfarin directly inhibits two enzymes that regenerate vitamin KH_2, vitamin K epoxide reductase and Vitamin K reductase (Figure 1.3) (10,31,32). In addition warfarin also inhibits carboxylation and subsequent activation of protein factors C and S (33).



1. **KO - reductase - warfarin sensitive**
2. **K - reductase - relatively warfarin resistant**

Figure 1.3: Indicating the molecular sites of action of Warfarin (32).

Warfarin is a racemic mixture with R-warfarin as the predominant component in a steady state ratio of 0.5:1 (S-warfarin: R-warfarin). However S-warfarin is 5 times more potent than R-warfarin and therefore mediates its hematologic effect predominantly through its S-enantiomer. The S-enantiomer is then oxidized to its main inactive form S-7-hydroxywarfarin (Figure 1.4) (21,34) by two subfamilies of the cytochrome P450 complex (CYP2C9 and CYP2A6) to the inactive metabolite that is excreted in the bile (5). The CYP2C9 accounts for 80-85% of the metabolism of the S-enantiomer (16). Therefore any genetic polymorphism that would affect enzyme activity would have a profound effect on the metabolism of warfarin.

Minor pathways (Figure 1.4) of S-warfarin metabolism includes 6-hydroxylation by CYP2C9 and 4-hydroxylation by CYP2C8, CYP2C18 or CYP2C19. The metabolism of R-warfarin is more complicated where CYP1A1, CYP1A2, CYP2C9, CYP2C18 and CYP2C19 are responsible for a number of hydroxylation reactions (35).

Warfarin plasma concentration is inversely proportional to plasma clearance, and the plasma clearance of S-warfarin (0.0744 ml/minute/kg) is twice that of R-warfarin (0.0357 ml/minute/kg) (36). Reduction in the clearance of S-warfarin would increase the S-warfarin:R-warfarin ratio and have an influence on the INR. Individuals with the CYP2C9*2 and CYP2C9*3 polymorphisms have decreased clearance of S-warfarin and therefore also an increased plasma concentration ratio when compared to R-warfarin. In patients homozygous for CYP2C9*1 the steady state for the S-warfarin: R-warfarin ratio is 0.5:1(37). In patients with allelic variants the relative concentration of S-warfarin increases and a ratio of 0.87:1(37) is observed for a patient homozygous for CYP2C9*2 polymorphism and 4:1(37) for a patient homozygous for the CYP2C9*3 polymorphism.

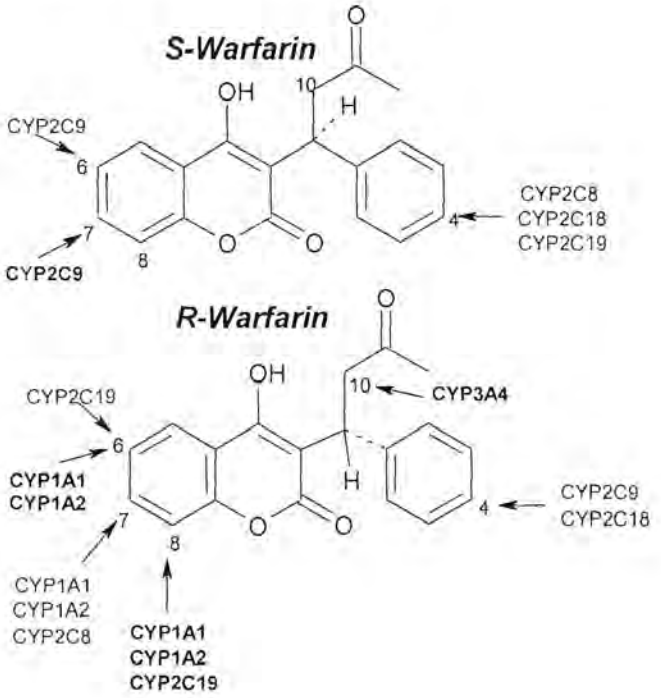


Figure 1.4: Sites of hydroxylation of R- and S-warfarin catalyzed by human cytochrome P450. Bold = major metabolic pathway (35).

1.5 The CYP2C9 genes, enzymes and polymorphisms.

Enzymes in the human CYP2C gene subfamily are clustered on chromosome 10q24 and are contained within a 300 kb region (38). Four CYP2C genes with polymorphisms have been identified CYP2C8, CYP2C9, CYP2C18 and CYP2C19 (39). The CYP2C9 gene is expressed the highest concentration in the human liver (35). The transcribed protein of the CYP2C9 gene is known as mephenytoin 4-hydroxylase (Figure 1.2) and is responsible for the metabolism of 80% of the S-warfarin (4,5,6). Therefore polymorphisms within this region would make the largest contribution to impaired warfarin metabolism (16,40).

MDSLVVLVLC[10] LSCLLLLSLW[20] RQSSGRGKLP[30] PGPTPLPVIG[40]
NILQIGIKDI[50] ...IRNFGMGKRS[140] IED **R** VQEEAR[150] CLVEELRKTK[160]
...DRSHMPYTDA[350] VVHEVQRY **I** D[360] LLPTSLPHAV[370] ...
PFYQLCFIPV[490]

Figure 1.5: Regions of CYP2C9*1 amino acid sequence showing the positions of the CYP2C9*2 and *3 polymorphisms in S-mephenytoin 4-hydroxylase enzyme (28).

Point mutations of the CYP2C9 gene, CYP2C9*2 (cysteine (C) substituted for arginine (R)) (Bold red, Figure 1.5) at amino acid 144 (R144C)) and CYP2C9*3 (leucine (L) is substituted for isoleucine (I)) (Bold red, Figure1.5) at amino acid 359 (I359L)) are associated with impaired warfarin metabolism (4-10) (Table 1.1).

Table 1.1: Classification of CYP2C9 Genotypes (41).

Amino acid	Nucleotide	Polymorphism	Genotype
Arg-144 to Cys	CGT/CGT	CYP2C9*1 Homozygous, normal	2C9*1/2C9*1#
	CGT/TGT	CYP2C9*2 Heterozygous	2C9*1/2C9*2
	TGT/TGT	CYP2C9*2 Homozygous, polymorphic	2C9*2/2C9*2
Ile-359 to Leu	ATT/ATT	CYP2C9*1 Homozygous, normal	2C9*1/2C9*1#
	ATT/**C**TT	CYP2C9*3 Heterozygous	2C9*1/2C9*3
	CTT/**C**TT	CYP2C9*3 Homozygous, polymorphic	2C9*3/2C9*3
Compound heterozygous	TGT/**C**TT	CYP2C9*2/CYP2C9*3 Heterozygous	2C9*2/2C9*3

Absence of both mutations within this allele.

In addition CYP2A6*2 (histidine is substituted for leucine at amino acid 160 (H160I) and CYP2C6*3 (multiple mutations) also contributes to impaired warfarin metabolism. Individuals that are homozygous for the CYP2C9*2 and CYP2C9*3 polymorphisms present with reduced enzymatic activity of 12% and 5% respectively when compared to the wild type CYP2C9*1, homozygous normal genotype (6). An individual that is heterozygous for CYP2C9*2 and *3 will present with 56% and 52% activity respectively. Compound herozygosity with a single copy of the polymorphism on each allele will present 9% of normal enzyme activity (6,42).

The effect that the CYP2C9*2 and CYP2C9*3 polymorphisms have on warfarin dosage has been studied by several researchers (Table 1.2) (43). A single copy of the CYP2C9*2 polymorphism, will result in a 13%-22% decrease in the mean warfarin maintenance dosage. Heterozygosity for the CYP2C9*3 polymorphism also results in a 21-42% decrease in the mean warfarin maintenance dosage. In these population studies (43) only a few individuals would have been expected to be homozygous for the CYP2C9*2 and *3 polymorphisms and therefore care must be taken when interpreting the results. However from Table 1.2 it can be seen that homozygosity or compound heterozygosity does results in a decrease in the mean maintenance dosage of 18-87% (43).

The CYP2C9*3 polymorphism occurs in a very important site, the binding receptor site 5 (SRS-5) (44). Secondary modelling predicts that this region falls in a β-strand aligned with CYP101 (Section 1.2). The Ile-359 to Leu variant disrupts this part of the structure. Kaminsky *et al* (45) found variable results for S-warfarin 7-hydroxylation V_{max} with similar activities for the variants with microsomes from the yeast expression system, but lower activity for the Ile-359 to Leu variant using purified and reconstituted CYP, while Sullivan-Klose (41) found the K_m to be higher for the Ile-359 to Leu variant.

The CYP2C9*2, Arg-144 to Cys (Table 1.1) residue is not within the defined (44) substrate recognition sites (SRSs). It is however in close proximity of SRS-1. Structural modeling of this region of the protein predicts a stretch of α-helix consistent with alignments with CYP101 (44,46). The CYP2C9*2 polymorphism will disrupt the α-helix and alter the functionality of the protein. The functional effect on this protein has been described by a number of studies. The results were conflicting. Rettie *et al* (42) and Sullivan-Klose *et al* (41) reported that the K_m for S-warfarin 7-hydroxylation remained unchanged in the CYP2C9*2

polymorphic enzyme. In contrast, the V_{max} value was reported to be unchanged (41), decreased ten-fold (42) and increased three-fold (45) when compared to wild type activity. Rettie (42) reported a five to six fold lower V_{max}/K_m relationship for the Arg-144 to Cys variant enzyme. Individuals who express the variant enzyme would therefore be expected to be slower metabolizers of warfarin.

Bhasker *et al* (1) studied the RNA expression of alleles in individuals heterozygous for R144C. Using reverse transcription Polymerase Chain Reaction (PCR), the authors found that the R144C had a 5-10 times higher level of expression than in the normal Arg 144 allele. Although the actual effect of the R144C substitution is not known, it is however accepted that patients either heterozygous or homozygous present with altered warfarin metabolism (Table 1.2).

Table 1.2: The effect of CYP2C9 genotypes on mean warfarin maintenance dose as mg per day *(percentage reduction in comparison to prescribed doses for wild type genotype 1*/1*)* (43). N=size of the population studied.

Population	N	*1/*1	*1/*2	*1/*3	*2/*2	*2/*3	*3/*3
Caucasian	94	4.7	3.8 *(19%↓)*	nd	nd	nd	nd
Caucasian	52	4.25	3.5 *(18%↓)*	2.5 *(41%↓)*	3.5 *(18%↓)*	nd	nd
Caucasian	180	6.7	5.2 *(22%↓)*	3.8 *(43%↓)*	5.2 *(22%↓)*	1.8 *(73%↓)*	nd
Caucasian	561	5.01	4.3 *(14%↓)*	3.97 *(21%↓)*	3.0 *(39%↓)*	4.1 *(18%↓)*	0
Caucasian	185	5.6	4.9 *(13%↓)*	3.3 *(41%↓)*	4.1 *(27%↓)*	2.3 *(59%↓)*	1.6 *(71%↓)*
Asian	47	3.0	nd	1.8 *(42%↓)*	nd	nd	0.4 *(87%↓)*

nd= *not determined*

1.6 Prevalence of the CYP2C9*2 and *3 polymorphisms in different populations.

The distribution of the CYP2C9*2 and *3 polymorphisms differ between different populations. Not only can differences in the prevalence of each polymorphism occur but one type may be more predominant than the other. In the Caucasian population the prevalence of the CYP2C9*2 polymorphism varies from 8-19.1% (4,6) and is higher than that found in the indigenous Canadian, Afro-American and Asian populations (Table 1.3). In contrast the CYP2C9*3 polymorphism is more common in the Asian population (1.7-5%) than the CYP2C9*2 (0%) polymorphism (43). The prevalence of the CYP2C9*3 polymorphism in Asians is however still lower than that found in the Caucasian population (6-10%). Two polymorphisms, namely CYP2C9*4 and *5 have been identified in the Afro-American and the Asian populations, however no studies have been undertaken to determine whether these polymorphisms are specific for these populations (48).

Table 1.3: Prevalence of CYP2C9 polymorphisms among different ethnic groups (43).

Polymorphism	CYP2C9*1	CYP2C9*2	CYP2C9*3	CYP2C9*4	CYP2C9*5
Point mutation	Arg_{144}/Ile_{359}	Cys_{144}/Ile_{359}	Arg_{144}/Leu_{359}	Arg_{144}/Thr_{359}	Arg_{144}/Glu_{360}
Caucasians	79%-86%	8%-19.1%	6%-10%	nd	nd
Indigenous Canadian	91%	3%	6%	nd	nd
African American	98.5%	1%-3.6%	0.5%-1.5%	nd	2.3%
Asians	95%-98.3%	0%	1.7%-5%	0%-1.6%	0%

nd=*not determined*

The South African Caucasian population is of European origin and is mostly of English, French and Dutch descent. The prevalence of these mutations in the French and Dutch populations has not been documented. Many studies have been undertaken in the British population and will be discussed in greater detail due to their relevance to this study. Aithal *et al* (4) in 1999 studied the prevalence of the CYP2C9*2 and *3 polymorphisms in a population from North-East England. The population consisted of three groups, a community control (a random control group of 100 individuals), a clinic control group (52 patients attending the anticoagulation clinic) and a low dosage clinic control group (patients (n= 36))

receiving a warfarin maintenance dosage of 1.5 mg/day or less for 3 consecutive days. These results are tabulated in Table 1.4a-c. Although a small population study, some information can be obtained regarding the allele and gene frequencies of the CYP2C9*2 and *3 polymorphisms in this particular British population.

In the community control group studied by Aithal (4), none of the individuals screened were homozygous for any of the polymorphisms while two percent of the population was compound heterozygous (Table 1.4a). From this data an allele frequency and a gene frequency can be calculated. The allele and gene frequency for the CYP2C9*2 polymorphism was 0.19 and 0.014 respectively. For the CYP2C9*3 polymorphisms the frequencies were 0.18 and 0.01 respectively while the frequency of compound heterozygosity was 0.034. Both polymorphisms were present in the same population at the same gene and allele frequency.

Table 1.4a: The prevalence of the CYP2C9*2 and *3 polymorphisms in the community control group of North East England (4).

		CYP2C9*2		
CYP2C9*3		-/-	+/-	+/+
	-/-	60%	20%	0%
	+/-	17%	2%	0%
	+/+	1%	0%	0%

In the clinic control group 2% of the population was homozygous for the CYP2C9*2 polymorphism. No individuals homozygous for the CYP2C9*3 polymorphism or compound heterozygous for the CYP2C9 polymorphisms were found (Table 1.4b). In this population the calculated CYP2C9*2 gene and allele frequencies was 0.011 and 0.187 respectively. The gene frequency for the CYP2C9*3 polymorphism was lower at 0.009 and the allele frequency was 0.17. The calculated frequency for compound heterozygosity was the same as for the community control group. The allele frequencies was the same for the community control and random clinic control group.

Table 1.4b: The prevalence of the CYP2C9*2 and *3 polymorphisms in the random clinic control group of North East England (4).

		CYP2C9*2		
CYP2C9*3		-/-	+/-	+/+
	-/-	62%	17%	2%
	+/-	19%	0%	0%
	+/+	0%	0%	0%

Likewise the gene frequencies and the allele frequencies can be determined for the clinic low dosage group. In the clinic low dosage group 6% of the population was CYP2C9*2 homozygous, 0% were CYP2C9*3 homozygous while 14% of the population was compound heterozygous (Table 1.4c). This translates into a gene and allele frequency of 0.08 and 0.41 respectively for the CYP2C9*2 polymorphism. Similarly the prevalence of the CYP2C9*3 polymorphism is also increased at a gene and allele frequency of 0.044 and 0.33 respectively. The incidence of compound heterozygosity at 0.135 is higher than found in the community and clinic control groups. As shown in Table 1.2 the presence of a single copy of either the CYP2C9*2 and *3 polymorphisms is associated with a decrease in the amount of warfarin required. For the random community and clinic control group, the percentage of the population presenting with polymorphisms associated with impaired warfarin metabolism is 40% and 38% respectively.

Table 1.4c: The prevalence of the CYP2C9*2 and *3 polymorphisms in the low dose clinic group of North East England (4).

		CYP2C9*2		
CYP2C9*3		-/-	+/-	+/+
	-/-	19%	33%	6%
	+/-	28%	14%	0%
	+/+	0%	0%	0%

In the low dosage warfarin group the gene frequency of the CYP2C9*2 polymorphism was eight times higher that found in the community and the clinic control groups. The gene frequency for the CYP2C9*3 polymorphism and compound heterozygosity was four times

greater than the control populations. The prevalence of genotypes associated with altered warfarin metaboism increases from 38-40% for the control populations to 81% in the low warfarin dosage group.

Redman (49) reported that the frequency of homozygotes for the CYP2C9*2 polymorphism in the American Caucasian population was about 20% (14,15) and 8.5% for the CYP2C9*3 polymorphism. Freeman *et al* (5) studied the prevalence of the CYP2C9 and CYP2A6 mutations in 38 patients either chronically receiving warfarin therapy or undergoing the initiation of warfarin therapy. Within this group an allele frequency of 21% was observed for CYP2C9*2 and 8% for the CYP2C9*3 mutations respectively. Overall 34% of the patients possessed a genotype, associated with impaired warfarin metabolism.

In a larger population study, Taube *et al* (6) undertook to determine whether the presence of CYP2C9 polymorphisms would influence warfarin sensitivity and risk of over-anticoagulation in patients on long-term treatment as prescribed by the Cambridge anticoagulant clinic. A total of 561 patients with a target INR of 2.5 who had been treated with warfarin for more than 2 months were screened. The prevalence of the CYP2C9*2 and *3 polymorphisms in the population studied are presented in Table 1.5.

Table 1.5: The prevalence of the CYP2C9*2 and *3 polymorphisms in a clinic population of 561 patients with a target INR of 2.5 (6).

		CYP2C9*2		
CYP2C9*3		-/-	+/-	+/+
	-/-	69.9%	19.1%	0.5%
	+/-	9.5%	1%	0%
	+/+	0%	0%	0%

In the population 0.5% and 0% were homozygous for the CYP2C9*2 and *3 polymorphisms respectively, while 1% of the population was compound heterozygous. The calculated gene and allele frequencies for the CYP2C9*2 polymorphism were 0.011 and 0.19 respectively. For the CYP2C9*3 polymorphism the gene and allele frequencies were 0.0025 and 0.095 respectively. The frequency for compound heterozygosity was 0.018. Furthermore 30% of the population were found to have polymorphisms associated with altered warfarin metabolism.

The gene frequency for the CYP2C9*2 polymorphism is similar to that found by Aithal *et al* (4) in the community and clinic control population. However these results are considerably lower than that reported by Aithal *et al* (4) for the clinic low dosage group. These differences could be due to the criteria used to select patients where Aithal *et al* (4) selected patients with a warfarin maintenance dosage of 1.5 mg/day while the selection criteria used by Taube *et al* (6) was 2.5 mg/day. However both studies do show that at least 20% of the population studied have a CYP2C9*2 polymorphism that is associated altered warfarin metabolism. The gene frequency of compound heterozygotes varied greatly, 0.034 vs 0.018, for the clinic control groups. In the study of Aithal *et al* (4) the frequency of both CYP2C9*2 and *3 were similar while in the study of Taube *et al* (6) the CYP2C9*2 is the predominant polymorphism. The study of Taube *et al* (6) is a clinic-based study, and the 30% prevalence of CYP2C9 polymorphisms associated with altered warfarin metabolism, is considerably less than the 40% reported by Aithal *et al* (4) for the general British population.

Redman (49) incorrectly reported from studies undertaken by Miners *et al* (15), Yamazaki *et al* (16) and Kimura *et al* (15) a frequency of 20% homozygosity for CYP2C9*2 and 8.5% homzygosity for the CYP2C9*3 polymorphism in the Caucasian population. However these findings coincide with the allele frequencies and not gene frequencies. Factors such as population size and patient selection criteria are critical when determining the frequencies of a polymorphism in a population, especially if it is rare.

1.7 Detection of the CYP2C9*2 and CYP2C9*3 polymorphisms.

The presence of the CYP2C9*2 and *3 polymorphisms are determined using restriction fragment length polymorphism (RFLP) polymerase chain reaction (PCR). PCR has a wide range of applications in the field of medical diagnostics. The major advantages of PCR is that once correctly optimised it is highly specific, requires small amounts of DNA, is a rapid, cost effective method that can be used to screen large populations for specific mutations or polymorphisms. DNA is usually isolated from the leukocyte population of the blood. This involves the collection of a single uncoagulated blood sample using the established phlembotomy services within a clinical laboratory. DNA can be isolated immediately, the blood can be frozen and the DNA isolated at a later stage or the DNA can be isolated and stored for extended periods of time. All DNA isolation procedures include the following

basic steps: cell lysis, solublization of protein and DNA, and selective isolation of DNA from all other cellular components. The most frequently used methods are phenol/chloroform extraction, selective precipitation or selective binding to a membrane, column or particulate matrix. The method of choice should fulfil the following requirements; the isolation of pure intact DNA, limited exposure to chemical and biological interference, rapid, high throughput capacity, reproducibility and cost effectiveness (50).

The method of Stubbins *et al* (51) and Sullivan-Klose *et al* (41) for the detection of the CYP2C9*2 and *3 polymorphisms respectively involves the amplification of a specific DNA fragment that contains the region where the single nucleotide substitution has occurred. Two primers, short nucleotide sequences (18-24 bp) that are highly specific and complementary to the region to be amplified are required. Several variables must be taken into account when designing PCR primers. Among the most critical are: primer length, melting temperature I, specificity, complementary primer sequences, G/C content and polypyrimidine (T, C) or polypurine (A, G) stretches at the 3'-end (19). Amplification consists of three basic steps. First, the target genetic material must be denatured-that is, the strands of its helix must be unwound and separated-by heating to 90-96°C. The second step is hybridisation or annealing, in which the primers bind to their complementary bases on the now single-stranded DNA. The third is DNA synthesis by a DNA Taq polymerase (isolated from the thermophilic eubacterium *Thermus aquaticus BM*, a strain that lacks Taq I restriction endonuclease activity) (52). Taq polymerase Lacks 3' to 5' exonuclease activity, But has a 5' to 3' polymerisation-dependant exonuclease activity and removes nucleotides ahead of the newly synthesized strand. Starting from the primer, the Taq DNA polymerase can rapidly read a template strand and match it with complementary nucleotides (53). The result is two new helixes in place of the first, each composed of one of the original strands plus its newly assembled complementary strand. The entire process is repeated for a number of cycles with an exponential increase in the number of new strands that are synthesised either using the DNA or newly synthesized strands as template (54). PCR amplification methodologies usually start as techniques that developed in a research environment to determine whether a mutation or polymorphism is associated with a disorder. Once a positive correlation is obtained and testing is applicable to the population for which a service is provided it will become a part of the repertoire of testing found in a molecular diagnostic laboratory.

To achieve this end, several parameters need to be optimised such as $MgCl_2$, primer, Taq DNA polymerase and DNA concentration. The optimal Taq DNA polymerase enzyme activity is between 0.5 and 2.5 units. Increased enzyme concentration sometimes leads to decreased specificity (52,54). The concentration of Mg^{2+} influences Taq DNA polymerase activity. It increases the Tm of double stranded DNA and forms soluble complexes with dNTP's to produce the actual substrate that the Taq DNA polymerase recognizes. The concentration of free Mg^{2+} depends on the concentration of compounds that bind the ion, including dNTP's, free pyrophosphate and EDTA. The concentration of magnesium in the PCR reaction is critical to the success of the reaction (55). The final magnesium concentration in the reaction mixture may be optimised by the user according to individual requirements and is usually between 1.0 and 3.5 mM. Higher Mg^{2+} increases the PCR yield but decreases the specificity of the reaction (increases the formation of primer dimers). A lower Mg^{2+} concentration increases the specificity but reduces the yield (52,54). To improve specificity and yield of difficult targets in PCR amplifications, researchers often include enhancing agents in the reaction (56). The two PCR enhancing agents that deserve particular mentioning are betaine (N, N, N-trimethylglycine) and dimethyl sulphoxide (DMSO). DMSO is probably the most commonly used enhancing agent and is frequently included as part of standard optimising of PCR amplifications (47). Both of these agents facilitate strand separation; DMSO disrupts base pairing whereas betaine, an iso-stabilizing agent, equalizes the contribution of GC- and AT-bases pairing to the stability of the DNA duplex (52,57).

Above mentioned parameters must be optimised for a standard PCR methodology such as that used for the detection of the CYP2C9*2 polymorphism. This becomes even more important for methodologies where a mismatch primer is used to create a specific restriction site. The method of Stubbins *et al* (51) is a standard RFLP method where two primers following amplification lead to the formation of a 190 bp amplification product (41). The CYP2C9*3 nucleotide substitution does not naturally create or destroy a restriction site. A restriction site was forced into the forward primer '5- aat aat aat atg cac gag gtc cag aga tgc-3'(where g is the site of the forced mismatch), such that A1061 in combination with the mismatch creates a restriction site (a tgc at)(a = A1061) for the Nsi I restriction enzyme.

The 3' terminal position in PCR primers is essential for the control of mis-priming (5). The inclusion of a G or C residue at the 3' end of primers usually reduces the effect of primer mispriming. Not only does the inclusion of a g at at the site of the forced mismatch creates a

site for the restriction enzyme Nsi I but it provides an additional "GC Clamp" that helps to ensure correct binding at the 3' end due to the stronger hydrogen bonding of G/C residues (58).

Restriction enzymes recognise specific sites of different length and base composition. The typical restriction site is an exact palindrome of 4-8 bp with an axis of rotational symmetry. The number and the size of the fragments generated by a restriction enzyme depend on the frequency of occurrence of the restriction enzyme site in the DNA to be cut. In RFLP a restriction enzyme is used that recognizes the sequence where the point mutation has occurred (52). If the substitution has occurred, the restriction enzyme will recognise the sequence, bind and cut the site resulting in two fragments, usually of different size, that can be resolved by gel electrophoresis. If the mutation is absent the DNA fragment is not cut. Often if no restriction enzyme is found that recognises the mutation site, then an enzyme is chosen that recognizes the normal sequence. Subsequently the reverse is observed the DNA fragment without the mutation is cut and the fragment with the mutation is not. The T-C substitution at codon 144 for the CYP2C9*2 polymorphism results in the loss of a restriction site for Ava II (41). Therefore Ava II will digest the 190 bp amplification product from an individual that has the normal allele. As the normal genotype, where the polymorphism is absent, is the most common genotype, this type of RFLP-PCR design is ideal as it serves as an internal control ensuring that complete and effective digestion has occurred.

The RFLP digestion products are usually separated according to size by either polyacrylamide or agarose gel electrophoresis. The RFLP digestion products produce different banding patterns following gel electrophoresis and these are used to identify whether an individual is homozygous normal, heterozygous or homozygous for either the CYP2C9*2 or *3 polymorphisms.

1.8 Nucleotide sequencing.

To verify that the PCR product contains the point mutation of interest and represents the correct nucleotide region where the polymorphism occurs, the PCR product needs to be sequenced either by manual or automated sequencing. Chain termination is the predominant method used for DNA sequencing. For DNA sequencing all the components needed for in

vitro DNA synthesis are combined in a reaction. This includes Taq DNA polymerase and additionally adding a 2',3'-dideoxynucleotide (ddNTP's). The ddNTP's can be incorporated into a DNA structure with the usual elongation step of DNA synthesis through a bond formation between the 3' carbon of the last incorporated 3' terminal sugar and the 5' phosphate group of the ddNTP. The ddNTP however lacks the 3'-OH group necessary for phosphodiester bond formation, which is needed for further elongation. This results in the termination of extension at the precise point of ddNTP incorporation (59).

Four sets of reactions are used to complete one sequencing reaction. The four reactions differ only in the specific ddNTP that is added. The dNTP:ddNTP ratio is specifically selected so that the resulting labelled strands from a nested set of molecules up to several thousand bases long, each terminating at a specific base. These strands are separated according to size by high-resolution denaturing gel electrophoresis. This provides a ladder of bands from which the DNA sequence can be read. An important limitation of the standard chain termination method is that only a single labelled DNA molecule is produced from each primer template complex and this limits the sensitivity of the method. The sensitivity of the method is limited by the molar quantity of DNA template that can be used in the reaction. This becomes a serious problem when sequencing large DNA templates and purified PCR product. This limitation can be overcome by repeating the reactions in a number of cycles. The described steps can be repeated through thermal cycling, these cycles of optimal denaturation, primer binding and elongation temperatures allows linear amplification of the amount of labelled material present in the reaction. Thus less DNA template is required. In combination with fluorescent label and semi-automated sequencers, this is the method used to generate most new sequencing information, for example genome sequencing projects (60).

Some studies prefer to use sequencing methods to screen all their patients in the group studied (17). This approach is definitely more accurate but also more expensive and time consuming. In diagnostic molecular pathology, sequencing is usually used as an initial step to identify the site of a genetic mutation or polymorphism and to verify the sequence. The sequence of about 100-300 bp around the region is used to develop primers for a PCR method that will allow the rapid screening of individuals for a mutation (52). When using a PCR method to screen patients for a mutation (or polymorphism), it is necessary to include two controls, namely a negative and positive control. The negative control is an individual that

does not present with the polymorphism or mutation, while the positive control is homozygous for the polymorphism or mutation and both should be sequenced.

1.9 Screening of a population for a specific mutation.

Population screening involves the offer of genetic testing on an equitable basis to all individuals in a defined population. Its primary objective is to enhance autonomy by enabling individuals to be better informed about genetic risks and in some instances reproductive options. There are certain criteria for evaluating the validity of a screening programme; these can be subdivided into criteria for the disease, test and programme. Criteria for a specific disease are a high prevalence in a target population where the presence of a specific mutation has a serious effect on health. The disease should be a treatable or preventable disorder. Criteria for the test are that it should be non-invasive and easily carried out, accurate, reliable and inexpensive. The programme should be equitable, available, involve voluntary participation and be acceptable to the target population and full information and counselling should be provided (61).

1.10 The Founder Effect and its impact on population screening in the South African population.

The founder effect is a genetic phenomenon caused by populations starting at or proceeding through small numbers. This phenomenon is characterised by a relatively high incidence of alleles associated with a specific phenotype, found within a certain population. When a population is initiated by a small, and therefore genetically unrepresentative, sample of the parent population, the genetic drift observed in the sub-population is referred to as the founder effect (61). The founder effect can result in an increased prevalence of the mutations found in the parent population or an increased prevalence of mutations that are only found in this specific population. Examples of the latter are the D206E, V408M and the D154N familial hypercholesterolaemia mutations specific for the South African Caucasian population (62). An increased prevalence of the C282Y and the H63D mutation associated with hereditary heamochromatosis also found in other Caucasian populations is found in the South African population. The CYP2C9*2 and CYP2C9*3 polymorphisms have been found in numerous other Caucasian populations and are therefore not unique to South Africa.

However, it is necessary to determine whether the prevalence of these polymorphisms is greater or less than that described for other populations.

Several factors should be taken into account when screening a population for a mutation or polymorphism. These factors include the origin of the population group as this may lead to discrepancies as seen in the studies of Taube *et al* (6) and Aithal *et al* (4) for the CYP2C9*2 and *3 polymorphisms. Furthermore the size of the population could greatly influence the results obtained specially if the incidence of the mutation and the polymorphism is low as has been shown for the CYP2C9*3 polymorphism. The equilibrium of the alleles tested in the population is also important and is analysed using the Hardy-Weinberg principle. The Hardy-Weinberg (63) principle states that expected genotypic proportions for an autosomal, diploid locus can be calculated in terms of allelic frequencies by a binomial or multinomial function (63):

$$(p+q)^2 = p^2 + 2pq + q^2$$

Where there is a gene locus with two alleles A and a, which have frequencies of p and q respectively and where these are the only alleles found at this locus so that p+q = 100% or 1 (54). From this data the gene and allele frequencies for a specific mutation or polymorphism can be calculated for a specific population. The Hardy-Weinberg model was formulated for autosomal loci in sexually reproducing diploid species. The system makes the following assumptions: that mating is random among phenotypes, generations do not overlap, there is a large population size, there is no variation from random drift, no migration, no mutation or natural selection favouring some genotypes and that the allele frequencies are the same in both sexes (61). If these assumptions hold, Hardy-Weinberg proportions are reached in a single generation. The population will remain indefinitely in these proportions until some event, such as random genetic drift or natural selection changes allelic frequencies. Hardy Weinberg model can be used to determine the frequencies of one or two polymorphisms (63). This is important when two different polymorphisms or mutations contribute to the development of the same disorder such as seen in Herediatry Haemochromatosis where compound heterozygosity is associated with disease. Likewise it has been shown that individuals that are homozygous for the CYP2C9*2 and *3 alleles require a lower warfarin maintenance dose.

1.11 The aims of this study.

- **To develop screening methods for the detection of the CYP2C9*2 and *3 polymorphisms.**

To achieve these aims the optimal concentrations of $MgCl_2$, DNA and primers were determined. The effect of adjuvants such as glycerol, BSA and DMSO was also investigated. The sequences of the CYP2C9*2 and *3 PCR products was determined to verify that the correct region was amplified and to verify the presence of the site of restriction enzyme digestion.

- **To determine the prevalence of these polymorphisms in the South African Caucasian population.**

Following this, a random Control population and Anticoagulation clinic population were screened. The data obtained was statistically analysed and the gene and allele frequencies of the CYP2C9*2 and *3 polymorphisms in the South African population calculated. The possibility of a founder effect in the South African Caucasian population is discussed in Chapter 4.

- **To illustrate the association between the polymorphisms and decreased daily warfarin maintenance dose and to discuss recommendations regarding the implimentation of screening methodologies for patients initiating warfarin therapy.**

The presence of an allele associated with altered warfarin metabolism was correlated with warfarin maintenance dosage as part of a collaboration study with the Department of Haematology ant the University of Pretoria. Information obtained will be used to develop a screening strategy for patients of the South African Caucasian population commencing warfarin therapy.

Chapter 2:

Materials and Methods.

2.1 Materials.

All primers used in this study were purchased from Whitehead Scientific, Cape Town, South Africa, synthesized by Integrated DNA Technologies, Coralville, USA. All other PCR reagents and restriction enzymes used were from Promega Corporation, Madison, USA and supplied by Whitehead Scientific, Cape Town, South Africa. All PCR reactions were carried out using a Thermal Cycling System from Hybaid Limited, Teddington, Middlesex, UK, supplied by Scientific Group, Cape Town, South Africa. PCR tubes and tips were manufactured in the USA by Quality Scientific Plastics and supplied by Sterilab Services, Johannesburg, South Africa. The QIAmp DNA Blood Mini Kit used was manufactured by Qiagen and supplied by Southern Cross Biotechnology, Cape Town, South Africa.

Water was double distilled and deionized with a Continental Water System and sterilized by filtration through a Millex 0.22µm filter. Glassware and PCR tubes were sterilized at 140°C for 20 minutes in a Speedy Autoclave HL-341.

2.2 Subjects.

DNA previously isolated from 200 white male medical students attending the University of Pretoria was used as the Control population and represented a normal healthy South African Caucasian population. A single tube of citrate blood was collected from 144 patients attending a Anticoagulation Clinic and this represented the Anticoagulation Clinic population. Informed consent was obtained from each of these individuals. The blood samples were stored at –20°C.

2.3 Isolation of DNA from frozen whole blood samples.

The DNA was isolated from the Control and the Thrombosis Clinic population using the QIAamp DNA Blood Mini Kit (50). The isolation procedure was a modification of that recommended by the manufacturers.

The frozen blood samples were removed from the freezer and placed on the bench top to thaw. The procedure was as follows: a 20µl volume of the QIAGEN Protease solution was pipetted into the bottom of 1.5ml micro-centrifuge tube. A 200µl volume of the thawed blood sample was added followed by 200µl Buffer AL (the composition of the buffers and solutions was not provided by the manufacturers). The sample was mixed by pulse-vortexing for 15 seconds. The sample was incubated at 56°C for 10 minutes in a water-bath. Following incubation the sample was centrifuged at 6 000 x g (8 000rpm) for 30 seconds in a Hereus Biofuge 15 Centrifuge to remove droplets from the inside of the lid. The tube was opened and 200µl ethanol was added to the sample, the sample was mixed again by pulse-vortexing for 15 seconds and the centrifugation step was repeated.

The lysed blood sample was carefully applied with a pipette to the QIAamp spin column ensuring that the rim remained dry. The cap was closed, and the tube with the spin column was centrifuge at 6 000 x g for 1 minute. The spin column was placed into a clean 2 ml collection tube and the tube containing the filtrate was discarded. The QIAamp spin columns were opened carefully and 500µl of Buffer AW1 were added without wetting the rim. The cap was closed and the sample was centrifuged at 6000 x g (8 000 rpm) for 1 minute. The spin column was removed, placed into a clean 2ml collection tube, and the collection tube containing the filtrate was discarded. The spin column was carefully opened and 500µl Buffer AW2 was added without wetting the rim. The cap was closed and the sample was centrifuged at full speed (20 000 x g; 14 000rpm) for 3 minutes. The spin column was removed from the collection tube, dried with a paper towel to reduce possible buffer carryover. The collection tube was replaced with a new tube and centrifuged at 8000 x g for 1 minute.

The spin column was placed in a clean 1.5ml micro-centrifuge tube and the collection tube containing the filtrate was discarded. The column was carefully opened and 200µl AE Buffer was added. The spin column was incubated at room temperature for 1 minute, then centrifuged at 6000 x g for 1 minute to extract the DNA. The tube containing the DNA sample was closed, labeled and stored at 4°C.

2.3.1 Determination of DNA concentration and quality.

The DNA quality was determined by electrophoretic analysis (Section 2.3.1.1) and DNA purity and concentration was determined by spectrometric analysis (Section 2.3.1.2).

2.3.1.1 Electrophoretic analysis of isolated DNA.

A 1% agarose gel was prepared in 1x TBE buffer (0.089 M Tris, 0.079M Boric acid, 0.002M EDTA at pH 8.3) containing 0.01% ethidium bromide (EtBr). Electrophoresis was carried out using a Hoefer Submarine Gel Electrophoresis System coupled to a Pharmacia PS 3000 DC power supply at 140 V. The gel was then visualized by ultraviolet radiation and photographed using UVIdoc Gel Documentation System manufactured by UVItec Limited, St John's Innovation Centre, Cambridge, UK and supplied by Whitehead Scientific, Cape Town South Africa. The DNA samples (either stored Control or freshly prepared Anticoagulation Clinic DNA) were diluted 1:1 with loading buffer (60% sucrose in 1x TBE buffer containing 0.001% bromophenol blue (BPB) and loaded onto the gel.

2.3.1.2 Spectrometric analysis of isolated DNA.

A 20µl aliquot of the isolated DNA was diluted 1:100 with 1x TBE buffer and measured against 1x TBE buffer. The absorbency at 260 and 280 nm was measured using a UV/VIS Lambda 2 Spectrophotometer (supplied by Perkin Elmer, Johannesburg, South Africa). The concentration of the sample was determined from the 260 nm absorbance reading and the purity from the 260/280 nm ratios. A DNA stock solution of 100 ng/ml

Calf Thymus DNA from Boehringer Mannheim and supplied by Roche, Johannesburg, South Africa was prepared. A concentration series of 0-6 ng/ml was used to prepare a standard curve for the determination of DNA concentration.

2.4 PCR strategy for the detection of the CYP2C9*2 polymorphism.

The CYP2C9*2 polymorphism is a C416T transition in exon 3 results in an arginine (R) to cysteine (C) substitution at position 144 of the amino acid sequence of CYP2C9. This results in the loss of an Ava II restriction digestion site (1) due to the presence of the single nucleotide substitution, C to T at nucleotide 416. PCR amplification of this region with the primers (sequences in Table 2.1) produces a 190 bp fragment (51) containing the CYP2C9*2 polymorphism. Ava II restriction enzyme digestion and subsequent gel electrophoresis of the amplified fragment allows the identification of the different CYP2C9*2 genotypes. The banding patterns that can be identified following gel electrophoresis represents the genotypes, CYP2C9*2 homozygous normal (-/-), CYP2C9*2 heterozygous (+/-) and CYP2C9*2 homozygous polymorphic (+/+)(Figure 2.1).

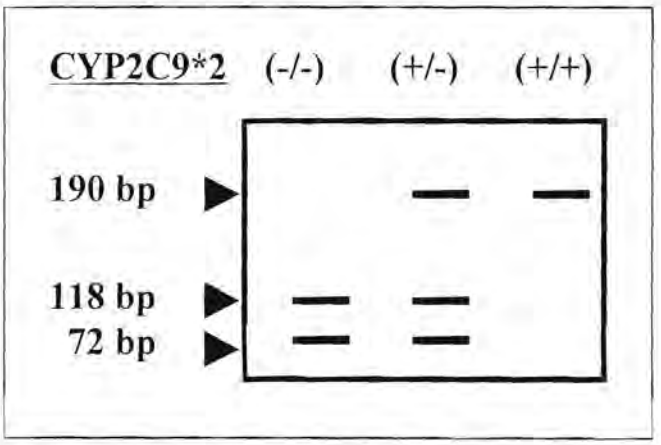


Figure 2.1: The predicted banding patterns of the digested PCR products for detection of individuals that are (i) Homozygous, normal (CYP2C9*2, -/-) (ii) Heterozygous (CYP2C9*2, +/-); and (iii) Homozygous (CYP2C9*2 +/+) for the CYP2C9*2 polymorphism.

The presence of the CYP2C9*2 polymorphism is identified by an undigested band at 190bp. If a single nucleotide substitution has not occurred the 190 bp DNA fragment is digested and two bands of 118 and 72 bp are observed following gel electrophoresis (51). Heterozygotes have a banding pattern that is a combination of both patterns.

2.4.1 Optimization of the PCR methodology for the detection of the CYP2C9*2 polymorphism.

To be able to identify the CYP2C9*2 polymorphism it was necessary to optimize the PCR methodology of Stubbins *et al* (51). Four parameters namely the $MgCl_2$ concentration, primer concentration, template concentration and the effect of adjuvents such as DMSO, glycerol and BSA were optimized. The presence of a single band of the correct size of 190 bp with no secondary product formation, were considered as criteria for successful PCR amplification. Besides the compound being tested the concentration of all other components were the same as presented in Section 2.4.2. Cycling conditions were also kept as described in Section 2.4.2.

2.4.1.1 Optimization of $MgCl_2$ concentration.

Different concentrations of $MgCl_2$ prepared from a 25 mM $MgCl_2$ stock solution were evaluated. Volumes of 1 to 3µl representing a concentration ratio of 1-3 mM, were evaluated in a final reaction volume of 25 µl (or 0.04-0.12 mmol/µl).

2.4.1.2 Optimization of DNA concentration .

The optimal DNA concentration was determined using different volumes of 1- 5 µl DNA that represents a DNA concentration range of 39-195 ng in a reaction volume of 25 µl (or 1.56/7.8 ng/µl).

2.4.1.3 Optimization of primer concentration.

A stock PCR primer solution of 100 pmol/µl was prepared by adding sterile double distilled water to the purchased freeze-dried primers. Volumes of 10µl were aliquoted into micro-centrifuge tubes and stored at –20 °C. Working primer solution of 10 pmol/µl were prepared by making a further 10 times dilution of the primer stock solution. A primer concentration range of 10-50 pmol in a reaction volume of 25µl was evaluated (or 0.4-2 pmol/µl).

2.4.1.4 The effect of BSA, DMSO and glycerol on PCR amplification.

The effect of different adjuvants such as DMSO, BSA and glycerol in a reaction volume of 25 µl was also evaluated. A BSA concentration range of 0.2-1µg/µl prepared from a BSA stock solution of 10mg/ml was evaluated. For DMSO and glycerol a concentration range of 2-10% and 1-5%, respectively, was evaluated.

2.4.2 Amplification of the 190bp region containing the CYP2C9*2 polymorphism.

PCR amplification was carried out in total volume of 25µl consisting of a 2.5µl volume of genomic DNA and 22.5µl PCR reaction mix. The primer sequences obtained from Stubbins *et al* (51) and are shown in Table 2.1.

Table 2.1: Primer sequences for the detection of the CYP2C9*2 polymorphism (51).

Primer	Sequence
CYP2C9*2 forward primer	**5'-acg tga att ctt cct gtt agg aat tgt t-3'**
CYP2C9*2 reverse primer	**5'-tca ggg atc cgg tca ccc acc ctt ggt t-3'**

The dNTP working solution of 10 mM of each primer was prepared from a dATP, dTTP, dCTP and dGTP stock solutions of 100 mM. The working dNTP solution was prepared by mixing together 10µl of each of the 100 mM dGTP, dCTP, dATP and dTTP and 60µl H_2O. Aliquots of 10µl were prepared and stored at –20°C. All other reagents were also stored at –20°C and thawed to room temperature.

For CYP2C9*2 PCR amplification of 10 tubes the following mixture of all PCR reagents was prepared and reagents were added in the following order into a 1.5 ml centrifuge tube: 173µl water, 25µl of 10 x PCR buffer, 15µl of 25 mM $MgCl_2$, 5µl of 10 mM dNTP, 3µl of the two 10pmol/µl primers and 1µl of 5U/µl Taq DNA polymerase. A volume of 22.5µl of the PCR mixture was transferred into each of the ten 600µl thin walled PCR tubes. A 2.5µl volume of genomic DNA, which had been mixed by gentle vortexing, was added to the tubes. The caps of the tubes were closed, the contents were mixed by vortex and the samples were centrifuged for 30 seconds at 8000rpm.

The final PCR reaction consisted of 200µM each of the nucleotides dATP, dGTP, dCTP, and dTTP, 10mM Tris-HCl (pH 9.0 at 25°C), 50mM KCl and 0.1% Triton X-100, 1.5 mM $MgCl_2$, 0.5U Taq polymerase and 30 pmol of each primer.

A PCR tube containing all PCR components except DNA was included as a negative control. An additional two tubes, a positive (CYP2C9*2 (+/+)) and a negative control (CYP2C9*2 (-/-)) were also included. The controls were identified by restriction enzyme digestion and electrophoretic analysis. DNA sequencing further confirmed the presence or absence of the CYP2C9*2 polymorphisms (Section 2.7). These controls were included in each batch of samples that were subjected to PCR amplification and restriction enzyme digestion.

The samples were placed in a Hybaid Touchdown Thermocycler and subjected to the following cycling conditions. One cycle of denaturation (94°C for 5 minutes), 30 cycles of denaturation (94°C for 1 minute), annealing (60°C for 1 minute) and extension (72°C

for 1 minute). A final extension step was performed at 72°C for 5 minutes. No oil overlay was used during amplification.

2.4.2.1 Ava II digestion of the CYP2C9*2, 190 bp product.

The CYP2C9*2 PCR 190bp product was digested with 4 U Ava II. The CYP2C9*2 restriction enzyme cocktail consisted of 0.4µl restriction enzyme Ava II (10U/µl), 3µl of 10 x Buffer C (100mM Tris/HCl (pH 7.9), 500mM NaCl, 100mM $MgCl_2$, and 10mM DTT at 37°C), 0.3µl acetylated BSA (0.1mg/ml) and 1.3 µl H_2O to a final volume of 5 µl. After adding the restriction enzyme cocktail (5µl) to the PCR product (25µl), the sample was mixed well centrifuged at 800 rpm for 10 seconds and incubated overnight at 37°C in a Merck D-64271 Incubator.

2.4.2.2 Polyacrylamide Gel Electrophoresis (PAGE) of PCR amplification products and restriction enzyme digests.

The 190bp CYP2C9*2 PCR product and PCR digests were analyzed by polyacrylamide gel electrophoresis (PAGE). A 5µl volume of loading buffer (70% sucrose containing 0.01% BPB) was added to each sample. Each were mixed, centrifuged and loaded onto a 10% PAGE (10% T, 2.9%C) gel and separated at room temperature on a Hoefer Mighty Small Electrophoresis System in 1x TBE buffer at 140V for 1 hour.

A 30% stock solution of acrylamide: methylene bisacrylamide (T=30%, C= 8.7%) was prepared. Ten percent PAGE gels were prepared by adding 3.3ml of 30% polyacrylamide and 1ml 10 x TBE buffer, to 5.6ml H_2O in a small beaker. A 40µl aliquot of a 20% ammonium persulphate solution and 40µl TEMED was added, and the solution was mixed well. Gels with well size of 64 mm^3 were prepared using the Hoeffer Mighty Small SE 245 dual gel caster. The gels were allowed to polymerize for 10 minutes at room temperature. The combs were removed and the wells rinsed with 1XTBE, the gels were removed from the casting apparatus and placed onto the Hoeffer Mighty Small

Electrophoresis Unit. The samples were loaded and the gels were run using a Pharmacia Biotech Power Pack set at 140V, 42mA and 6W. The gels were removed from between the glass plates placed in a 0.0001% EtBr staining solution prepared from a 10mg/ EtBr stock solution for 4 minutes. The bands were visualized with ultra violet (UV) light and photographed with an UVIdoc Gel Documentation System.

A molecular weight size marker (phiX174/Hinf 1) (726-24 bp) was included in each electrophoretic run. The DNA size marker was prepared by digesting 1µg ØX174 with 10U Hinf 1 and 3 µl 10x Buffer B [6mM Tris-HCl, 6mM $MgCl_2$, 50mM NaCl, 1mM dithiothreitol (DTT), (pH 7.5 at 37°C)] in 25 µl ddH_2O at 37°C overnight.

2.5 PCR strategy for the detection of the CYP2C9*3 polymorphism.

The CYP2C9*3 nucleotide substitution does not naturally create or destroy a restriction site. A restriction site was forced into the forward primer '5- aat aat aat atg cac gag gtc cag aga tgc-3'(where **g** is the site of the forced mismatch), such that A1061 in combination with the mismatch creates a restriction site (a tgc **at**)(**a** = A1061) for the Nsi I restriction enzyme (41). The CYP2C9*3 A1061C substitution removes this restriction site and the 165 bp PCR amplification product is undigested. Restriction enzyme digestion of the normal allele at the site of the forced mismatch results in two fragments of 135 and 30 bp. From the banding patterns three different genotypes could be identified, CYP2C9*3, homozygous normal (-/-), heterozygous (+/-) and homozygous (+/+) for this polymorphism (Figure 2.2).

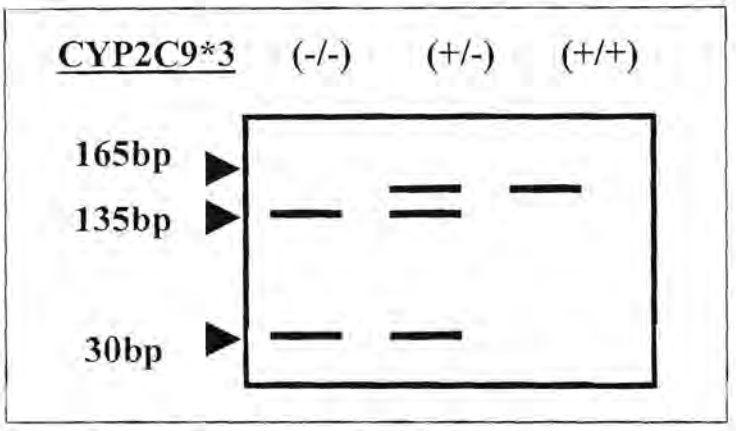


Figure 2.2: The predicted banding patterns of the digested CYP2C9*3 PCR product (165bp) for detection of individuals that are (i) Homozygous, normal (-/-) (ii) Heterozygous (+/-); and (iii) Homozygous (+/+) for the CYP2C9*3 polymorphism.

2.5.1 Optimization of PCR methodology for the detection of the CYP2C9*3 polymorphism.

To be able to identify the CYP2C9*3 polymorphism it was necessary to optimize the PCR methodology of Sullivan-Klose *et al* (41). Four parameters namely $MgCl_2$ concentration, primer concentration, DNA concentration and the effect of adjuvants such as DMSO, glycerol and BSA were optimized. Following gel electrophoresis (Section 2.4.2.2) the presence of a single band of the correct molecular mass of 165 bp and no secondary product formation, were considered as criteria for successful PCR amplification. Besides the component being optimized the concentration of all other components were the same as in Section 2.5.2. Cycling conditions were also the same as described in Section 2.5.2.

2.5.1.1 Optimization of $MgCl_2$ concentration.

The effect of $MgCl_2$ on the formation of a CYP2C*3 amplification was evaluated as described in Section 2.4.1.1.

2.5.1.2 Optimization of DNA concentration.

The optimal DNA concentration for successful amplification of the 165 bp CYP2C9*3 product was determined. In a final reaction volume of 24µl a DNA concentration range of 1.62-11.3 ng/µl was evaluated.

2.5.1.3 Optimization of primer concentration.

A 10 pmol/µl primer working solution was prepared as described in Section 2.4.1.3. A primer concentration range of 10-50 pmol or 0.42-2.1 pmol/µl per 24µl reaction volume was evaluated.

2.5.1.4 The effect of BSA, DMSO and glycerol on PCR amplification.

The effect of different adjuvants on the amplification of the CYP2C9*3 PCR product was evaluated in a reaction volume of 24µl. A BSA concentration range of 10 - 50µg/µl, prepared as described in Section 2.4.1.4 was evaluated. Different concentrations of DMSO (2-10%), and glycerol (1-5%) were also evaluated.

2.5.2 Amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism.

The PCR amplification was carried out in a total volume of 24µl consisting of 4µl genomic DNA and 20µl PCR reaction mix using the primer sequences of Sullivan-Klose *et al* (41) (Table 2.2).

All controls; PCR components except DNA, PCR components and DNA either from an individual that had previously tested homozygous polymorphic (+/+) or homozygous, normal (-/-) for the CYP2C9*3 polymorphism were included in each batch of samples that were subjected to PCR amplification and restriction enzyme digestion.

Table 2.2: Primer sequences for the detection of the CYP2C9*3 polymorphisms (41).

Primers	Sequence
CYP2C9*3 forward primer	**5'-aat aat aat atg cac gag gtc cag aga tgc-3'**
CYP2C9*3 reverse primer	**5'-gat act atg aat ttg gga ctt c-3'**

For a PCR amplification of 10 tubes the following mixture of all PCR reagents was prepared and reagents were added in the following order into a 1.5 ml micro centrifuge tube: 140 µl H_2O, 25µl of 10 x PCR buffer, 20µl of 25 mM $MgCl_2$, 5µl of 1mM dNTP, 2µl each of 10 pmol forward and reverse primers, 1µl of 5U/µl Taq DNA polymerase, and 5µl BSA. A volume of 20µl of the PCR mixture was aliquoted into ten 600µl PCR micro-centrifuge tubes. A 4µl volume of genomic DNA, which had been mixed by gentle vortexing, was added to the tubes. The caps of the tubes were closed, the contents were mixed and the samples were centrifuged for 30 seconds at 8000rpm.

The final PCR reaction in a volume of 25 µl consisted of 200µM each of dATP, dGTP, dCTP, and dTTP, 10mM Tris-HCl (pH 9.0 at 25°C), 50mM KCl, 0.1% Triton X-100, 2 mM $MgCl_2$, 0.5U Taq polymerase, and 2 pmol/µl of each primer.

The samples were placed in a Hybaid Touchdown Thermocycler and subjected to the following cycling conditions. One cycle of denaturation (94°C for 5 minutes), 35 cycles of denaturation at 94°C for 40 seconds, annealing at 50°C for 20 seconds and extension at 72°C for 20 seconds. A final extension step was performed at 72°C for 5 minutes.

2.5.2.1 Nsi I, restriction enzyme digestion and PAGE of the CYP2C9*3 digests.

The 24 µl CYP2C9*3 PCR product was digested with 2.5 U/µl Nsi I. The restriction enzyme cocktail consisted of 0.4µl restriction enzyme NsiI (10U/µl), 3.0µl of the 10x

Buffer D (60mM Tris-HCl (pH 7.9), 1.5M NaCl, 60mM $MgCl_2$ and 10mM DTT at 37°C), 0.3µl BSA (0.1mg/ml) and 1.3µl H_2O. A total volume of 5µl restriction enzyme cocktail was added to the PCR product and then incubated overnight at 37°C in a Merck D-64271 Incubator. The digested products and size marker (ØX174 DNA/Hinf 1) were separated by electrophoresis at 140V for 60 minutes on a polyacrylamide (10%T, 2.9%C) gel and visualized as described in Section 2.4.2.2. The smaller fragment of 30 bp is eluted from the gel and individuals that are CYP2C9*3 homozygous, normal (-/-) are identified by the presence of a 135 bp band. The 165bp product is undigested in individuals homozygous for the CYP2C9*3 polymorphism. In heterozygouss, where only one allele is polymorphic, two bands of 165 and 135 bp are present.

2.6 The stability of a CYP2C9*2 and *3 PCR master mix after storage at -70°C.

Twenty-two comma five micro-liter aliquots of the PCR reaction mix for the CYP2C9*2 and 20µl aliquots for the CYP2C9*3 polymorphism were prepared as described in Section 2.4.2 and 2.5.2. and stored at –70°C. The stability of the PCR reaction mixes were evaluated by visualizing the PCR product formed, (190bp and 165bp for the CYP2C9*2 and CYP2C9*3 polymorphisms respectively) following polyacrylamide gel electrophoresis (Section 2.4.2.2) and EtBr staining. Specific attention was given to factors such as band intensity, the presence of smears and primer dimers. This evaluation was conducted 6 and 11 days after storage of the prepared PCR mix at –70°C.

2.7 Sequencing of CYP2C9*2 and CYP2C9*3 positive and negative controls.

The PCR product from an individual that tested homozygous positive (+/+) and another that tested homozygous negative (-/-) for the CYP2C9*2 and CYP2C9*3 polymorphisms was sequenced at the DNA Sequencing Unit at the University of Stellenbosch. DNA was amplified as described in Section 2.4.2 and 2.5.2 for the CYP2C9*2 and CYP2C9*3 polymorphisms respectively. The product of PCR amplification was loaded on a 1%

agarose gel with 1x Tris/Acetate/EDTA (TAE) as running buffer. Ten times TAE buffer consisting of 0.4M Tris, 0.02M EDTA and 20% glacial acetic acid was prepared. The PCR product was subjected to electrophoresis until the band (190 and 165bp respectively) was separated from other components. The bands containing the PCR product were dissected from the gel using a sterile scalpel. The DNA fragments were isolated using a Roche High Pure PCR Product Purification kit (64). The weight of the excised agarose slice was determined by weighing the slice in a pre-weighed tube. A volume of 300µl of Binding Buffer was added to the tube for every 100mg of agarose. The agarose gel slice was dissolved by vortexing the tube containing the agarose gel with Binding Buffer for 30 seconds. The suspension was incubated for 10 minutes at 56 °C.

Once the gel slice was completely dissolved, 150µl of isopropanol was added for every 100mg agarose and the contents of the tube was again vortexed thoroughly. A High Pure filter tube was inserted into a collection tube and the content of the micro-centrifuge tube was added to the upper reservoir of the filter tube. Care must be taken not to exceed the specified volume of 700µl per filter tube. The filter tube was then centrifuged for 20 seconds at 8000 rpm in a Hereus Biofuge 15 Centrifuge. The flow through was discarded and the filter tube re-connected to the same collection tube. As part of the first washing step 500µl of Washing Buffer 1 was pippeted into the upper reservoir of the filter tube and centrifuged for one minute at 6000 x g (8000 rpm). The flow through was again discarded and the filter tube re-connected to the same collection tube. As part of the second washing step 200µl of Washing Buffer 2 was pippeted into the upper reservoir of the filter tube and centrifuged for one minute at 6000 x g. After discarding the flow through and collection tube, the filter tube was inserted into a clean 1.5 ml micro-centrifuge tube. The purified fragments were eluted by adding 80µl elution buffer to the upper reservoir of the filter tube followed by centrifugation at 6000 x g for one minute. The micro-centrifuge tube now contained the purified DNA fragments, which were stored at 4°C.

2.7.1 Stability of the primers and DNA fragments used for sequencing.

For sequencing of the isolated PCR fragments the samples and primers were usually packed into a cool box together with an ice brick. The samples were then sent by an overnight courier service to the DNA Sequencing Unit at the University of Stellenbosch. In an effort to reduce the cost of transport, the possibility of sending the fragments and primers ordinary registered mail was considered. The stability of the DNA and primers had to first be determined. The isolated PCR product and primers were left on the laboratory bench for a period of 5 days at room temperature. After this period the primers and fragments were used together in a nested PCR reaction. The same primers and an aliquot of the PCR product stored at 4°C served as a control for the nested PCR. All conditions for amplification were the same as described in Section 2.4.2 and 2.5.2.

2.7.2 The sequences of the 190 and the 165 bp fragments containing the CYP2C9*2 and *3 polymorphisms.

The sequence for the regions containing the CYP2C9*2 and CYP2C9*3 polymorphisms were obtained from Blast (65) placed by Romkes (23) and de Morais *et al* (66) for each polymorphism respectively. The data obtained from sequencing was compared to these DNA regions.

2.8 Population screening for the CYP2C9*2 and 3 polymorphisms in the Control and the Thrombosis Clinic population group.

The number of individuals that were homozygous normal (-/-), heterozygous, (+/-) homozygous polymorphic (+/+) was determined for the CYP2C9*2 and *3 polymorphisms respectively. Both the Control and the Thrombosis Clinic population were screened. If no clear result were obtained for an individual, the test was repeated. If the result was still unclear the original blood sample was thawed and a new sample of genomic DNA prepared and the entire procedure repeated. The carrier frequency was

determined for each genotype on the basis of the observed genotype distribution that was calculated. The expected occurrence rates of the different genotypes within each group was calculated according to the Hardy-Weinberg principle and compared with the observed occurrence rates by the Chi-square test.

As compound heterozygosity may occur the frequency of compound heterozygouss CYP2C9*2/CYP2C9*3 was determined, tabulated and the observed genotype distribution was calculated and compared to the observed frequency by the Chi-square test. The prevalence of the polymorphisms in the Control and the Thrombosis Clinic population was also compared. Differences between the populations were determined using the Chi Square test.

2.9 Correlation between CYP2C9*2 and *3 genotypes and warfarin maintenance dosages.

The Anticoagulation Clinic group was divided into two groups patients (i) homogenous normal and (ii) those with one or two copies of either the CYP2C9*2 and *3 polymorphism. Dr. de Wet (67) from the Department of Haematology determined the mean maintenance dosage for each group as part of a collaboration study with the Department of Chemical Pathology.

Chapter 3:

Results.

The prevalence of the CYP2C9*2 and CYP2C9*3 polymorphisms in the Control and the Anticoagulation clinic population was determined. To enable population screening each of the screening methodologies needed to be optimized. The nucleotide sequence of the DNA PCR amplification products of the positive and negative controls, for the CYP2C9*2 and CYP2C9*3 polymorphisms was determined.

3.1 Subjects.

DNA from 200 white male medical students served as the Control population that represents the normal healthy South African Caucasian population. The DNA from this population group had been isolated 18 months previously and had been stored at 4°C. A single tube of citrate blood was collected from 144 patients attending an Anticoagulation Clinic. Informed consent was obtained from each of these individuals. DNA was isolated from each blood sample and was stored at 4°C.

3.2 Determination of DNA concentration and quality.

3.2.1 Electrophoretic analysis of isolated DNA.

The DNA used to determine the frequency of the CYP2C9*2 and *3 polymorphisms in the Control population had been isolated 18 months previously. It was necessary to determine whether this DNA was still of sufficient quality for successful PCR amplification. Recently isolated DNA from the Anticoagulation clinic population served as the control. The integrity of five DNA samples of the Control population group and three DNA samples from the Anticoagulation clinic population were compared. The samples were loaded onto a 1% agarose gel containing EtBr (Figure 3.1), following

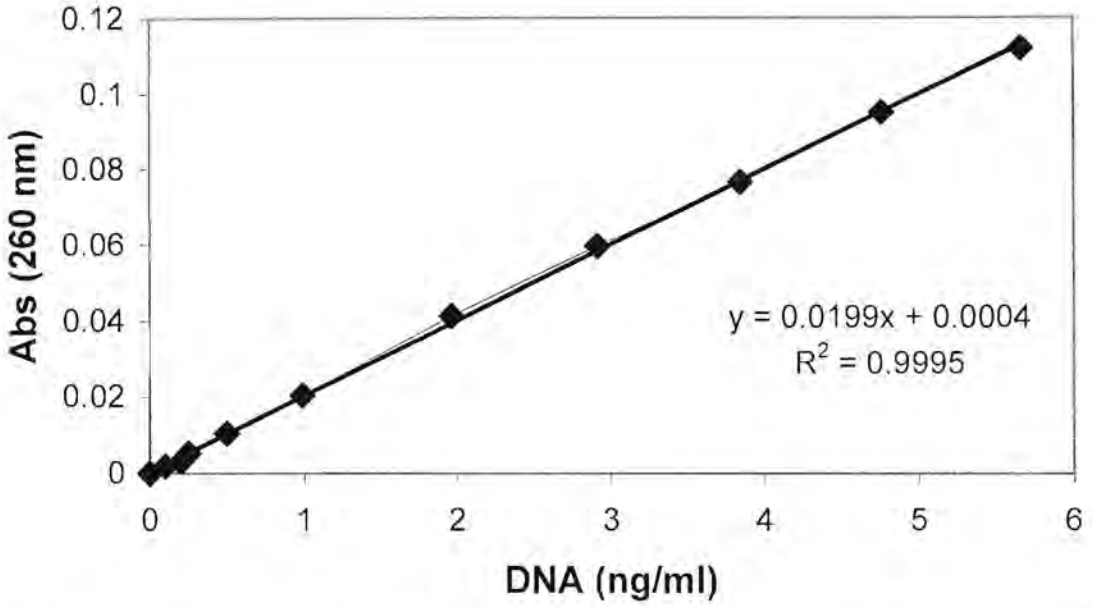


Figure 3.2: DNA standard curve used to determine the concentration of the isolated DNA samples.

3.3 Optimization of the PCR methodology for the detection of the CYP2C9*2 polymorphism.

The CYP2C9*2 polymorphism was determined following the amplification of the cytochrome P450 region of exon 3. Before the frequency of the CYP2C9*2 could be determined, it was necessary to optimize various parameters such as $MgCl_2$, primer and DNA concentration. The effect of adjuvants such as BSA, DMSO and glycerol was also determined. A single band of 190 bp without secondary product formation or streaking was considered as criteria for successful PCR amplification. To enable the rapid screening of large populations the possible storage of the prepared mix of all reagents at –70°C was evaluated.

3.3.1 Optimization of $MgCl_2$ concentration.

The effect of different $MgCl_2$ concentrations on the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism was determined. The final PCR reaction in a volume of 25 µl consisted of 200µM each of dATP, dGTP, dCTP, and dTTP, 10mM

3.4.1 Restriction Enzyme Digestion, Polyacrylamide Gel Electrophoresis and Genotype Analysis for the CYP2C9*2 polymorphism.

The CYP2C9*2 190 bp PCR product was digested with different concentrations (3U, 4U and 5U per reaction volume) of AvaII. The restriction enzyme digests were separated on a 10% polyacrylamide gel, stained with EtBr so that the different banding patterns could be evaluated. At the lower AvaII concentration of 3U in an individual where the CYP2C9*2 polymorphism is absent on both alleles, digestion of the 190 bp fragment was incomplete (Figure 3.6, Lane 5). At a higher Ava II concentration (4U) digestion was complete without any detectable undigested product at 190bp.

The presence of the CYP2C9*2 polymorphism in an allele, results in the loss of a restriction site for Ava II. In an individual homozygous for this polymorphism a single band at 190bp is observed (Figure 3.6, Lane 4). If the CYP2C9*2 polymorphism is absent, there will be a restriction site for Ava II, and the 190bp fragment will be digested, forming two bands of 118bp and another band at 72bp.Therefore in an individual where the polymorphism is absent from both alleles two bands of 118 and 72bp are observed following gel electrophoresis (Figure 3.6, Lane 5). In an individual with the polymorphism on one of the two alleles (Heterozygous, Figure 3.6, Lane 2 and 3), half of the PCR product (normal allele) will be digested and be visualized as bands at 118bp and 72bp. The remaining PCR product will not be digested and a band at 190bp will be observed.

A DNA size marker (ØX174/Hinf 1) (726-24bp) was included in each electrophoretic run (Figure 3.6, Lane 1). The marker was digested by digesting ØX174 DNA with the restriction enzyme Hinf1. The DNA size marker produced following electrophoresis a marker ladder comparable with commercially available products such as, ØX174/Hinf 1 dephosphorilated markers, produced by Promega and supplied by Whitehead Scientific. This reduced costs significantly. The inclusion of a negative control that contains all components of the PCR except genomic DNA was used to exclude the possibility of contamination of the cocktail with endogenous DNA.

3.7 Method development for screening for CYP2C9*3 polymorphisms.

3.7.1 Optimization of PCR methodology for the detection of the CYP2C9*3 polymorphism.

The CYP2C9*3 polymorphism was determined following the amplification of the cytochrome P450 region of exon 7. Before the frequency of the CYP2C9*3 polymorphism could be determined in the Control and the Anticoagulation clinic population, it was necessary to optimize the concentration of various constituents of the CYP2C9*3 PCR reaction such as $MgCl_2$, primer and DNA concentration. Furthermore the effect of adjuvants such as BSA, DMSO and glycerol was also determined. As in Section 3.3 for the CYP2C9*2 polymorphism a single band without secondary product formation or streaking will be considered as criteria for successful PCR amplification. To enable the rapid screening of large population groups the storage conditions for aliquots of prepared PCR mix was determined.

3.7.1.1 Optimization of $MgCl_2$ concentration.

The effect of different $MgCl_2$ concentrations on the amplification of the 165 bp fragment containing the CYP2C9*3 polymorphism was determined. In a PCR reaction mix of 24 µl containing 2.5 µl PCR buffer, 200µM dNTP's, 2 pmol (1.2 pmol/µl) of each primer and 0.5 U Taq DNA polymerase different concentrations of $MgCl_2$ was added representing a concentration range of 1.0-3.0 mM $MgCl_2$ (Figure 3.9).

At all concentrations of $MgCl_2$ evaluated a single band of 165bp was observed. A final $MgCl_2$ concentration of 2 mM was chosen for all subsequent amplification reactions.

3.11 Reduction of PCR reaction volume for the detection of the CYP2C9*3 polymorphism.

To allow the more cost effective screening of a large population for the CYP2C9*3 polymorphism the PCR reaction volume was reduced from 24 µl to15 µl including 2.4µl DNA as substrate. Sufficient PCR product formed for successful amplification and identification of the different genotypes similar to that shown in Figure 3.13.

3.12 Sequencing of CYP2C9*2 and CYP2C9*3 negative controls.

The 190 and 165 bp PCR fragment containing the CYP2C9*2 and CYP2C9*3 polymorphisms was sequenced to determine the exact nucleotide sequence and to verify the position of each nucleotide substitution.

3.12.1 Preparation of samples for sequencing.

The samples were sent to the DNA Sequencing Unit at the University of Stellenbosch where the samples were sequenced using an ABI Prism Model 3100 version 3.7, DNA sequencer. DNA was amplified as described in Section 2.4.2 and 2.5.2 for the CYP2C9*2 and CYP2C9*3 polymorphisms respectively. The PCR amplification products were loaded on a 1% agarose gel (Section 2.3.1.1) in TBE buffer. The samples were subject to gel electrophoresis and the bands containing either the CYP2C9*2 or *3 PCR products (Figure 3.15) were cut from the gel.

Table 3.1: Nucleotide sequence for the 190bp DNA fragment containing the CYP2C9*2 polymorphism.

CYP2C9*2	Nucleotide Sequence
Homozygous, -/-	
Forward	ttn cnc act cct ccc aag gca gcg cgc ttc ctc ttg aac ac***g*** gtc ctc aat gct cct ctt ccc cat ccc aaa att ccg cag cgt cat gag gga gaa acg ccg gat ctc ctt cca ttt ctt tcc att gct gaa **aac aat tcc taa cag gaa gaa ttc acg t**
Reverse	ctn cag caa tgg aaa aga aat gga agg ccc atc cgg cga ttt ctc cct cat gac gct gcg gaa ttt tgg gat ggg gaa gag gag cat tga gga c***c***g tgt tca aga gga agc ccg ctg cct tgt gga gga gtt gag aaa **aac caa ggg tgg gtg acc gga tcc ctg** aan

Bold red: Primer sequence
Bold Italic blue: Site of polymorphism

Table 3.2 Nucleotide sequences for the 165bp DNA fragment containing the CYP2C9*3 polymorphism.

CYP2C9*3	Nucleotide Sequence
Homozygous, -/-	
Forward	tgg gaa tka kna wag ttt ctg aat tta atg tca cag gtc act gca tgg gcg cag gct ggt ggg gag aag gtc ***atgc atc*** **tct gga cct tcg tgc ata tta tta tt**
Reverse	ctg ccc cag tga cct gtg aca tta aat tca gaa act atc tca ttc cca agg taa gtt tgt ttc tcc tac tgc aac tcc atg ttt tcg **aag tcc caa att cat agt atc a**

Bold red: Primer sequence
Bold Italic blue: Site of polymorphism
Bold Italic green: Site of forced mismatch

The site of the primers were identified and the sequences were compared to sequences obtained by the NCBI Annotation project (65) for the CYP2C9*2 gene in exon 3 and the CYP2C9*3 gene in exon 7. For the CYP2C9*3 PCR method a restriction site is created

with the insertion of a mismatch and therefore only the reverse sequence can be used to identify the site of the polymorphism.

3.13 Genotype distribution, gene and allele frequencies of the CYP2C9*2 and *3 alleles in a Control population.

3.13.1 Genotype distribution, gene and allele frequencies of the CYP2C9*2 allele in the Control population.

A population of 200 Caucasian students were screened for the CYP2C9*2 and the CYP2C9*3 polymorphism. The frequency of each polymorphism was tabulated and statically evaluated. The frequency of compound heterozygosity, the presence of both alleles within a single individual was also determined.

In this population group (Table 3.3a) 150 were homozygous normal, 40 were heterozygous and 10 homozygous for the CYP2C9*2 polymorphism. The CYP2C9*2 polymorphism in this population was found to be in Hardy-Weinberg equilibrium. The data was evaluated using the Chi-square goodness of fit and an X^2 of 9.3 (Table 3.3b) was obtained. A carrier and a gene of 0.255 and 0.023 were obtained respectively. A gene frequency of 0.023 implies 2.3% of the population screened is CYP2C9*2 homozygous.

Table 3.3a Distribution of CYP2C9*2 in the Control population (n=200).

Genotype	Prevalence	
	Observed	Expected
Homozygous CYP2C9*1/*1 (-/-)	150	144.5
Heterozygous CYP2C9*1/*2 (+/-)	40	51.0
Homozygous CYP2C9*2/*2 (+/+)	10	4.5

Carrier Frequency= 0.255, Gene Frequency= 0.023

Table 3.3b Chi-square for goodness of fit for the CYP2C9*2 in the Control population (n=200).

Genotype	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
Homozygous CYP2C9*1/*1 (-/-)	150	144.5	5.5	30.25	0.21
Heterozygous CYP2C9*1/*2 (+/-)	40	51.0	11.0	121.0	2.37
Homozygous CYP2C9*2/*2 (+/+)	10	4.5	5.5	30.25	6.72

X^2_{2df}=9.3, P($X^2_{0.995}$)= 10.60 (Difference not significant)

All differences were not significant at a X^2 of 0.99, whereas the CYP2C9*2 polymorphism with a chi-square of 9.3 was not significant at a X^2 value of 0.995.

3.13.2 Genotype distribution, gene and allele frequencies of the CYP2C9*3 allele in the Control population.

The same population of 200 individuals were also screened for the CYP2C9*3 polymorphism. In this population group (Table 3.4a) 161 were homozygous normal, 38 were heterozygous and 1 was homozygous for the CYP2C9*3 polymorphism. This population for the CYP2C9*3 polymorphism is in Hardy Weinberg equilibrium. The data was evaluated using the Chi-square goodness of fit and a X^2 of 0.616 (Table 3.4b) was obtained. A carrier frequency of 0.18 and a gene frequency of 0.01 were found in this Control population. Consequently 1 in 100 (1%) individuals in the population screened are homozygous for the CYP2C9*3 polymorphism.

Table 3.4a Distribution of the CYP2C9*3 polymorphism in the Control population (n=200).

Genotype	Prevalence	
	Observed	Expected
Homozygous CYP2C9*1/*1 (-/-)	161	162.0
Heterozygous CYP2C9*1/*3 (+/-)	38	36.0
Homozygous CYP2C9*3/*3 (+/+)	1	2.0

Carrier Frequency: 0.18, Gene Frequency: 0.01

Table 3.4b Chi-square for goodness of fit for the CYP2C9*3 polymorphism in the Control population (n=200).

Genotype	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
Homozygous CYP2C9*1/*1 (-/-)	161	162.0	1.0	1.0	0.006
Heterozygous CYP2C9*1/*3 (+/-)	38	36.0	2.0	4.0	0.11
Homozygous CYP2C9*3/*3 (+/+)	1	2.00	1.01	1.00	0.5

$X^2_{2df}=0.616$, $P(X^2_{0.99})=9.21$ *(Difference not significant)*

3.13.3 Distribution of both CYP2C9*2 and *3 polymorphism in the Control population.

The distribution of both the CYP2C9*2 and *3 alleles was determined within the same population control group. In this population group 10 individual were homozygote for the CYP2C9*2 and 1 individual for the CYP2C9*3 polymorphism (Table 3.5 a and b). Ten individuals were compound heterozygous with a copy of both the CYP2C9*2 and the CYP2C9*3 polymorphism. Two individuals were CYP2C9*2 homozygous and CYP2C9*3 heterozygous.

Table 3.5a Distribution of both the CYP2C9*2 and CYP2C9*3 polymorphisms in the Control population (n=200).

Genotype		Prevalence	
CYP2C9*2	CYP2C9*3	Observed	Expected
-/-	-/-	124	116.9
-/-	+/-	25	25.99
-/-	+/+	1	1.44
+/-	-/-	29	41.31
+/-	+/-	11	9.18
+/-	+/+	0	0.51
+/+	-/-	8	3.64
+/+	+/-	2	0.81
+/+	+/+	0	0.045

Table 3.5b Chi-square for goodness of fit for both the CYP2C9*2 and *3 in the Control population (n=200).

Genotype		Prevalence				
CYP2C9*2	CYP2C9*3	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
-/-	-/-	124	116.9	7.1	50.41	0.43
-/-	+/-	25	25.99	0.99	0.98	0.038
-/-	+/+	1	1.44	0.44	0.19	0.13
+/-	-/-	29	41.31	12.31	151.5	3.67
+/-	+/-	11	9.18	1.82	3.31	0.36
+/-	+/+	0	0.51	0.51	0.26	0.51
+/+	-/-	8	3.64	4.36	19.0	5.22
+/+	+/-	2	0.81	1.19	1.41	1.74
+/+	+/+	0	0.045	0.045	0.0020	0.045

X^2_{4df}=12.14, P($X^2_{0.99}$)=13.28 (Difference not significant)

3.14 Genotype distribution, gene and allele frequencies of the CYP2C9*2 and *3 alleles in the Anticoagulation clinic population.

3.14.1 Genotype distribution, gene and allele frequencies of the CYP2C9*2 allele in the Anticoagulation clinic population.

A population of 144 individuals attending an Anticoagulation Clinic were screened for the CYP2C9*2 polymorphism. In this population (Table 3.6a) 114 were homozygous normal, 29 were heterozygous and 1 homozygous for the CYP2C9*2 polymorphism. A carrier frequency of 0.196 and a gene frequency of 0.012 were found in this population. For this specific population 1.2% were homozygous for the CYP2C9*2 polymorphism. The population was found to be in Hardy Weinberg equilibrium in this population. The data was evaluated using the Chi-square goodness of fit (Table 3.6b) and a X^2 of 0.32 was obtained. No differences were obtained in the prevalence of the CYP2C9*2 polymorphism in the Control and the Anticoagulation clinic population.

Table 3.6a Distribution of the CYP2C9*2 polymorphism in the Anticoagulation clinic population (n=144).

Genotype	Prevalence	
	Observed	Expected
Homozygous CYP2C9*1/*1 (-/-)	114	114.4
Heterozygous CYP2C9*1/*2 (+/-)	29	28.4
Homozygous CYP2C9*2/*2 (+/+)	1	1.64

Carrier frequency =0.196, Gene frequency =0.012

Table 3.6b Chi-square for goodness of fit for the CYP2C9*2 polymorphism in the Anticoagulation clinic population (n=144).

Genotype	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
Homozygous CYP2C9*1/*1 (-/-)	114	114.4	0.4	0.16	0.0013
Heterozygous CYP2C9*1/*2 (+/-)	29	27.5	1.5	2.25	0.08
Homozygous CYP2C9*2/*2 (+/+)	1	1.64	0.64	0.40	0.24

X^2_{2df}=0.32, P($X^2_{0.99}$)= 9.21 (Difference not significant)

3.14.2 Genotype distribution, gene and allele frequencies of the CYP2C9*3 allele in the Anticoagulation clinic population.

The same population was screened for the CYP2C9*3 polymorphism (Table1 3.7a). One hundred and twenty nine individuals were homozygous normal, 15 were heterozygous and none were homozygous for the CYP2C9*3 polymorphism. The population is in Hardy-Weinberg equilibrium with a Chi-square of 0.398 (Table 3.7b). A carrier frequency of 0.095 and a gene frequency of 0.0025 means that 3 in 1000 of the population attending the Anticoagulation clinic are homozygous for the CYP2C9*3 polymorphism. The carrier and the gene frequency for the CYP2C9*3 polymorphisms in

the Anticoagulation clinic population was lower than the Control population 0.095 vs 0.18 and 0.0025 vs 0.01.

Table 3.7a Distribution of the CYP2C9*3 polymorphism in the Anticoagulation clinic population (n=144).

Genotype	Prevalence	
	Observed	Expected
Homozygous CYP2C9*1/*1 (-/-)	129	129.14
Heterozygous CYP2C9*1/*3 (+/-)	15	14.48
Homozygous CYP2C9*3/3* (+/+)	0	0.38

Carrier Frequency=0.095, Gene Frequency=0.0025

Table 3.7b Chi-square for goodness of fit for the CYP2C9*3 polymorphism in the Anticoagulation clinic population (n=144).

Genotype	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
Homozygous CYP2C9*1/*1 (-/-)	129	129.14	0.14	0.0196	0.00015
Heterozygous CYP2C9*1/*3 (+/-)	15	14.48	0.52	0.27	0.018
Homozygous CYP2C9*3/*3 (+/+)	0	0.38	0.38	0.144	0.38

X^2_{2df}=0.398, P($X^2_{0.99}$)= 9.21 (Difference not significant)

3.14.3 Distribution of both CYP2C9*2 and *3 polymorphism in the Anticoagulation clinic population.

The distribution of both the CYP2C9*2 and *3 alleles was determined within the same Anticoagulation clinic population. In this population group 1 individual was homozygote for the CYP2C9*2 and none for the CYP2C9*3 polymorphism (Table 3.8a and b). Two individuals were compound heterozygous with a copy of both polymorphisms, (CYP2C9*2 and CYP2C9*3).

Table 3.8a Distribution of both the CYP2C9*2 and CYP2C9*3 polymorphisms in the Anticoagulation clinic population (n= 144).

Genotype		Prevalence	
CYP2C9*2	CYP2C9*3	Observed	Expected
-/-	-/-	101	101.2
-/-	+/-	13	11.14
-/-	+/+	0	0.22
+/-	-/-	27	25.2
+/-	+/-	2	2.78
+/-	+/+	0	0.05
+/+	-/-	1	1.40
+/+	+/-	0	0.155
+/+	+/+	0	0.003

Table 3.8b Chi-square for goodness of fit for both the CYP2C9*2 and *3 in Anticoagulation clinic population (n=144).

Genotype		Prevalence				
CYP2C9*2	CYP2C9*3	Observed	Expected	(O-E)	$(O-E)^2$	$(O-E)^2/E$
-/-	-/-	101	101.2	0.2	0.04	0.0004
-/-	+/-	13	11.14	1.86	3.45	0.31
-/-	+/+	0	0.22	0.22	0.048	0.22
+/-	-/-	27	25.2	1.8	3.24	0.128
+/-	+/-	2	2.78	0.78	0.608	0.21
+/-	+/+	0	0.05	0.05	0.0025	0.05
+/+	-/-	1	1.40	0.4	0.16	0.114
+/+	+/-	0	0.155	0.155	0.024	0.155
+/+	+/+	0	0.003	0.003	0.0009	0.003

X^2_{4df}=1.19, P($X^2_{0.99}$)=13.28 (Difference not significant)

3.15 Evaluation of the sensitivity and specificity of the CYP2C9*2 and *3 polymorphism screening methods.

The reproducibility of the two developed screening methods were evaluated in the first 117 patients tested in the Anticoagulation clinic group. The first screening produced 111 results for the CYP2C9*2 method and 117 results for the CYP2C9*3 method. These tests were repeated to confirm the obtained results. The second screening gave 116 results for the CYP2C9*2 method and again 117 results for the CYP2C9*3 method. Only one patient that tested as a carrier of the CYP2C9*2 polymorphism in the first run tested normal in the second run. Thus one false positive result was obtained for the CYP2C9*2 method. No discrepancies were found between the first and second screening of the 117 patients for the CYP2C9*3 method.

3.16 Correlation between CYP2C9*2 and *3 genotypes and warfarin maintenance dosages.

The Anticoagulation clinic group was divided into two groups patients (i) homogenous normal (n=99) and (ii) those with one or two copies of either the CYP2C9*2 and *3 polymorphism (n=45) (Figure 3.18). Dr. de Wet (67) determined the mean maintenance dosage for each group as part of a collaboration study with the Department of Chemical Pathology. For individuals that were homogenous normal the mean warfarin dosage was 4.85 mg/day while that for patients with at least one copy of either the CYP2C9*2 or *3 polymorphism the mean warfarin dosage was 3.71 (Figure 3.18) (67).

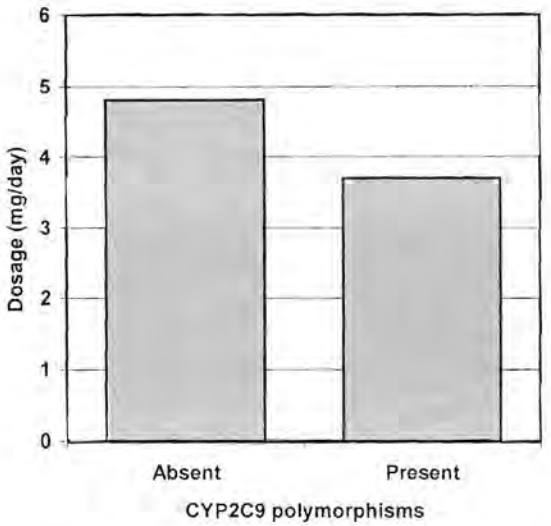


Figure 3.18: The effect of the polymorphisms on warfarin maintenance dosage (mg/day) of patients from the Anticoagulation clinic population was determined (67). A mean daily warfarin dose of 4.85 mg (SD=3.71) was calculated for patients belonging to the wild type group compared to 3.71 mg (SD=1.67) for polymorphic patients.

Chapter 4:

Discussion.

4.1 Introduction.

The South African population presents with genetic founder effects, either with unique mutations, or increased prevalence of mutations that are found in many other populations. Increased prevalence of unique mutations has been found in the South African population for the R95W mutation associated with porphyria variegate (21) and the D206E, V408M and the D154N mutations associated with familial hypercholesterolaemia (62). Other mutations that are found in many other populations, but occur in South Africa at an increased prevalence are the C282Y and H63D mutations associated with hereditary haemochromatosis (20). Confusion exists between the definition of mutations and polymorphisms. In this study a polymorphism is defined as follows: A polymorphism in the same way as many mutations, is a single nucleotide substitution. The occurrence of a polymorphism is usually high in a population and is not known to directly lead to disease, but is associated with increased risk for certain disorders. Homozygosity for MTHFR, characterized by a C677T nucleotide substitution, is a good example of a mutation. It has been found to be a risk factor for neural tube defect and cardiovascular disease (68). Whereas the CYP2C9*2 and CYP2C9*3 polymorphisms are associated with altered warfarin metabolism (4,5,6), but are not known to cause disease. A patient that is warfarin sensitive is six times more likely to have one or more of the described polymorphisms associated with impaired warfarin metabolism. In this study the gene frequency of these polymorphisms has been determined for the South African population. A high prevalence of CYP2C9 polymorphisms means the screening of patients attending an Anticoagulant Clinic receiving warfarin therapy should be implemented.

4.2 Optimization of analytical techniques.

The objectives of this study was firstly to establish and optimize screening methodologies for both the CYP2C9*2 and CYP2C9*3 polymorphisms. Both methods were then used to screen a random Caucasian Control (n=200) and an Anticoagulation Clinic (n=144) population for both polymorphisms. The allele and gene frequencies and the prevalence of compound heterozygosity for the CYP2C9*2 and *3 polymorphisms was determined for both

populations studied. The correlation between the presence of a CYP2C9 polymorphic allele and warfarin maintenance dosage was determined as part of a collaboration study with the Department of Haematology, University of Pretoria. Information obtained was used to develop testing strategies for patients that do not conform to the standard protocols of warfarin therapy.

The presence of the CYP2C9*2 and *3 polymorphisms were detected using restriction length polymorphism (RFLP) polymerase chain reaction (PCR) methodologies. Various RFLP PCR methods have been described for the detection of the CYP2C9*2 polymorphism (6,33,41). The CYP2C9*2 polymorphism is a C416T transition in exon 3 and results in an arginine (R) to cysteine (C) substitution at position 144 of the amino acid sequence of CYP2C9. Wang *et al* (33) described a method in which the forward primer bound up-stream at nucleotides 221-241 while the reverse primer bound a region further down stream at 469-449. Amplification results in a product of 420 bp. Restriction enzyme digestion of an individual homozygous for the CYP2C9*2 polymorphism results in two fragments of 363 and 57 bp. Taube *et al* (6) described a method where the forward primer bound at a location of 331-354 and a reverse primer at nucleotide position 555-581 and results in the amplification of a 190 bp fragment. The presence of the single nucleotide substitution, C to T at nucleotide 416 results in the loss of an Ava II restriction digestion site. The latter method was chosen as it results in a smaller fragment, which simplifies amplification optimization and the restriction enzyme recognizes the normal allele. A smaller fragment is also easier to resolve by polyacrylamide gel electrophoresis. Gel electrophoresis of the AvaII digested PCR product allows the identification of the different CYP2C9*2 genotypes.

Likewise methods for the detection of the CYP2C9*3 polymorphism have been described (4,33,41). A method adapted by Aithal *et al* (4) from Wang *et al* (33) involved the detection of both the normal and the CYP2C9*3 alleles. Similarly Sullivan–Klose (41) used two separate mismatched forward primers and an intron 7 reverse primer developed by de Morais *et al*. (66). The first mismatched primer introduced an Nsi I cleavage site in the CYP2C9*3 allele and the second mismatched primer introduced a Kpn I restriction site in the CYP2C9*1 allele. These primers were modified with the addition of 10 bases at the 5' end to aid in the detection of the fragments when resolved by gel electrophoresis. The reverse primer was intron-specific to eliminate unspecific binding. Amplification results in a 141bp PCR product. This method was very sensitive as it was used to identify the normal and the allele

containing the CYP2C9*3 polymorphism. An alternative method developed by Taube *et al* (6) uses a forward primer to create a restriction site for Nsi I in the normal allele. The CYP2C9*3 A1061C substitution removes this restriction site and the 165 bp PCR amplification product is undigested. This method was used for the screening of the Control and Anticoagulation Clinic populations for the CYP2C9*3 polymorphism. The restriction enzyme can identify a unique sequence either in the normal allele or in the allele where the point mutation has occurred. In both of the studied polymorphisms the restriction enzyme recognizes the sequence in the normal allele. Often the restriction enzyme will recognize several sites; one a common site to both alleles and a site where the point mutation has occurred. Such a common site serves as an internal control whereby the effectiveness of restriction digestion can be determined. Alternatively the restriction enzyme recognizes the region where the substitution occurs in the normal allele, where the polymorphism is absent, again this will also serve as a control to evaluate successful digestion as there are usually more individuals within a population that are negative for a mutation or a polymorphism. This is of great importance, especially if the presence of a rare mutation or polymorphism is being investigated.

In PCR the primers define the region to be amplified, which is usually the region containing the polymorphism of interest. Several factors play an important role in the success of a PCR reaction. This includes the quality and concentration of the DNA, Taq DNA polymerase, $MgCl_2$ and primer concentration.

The DNA used for the Control population had been stored for a period of 18 months at 4°C. For successful PCR amplification DNA should be intact, no fragmentation should have occurred (50). Agarose gel electrophoresis revealed that the stored DNA was intact without any fragmentation and was of the same quality as the freshly isolated DNA obtained from samples collected for the Anticoagulation Clinic group (Figure 3.1). The concentration of the DNA in the samples was determined spectrophotometrically from the 260 nm absorbance reading and the purity from the 260/280 nm ratios. Both sources of DNA had a ratio of 1.85, which is within the range of 1.70 – 1.90 and is acceptable (50) for PCR amplification. The purity and the concentration of five DNA samples from the Control and three isolated from the Anticoagulation Clinic population were determined. Rather than determining these parameters for all samples the assumption was made that when using the same isolation methods, the DNA concentrations of the other samples would be similar. However if a patient

has an infection such as a bacterial infection, the leukocyte levels are raised and larger amounts of DNA will be isolated. The effect of different DNA concentrations on the amplification of the 190 and 165 bp PCR amplification product for the CYP2C9*2 and *3 polymorphism was therefore determined. Conditions in which the amount of DNA available is very low, reaction or cycling conditions can be adapted and modified to allow the reaction to work efficiently. Within limits, increasing the template concentration may improve the outcome of the PCR reaction. The concentration of DNA was optimized for each PCR reaction. Both methods were not affected by different concentrations of DNA and a clear band without streaking from the well was observed. Different concentrations of 4.2 ng/µl and 6.2 ng/µl DNA was used for the amplification of the 190 and 165bp PCR fragment used for the identification of the CYP2C9*2 and *3 polymorphisms respectively.

Taq DNA polymerase is used to synthesize new strands of DNA and the optimal Taq DNA polymerase enzyme activity is between 0.5 and 2.5 units. In the methods developed, the aim was to use as little as possible of the Taq DNA polymerase to reduce screening costs. The effect of Taq DNA polymerase on the amplification of the CYP2C*2 and *3 polymorphisms was evaluated. At a Taq DNA polymerase concentration of 0.5U (0.025 and 0.022U/µl for the CYP2C9*2 and *3 polymorphism respectively) sufficient product for successful restriction enzyme digestion and the identification of the different genotypes was obtained.

For effective incorporation of nucleotide bases (dNTP's), Taq DNA polymerase requires free magnesium (57). At a dNTP concentration of 200µM a concentration of 1.5mM $MgCl_2$ is usually recommended. At a constant dNTP concentration of 200µM the effect of various concentrations of $MgCl_2$ was investigated. Criteria for successful amplification include a clearly defined band without streaking and nonspecific amplification product formation. At 1.5 mM $MgCl_2$ for the CYP2C9*2 polymorphism and 2 mM $MgCl_2$ for the CYP2C9*3 polymorphism this criteria was met.

Various authors (47,57) recommend DMSO and glycerol to improve amplification efficiency (greater amount of product) and specificity (no unspecific products) of PCR, when used in concentrations varying between 2-10% (57). Often the use of these adjuvants gives conflicting results. In some instances DMSO improve the amplification of some products, but decrease the product formation of others whereas other PCR amplification reactions are not

influenced at all. Similar results were obtained with 5% glycerol. BSA, in concentrations of up to 0.8µg/µl, increases the efficiency of the PCR reaction to a greater extent than either DMSO or glycerol. Therefore, the value of these adjuvants was tested for each PCR reaction. Adjuvants BSA, DMSO, and glycerol had no effect on the amplification of the 190 bp fragment containing the CYP2C9*2 polymorphism. For the CYP2C9*3 amplification reaction DMSO and glycerol had no effect while 5µg BSA resulted in an increase in amplification efficiency.

In order to screen a large population for the CYP2C9*2 and *3 polymorphisms cost-effectively, it was necessary to determine the minimum reaction volume required for successful amplification. The original PCR reaction volume was decreased to 60% of the volume originally used. A reaction volume of 15µl for both polymorphisms was sufficient. Often restriction enzyme digestion is the most expensive step in the identification of a point mutation and therefore it is necessary to optimize the amount of restriction enzyme used. This can either be achieved by decreasing the total PCR reaction volume, or by decreasing the actual amount of restriction enzyme added. In this study, decreasing the total reaction volume reduced the amount of restriction enzyme used by 40%. The amount of restriction enzyme used (units/volume) could only be reduced in the digestion of the CYP2C9*3 PCR product. Reduction of the amount of restriction enzyme added in the digestion step of the CYP2C9*2 analysis, compromised on the efficacy of the enzyme and the different genotypes could no be identified.

Adelekan (68) described a method for the storage of PCR reaction mixtures (that included PCR buffer, nucleotides, Taq polymerase, $MgCl_2$ and primers) for extended periods at –70°C. Subsequent thawing, addition of DNA and PCR amplification resulted in the successful amplification and detection of the C282Y and H63D, Hereditary Haemochromatosis mutations. However Adelekan (68) found that this method could not be used for ARMS PCR based methodologies. The methods used for the detection of the CYP2C9*2 and *3 polymorphisms are RFLP methodologies and therefore the storage of mixtures of PCR reagents including the primers, dNTP, Taq DNA polymerase in the presence of $MgCl_2$ and buffer should be considered. Aliquots of the PCR mixtures of both polymorphisms, including all reagents except template DNA was stored at –70°C. These reaction mixtures were tested after 5, 6, 11 days. Subsequent thawing of the PCR reaction mixture, followed by the addition

of genomic DNA, resulted in no amplification of the 190 and 165 bp fragments of the CYP2C9*2 and *3 polymorphisms after only 11 days. Therefore freshly prepared PCR mixture had to be prepared for the analysis of each polymorphism.

The optimized PCR method that resulted in the amplification of a 165bp fragment, followed by Nsi I digestion resulted in two fragments of 30 and 135bp. These results differed from that of Sullivan-Klose *et al* (41) who were using the same primers and reported an amplification product of 141bp, which was digested into two fragments of 29 and 112bp by Nsi I. All indications were that this product was very specific and optimization of several parameters such as annealing temperature, $MgCl_2$ and primer concentration still resulted in a single band of 165 bp. Restriction enzyme digestion of the 165bp fragment with Nsi I resulted in two fragments of 30bp and 135bp indicating that the restriction enzyme recognized the mismatch and therefore digestion was occurring at the correct site. The sequence for the region containing the CYP2C9*3 polymorphism was obtained from BLAST (Basic Local Alignment Search Tool) (65). The primer binding sequences were identified and the expected size of the PCR product was calculated. The calculated size of the PCR product correlated with a size of 155bp. The AAT tail of the forward primer added another 10bp to the fragment size and the final fragment length was 165bp. To further ensure that the fragment obtained was the correct region, the 165 bp fragment containing the CYP2C9*3 polymorphism was sequenced. The sequencing results obtained from the DNA Sequencing Unit at the University of Stellenbosch showed that the 165bp band was the region containing the CYP2C9*3 polymorphism (Section 3.12.4). Neither Wang (33) nor Sullivan-Klose (41) presented sequencing results for the obtained fragment of 141bp. Sequencing is the only way to confirm beyond doubt that the correct fragment has been amplified. The method developed to screen the South African Caucasian population for the CYP2C9*3 polymorphism, yielded a fragment of 165bp. The size of this fragment has been confirmed by sequencing (Table 3.2) and can be accepted as correct.

For the identification of a specific polymorphism a number of controls should be included in each batch of PCR reactions. The first of these are a negative control that includes all reagents except genomic DNA, which determines whether the prepared PCR mixture is contaminated with foreign DNA during any of the preparation steps. The next two controls are patient negative and positive controls i.e. DNA of a patient that has tested negative and DNA of another patient that has tested positive for the polymorphism in the past. To verify

whether in these control patients the polymorphism is absent on both alleles (homozygous normal (-/-)) or present on both alleles (homozygous polymorphic (+/+)) it is necessary to determine the exact nucleotide sequence for each control. This was achieved by determining the sequence of the PCR products. In order to sequence a fragment of DNA the PCR product, 190 and 165 bp fragments for the CYP2C9*2 and *3 polymorphisms respectively, must be isolated and purified. The fragments are separated from the other components in the PCR reaction by agarose gel electrophoresis. Low melting point agarose is used instead of the inert polymerized polyacrylamide gels as the PCR fragments can be eluted from agarose gel.

A sample of the PCR product isolated from the agarose gel usually is reloaded onto an electrophoresis gel. This is an expensive process as a large amount of isolated product is used. Therefore it was decided to use a small volume of the isolated sample as substrate for a nested PCR using the same primers as used in its initial amplification. Subsequently a single band of the correct size verified that the fragment isolated was the product of interest. A band of either 190 or 165 bp for the CYP2C9*2 and *3 polymorphism respectively, indicated that the correct fragments had been isolated.

The concentration of the isolated 190 and 165 bp fragments was determined by measuring their absorbance at 260 nm. The samples containing the CYP2C9*2 and *3 fragments were diluted to a final concentration of 1.0 ng/μl. The two primers used in the amplification of each polymorphism were diluted to a final concentration of 1.1pmol/μl. The fragments and primers were sent to the DNA Sequencing Unit at the University of Stellenbosch for sequencing.

Initially sequencing was not successful due to the low concentrations of the PCR products. It was recommended by the DNA Sequencing Unit at the University of Stellenbosch that the centrifugation speed should be reduced at the step where the fragments are separated in a resin column. When using a resin DNA isolation procedure, the resin beads move from the column into the end product when subjected to high centrifugal force. These beads falsely increase absorbance and the incorrect concentration is calculated.

To send DNA samples to Stellenbosch using a courier service is expensive. Sending the samples via registered mail would be more cost effective. It was therefore necessary to determine the stability of the DNA and primers at room temperature for a period of four days.

The integrity of the fragments and primers, after subjecting it to conditions similar to those expected from the 2000 km journey for four days, was found to be stable. The CYP2C9*2 and *3 fragments were sequenced using the Big Dye Terminator Cycle sequencing kit for automated sequencing on an ABI Prism 377 DNA sequencer. The sequences for both PCR fragments were the same as reported by Romkes *et al* (23) for exon 3 and by de Morais *et al* (66) for exon 7.

Both the homozygous normal (-/-)) controls for both polymorphism methods were sequenced. In the developed screening methods for the CYP2C9*2 and 3 polymorphisms, the normal allele is identified by the restriction enzyme. The sequences are presented in Table 3.1 and 3.2 respectively. The nucleotide sites of restriction, polymorphism and primer binding are indicated on these sequences. Both the forward and reverse sequences are present and once overlaid the complete sequence for the CYP2C9*2 (190bp) and CYP2C9*3 (165bp) fragment is obtained. To ensure that complete forward and reverse sequence is obtained, the sequencing cycling conditions should be reevaluated and possibly a longer extension time should be considered.

4.3 Population study.

Once all methodologies were optimized and well established, screening of the Control and Anticoagulation Clinic population was undertaken. The reproducibility of the methods was evaluated. All the patients in the Anticoagulation Clinic group were screened in duplicate. One individual tested heterozygous for the CYP2C9*2 polymorphism in the first screening but tested homozygous normal in the second run. Thus one false positive was detected for 288 tests performed. The false positive was probably detected due to incomplete digestion by the restriction enzyme Ava II in the first analysis.

4.3.1 The allele frequencies of the CYP2C9*2 and *3 polymorphisms in different normal populations.

The frequencies of the CYP2C9*2 and CYP2C9*3 polymorphisms have been determined in a number of populations and these together obtained for the South African Caucasian control population is summarized in Table 4.1.

Table 4.1: The allele frequency of the CYP2C9*2 and *3 polymorphisms in different populations.

Polymorphism	CYP2C9*1	CYP2C9*2	CYP2C9*3
Point mutation	Arg_{144}/Ile_{359}	Cys_{144}/Ile_{35}	Arg_{144}/Leu_{359}
Caucasians (41)	79%-86%	8%-19.1%	6%-10%
Indigenous Canadian (43)	91%	3%	6%
African American (41,43)	98.5%	1%-3.6%	0.5%-1.5%
Asians (43)	95%-98.3%	0%	1.7%-5%
Taiwanese (33,41)	97.4 %	0%	2.6%
British Aithal *et al* (4)	60%	22%	18%
British Taube *et al* (6)	69%	20.6%	10.4%
South African #	55.5%	25%	19.5%

 # South African figures are from this thesis.

The CYP2C9*2 and CYP2C9*3 polymorphisms is rarely found in African American, Asian and Taiwanese populations (Table 4.1). It is predominantly found in Caucasian populations. Wittkowsky (43) reported at the Anticoagulation Forum in 2002 that there is also a vast range in prevalence between different Caucasian populations with 8-19.1% and 6-10% for the CYP2C9*2 and *3 polymorphism respectively. It is not very clear which Caucasian populations were studied or how many patients were included in the study. The CYP2C9*2 (+/-) polymorphism was found in 22% of the Australian Caucasian individuals (n=18) studied by Bhasker *et al* (1). Furuya *et al* (13) found the CYP2C9*2 polymorphism to be present in 38% of the 94 British Caucasians that were studied. Aithal *et al* (4) performed a regional specific study in North-East England and found the CYP2C9*2 polymorphism in 22% and the CYP2C9*3 polymorphism in 18% of the individuals (n=100) studied. Taube *et al* (6) studied a group of 561 British Caucasian individuals and reported the CYP2C9*2 polymorphism in 20.6% of the individuals and the CYP2C9*3 polymorphism in 10.4% of the individuals. In the South African Caucasian population the CYP2C9*2 polymorphism was found in 25% and the CYP2C9*3 polymorphism in 19.5% of the individuals screened. The prevalence of the CYP2C9*2 polymorphism is in the higher range of that described by Wittkowsky (43) for various Caucasian populations. The allele frequency of the CYP2C9*3 polymorphism is almost double the maximum of the range. The allele frequency of the CYP2C9*2 polymorphism is similar to that found by Aithal *et al* (4) and Taube *et al* (6) for

the British population. The allele frequency of the CYP2C9*3 polymorphism is similar to that of Aithal *et al* (4) which is the highest prevalence shown is Table 4.1.

The South African Caucasian population is of European origin with many South Africans being of British, French and Dutch descend. No data could be found concerning the allele frequency of these polymorphisms in the Dutch and French populations. Therefore the South African prevalence's can only be compared to the two studies of reference performed by Aithal *et al* (4) and Taube *et al* (6).

4.3.2 Distribution of the various CYP2C9*2 and *3 genotypes in the British and South African Control and Anticoagulation clinic populations.

Aithal (4) determined the effect of the CYP2C9*2 and *3 polymorphisms in the CYP2C9 gene on warfarin dose requirements and risk of bleeding complications. The prevalence of these polymorphisms was determined in a community control (n=100), a clinic control population (n=52) and low dosage group (n=36). The South African Caucasian population in this study comprised a Control population (n=200) and an Anticoagulation Clinic (n=144) population. In the community control group of Aithel *et al* (4) 20% of the population was found to be heterozygous and none homozygous for the CYP2C9*2 polymorphism. In the South African Control population with the same selection criteria as the community control population of Aithel *et al* (4), 20% were heterozygous and 5% CYP2C9*2 homozygous. In the same population Aithal *et al* (4) reported 17% individuals heterozygous for the CYP2C9*3 polymorphism and only 1% homozygous. In this study 19% of the population were CYP2C9*3 heterozygous and 0.5% homozygous. Aithal *et al* (4) reported 2% compound heterozygotes compared to the 5.5% compound heterozygotes found in this South African Caucasian population. In conclusion, the presence of individuals that are CYP2C9*2 homozygous and compound heterozygous are greater than that found by Aithel *et al* (4) in the British Caucasian population. The homozygous normal genotype was evenly distributed between the two populations. Sixty percent (4) of the individuals in the UK population and 61% of the 200 individuals in the SA population tested negative for both of the polymorphisms.

4.3.3 Distribution of various CYP2C9*2 and *3 genotypes in the British and South African Anticoagulation Clinic populations.

The clinic control group of Aithal (4) consisted of 52 patients that were randomly selected from a wide range of warfarin dose requirements and attending an anticoagulant clinic in northeast England. The clinic control group of Aithal *et al* (4) was small and the population prevalences of the polymorphisms cannot be accurately predicted. Nevertheless, the gene distributions in Aithal's study and this study can be compared. Seventeen percent of Aithal *et al*'s clinic control group (4) was heterozygous for the CYP2C9*2 polymorphism and 19% were heterozygous for the CYP2C9*3 polymorphism. In the South African Anticoagulation Clinic group 20% of the population was CYP2C9*2 heterozygous and 10% were CYP2C9*3 heterozygous. The percentage of individuals in the clinic control group (4) heterozygous for the CYP2C9*3 polymorphism was almost twice that found in the SA Anticoagulation Clinic population. In the clinic control population 2% of the patients were CYP2C9*2 homozygous but none CYP2C9*3 homozygotes, or compound heterozygotes were found.

In the Anticoagulation Clinic population a single patient (0.6%) was homozygous for the CYP2C9*2 polymorphism but no CYP2C9*3 homozygous patient was detected. No conclusion can really be made when comparing these populations due to the small size of the clinic control population. Although 10% of the clinic control group of Aithel *et al* (4) were CYP2C9*3 heterozygous, no patients were homozygous. The explanation for this by the authors was that these patients might have been excluded from the group due to the fact that they could not comply with the usual dosing strategies. However, absence of homozygosity is often a common phenomenon when small populations are screened for a polymorphism that occurs at a low prevalence. The method of detection may be another possible reason for the lower than expected prevalence of the CYP2C9*3 polymorphism. The CYP2C9*3 polymorphism is detected with a mismatch RFLP method. The only study where more CYP2C9*3 homozygotes were detected than expected according to the calculated Hardy-Weinberg value, was done by Higashi *et al* (17). The authors used direct sequencing as detection method for the described polymorphisms. Direct sequencing is much more expensive when screening large populations but also much more accurate. The only way to confirm the accuracy of the mismatch primer RFLP screening method used for the detection of the CYP2C9*3 polymorphism in the South African population, is to do a comparative

sequencing study. This may be the reason why Wang *et al* (33) used two separate methods for detection of the CYP2C9*1 and CYP2C9*3 polymorphisms.

4.3.4 The impact of the CYP2C9*2 and *3 on warfarin therapy.

In a larger study, Taube *et al* (6) studied the influence of CYP2C9*2 and *3 polymorphisms on warfarin sensitivity and risk of over-anticoagulation in patients on long-term treatment prescribed by the Cambridge anticoagulant clinic. Five hundred and sixty one patients with a target INR of 2.5 who had been treated with warfarin for more than 2 months were screened. In this population 0.5% and 0% were homozygote for CYP2C9*2 and CYP2C*3 respectively. Although the selection criteria used in this study differs from the Anticoagulation Clinic population due to the large population used in the latter study some comparisons can be made.

Aithal *et al* (4) also screened a low dose group that consisted of 36 patients that received 1.5 mg or less warfarin. The prevalence of the polymorphisms in the low dose group was 6% and 0% for the CYP2C9*2 and *3 polymorphisms respectively. Patients in the low dose group were six times more likely to have variant alleles in the CYP2C9 region compared to the general population. In 20% of these patients doctors experienced difficulty establishing optimum anticoagulation doses (4). These polymorphisms are associated with adverse clinical outcomes including over-anticoagulation during initiation (4) and major and minor bleedings (8). In a group of 185 patients studied over 2.3 years serious and life threatening incidents of bleeding occurred 2.2 times more often in patients with the polymorphisms than in patients with the wild type genetic composition (17). According to Margaglione *et al* (8), carriers of the CYP2C9*2 and *3 polymorphisms had a calculated rate of bleeding complications of 27.9 per 100 patient years (8).

In this study as a part of a collaboration study with the Department of Haematology, the Anticoagulation Clinic population was divided into two groups; those without a CYP2C9*2 or *3 polymorphism and those with a polymorphism either homozygous or heterozygous. The first group presented with a mean warfarin maintenance dosage of 4.85 mg/day while patients with the CYP2C9 had a mean maintenance dosage of 3.71mg/day. De Wet (67) also evaluated age, sex, and indication for warfarin therapy, concurrent medication and INR. From

this study it was found that genotype and age was very significant determinants of daily warfarin dose.

4.4 Recommendations for warfarin therapy.

Unpredictable pharmacological behavior occurs with many commonly used medications. It is a common source of excessive or inadequate drug dosing, and, for a given patient may cause a specific medication to be toxic or ineffective (41). Although dosing problems have historically been attributed to factors such as patient demography, diet, age, hepatic dysfunction, underlying disease or drug interactions, it is now apparent that an individual's pharmacological profile for a given medication may be largely genetically determined (32). Functionally significant; drug metabolizing enzyme polymorphisms, may partly underlie variability in drug response and determine efficacy.

Warfarin (Coumadin) is the most frequently prescribed oral anti-coagulant, the fourth most prescribed cardiovascular agent and the overall eleventh most prescribed drug in the USA (69) with annual sales of $500 million (70). Despite this the Agency for Healthcare Policy and Research (AHCPR) stated that the drug was still greatly under utilized for stroke prevention. Although warfarin prevents 20 strokes for every bleeding episode, physicians are still reluctant to prescribe it because of unfamiliarity with techniques for safe administration (7).

Bleeding is the major complication of warfarin therapy. The estimated average annual frequency of fatal, major and minor bleedings is 0.6%, 3.0% and 9.6% respectively (71). The threat of bleeding in some studies was more prominent early in therapy. The prevalence of major bleeding in a population followed prospectively after beginning therapy was 3.0% during the first month, 0.8% per month during the remainder of the first year, and 0.3% per month after the first year (72). Prediction of individual dosage requirements would help decrease the risk of bleeding. Several methods are used to anticipate the maintenance dosage of warfarin based on early response, such as algorithms, measurement of protein C and S levels, and computer assisted models (4).

Other factors that may influence warfarin therapy are concomitant drugs, hepatic impairment, diet and compliance. Other drugs may influence the pharmacodynamics of warfarin by

inhibiting synthesis or increasing clearance of vitamin K-dependant coagulation factors or by interfering with other pathways of hemostasis. The anticoagulant effect of warfarin is augmented by second and third generation cephalosporins, which inhibit the cyclic interconversion of vitamin K (73,74); by thyroxine, which increases the metabolism of coagulation factors (75); and by clofibrate, through an unknown mechanism (76). Doses of salicylates > 1.5 g per day (77), and acetaminophen (78) also augment the anti-coagulant effect of warfarin. Drugs such as aspirin (79), nonsteroidal anti-inflammatory drugs (80), penicillins (in high doses) (81,82) and moxolactam (74) increase the risk of warfarin-associated bleeding by inhibiting platelet function. Of these aspirin is the most important because of its widespread use and prolonged effect (83).

Age and liver volume account for about 34% of the inter-individual difference in warfarin dosage requirements (84). A noteworthy negative association was established between age and liver volume and a positive correlation between liver volume and dose. This implies that a fall in liver mass contributes to the age-related increase in sensitivity to warfarin (85), which might be the result of a fall in total hepatic content of vitamin K epoxide reductase (86).

In a number of studies as well as this study it has been shown that there is there is a strong association between CYP2C9 genotype and warfarin sensitivity (4,5,6). An individual requiring a low warfarin dose is six times more likely to be positive for one or more of the variant alleles (CYP2C9*2 and CYP2C9*3) associated with impaired warfarin metabolism when compared with the general population. Identifying these patients before prescribing warfarin may be the solution to bleeding complications associated with this anticoagulant.

Horton (7) describes in detail different strategies for prescribing warfarin. A very simplified approach is to administer 5mg/day and to monitor the INR very closely until the desired range has been achieved and maintained (10). The frequency of monitoring can then be reduced to every 4 weeks (10,87,88). These models oblige patients to have taken numerous doses of the drug, thus exposing the patient to the risk of bleeding from the onset. The capability to predict dosage requirements prior to starting therapy would be ideal. Identifying patients with an inherited impaired ability to metabolise warfarin is one proposed method. By screening patients for genetic polymorphisms such as CYP2C9*2 and *3 prior to the

initiation of warfarin therapy may reduce the number of dosage adjustments required until the maintenance dosage is achieved.

However, testing is usually only available in certain large molecular diagnostic laboratories. Furthermore the screening methodologies described in this study are not ideal for mass routine screening. Rapid CYP2C9*3 assays have been developed by Hiratsuka *et al* (89) in 1999 that shows promise for high volume use in clinical laboratories with the availability of results in less than 3 hours.

As point-of-care genetic testing becomes clinically feasible, the use of genotyping may have practical implications for the management of warfarin patients. The genetic basis of dose-response is so profound that genetic testing before administering warfarin might rule out episodes of hemorrhagic and thrombotic complications to a large extent according to USA data. Treating a major episode of bleeding in 1997 costed an estimated amount of $1245 for patients already admitted to a hospital and $4149 for outpatients (90). In comparison the cost of screening for the CYP2C9*2 and CYP2C9*3 polymorphisms is merely R780.00 ($100), according to the "Scale of benefits" as used by the Bureau of Health Care Funders. No information is available regarding the cost in South Africa. Any added information that would improve warfarin administration and reduce the prevalence of bleeding events would reduce the costs to an already overextended health care system in South Africa. Difficult induction has financial implications that lead to delayed discharges and numerous visits to the clinics. A screening test for CYP2C9*1 *2 and *3 is currently available in America at a cost of approximately $250 (91). The Molecular Diagnostic Laboratory, University of Pretoria performs the same test for approximately R400 for both polymorphisms.

The South African population has diverse ethnicity and population screening using the above methodologies are only limited to the Caucasian population. However recent studies have identified other polymorphisms such as the CYP2C9*4 and *5 that may be specific to other population groups present in South Africa. A new allelic variant in exon 7 of the CYP2C9 gene has been found in African Americans as described by Dickmann *et al* (48). The polymorphism is characterized by a C360G substitution at nucleotide 1080 that causes a mis-sense mutation. This polymorphism is known as CYP2C9*5 (48). The CYP2C9 polymorphisms is found in about 10% of African Americans and prevalences are roughly

equally devided between the CYP2C9*2, CYP2C9*3 and CYP2C9*5 variants. The CYP2C9*5 polymorphism is associated with a 31 fold reduction in warfarin clearance (48).

As genomic information becomes more readily available, clinicians will need to establish new guidelines for patient management, especially when administering drugs with a narrow therapeutic index like warfarin. Variant genotype, CYP2C9 in this instance, could in future be considered a "sensitivity factor" when initiating therapy, or might direct the clinicians in prescribing alternate therapies.

In conclusion both the CYP2C9*2 and *3 polymorphisms are present in the South African Caucasian population. The prevalence of the CYP2C9*2 and *3 polymorphisms is similar to that of the British population. Indications are that the prevalence of the CYP2C3 polymorphism may be higher. However this can only be verified in a larger population study. No founder effect has been found for these polymorphisms. The screening of patients attending an Anticoagulation Clinic should be implemented, as this will have a significant effect on the treatment strategy utilized.

4.5 Conclusion.

The objectives for this thesis have been met:

1. PCR screening methods for detecting the CYP2C9*2 and *3 polymorphisms have been established and optimized.
2. The prevalence of these polymorphisms has been determined in the South African Caucasian population.
3. There seems to be no Founder Effect compared to the British population; however, figures for the Dutch and French populations are not known and the Founder Effect may not yet be excluded.
4. It is clear that the prevalence of the CYP2C9*2 and *3 polymorphisms are high enough in the South African Caucasian population to consider screening patients before initiating warfarin therapy.

5. References:

1. Bhasker CR, Miners JO, Coulter S. Allelic and functional variability of cytochrome P4502C9. Pharmacogenetics 1997;7:51-58.
2. Nebert DW, Russell DW. Clinical importance of the cytochromes P450. Lancet 2002;360:1155-1162.
3. Horai Y, Ishizaki T. Pharmacogenetics and its clinical implications: N-acetylation Polymorphism. Pharmacology for Physicians 1987;21:1-7.
4. Aithal GP, Day CP, Kesteven PJL *et al.* Association of polymorphisms in cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 1999;353:717-719.
5. Freeman BD, Zehnbauer BA, McGrath S *et al.* Cytochrome P450 Polymorphisms are associated with reduced warfarin dose. Surgery 2000;128:281-285.
6. Taube J, Halsall D, 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;96:1816-1819.
7. Horten JD, Bushwick BM. Warfarin Therapy: Evolving Strategies in Anticoagulation. American Family Physician 1999;59:635-646.
8. Margaglione M, Colaizzo D, D'Andrea G et al. Genetic Modulation of Oral Anticoagulation with Warfarin. Thromb Haemost 2000;84:775-778.
9. Budavari S, O'Neil MJ, Smith A *et al*, editors. The Merck Index, an encyclopedia of chemicals and drugs. 12th ed. Whitehouse Station (NJ): Merck Research Laboratories; 1990.
10. Hirsh J, Dalen JE, Anderson DR, *et al.* Oral anticoagulants: mechanism of action, clinical effectiveness, and optimal therapeutic range. Chest 1998;114:4458-69.
11. Steward DJ, Haining RL, Henne KR *et al.* Genetic association between sensitivity to Warfarin and expression of CYP2C9*3. Pharmacogenetics 1997;7:361-367.
12. Mannuccio MP. Genetic control of anti-coagulation. Lancet 1999;353:688-689.
13. Furuya H, Fernandez-Salguero P, Gregory W *et al.* Genetic polymorphisms of CYP2C9 and its effect on warfarin maintenance dose requirement in patients undergoing anti-coagulation therapy. Pharmacogenetics 1995;5:389-392.

14. Yamazaki H, Inoue K, Chiba K, *et al.* Comparative studies on the catalystic roles of cytochrome P450 2C9 and its Cys- and Leu variants in the oxidation of warfarin, flurbiprofen and diclofenac by human liver microsomes. Biochem pharmacol 1998;56:243-251.
15. Kimura M, Ieiri I, Mamiya K *et al.* Genetic polymorphism of cytochrome P450s CYP2C19 and CYP2C9 in a Japanese population. Ther Drug Monit 1998;20:243-247.
16. Miners JO, Birkett DJ. Cytochrome P4502C9: an enzyme of major importance in human drug metabolism. Br J Clin Pharmacol 1998;45:525-528.
17. Higashi MK, Veenstra DL, Kondo ML *et al.* Association between CYP2C9 Genetic Variants and Anticoagulation-Related Outcomes During Warfarin Therapy. Medical Periodicals 2002;287:1690-1698.
18. Fernandez-Salguero P, Hoffman SMG, Cholerton S *et al.* A genetic polymorphism in Caumarin 7-Hydroxylation: Sequence of the Human CYP2A genes and identification of variant CYP2A6 alleles. Am J Hum Genet 1995;57:651-660.
19. Dieffenbach CW, Lowe TMJ, Dveksler GS. General Concepts for PCR Primer Design, in PCR. A Laboratory Manual. Dieffenbach CW, Dveksler GS, Editors. New York: Cold Spring Harbor Laboratory Press; 1995. p. 133-155.
20. Merryweather-Clarke AT, Pointon JJ, Shearman JD *et al.* Global prevalence of putative haemochromatosis mutations. Journal of medical genetics 1997;34:275-278.
21. Warnich L. Identification of three mutations and associated haplotypes in the protoporphyrinogen oxidase gene in South African families with variegate porphyria. Human Molecular Genetics 1996;5:981-984.
22. Zuber R, Anzenbacherova E, Anzenbacher P. Cytochromes P450 and experimental models of drug metabolism. J Cell Mol Med 2002;6:189-198.
23. Romkes R, Faletto MB, Blaisdell JA *et al.* Cloning and expression of complementary DNA's for multiple members of the human cytochrome P450 2 C subfamily. Biochemistry 1991;30:3247-3255.
24. White, RE, Coon MJ Ann Rev Biochem 1980;49:315-356.

25. http://mhc.com/cytochrome/links.html
26. Krishna D, Klotz U. Extra hepatic metabolism of drugs in humans. Clin Pharmacokinet 1994;26:144.
27. Nelson DR, Koymans L, Kamataki T *et al* P450 Superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 1996;6:1-41.
28. http://us.expacy.org/cgi-bin/niceprot.pl
29. Phillips IR, Shephard EA editors. Cytochrome P450 Protocols. In: Nelson DR. Methods in Molecular Biology, Totowa, New Jersey: Humana Press; 1998. p. 15-24.
30. http://www.biozentrum.unibas.ch/schwede/
31. Takahashi H, Kashima T, Nomizo Y *et al*. Metabolism of Warfarin enantiomeres in Japanese patients with heart disease having different CYP2C9 and CYP2C19 genotypes. Clin Pharmacol Ther 1998;63:519-528.
32. Hirsh J, Fuster V, Halperin JL. American Heart Association/ American College of Cardiology Foundation Guide to Warfarin Therapy. Circulation 2003;107:1692.
33. Wang S, Huang J, Lai M *et al*. Detection of CYP2C9 polymorphism based on the polymerase chain reaction in Chinese. Pharmacogenetics 1995;5:37-42.
34. Rendic SD, Catlo EJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. Drug Metab Rev 1997;29:413-580.
35. Kaminsky LS, Zhang ZY. Human P450 metabolism of Warfarin. Pharmacol Ther 1997;73:67-74.
36. Chan E, McLachlan AJ, Pegg M *et al*. Disposition of warfarin enantiomers and metabolites in patients during multiple dosing with rac-warfarin. Br J Clin Pharmacol 1994;37:563-569.
37. Henne KR, Gaedigk A, Gupta G *et al*. Chiral phase analysis of warfarin enantiomers in patient plasma in relation to CYP2C9 genotype. J Chromatogr 1998;710:143-148.
38. http://drnelson.utmem.edu/human.genecount.html

39. Nelson DR, Kamataki T, Waxman DJ *et al.* The CYP P450 superfamily: update or new sequences, gene mapping, accession numbers, early trivial names of enzymes and nomenclature. DNA Cell Bio 1993;12:1-51.
40. Gray IC, Nobile C, Moresu R *et al.* A 2.4-megabyte physical map spanning the CYP2C gene cluster on chromosome 10q24. Genomics 1995;28:328-332.
41. Sullivan-Klose TH, Ghanayem BI, Bell DA *et al.* The role of the CYP2C9-Leu^{359} allelic variant in the tolbutamide polymorphism. Pharmacogenetics 1996;6:341-349.
42. Rettie A, Wienkers LC, Gonzales FJ *et al.* Impaired (S)-warfarin metabolism catalyzed by the R144C allelic variant of CYP2C9. Pharmacogenetics 1994;4:39-42.
43. Wittkowsky AK. Pharmacogenomics of Warfarin. Anticoagulation Forum 2002;6.
44. Gotoh S. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J Biol Chem 1991;267:83-90.
45. Kaminsky LS, de Morais SMF, Faletto MB *et al.* Correlation of human P4502C substrate specificities with primary structure – warfarin as a probe. Mol Pharmacol 1993;43:234-239.
46. Polous TL, Finzel BC, Howard AJ. High-resolution crystal structure of P450cam. J Mol Biol 1989;195:687-700.
47. Winship PR. An improved method for directly sequencing PCR amplified material using DMSO. Nucleic Acids Res 1989;17:1266.
48. Dickmann LJ, Rettie AE, Kneller MB *et al.* Identification and functional characterization of a new CYP2C9 variant (CYP2C9*5) expressed among African Americans. Mol Pharmacol 2001;60:382-87.
49. Redman AR. Implications of Cytochrome P450 2C9 Polymorphism on Warfarin Metabolism and Dosing. Pharmacotherapy 2001;21:235-242.
50. QIAamp DNA mini kit and QIAamp DNA blood mini kit handbook.US Patent 4,683,195 and 4,683,202.1999 Jan QIAGEN.
51. Stubbins MJ, Harries LW, Smith G, *et al.* Genetic analysis of human cytochrome P450 CYP2C9 locus. Pharmacogenetics 1996;6:429-439.

52. By Tabitha M. Powledge From: http://www.faseb.org/opar/bloodsupply/pcr.html
53. Innis MA, Gelfand DH. Optimization of PCRs. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR protocols, A Guide to Methods and Applications. London: CRC Press; 1994. p. 5-11.
54. Strachman T, Read AP. Human Molecular Genetics. Bios Scientific Publishers Limited, 1996.
55. Sharrocks AD. The design of primers for PCR. In: Griffin HG, Griffin AM, editors. PCR Technology, Current Innovations. London: CRC Press; 1994. p. 5-11.
56. Furrer B, Candrian U, Wieland P, Luthy J. Improving PCR efficiency. Nature 1990;346:324.
57. Promega Corporation. Betain and DMSO: Enhancing agent for PCR. Promega Note: Promega Corporation Publisher 1998;65:27-29.
58. Kwok S, Kellog DE, McKinney N *et al.* Effects of primer-template mismatches on the polymerase chain reaction: Human Immunodeficiency Virus 1 model studies. Nucleic Acids Res 1990;18:999-1005.
59. Alphey L. Strategies for New Sequence Determination. In: Alphey L. DNA Sequencing, from experimental methods to bioinformatics. Manchester: Bios Scientific Publishers; 1997. p. 111-123.
60. Chain Termination (Sanger Dideoxy) Method. In: Alphey L. DNA Sequencing, from experimental methods to bioinformatics. Manchester: Bios Scientific Publishers Limited; 1997. p.15-25.
61. Mueller RF, Young ID. Emery's Elements of Medical Genetics. 10th ed. Hong Kong: Churchill Livingstone; 1998.
62. Jooste PL, Benade AJS, Rossouw JE. Prevalence of familial hypercholesterolaemia in three SA communities. S Afr Med J 1986;69:548-551.
63. Troendle JF, Yu KF. A note on testing the Hardy-Weinberg law across strata. Annals of Human Genetics 1994;58:397-402.
64. Roche, High Pure PCR Purification kit. Instruction Manual, version 2, Dec 2000; p. 9-11.
65. http://www.ncbi.nlm.nih.gov

66. de Morais SM, Scweikl H, Blaisdell JA, Raucy JL, Goldstein JA. Gene structure and upstream regulatory regions of human CYP2C9 and CYP2C18. Biochem Biophys Res Commun 1993;194:194-201.
67. de Wet L. The role of Cytochrome P450 CYP2C9 polymorphisms on maintenance warfarin dose requirement in South African Caucasians [dissertation]. Pretoria: Pretoria Univ.; 2003.
68. Adelekan AM. Establishment of Screening Procedures for Genetic Disorders and risk factors in the South African Caucasian Population [dissertation]. Pretoria: Pretoria Univ.; 2002.
69. National prescription audit: Physician speciality report, dispenced data. Plymouth meeting, Pa.:IMS America, 1998.
70. Retail perspective and provider perspective audit. Plymouth meeting, Pa.:IMS America, 1998.
71. Landefeld CS, Beyth RJ. Anticoagulant-related bleeding: Clinical epidemiology, prediction and prevention. Am J Med 1993;95:315-28.
72. Landefeld CS, Beyth RJ. Major bleedings in outpatients treated with warfarin: incidence and prediction of factors known at the start of outpatient therapy. Am J Med 1989;87:144-52.
73. Bechtold H, Andrassy K, Jahnchen E, *et al.* Evidence for impaired hepatic vitamin K1 metabolism in patients treated with N-methyl-thiotetrazole cephalosporins. Thromb Haemost 1984;51:358-361.
74. Weitekamp MR, Aber AC. Prolonged bleeding times and bleeding diathesis associated with moxalactam administration. JAMA 1983;249:69-71.
75. Owens JC, Neely WB, Owen WR. Effect of sodium Dextrothyroxine in patients receiving anticoagulants. N Engl J Med 1962;266:76-79.
76. O'Reilly RA, Sahud MA, Robinson AJ. Studies on the interaction of warfarin and clofibrate in man. Thromb Diath Haemorrh 1972;27:309-318.
77. Rothschild BM. Hematologic perturbations associated with salicylate. Clin Pharmacol Ther 1979;26:145-152.
78. Hylek EM, Heiman H, Skates SJ, *et al.* Acetaminophen and other risk factors for excessive warfarin anticoagulation. JAMA 1998;279:657-662.

79. Dale J, Myhre E, Loew D. Bleeding during acetylsalicylic acid and anticoagulant therapy in patients with reduced platelet reactivity after aortic valve replacement. Am Heart J 1980;99:746-752.
80. Schulman S, Henriksson K. Interaction of ibuprofen and warfarin on primary haemostasis. Br J Rheumatol 1989;28:46-49.
81. Cazenave J-P, Pacham MA, Guccione MA, *et al.* Effects of Penicillin G on platelet aggregation, release and adherence to collagen. Proc Soc Exp Med 1973;142:159-166.
82. Brown CH II, Natelson EA, Bradshaw W, *et al.* The haemostatic defect produced by carbenicillin. N Engl J Med 1974;291:265-270.
83. Roth GJ, Majerus PW. The mechanism of the effect of aspirin on human platelets, I: acetylation of a particulate fraction protein. J Clin Invest 1975;56:624-632.
84. Wynne HA, Cope LH, Kelly P, Whittingham T, Edwards C, Kamali F. The influence of age, liver size and enantiomer concentrations on warfarin requirements. Br J Clin Pharmacol 1995;40:203-7
85. Shepherd AM, Hewick DS, Moreland TA *et al.* Age as a determinant of sensitivity to warfarin. Br J Clin Pharmacol 1977;4:315-20
86. Kamali F, Edwards C, Butler. The influence of (R)- and (S)- warfarin, Vitamin K and Vitamin K Epoxide upon Warfarin Anticoagulation. Thromb Haemost 2000;84:39-42
87. Ansell JE, Buttaro ML, Thomas OV *et al.* Consensus guidelines for coordinated outpatient oral anticoagulation therapy management. Ann Pharmacother 1997;31:604-615.
88. Bussey HI. How to monitor the dosage of warfarin. Heart Dis Stroke 1993;2:388-392.
89. Hiratsuka M, Agatsuma Y, Mizugaki M. Rapid detection of CYP2C9*3 alleles for real-time fluorescence PCR based on CYBR green. Mol Genet Metab 1999;68:357-362.
90. Gould MK, Dembitzer AD, Sanders GD *et al.* Low molecular weight heparins compared with unfractionated heparin for treatment of acute deep venous thrombosis. Ann Intern Med 1999;130:789-799.

91. www.genelex.com