CHAPTER V

AN INVESTIGATION OF GENETIC DETERMINANTS OF WARFARIN DOSE REQUIREMENT IN 1500 SWEDISH PATIENTS (WARG COHORT)

In the previous chapter, a study of 35 candidate genes in 201 warfarin-treated patients (Uppsala) was described. This study identified SNPs in *VKORC1*, *CYP2C9-CYP2C19*,-*CYP2C18*, *PROC*, *APOE*, *EPHX1*, *CALU*, *PDIA2*, *GGCX* and *ORM1-2* being associated with variability in warfarin dose requirement. The strength of these associations above nominal significance varied and only *VKORC1* and *CYP2C9* reached p-values below 1.65×10^{-4} (experiment-wise).

Our first goal was to replicate the findings from the Uppsala study in a large independent sample. However, the sample size of 201 patients in the discovery study is small and had statistical power to detect only 'strong' effects. We therefore devised a strategy to combine the replication experiment with a *de novo* investigation of all the 35 candidate genes in the much larger sample of the WARG study.

5.1 THE NATIONAL WARFARIN GENETIC (WARG) STUDY

The national warfarin genetic (WARG) study in Sweden (http://www.druggene.org/) is a prospective study to understand the genetic determinants of inter-individual warfarin dosing and warfarin-induced bleeding complication. Patients were recruited in 40 outpatient clinics in Sweden, and the criteria of patient recruitment are described in chapter II, section 2.1.2. Age distribution of WARG patients was comparable to that in the Uppsala study (Figure 5.1). The patients in WARG were 18-92 years old, with the majority of them been aged between 50 and 80. There is no patient younger than 18 years old enrolled in the study; only one recruited patient was aged 18.

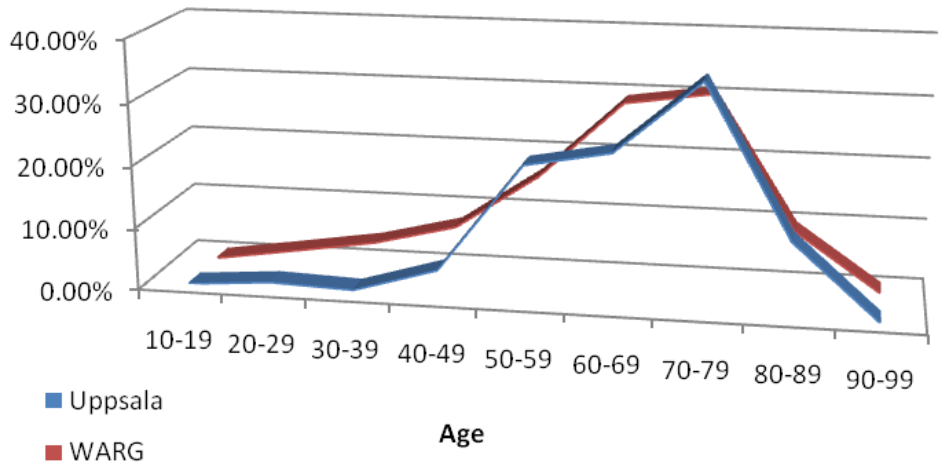

Figure 5.1. Age distribution of patients in the Uppsala and WARG studies. The majority of patients treated with warfarin was aged from 50-80 in both studies.

A total of 1523 patients were eventually enrolled in the WARG study of which 1496 were analysed in this study. The basic characteristics of these 1496 patients are described in Table 5.1. Patients have an average age of 66 years (95% Confidence interval (CI): 57 years and 74 years) and were predominantly male (947 male). The constraints of the approved consent

form and recent practises in Sweden, prevented patient's body weight being recorded. Various medical indications were recorded electronically in each outpatient clinic and thereafter submitted to a central database through the internet (Lindh et al. 2004). More than half of the patients (51%) suffered from atrial fibrillation, which resembles the patient composition of the Uppsala study (56%). The second most diagnosed indication is deep vein thrombosis and pulmonary embolism with 566 patients (38%). This is in contrast to the Uppsala study which comprises only 4.5% of such patients. The Uppsala study in return comprised a bigger portion of patients with heart valve prosthesis (24%) against a merely 4% in the WARG cohort.

Most patients in the WARG study had been monitored for at least three months and relevant medical information was registered into the database. The majority of patients were planned for life-long treatment or for treatment without predefined period. The maintenance dose used to test genetic association was defined as the mean of all doses given to a patient during a minimum series of three consecutive INR measurements between 2 and 3 (therapeutic INR). The stable dose was calculated from all doses that were unchanged over a minimum of three consecutive visits and that lead to a therapeutic INR. Patients outside target INR and the above criteria were removed from further analyses. In the WARG study there is 19-fold variation in warfarin dose requirement (6.0 to 113.75 mg/week) compared to 17-fold in the Uppsala study (4.5 to 77.25 mg/week). The recruitment of the severe bleeding patients is according to the World Health Organisation (WHO) definition, which is: lethal, life-threatening, permanently disabling, or leading to hospital admission (emergency room admissions excluded) or prolongation of hospitalisation (Lindh et al. 2007). During treatment, 146 patients (9.8%) were subject to bleeding complications whereas 28 had severe bleeding episodes (1.9%). The analysis of genetic association with severe bleeding complication is

described in chapter VI.

Warfarin doses and PT INR measurement intervals were chosen at the discretion of the treating physicians (Lindh et al. 2007). The targeted therapeutic range of PT INR is 2.0-3.0 which was measured in each visit to clinics. During the period of a total of 1276 patient-years, 63.8 % had targeted within therapeutic PT INR interval and an average of 24.5 PT INR values were recorded for each patient. The statistics of the number of INR recorded for each patient was shown in Figure 5.2. Information as to patients who did not consent to join the study was not recorded in the study design.

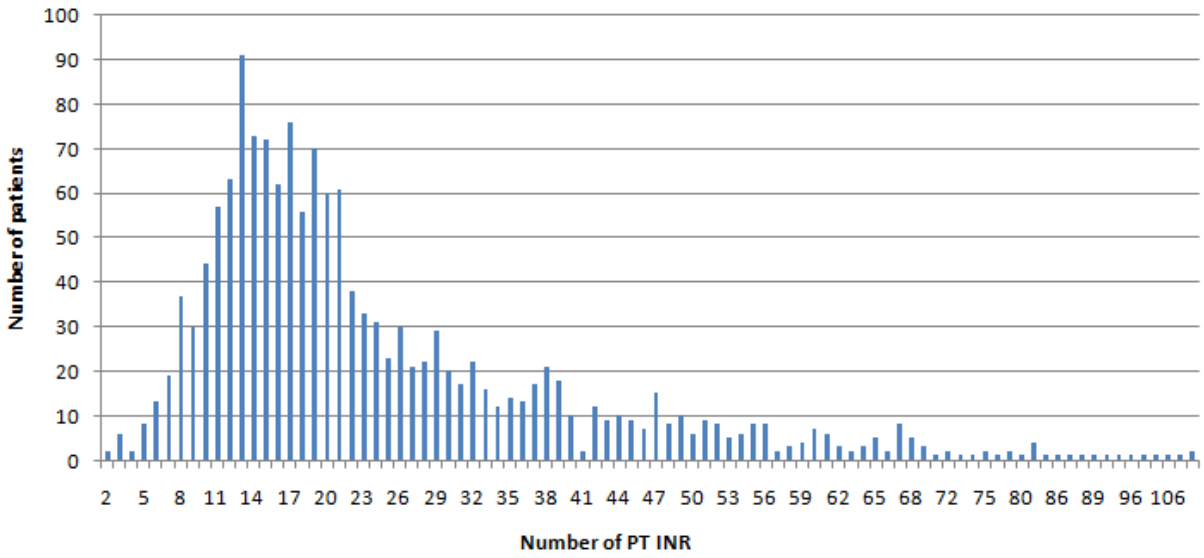

Figure 5.2. The number of PT INR recorded to each patient.

Most of the patients in WARG were receiving other medication with a total of 1528 different drugs being recorded (Table 5.1). Among the prescribed drugs, 781 are known to have no-interaction with warfarin; 56 as having decreasing effect in warfarin administration and 691 drugs potentiating the effect of warfarin which in turn increases anticoagulation and the risk for bleeding complications.

Table 5.1. Medical information statistics of patients in WARG cohort.

Characteristics	
	Patients (%)
Indication for warfarin treatment	
Atrial fibrillation	762 (51 %)
Deep venous thrombosis	378 (25 %)
Pulmonary embolism	188 (13 %)
Heart valve transplant, artificial	34 (2 %)
Cerebral infarction/transient ischemic attack	28 (2 %)
Heart valve transplant, biologic	25 (2 %)
Cardiomyopathy/cardiac failure	23 (2 %)
Other indications†	58 (4 %)
Planned treatment duration	
<3 months	13 (1 %)
3-5 months	214 (14 %)
6-12 months	387 (26 %)
>12 months, specified	14 (1 %)
Infinite or not predefined	868 (58 %)
Patients experiencing bleeding events (%)	
Serious bleeding	28 (1.9%)
All bleeding	146 (9.8%)
INR, international normalized ratio	
Lower limit of therapeutic interval	2.1 (1.5;3.1)
Target value	2.5 (1.8;3.0)
Upper limit of therapeutic interval	3.0 (2.0;3.6)
Concomitant medication, No. of drugs	2 (0; 18)
Time within therapeutic INR interval (%)	63.8 (50.9; 73.5)
Use of drugs interacting with warfarin	
No interacting drug	781 (52 %)
Drugs *decreasing* warfarin effect	56 (4 %)
Drugs *potentiating* warfarin effect	691(46 %)
Gender	
Men	947 (63%)
Women	549 (37%)
Age (95% CI)	66 (57; 74)

A male preponderance was observed in both the Uppsala and the WARG studies with a 2:1 ratio of male/female. Table 5.2 lists the number of male and female patients enrolled in a selection of warfarin genetic studies and only the studies enrolling mainly Caucasian subjects were included. Beside the two study cohorts described in this thesis, the study published by

Bodin et al (2005), which recruited normal subjects, also has male preponderance. However, other studies recruited slightly more male than female patients (5-10%) except the study by Aquilante et al (2006), which recruited a large percentage of patients from two Veterans Administration clinics. The male preponderance in both Swedish studies might be as a result of the earlier onset of heart related diesease in males or a potential bias in recruitment, for exmaple: male is more willing to participate clinical studies.

Table 5.2. Male/female ratio in a selection of warfarin genetic studies.

Year	REF	n	Male	%	Female	%	Ethnicity	Note
2005	Bodin L et al	222	145	65.3	77	34.7	French	Normal subject
2005	*Uppsala study*	*201*	*135*	*67.2*	*66*	*32.8*	*Swedish*	
2005	Sconce EA et al	297	160	53.9	137	46.1	British	
2005	D'Andrea GR et al	180	100	55.6	80	44.4	Italian	
2006	Aquilante CL et al	350	306	87.4	44	12.6	American Caucasian (91%) Afro-American (7%) Hispanic 1% Asian (0.3%)	A large percentage of patients were recruited from 2 Veterans Administration
2006	Carlquist JF et al	213	104	48.8	109	51.2	Predominantly white, of Northern European	
2006	Schalekamp T et al	231	133	57.6	98	42.4	European	
2007	Borgiani P et al	148	78	52.7	70	47.3	Italian	
2008	Schwarz UI et al	297	160	53.9	137	46.1	White (89.2%) Black (9.8%) Hispanic (1.0%)	
2008	*WARG*	*1496*	*947*	*63.3*	*549*	*36.7*	*Swedish*	

5.2 GENOTYPING APPROACH

The aims of this study were to: (i) replicate in WARG the association signals we obtained in the Uppsala study and (ii) investigate the smaller effects in the 35 candidate genes. The strategy of choice involved following a haplotype tagging approach with inclusion of the lead SNPs and important functional variants from the discovery study. Patients recruited in the WARG study were predominantly Swedish; the same as the patients collected in the Uppsala study. In chapter III the construction of detailed LD maps for each of the 35 candidate genes was described. Selection of haplotype tag SNPs was performed using Tagger (de Bakker et al. 2005) built in Haploview (Barrett et al. 2005). This approach reduced the genotyping and labour costs.

5.2.1 SNP selection

All SNPs with $P \leq 0.1$, in either the univariate or multivariate regression analyses of warfarin dose association in the Uppsala study, were selected for inclusion in the WARG study. Tagging was undertaken in all 35 candidate genes with the software Tagger (de Bakker et al. 2005) setting the pairwise r^2 thereshold at 0.8 and MAF $\geq$ 5%. If a tag SNP selected by Tagger was in the same bin (tightly linked SNPs) with any of the pre-selected SNPs, it was then replaced by the pre-selected SNP to represent that bin. A total of 216 SNPs were selected to be tested across the WARG cohort. Any failed genotyping assay was redesigned using different chemistry, i.e. iPLEX to MassEXTEND or vice versa.

5.2.2 Taqman genotyping *CYP2C9**2 and *3

Since the effect of *CYP2C9**2 and *3 alleles on warfarin dose requirement has been confirmed in the Uppsala cohort and other studies (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005), it was important to obtain accurate genotypes for these two variants in all individuals. However the respective SNP assay designed for the *CYP2C9**2 and *3 allele was shown to suffer from allele drop-out on the Sequenom platform regardless of chemistry used, i.e. both iPLEX and MassEXTEND assays

The *CYP2C9**2 and *3 SNPs were genotyped using Taqman Pre-developed genotyping assays (Applied Biosystems). Experimental setup was as per manufacturer's instructions in 5 μl reaction volume with slight modification. The addition of five amplification cycles increased the genotyping call rate of 15%.

5.2.3 Genotyping summary

Although SNPs which had MAF below 5% were rejected from analysis in the Uppsala study because of the small sample size, a few of them were included with previous evidence of being functionally important in the WARG study. SNPs genotyped in the WARG study, out of Hardy-Weinberg Equilibrium (HWE, $P < 0.001$) and with call rate below 70%, were removed from analyses. 216 SNPs passed the study criteria with an average of genotype call rate of 95% (see Appendix III for call rate of each SNP). The MAF of the 216 SNP is shown in Figure 5.3.

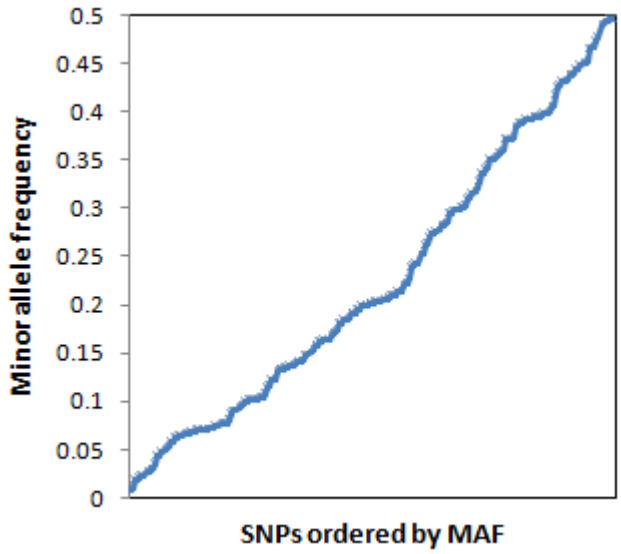

Figure 5.3. Minor allele frequency distribution of 216 SNPs genotyped in WARG study. The SNPs (X-axis) are ordered and plotted from low (1%) to high MAF (50%).

5.3 DOSE VARIATION

The univariate and multivariate linear regression model analyses were performed, with the R statistical package, to examine the association of dose and genetic and environmental factors. To correct for multiple testing, Bonferroni correction based on the effective number of independent tests (Meff) (Cheverud 2001; Li 2001; Nyholt 2004) was applied. The sum of effective tests is 190 in this study and the experiment-wise significance is P< 2.62E-04.

5.3.1 Gender, age, and dose

An unequal number of male and female individuals were recruited in the WARG study. In the Uppsala study, which comprises 201 patients, gender is not significantly associated with the dose when body weight is included in the multiple model. However, it is interchangeable when body weight information is not available. In the WARG study, gender is significantly associated with dose variation (p=6.89E-05). As mentioned earlier body weight information is not available in the WARG patients. Females required ~4 mg (10.7%) less than male patients in mean weekly dose, and the median weekly dose for male and female is 35.51 and 30.83 mg, respectively (Table 5.3; where N is the number of male or female patients).

Table 5.3. Gender effect in warfarin dose association.

Varable	N	Mean	Min	Median	Max
Sex					
M	947	37.42	5.97	35.51	113.75
F	549	33.48	6.45	30.83	103.95

Figure 5.4 is a plot of the weekly dose against age which shows a linear trend for both males

(A) and females (B). The majority of male patients were slightly younger than the female patients; which is in agreement with the earlier onset of heart related diesease in males. The linear trend line in female is more steep than in male suggesting that a smaller amount of warfarin is required for elder, as well as younger, female than male patients. Except for a few outliers, seven (1.27%) women out of 549 female patients were prescribed a dose of >70 mg/week and so were 28 men (2.96%) require a dose of >70 mg/week among 947 males.

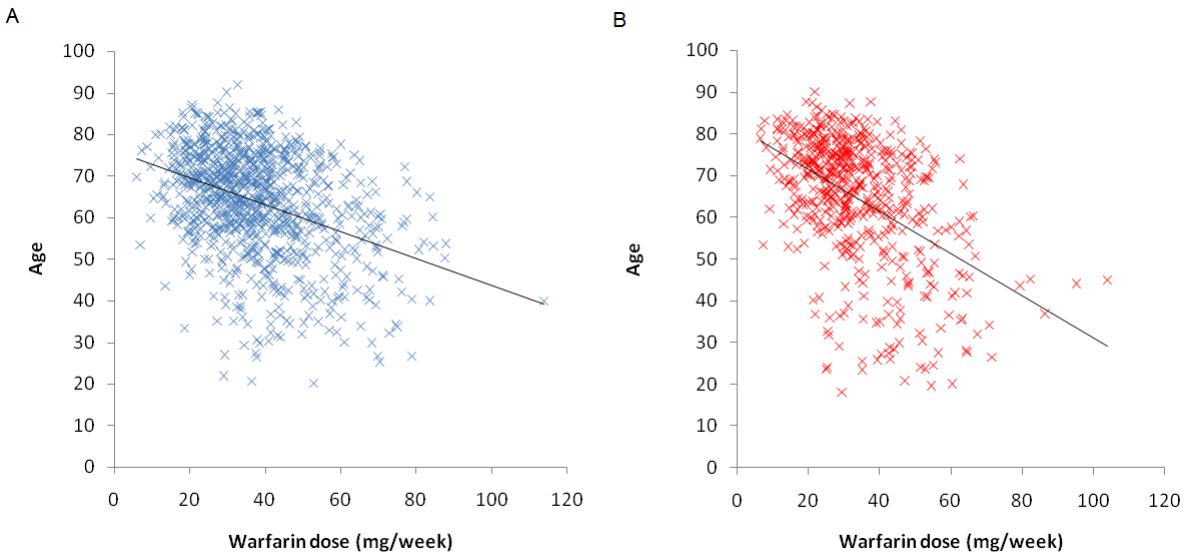

Figure 5.4. Age effect on warfarin dose association in (A) male and (B) female patients Patient's age and maintenance dose is plotted and a linear trend line is shown accordingly.

5.3.2 Univariate regression

Although genotypes of 1496 treated patients were obtained, only those with a stable PT INR between 2 and 3 were included in order to reduce any potential complex and confounding effect (n=1324). Table 5.4 lists the number of patients, if different exclusion criteria were concerned.

Among the 1496 patients, 88% (1324) of them achieved stable therapeutic PT INR of 2-3, while a set of 850 (64.2%) patients achieved stable dose during treatment. Although the inclusion of both stable INR and dose is more stringent, it will lead to 474 patients to be excluded from further analyses reducing the statistical power dramatically. We decided to carry out the analyses with consideration of stable INR criterion.

Table 5.4. Patient in the analyses with different inclusion criteria.

Inclusion criteria	Patients	Average dose	95% C.I.
All pateints	1496	33.8	(25.5, 43.6)
INR*	1324	33.8	(25.6, 44.5)
INR and dose#	850	32.5	(24.6,42.5)

For minimal three consecutive visit in clinic, *a stable INR between 2 and 3, plus #without dose variation

Univariate regression analysis was performed to test the association of inter-individual dose variation and the 216 tag SNPs on the 35 candidate genes. Figure 5.5 shows a Quantile-quantile plot of the observed versus expected p-values. Both expected and observed p-values are plotted in logarithm scale for the SNPs and the black line represents the close adherence of p-value which corresponds to the null hypothesis. The dots deviating from the black line suggest potential significant associations. The two black arrows indicate the result from the two SNPs in *VKORC1*, rs2359612 and rs9923231. The red arrow indicates the result of combined *CYP2C9**2 and *3 in association with dose. A total of 13 SNPs behave differently to the null hypothesis (above the black line) and are significant after correction for multiple tests; listed in Table 5.5.

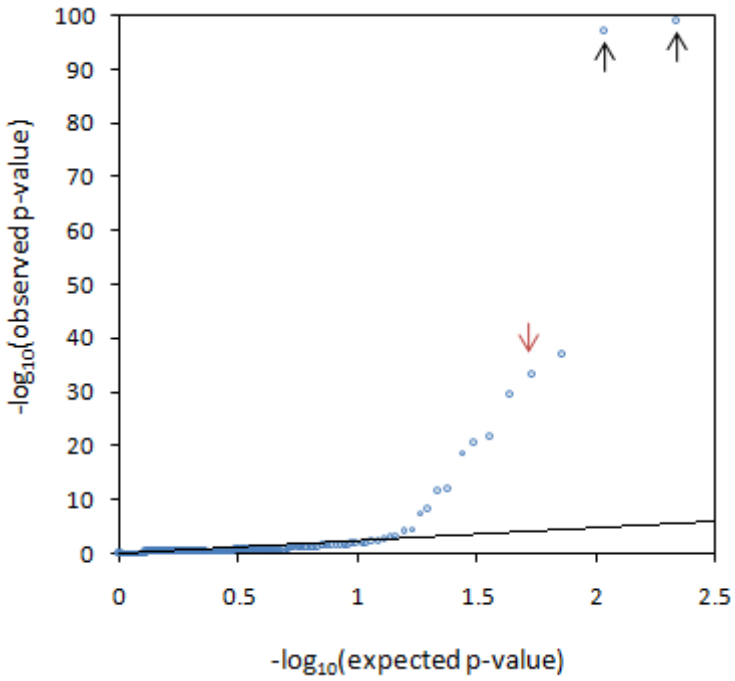

Figure 5.5. Quantile-quantile plot for univariate analysis in dose association. The black line indicates the close adherence of p-value corresponding to the null hypothesis.

A total of thirteen SNPs in *VKORC1* and the *CYP2C9-CYP2C19-CYP2C8* cluster were significantly associated with warfarin after Meff correction for multiple tests ($P \leq 2.62E-04$). The value of R^2 indicates the ratio of dose variation explained in the univariate regression model. The strongest associations with dose were observed for two nearly perfectly concordant SNPs, rs2359612 and rs9923231, in *VKORC1* which explained 29.8% and 29.3% of the inter-individual dose variance, respectively (Table 5.5).

Although rs2359612 gave a slightly better p-value associated with dose, rs9923231 has been demonstrated to potentially influence *VKORC1* mRNA expression (Rieder et al. 2005; Yuan et al. 2005). The slight difference in p-value is down to a small difference in call rate between the two SNPs. Based on the available functional evidence (see Figure 4.6), and the fact that the two SNPs are perfectly correlated, only rs9923231 was included in further analyses to represent the effect of *VKORC1*. In the group of patients with not only stable coagulation but

constant dose, SNPs rs2359612 and rs9923231 explained 33.3% ($P = 5.78x10^{-73}$) and 32.8% ($P = 3.97x10^{-72}$) of dose variation, respectively, which provides strong evidence that *VKORC1* is associated with the stable dose.

Table 5.5. SNPs showing significant association after Meff correction.

Gene	SNP	MAF	Patients	%	R^2	P-value
VKORC1	rs2359612	0.394	1292	97.58%	29.8	9.82E-100
VKORC1	rs9923231	0.393	1293	97.66%	29.3	1.03E-97
VKORC1	rs7294	0.393	1294	97.73%	12.3	1.21E-37
CYP2C9	CYP2C9 (*2 & *3)	n/a	1318	99.55%	11.8	6.63E-34
CYP2C9	rs4917639	0.209	1092	82.48%	11.7	4.61E-30
CYP2C9	rs2860905	0.214	1296	97.89%	7.4	3.05E-22
CYP2C19	rs3814637	0.073	1284	96.98%	7.1	3.14E-21
CYP2C9	rs1057910 (*3)	0.071	1321	99.77%	6.3	1.82E-19
CYP2C9	rs1799853 (*2)	0.114	1321	99.77%	4.1	1.34E-12
CYP2C8	rs11572080	0.1	1256	94.86%	4.1	3.16E-12
CYP2C9	rs1856908	0.433	1291	97.51%	2.9	6.71E-09
CYP2C8	rs10509681	0.093	1147	86.63%	3	3.08E-08
VKORC1	rs11150606	0.027	1309	98.87%	1.3	2.55E-05
CYP2C8	rs2275620	0.417	1204	90.94%	1.6	7.99E-05

Significant p-value after Bonferroni Meff correction is $P \leq 2.62x10^{-4}$.

As expected, the second best associated gene in this study is *CYP2C9*. In the univariate model, *CYP2C9**2 and *CYP2C9**3 individually explained 4.1% ($P = 1.34x10^{-12}$) and 6.3% ($P=1.82x10^{-19}$) of dose variation, respectively. However, when both alleles were taken into account simultaneously, *2 and *3 explained a sum of 11.8% of dose variation ($P = 6.63x10^{-34}$). A further analysis looking at the stable dose subgroup indicated these variants explained 12.3% of the variance ($P = 3.67x10^{-22}$).

Other than *CYP2C9*, SNPs in *CYP2C8* and *CYP2C19* displayed significant associations after correction for multiple testing (Table 5.5). However, LD analysis indicated that these

associations could be fully explained by either the *CYP2C9**2 or *CYP2C9**3 allele. In the multiple regression model (with either *CYP2C9**2 or *CYP2C9**3), none of the additional SNPs remains significant except rs3814637 in *CYP2C19*, and this shall be discussed in section 5.3.3.

The association of dose with *VKORC1* rs9923231 genotypes is illustrated in Figure 5.6. The average dose of patients having GG / AG / and AA genotypes is 42.5 mg (95 % CI: 23.3, 54.3), 32.1 mg (95 % CI: 25.8, 41.2), and 22.2 mg (95 % CI: 16.6, 27.0), respectively. The A allele has been reported to be associated with low mRNA expression (Rieder et al. 2005) which suggests lower warfarin dose could reach the therapeutic effect by blocking vitamin K recycling. In the WARG study, 529 patients are GG homozygotes whilst 714 are AG heterozygotes and 218 AA homozygotes. The patients homozygous for the G allele required a dose almost twice as much as patients homozygous for the A allele (Figure 5.6).

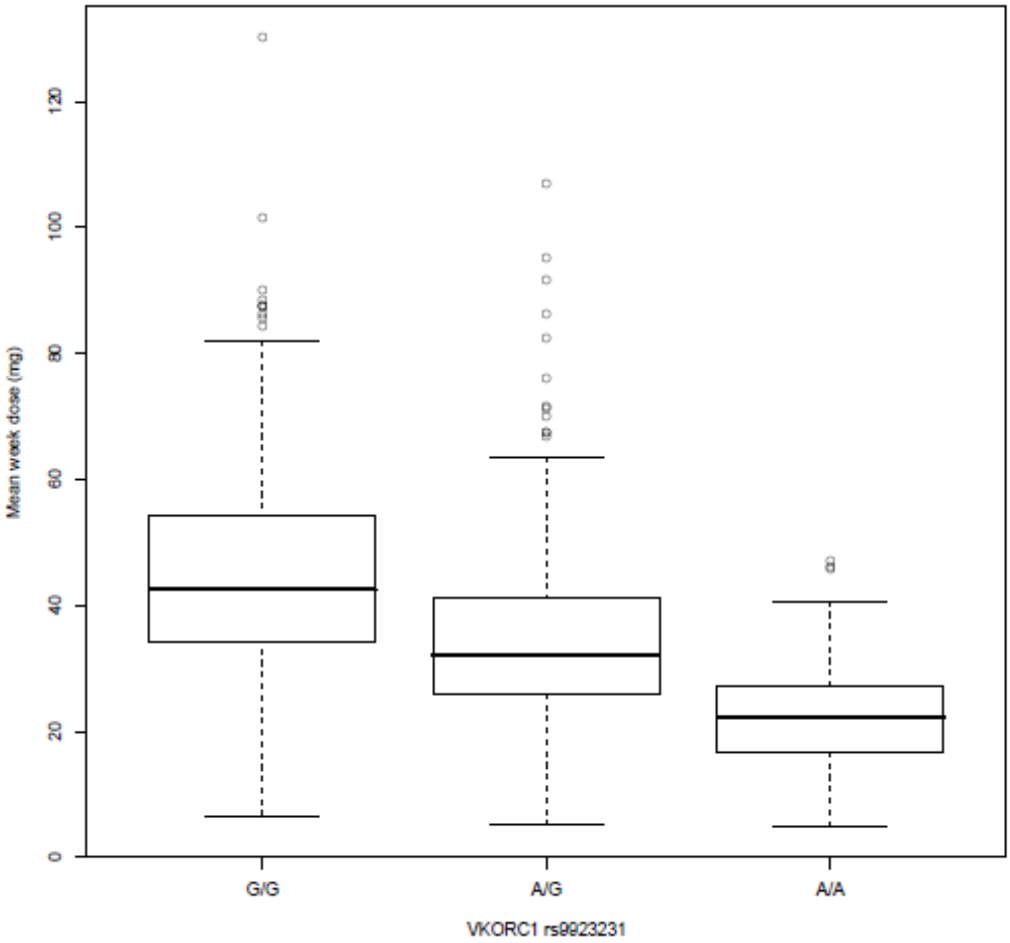

Figure 5.6. Mean weekly dose of rs9923231 genotype in WARG Swedish. The homozygous A / A patients (218) are associated with low maintenance dose whereas G / G homozygotes (529) need twice the dose compared to the A / A homozygotes.

*CYP2C9**2 and *3 alleles are known to impair the enzymatic activity of CYP2C9 in metabolising warfarin (Crespi and Miller 1997; Rettie et al. 1994; Sullivan-Klose et al. 1996). In the WARG study, patients not carrying the *2 and / or *3 alleles require an average warfarin dose of 36.79 mg/week (95% CI: 28.52, 47.56) whereas patients homozygous for the *3 allele (8 patients) required essentially low maintenance dose of 7.87 mg/week (95% CI: 5.48, 11.03) and those homozygous for the *2 allele (21 patients) needed 22.50 mg/week (95% CI: 17.23, 30.52). This result indicates how substantial the effect of *2 and *3 variants is. The patients being *1 / *2 and *1 / *3 heterozygotes require 6.74 and 8.25 mg less dose than *1/*1 homozygotes whereas *2/*3 homozygotes require a weekly dose of 19.86 mg

(Figure 5.7).

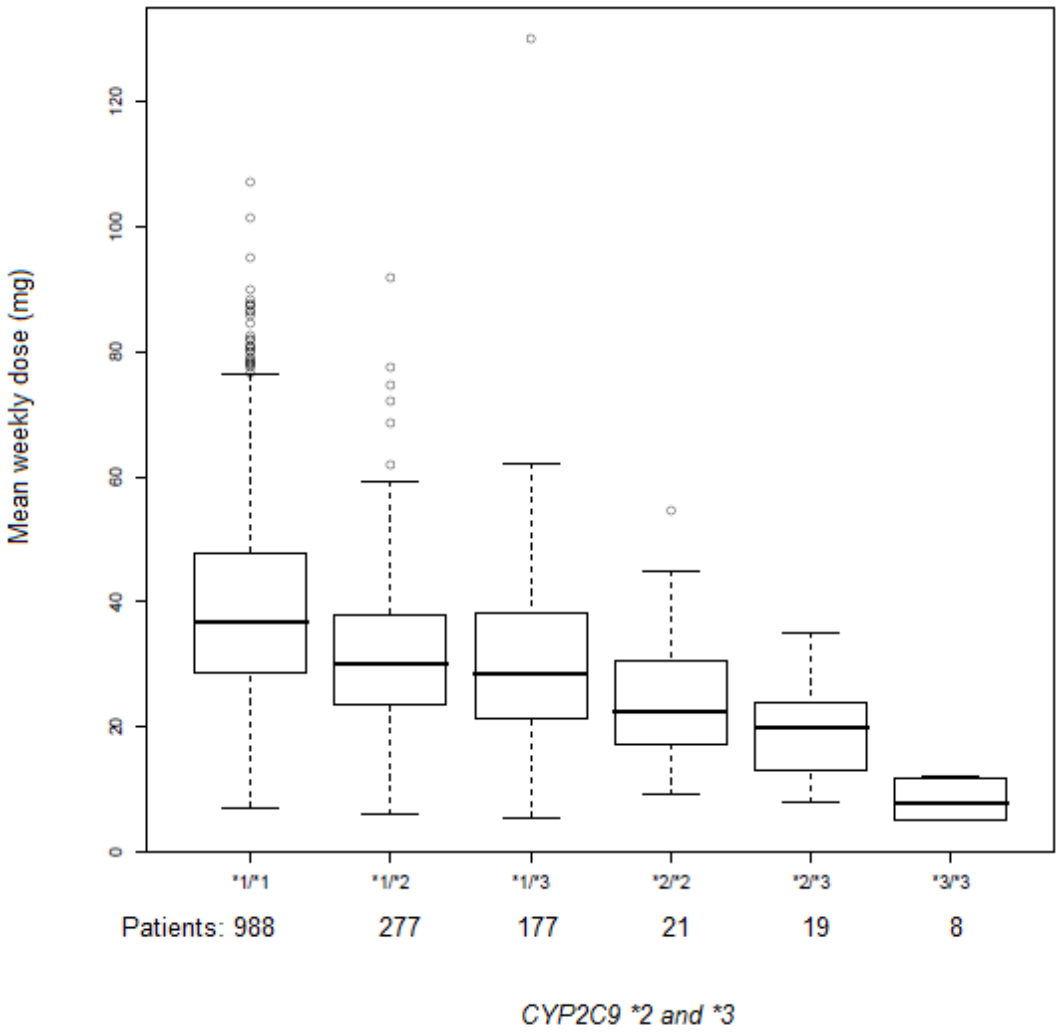

Figure 5.7. Mean weekly dose of rs9923231 genotype in WARG patients. The number of patients carrying different alleles is listed below each genotype.

5.3.3 Comparison of results in the Uppsala and WARG studies

In the Uppsala study (chapter IV), SNPs in *PROC*, *EPHX1*, *GGCX* and *ORM1-2* together explained an additional 17% of dose variation (section 4.6), and these SNPs were tested in the WARG study for replication. Table 5.6 summarises the comparison result in both Uppsala

and WARG studies of the SNPs substantially explaining dose variation in the Uppsala study.

Except *VKORC1* rs9923231 and *CYP2C9**2 and *3, other SNPs significant in the Uppsala study are not replicated in the WARG cohort (Table 5.6; R-square value indicates the explained ratio of inter-individual dose variation for each SNP). *VKORC1* explains an essential 29.3% of dose variation in WARG whilst *CYP2C9**2 and *3 explain 11.8% in univariate analysis. *PROC* SNP rs2069919 reached a gene-wise significant p-value in the Uppsala study explaining a substantial 9% of dose but reports nothing in WARG. SNPs in *EPHX1*, *GGCX* and *ORM1* reach nominal significance in Uppsala study and also explain substantial dose variation but yet do not reach significance in WARG.

The *GGCX* SNP rs4653436 was not directly typed in WARG, and instead SNP rs7568458 which is tightly linked to rs4653436 was tested. The finding in *GGCX* in the Uppsala study is not replicated, a recent Japanese study analysed 828 warfarin-treated patients and found no effect from *GGCX* (Cha et al. 2007). Meanwhile, Rieder and colleagues reported SNP rs11676382 in *GGCX* had a small but significant effect on warfarin maintenance dose (Rieder et al. 2007). Together with our finding in both Uppsala and WARG studies, the effects of GGCX are potentially population/sub-population dependent.

Although the lead SNPs in the Uppsala study for *EPHX1* and *ORM1*, rs4653436 and rs1687390, respectively, did not replicate in WARG (Table 5.6), other SNPs in these two genes, rs3817268 (P = 6.47E-03) and rs6426089 (P = 3.59E-02) in *EPHX1* and rs2787337 in *ORM1* (P = 1.08E-02) did report nominally significant associations (Table 5.7). Despite different SNPs associated with dose in the WARG cohort, however, the recurrence of the two genes suggests a possible minor effect caused by the two genes. Table 5.7 lists SNPs other

than those in the CYP2C cluster that reached nominal significance.

Table 5.6. Comparison between Uppsala and validation studies.

Genes in association	Uppsala study (n=201)		WARG study (n=1496)	
with warfarin dose	R-square	P-value	R-square	P-value
VKORC1 rs9923231	31.70%	$1.91x10^{-15}$	29.30%	$1.03x10^{-97}$
CYP2C9 *2 and *3	15.90%	$2.30x10^{-6}$	11.80%	$6.63x10^{-34}$
PROC rs2069919	9.00%	0.0002	0.20%	0.2073
EPHX1 rs4653436	4.80%	0.0084	0.10%	0.4247
GGCX rs12714145*	3.40%	0.0332	0.10%	0.4966*
ORM1 rs1687390	2.60%	0.0496	0%	0.9145

Table 5.7. Other SNPs showing nominal significant association with dose.

Gene	SNP	MAF	Patients	%	P-value
EPHX1	rs3817268	0.274	1285	97.66%	6.47E-03
PROS1	rs8178633	0.050	1218	97.73%	7.03E-03
ORM1	rs2787337	0.311	1227	99.55%	1.08E-02
CYP1A1	rs2470893	0.342	1271	82.48%	1.49E-02
CYP3A4	rs4986910	0.008	1299	97.89%	1.85E-02
PDIA5	rs1107377	0.496	1241	96.98%	2.46E-02
PROS1	rs9683303	0.353	1223	99.77%	3.26E-02
EPHX1	rs6426089	0.496	1288	99.77%	3.59E-02
PDIA3	rs11070411	0.164	1227	94.86%	4.24E-02
F10	rs2251102	0.192	1167	97.51%	4.82E-02

Apart from *EPHX1* and *ORM1* six more genes *PROS1*, *CYP1A1*, *CYP3A4*, *PDIA5*, *PDIA3* and *F10* also showed marginal association with warfarin dose. These genes are all involved in pharmacokinetics and pharmacodynamics of warfarin and require replication in other populations.

5.3.4 Multivariate

Outside *VKORC1,* SNPs showing significant association are tightly linked to either *CYP2C9**2 or *3 alleles in the WARG study. In the analysis of the Uppsala cohort, rs4917639 explained a substantial 10% and 2.5% of dose in the multivariate model including *CYP2C9**2 or *3 as covariates, respectively (section 4.4). The LD analysis indicated that this SNP is associated with both *2 and *3 and displayed a sum effect of both alleles. This association is also observed in the WARG study. The SNP rs4917639 explained approximately 11.7% ($P = 4.61x10^{-30}$) and 12.1% ($P = 1.99x10^{-20}$) of dose variation in models including either *2 or *3 alleles. It fails to explain any dose in the multiple model including both *2 and *3 allele. The LD between rs4917639 and *2 / *3 is present in a lesser extent in WARG ($r^2 = 0.836$) than in the Uppsala study but could still reflect the outcome of significant p-value and the substantial explanation of dose variation.

Apart from rs4917639, rs3814637 is the only SNP significant and moderately independent to *CYP2C9**2 and *3 in the WARG study ($P = 2.2x10^{-6}$) after Meff correction, although it is tightly linked to *CYP2C9**3 in the Uppsala study. However, in the multiple regression model including *VKORC1* rs9923231 and *CY2C9**2 / *3 alleles, *CYP2C19* SNP rs3814637 explained a merely 0.7% of dose variation and was therefore not considered for inclusion in the final regression model. However, this result does suggest a contribution of *CYP2C19* in warfarin dose requirement.

Except *VKORC1* and *CYP2C9**2 and *3, no SNP explains substantial dose variation in the WARG study. The promoter SNP, rs9923231, in *VKORC1* and *CYP2C9**2 / *3 are together as the only genetic determinants with broad utility for determining warfarin dose.

5.4 REGRESSION MODEL

In the Uppsala study, non-genetic factors including age, gender, body weight, medical indication, drug interaction, and PT INR value were tested and only age, body weight, and drug interaction were included in the final multiple regression model (section 4.6). In the WARG study, information of patient's body weight was not recorded and thus gender was included as a surrogate in the multiple model.

The predictors in the multiple model are listed in Table 5.8 including non-genetic factors age, gender, and drug interaction as well as *VKORC1* rs9923231 and *CYP2C9* *2 and *3 as the genetic determinants. Notably, the drug interaction is significantly associated with dose variation if an overall outcome is potentiating. The dose explained (R^2) and relevant p-value for each predictor is listed on Table 5.8.

Table 5.8 Predictors in multiple regression model.

Predictor		R^2	Patients	P-value
CYP2C9 *2 and *3		11.8	1318	6.63E-34
VKORC1 rs9923231		29.3	1293	1.03E-97
Age		14.5	1324	8.00E-47
Gender		1.2	1324	6.89E-05
Drug interaction	(potentiate)	1.5	1324	8.00E-06
	(decrease)	0	1324	9.03E-01

To achieve the best available coefficient for each predictor, cross validation was applied to test each model. The entire WARG database was thereby divided into two portions, 70% patients as the training dataset whilst 30% were tested for validation for performance. The

parameters for each predictor were determined after 1000 iterations of cross validation and the final model could explain 58.7% of warfarin dose variation (Figure 5.8A).

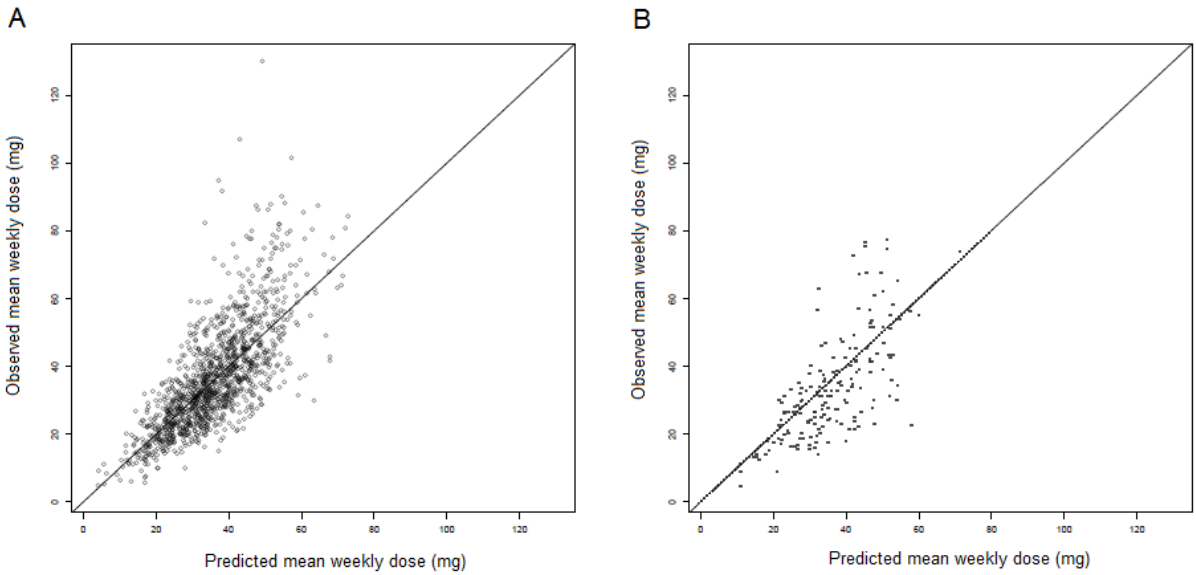

Figure 5.8. Multiple regression model developed accordingly in the WARG study. This model explains (A) 58.7% dose variation in the WARG cohort and (B) 53% in the Uppsala cohort.

The model was then used to estimate the patients in the Uppsala study. Information was available for all predictors in 181 individuals. The developed model predicts slightly higher dose for Uppsala patients, and 53% of dose variation was successfully predicted (Figure 5.8B). Interestingly, results from both WARG and Uppsala studies shows that some outliers required higher dose than the predictive amount, this suggests that some elements remain to be discovered.

5.5 THERAPEUTIC STABILISATION

Warfarin is difficult to administer because of its narrow therapeutic range of PT INR 2-3. When a patient is prescribed with warfarin, a trial dose is given and PT INR is closely monitored before increasing or decreasing the dose. Over anti-coagulation often leads to severe bleeding complication and hospitalisation. Therefore, it is important to investigate the factors which might contribute to the required time of reaching stable INR.

5.5.1 INR stabilisation

The INR values of each patient throughout the treatment period were analysed and plotted with LOESS regression model according to their *VKORC1* rs9923231 genotypes (Figure 5.9). LOESS regression is widely used in analysing localised data subsets and variation in the data point by point without specifying a global function and is, therefore, suitable in analysing PT INR value for each patient over different time points.

Patients with homozygous AA genotypes (green dotted line) had an INR peak at 2.7 before moving towards 2.5 of stable INR in the first week of therapy. Meanwhile, AG heterozygote patients showed slight fluctuation compared to the homozygous GG patients until they were stabilised. This result shows that AA homozygotes for *VKORC1* rs9923231 require lower warfarin dose, but patients were over dosed in the beginning of treatment resulting in over-anticoagulation and unstable PT INR in the first week. Our results suggest that this could be rectified if *VKORC1* genotype information is used. Once the INR is stabilised, *VKORC1* genotype has no effect on further anticoagulation stability.

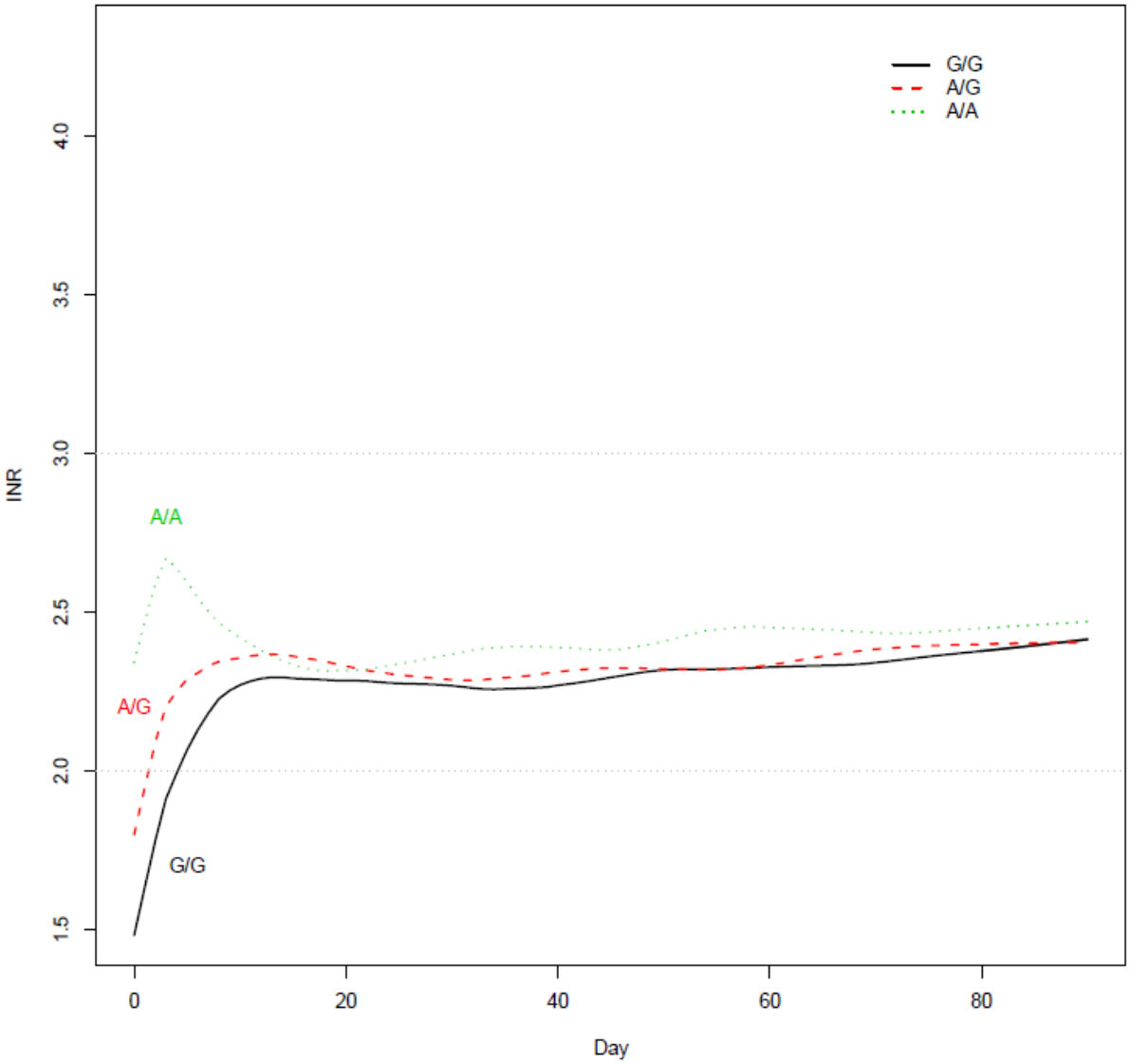

Figure 5.9. Lowess smoothed plot of PT INR values of patients treated with warfarin. Patients with different *VKORC1* rs9923231 genotype were plotted over treatment period.

*CYP2C9**2 and *3 alleles were analysed in LOWESS regression analysis for their effect on PT INR stability. Patients homozygous for *CYP2C9**3 had also extremely unstable INR values, with an average INR peak exceeding 4 during the first two weeks, and subsequent instability for three months (pink line, Figure 5.10). As a consequence of impaired enzymatic activity of CYP2C9, patients homozygous for the *3 allele were very sensitive in response to the increase or decrease in warfarin dose. This PT INR instability is also reflected in patients carrying *2 / *3 (light blue) and *2 / *2 alleles. Warfarin metabolism in heterozygotes

carrying *1 / *2 or *1 / *3 were compensated by the normal allele and they require a comparatively stable dose than those homozygous and heterozygous for *2 and *3 alleles. Only patients homozygous for the *1 allele (solid black line) reached stable anticoagulation of PT INR 2.5 in two weeks.

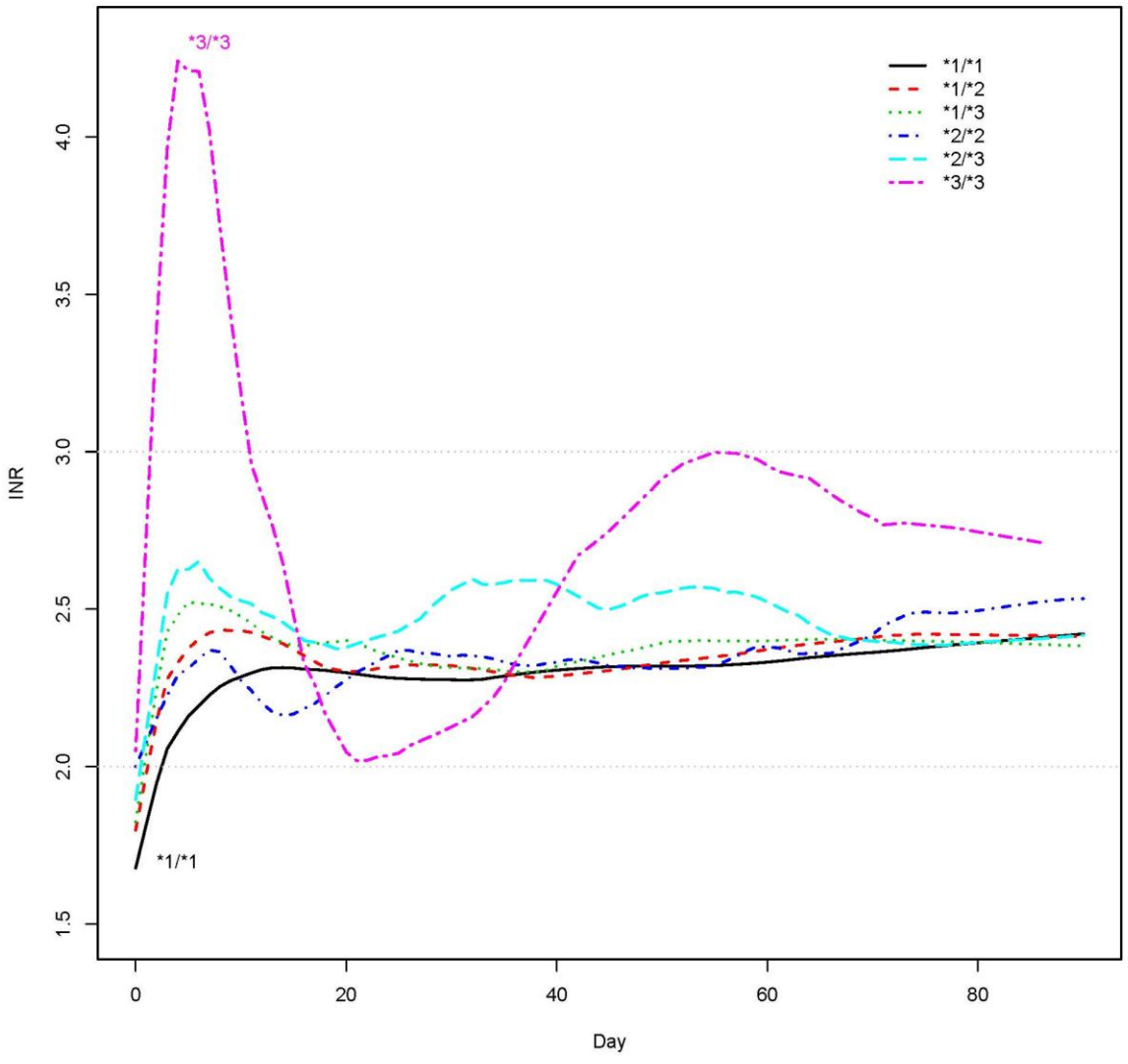

Figure 5.10. Lowess smoothed plot of PT INR values of patients treated with warfarin. Patients carrying normal or *CYP2C9**2 and / or *3 alleles were plotted over treating period.

5.5.2 Over-anticoagulation

To investigate how genetic variants influence over-anticoagulation in the first month and especially the first week treatment, all tag SNPs in the 35 candidate genes were also tested for association, with over-anticoagulation defined as PT INR > 4 within first five weeks of treatment, with log rank test. Log rank test (also called Mantel-Haenszel test or the Mantel-Cox test) is widely used clinically to test the survival distributions of different samples. Each SNP in the 35 candidate genes was tested and the type I error (p-value) was estimated. Table 5.9 only lists the SNPs showing significant p-value after correction for multiple tests.

Table 5.9. SNPs associated with over anti-coagulation in warfarin treatment.

Gene	SNP	MAF	Pateitns	P-value
CYP2C19	rs3814637	0.073	1447	1.11E-16
CYP2C9	rs1057910 (*3)	0.071	1493	1.11E-16
CYP2C9	CYP2C9*2 and *3		1490	1.11E-16
VKORC1	rs2359612	0.394	1455	4.19E-12
VKORC1	rs9923231	0.393	1461	6.56E-12
CYP2C9	rs4917639	0.209	1220	9.78E-06

Five SNPs are significantly associated with over-anticoagulation. The effect of rs3814637 in *CYP2C19* and rs4917639 in *CYP2C9,* are known to be well explained by *CYP2C9**2 and *3 (sections 4.4 and 5.3.4) and were therefore removed from further analysis. Despite the fact that it was not even nominally significant in the test, *CYP2C9**2 was analysed together with the *3 allele. As explained previously, rs9923231 was used for *VKORC1* (section 6.3.2). Therefore, only rs9923231 in *VKORC1* and *CYP2C9**2 / *3 are significantly associated with over-anticoagulation in the first five weeks treatment.

The effect of *VKORC1* rs9923231 and *CYP2C9**2 / *3 in over-anticoagulation was analysed in Cox regression model and illustrated with Kaplan-Meier curve plot. Cox regression model is a survival model for describing the risk changes over time, such as the choice of treatment or the effect of genotype. The advantage of Cox model is that it does not require the calculation of the hazard function. Kaplan-Meier curve is frequently used for illustration of survival function of life-time data. In medical statistics, this is often used to illustrate cumulative probability of clinical events, such as gain of tumour or survival after treatment.

Patients who are AA homozygotes for *VKORC1* SNP rs9923231 appeared to have a significantly increased probability of over-anticoagulation (dotted line, Figure 5.11A). In the first 5 weeks of treatment, 4.56 hazard ratio (95% CI: 2.85, 7.30) was observed in AA homozygous patients with PT INR above 4 comparing to GG homozygous patients (P = 2.4E-10). When the patients is heterozygous, i.e. carrying a G allele, a much lower hazard ratio of 1.74 (95% CI: 1.11, 2.71) was observed (P = 1.5E-02) in comparison with GG homozygotes (Figure 5.11A).

Figure 5.11B shows the effect of *CYP2C9**2 and *3 alleles. The over-anticoagulation is significantly severe in patients carrying homozygous CYP2C9*3 allele (pink dash line, Figure 5.11B). A significant increase of hazard ratio of 21.84 (95% CI: 9.457, 50.42) is observed in patients homozygous for *3 allele. Patients heterozygous for *2 / *3 alleles showed a hazard ratio of 2.98 (95% CI: 1.092, 8.15) comparing to *1 / *1 homozygotes (Figure 5.11B).

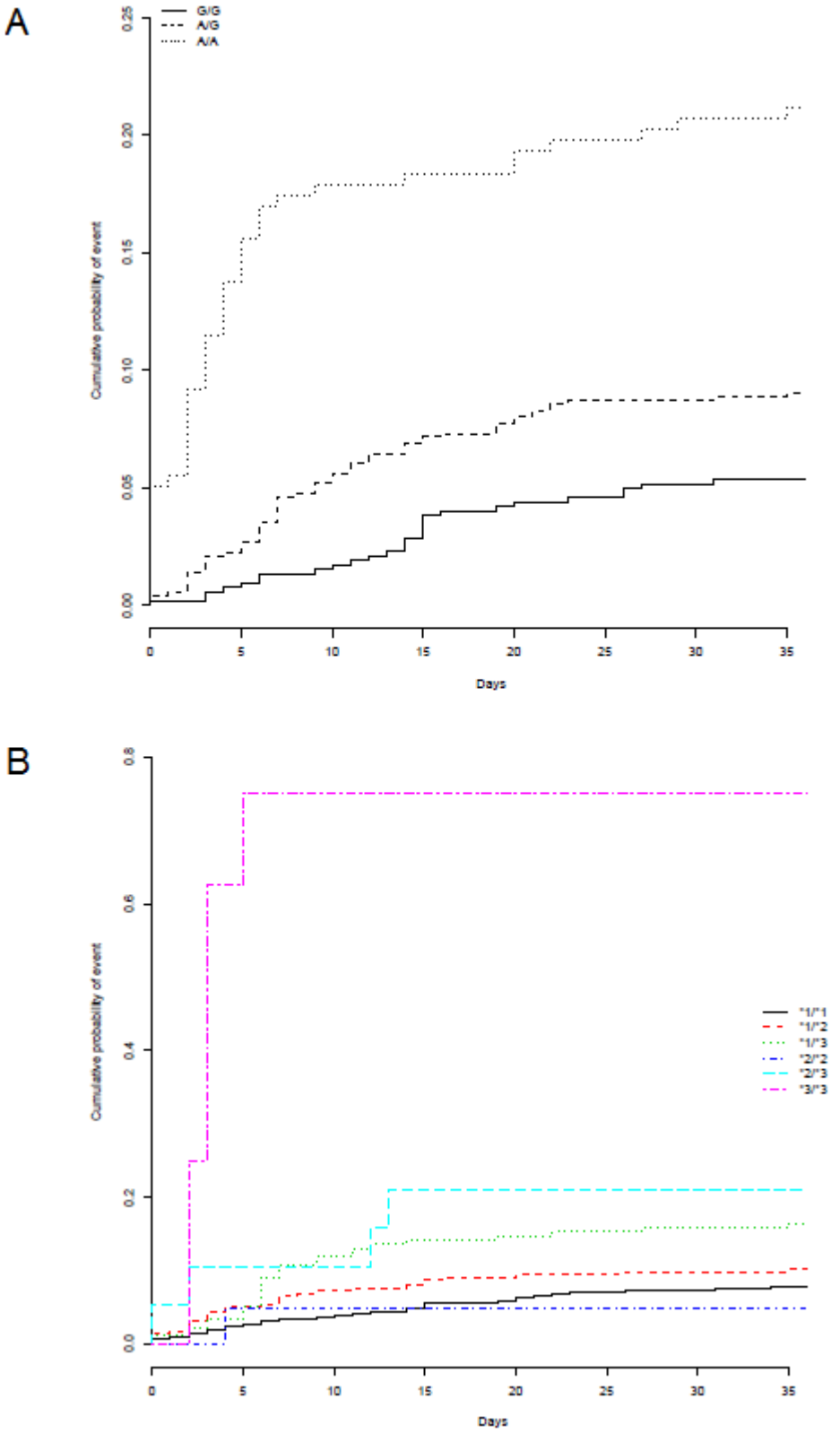

Figure 5.11. Survival (Kaplan-Meier) curve of cumulative probability in patients with PT INR > 4 related to (A) *VKORC1* rs9923231 and (B) *CYP2C9**2 and *3.

5.6 CONCLUSION

The WARG study provided a validation of the initial findings in the Uppsala study and a *de novo* investigation for the impact of the 35 candidate genes on warfarin dose requirement. Only genotypes of *VKORC1* and *CYP2C9* have a demonstrated effect on warfarin dose. Moreover therapeutic stability, that is stable INR and dose, is also influenced by *VKORC1* and *CYP2C9*. The inconsistent findings between the Uppsala and WARG studies suggest minor effects of these genes which may be due to study design/being treatment specific. These small effects were therefore excluded in the development of a global dosing algorithm. Meanwhile, sample size is inevitably important in identifying contributors in association studies.

The CYP2C gene cluster on chromosome 10 has been intensively studied for various drugs. Different methods have been suggested to tag the polymorphisms in this region (Ahmadi et al. 2005; Walton et al. 2005). In the WARG study, this region is comprehensively tagged with 25 SNPs and no effect other than the *CYP2C9**2 and *3 alleles is found.