A genome-wide association study of acenocoumarol maintenance dosage

Martina Teichert1,5, Mark Eijgelsheim1, Fernando Rivadeneira1,2, André G. Uitterlinden1,2,3,5, Ron H.N. van Schaik2,7, Albert Hofman1,6, Peter A.G.M. De Smet5,8, Teun van Gelder3,4, Loes E. Visser1,4 and Bruno H.Ch. Stricker2,6,9,*

1Department of Epidemiology, 2Department of Internal Medicine, 3Department of Clinical Chemistry and 4Department of Hospital Pharmacy, Erasmus Medical Center, Rotterdam, The Netherlands, 5Scientific Institute Dutch Pharmacists, The Hague, The Netherlands, 6Netherlands Genomics Initiative-Sponsored Netherlands Consortium for Healthy Aging, Rotterdam, The Netherlands, 7Star Medical Diagnostic Center, Rotterdam, The Netherlands, 8Department of Clinical Pharmacy, UMC St Radboud, Nijmegen, The Netherlands and 9Drug Safety Unit, Inspectorate for Health Care, The Hague, The Netherlands

Received April 21, 2009; Revised June 26, 2009; Accepted July 2, 2009

Several genome-wide association studies have been performed on warfarin. For acenocoumarol, the most frequently used coumarin in many countries worldwide, pharmacodynamic influences are expected to be comparable. Pharmacokinetics however might differ. We aimed to confirm known or identify new genetic variants contributing to interindividual variation on stabilized acenocoumarol dosage by a GWAS. The index population consisted of 1451 Caucasian subjects from the Rotterdam study and results were replicated in 287 subjects from the Rotterdam study extended cohort. Both cohorts were genotyped on the Illumina 550K Human Map SNP array. From polymorphisms tested for association with acenocoumarol dosage, 35 single nucleotide polymorphisms (SNPs) on chromosome 16 and 18 SNPs on chromosome 10 reached genome-wide significance. The SNP with the lowest P-value was rs10871454 on chromosome 16 linked to SNPs within the vitamin K epoxide reductase complex subunit 1 (VKORC1) (P = 2.0 × 10−123). The lowest P-value on chromosome 10 was obtained by rs4086116 within cytochrome P450 2C9 (CYP2C9) (P = 3.3 × 10−24). After adjustment for these SNPs, the rs2108622 polymorphism within cytochrome P450 4F2 (CYP4F2) gene on chromosome 19 reached genome-wide significance (P = 2.0 × 10−8). On chromosome 10, we further identified genetic variation in the cytochrome P450 2C18 (CYP2C18) gene contributing to variance of acenocoumarol dosage. Thus we confirmed earlier findings that acenocoumarol dosage mainly depends on polymorphisms in the VKORC1 and CYP2C9 genes. Besides age, gender, body mass index and target INR, one polymorphism within each of the VKORC1, CYP2C9, CYP4F2 and CYP2C18 genes could explain 48.8% of acenocoumarol dosage variation.

INTRODUCTION

Clinical management of anticoagulation with coumarins is difficult, as a target range has to be achieved by drugs with a narrow therapeutic index and high intra- and interindividual variability in pharmacokinetic and pharmacodynamic response. In Caucasians, single nucleotide polymorphisms (SNPs) in the vitamin K epoxide reductase complex subunit 1 (VKORC1) and cytochrome P450 2C9 (CYP2C9) genes are more strongly correlated with stabilized coumarin dosage than all other known patient-related factors (1–11). Other candidate genes that are part of the vitamin K pathway, vitamin K dependent clotting factors or minor metabolic or transport pathways have been systematically screened and showed weaker associations with warfarin maintenance dosage (12). Recently, a candidate gene study on warfarin dosing found...
that CYP4F2 genetic variation explained an additional 2% of dosage variation (13). This result was replicated (14). As genetic variation in the VKORC1 and CYP2C9 genes in combination with age, gender and body mass index (BMI) accounted for 30–50% of the variability in dosing of warfarin, the Food and Drug Administration in 2007 extended the product labeling information of warfarin mentioning that genetic data might be relevant to prescribing decisions (15). Most studies focused on warfarin, the worldwide mostly used coumarin drug. Acenocoumarol, which is the preferred coumarin in The Netherlands (16) and many other countries worldwide, has been studied for associations of polymorphisms in the VKORC1 and the CYP2C9 genes with outcomes in acenocoumarol treatment in four studies for initiation period (2,17–19) and in three studies for maintenance dosage (1,20,21). Both VKORC1 and CYP2C9 polymorphisms were not only independently correlated with acenocoumarol maintenance dose, but also with first international normalized ratio (INR) after initial standardized dosage, time to stabilized dose, time within the target therapeutic range and to bleeding events. It is expected that more reliable personalized coumarin dosing strategies could improve quality of care as well as enable safer treatment in more patients who would benefit from anticoagulation therapy (22,23). Therefore, it is important to determine whether there are common variants in genes which affect acenocoumarol dosage. Genome-wide association studies (GWAS) covering the majority of common variation in the human genome can be used in large samples to identify genetic variants that typically confer modest effect sizes for quantitative complex traits, defined in homogenous phenotypes, such as INR and coumarin dosage. The dose prediction models already developed for warfarin may not be applicable to acenocoumarol because of differences in pharmacokinetics between these coumarines (24). A GWAS of acenocoumarol dosage either confirming already known or identifying new genetic influences may supply important information for the clinical trials on genotype predicted dosage being under way. Therefore, we conducted a GWAS in two independent cohorts to identify polymorphisms that could explain a large fraction of the variance in acenocoumarol maintenance dosage.

RESULTS

Index and replication cohorts

For the index population from the total of 7983 subjects in the RS-I cohort, 5974 (75%) subjects were successfully genotyped and passed quality control (QC) for GWAS. Of these 5974 subjects, 1451 (24%) started acenocoumarol therapy at a standard initial dose and had a first INR measurement at the fourth day available. Of the 3011 subjects in the RS-II cohort 1895 persons (63%) were successfully genotyped for GWAS and 287 subjects (15%) started with acenocoumarol during the study period and had an INR measurement on day 4 after standard initial dose. Owing to study design, subjects in the replication cohort were on average 7 years younger and the percentage of male subjects was higher than in the RS-I cohort (50% versus 43%, Table 1). For BMI in the RS-I cohort, values for 39 persons (2.8%) had to be imputed and for the RS-II cohort, one value (0.3%) was missing. On average, BMI was slightly higher in the replication cohort. In our study populations, acenocoumarol was prescribed with a very low target range (2.0–2.5) for short-term prophylactic treatment, with a low target INR range (2.5–3.5) also in patients for short-term prophylactic treatment, for deep venous thrombosis, pulmonary embolism, atrial fibrillation and cerebrovascular ischemia, with medium target INR range (3.0–4.0) in patients with atrial fibrillation, myocardial infarction, vascular surgery, stroke, transient ischemic attacks and periphery artery disease and with a high target INR range (3.5–4.5) for patients with prosthetic heart valves (Table 1).

Polymorphisms associated with acenocoumarol dosage

Basic model. In the RS-I cohort, only SNPs on chromosome 16 and chromosome 10 were in genome-wide significant association with acenocoumarol dosage at the end of initiation period, adjusted for age, gender, BMI and target INR (Table 2 and Fig. 1). The 15 SNPs with P-values lower than 10\(^{-8}\) were all located on chromosome 16. Best associated, with a P-value of 2.0 × 10\(^{-12}\), was rs10871454, located in the Syntaxis 4 A placental (STX4A) gene, flanking the VKORC1 gene (Fig. 2). Each additional minor allele of this SNP showed a decrease of weekly acenocoumarol dosage with 5.2 mg/week. This SNP was in linkage disequilibrium (LD) (r\(^2\) = 0.99), one of the two SNPs that were able to predict the influence of VKORC1 on warfarin dosage (6,25) and possibly one of the two causal putative SNPs within the VKORC1 gene (5) (Fig. 3).

The second locus reaching genome-wide significance was located on chromosome 10 (Table 2). The SNPs with the lowest P-values were rs4086116 (P = 3.3 × 10\(^{-24}\), Fig. 4) and rs4917639 (P = 8.0 × 10\(^{-24}\)), both located within CYP2C9 and in complete LD with each other (r\(^2\) = 1.0). Our Illumina array included the CYP2C9+3 SNP (rs1057910), reaching a genome-wide significant association with acenocoumarol dosage at a P = 6.44 × 10\(^{-12}\). More SNPs within

### Table 1. Characteristics of the RS-I and RS-II cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>RS-I cohort</th>
<th>RS-II cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>1451</td>
<td>287</td>
</tr>
<tr>
<td>Age in years, average (SD)</td>
<td>75.9 (7.9)</td>
<td>68.7 (9.5)</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>620 (42.7)</td>
<td>144 (50.2)</td>
</tr>
<tr>
<td>Female</td>
<td>831 (57.3)</td>
<td>143 (49.8)</td>
</tr>
<tr>
<td>BMI, average (SD)</td>
<td>26.9 (3.8)</td>
<td>27.9 (3.9)</td>
</tr>
<tr>
<td>Number of patients within a target INR-level (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very low: 2.0–2.5</td>
<td>54 (3.7)</td>
<td>23 (8.0)</td>
</tr>
<tr>
<td>Low: 2.5–3.5</td>
<td>840 (57.9)</td>
<td>171 (59.6)</td>
</tr>
<tr>
<td>Medium: 3.0–4.0</td>
<td>555 (38.2)</td>
<td>90 (31.4)</td>
</tr>
<tr>
<td>High: 3.5–4.5</td>
<td>2 (0.1)</td>
<td>3 (1.0)</td>
</tr>
</tbody>
</table>
or flanking the CYP2C18, CYP2C19 and CYP2C8 genes reached genome-wide significance. In our replication cohort RS-II, all SNPs on chromosome 16 were convincingly replicated (Table 2, Supplementary Material online). The two SNPs on chromosome 10 within CYP2C9 scoring lowest for genome-wide significance in the RS-I cohort were replicated in the RS-II cohort with P-values less than 10^{-26}, which is not genome-wide significant but yet far below the threshold of significance for replication (Table 2, Supplementary Material online). Further SNPs for CYP2C18 and CYP2C19 were replicated (P-values 10^{-26}), but not for CYP2C8 (P = 0.11).

Subanalysis. For GWAS on the first INR after initial standard dosage of acenocoumarol, P-values of the basic model within the RS-I cohort were substantially higher, with rs10871454 reaching a P < 10^{-15} instead of P < 10^{-123} for dosage at the end of initiation period as an outcome. Genome-wide significant values were only obtained for 11 SNPs from chromosome 16 and thus no other genes besides VKORC1 were associated with INR after standard initial dosage of acenocoumarol. The other subanalyses on maintenance dosage adjusting for CYP2C9 co-medication or for vitamin K intake did not change our results.

Extended model. In the extended model, we included the SNPs with the lowest P-values from the basic model for VKORC1 (rs108714540) and CYP2C9 (rs4086116) as independent covariates (Table 3 and Fig. 5). After adjustment for the influence of these two genes on variation of acenocoumarol dosage, only three SNPs on chromosome 10 (rs1998591, rs2104543, rs2111578) reached genome-wide significance in the extended model (P-values 10^{-26}).

### Table 2. GWAS of acenocoumarol maintenance dosage, basic model

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>SNP</th>
<th>Gene</th>
<th>Base pair position</th>
<th>Subjects</th>
<th>Beta^a</th>
<th>R^2b</th>
<th>P-value</th>
<th>RS-II cohort P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>rs10871454</td>
<td>STX4A</td>
<td>30955580</td>
<td>1451</td>
<td>-5.2</td>
<td>0.3198</td>
<td>2.0 \times 10^{-123}</td>
<td>1.4 \times 10^{-24}</td>
</tr>
<tr>
<td>10</td>
<td>rs4086116</td>
<td>CYP2C9</td>
<td>96697192</td>
<td>1450</td>
<td>-2.9</td>
<td>0.06871</td>
<td>3.3 \times 10^{-24}</td>
<td>1.6 \times 10^{-6}</td>
</tr>
</tbody>
</table>

The SNPs with lowest P-values each within chromosome 16 and chromosome 10 are reported. For all 53 genome-wide significant SNPs from the basic model, we supply an online supplement for complete overview of all genome-wide significant SNPs from the basic model. Basic model: adjusted for age, gender, BMI and target INR, significant results are printed in bold. Abbreviations: STX4A (Syntaxin 4A, placenta), CYP2C9 (Cytochrome P450 2C9).

^aDosage change of acenocoumarol in mg/week per additional variant allele.

^bUnivariate r^2 between acenocoumarol maintenance dosage and the specific SNP.
rs4086116. Thus in the extended model, this SNP no longer achieved genome-wide significance. However, rs1322179 and rs12767584, two other SNPs within CYP2C19, reached P-values of $10^{-6}$, still showing some association with the outcome though not genome-wide significant. These SNPs just failed to be replicated in our second cohort with P-values = 0.06.

### Meta-analysis

A meta-analysis with pooling of data from both cohorts was performed for the extended model with age, gender, BMI, target INR, rs10871454 (VKORC1) and rs4086116 (CYP2C9). P-values for the SNPs flanking CYP2C18 improved from $P < 10^{-9}$ for the extended model within the index cohort to $P<10^{-12}$ in the meta-analysis (Table 3). There was also an improvement for rs2108622 within CYP4F2 from $P<10^{-8}$ to $P<10^{-10}$. Two more SNPs within CYP4F2 became genome-wide significant in the meta-analysis. P-values for SNPs on chromosome 3 remained at $P<10^{-7}$ and failed to reach genome-wide significance in the meta-analysis. Further P-values for CYP2C19 decreased from $10^{-6}$ with the extended model in index population to $10^{-7}$ in meta-analysis, but were still not genome-wide significant.

### Independence of signals within the CYP2C cluster

In order to elucidate in how far the influence on acenocoumarol dosage of the SNPs within CYP2C18 were independent from CYP2C9, in an extended model on the CYP2C cluster we further adjusted for the effect of CYP2C9*3 by adding rs1057910 as another independent co-variate to our extended model. We chose this SNP as linkage between rs4086116, the lowest scoring SNP for CYP2C9 from our basic model, was reasonable with CYP2C9*2 ($r^2 = 0.62$) and poor with CYP2C9*3 ($r^2 = 0.26$) (Fig. 7). Earlier it was shown that an association of rs4917639, being completely linked to rs4086116, could not fully be explained by CYP2C9*3 (12). Thus rs1057910, representing CYP2C9*3, was expected to provide extra information on CYP2C9 activity which could not fully be explained by rs4086116. The outcome of the analysis with the additional adjustment for CYP2C9*3 showed that P-values of the three SNPs on chromosome 10 flanking CYP2C18 increased from $10^{-9}$ up to $10^{-6}$, which means that CYP2C9*3 could partially but not completely explain variation on acenocoumarol dosage by SNPs flanking CYP2C18 (Table 3). These three SNPs were in complete LD with each other ($r^2 = 0.99$). Two other SNPs within CYP2C18 were independent from CYP2C9*3 and P-values in meta-analysis decreased from $P = 10^{-6}$ to $P = 10^{-7}$, still not reaching genome-wide significance. These two SNPs were completely linked to each other ($r^2 = 0.99$) and also in complete LD with the two lowest scoring SNPs within CYP2C19, rs1322179 and rs12767583 (Fig. 7).

### Clinical implications

A multivariate model, including the clinical factors age, gender, BMI and target INR, but without any genotypic data explained 12.6% of acenocoumarol dosage variation (Table 4). The addition of rs9934438 (VKORC1 C1173T) to this model dramatically increased the explained variance to...
40.6% and the CYP2C9*2/*3 polymorphism could add another 5.8%. Replacing the SNPs commonly used in literature for VKORC1 activity by rs10871454, the SNP with the lowest $P$-value from our results, slightly improved the percentage of explained variance from 46.4 to 46.6%. Replacing the commonly used combined CYP2C9*2/*3 genotype by the SNP with the lowest $P$-value from our results, rs4086116, nearly explained the same amount of dosage variation than the combined genotype (0.3% less). A model including age, gender, BMI, target INR, rs10871454 for VKORC1, rs4086116 for CYP2C9, rs2108622 for CYP4F2 and rs1998581 for CYP2C18 could explain 48.8% of acenocoumarol dosage variation. Adding rs1057910 for the CYP2C9*3 polymorphism only contributed an additional 0.1% to the explained variance. Each additional variant allele of the CYP4F2 polymorphism and also each additional variant allele of the CYP2C18 polymorphism increased the weekly acenocoumarol dosage by 1.2 mg. Genotype and allele frequencies of the top SNPs for VKORC1, CYP2C9, CYP4F2 and CYP2C18 were comparable between our study cohorts and to findings from literature (Table 5).

**DISCUSSION**

In this large population-based cohort study, VKORC1 and CYP2C9 were the main factors regulating acenocoumarol-induced anticoagulation, in line with similar findings for warfarin.

**Table 3.** GWAS of acenocoumarol dosage adjusted for VKORC1 and CYP2C9, the 15 SNPs with lowest $P$-values (extended models)

<table>
<thead>
<tr>
<th>Chr</th>
<th>SNP (rs)</th>
<th>Gene</th>
<th>Base pair position</th>
<th>RS-I cohort Extended model$^a$ ($P$-value)</th>
<th>Extended CYP2C model$^b$ ($P$-value)</th>
<th>RS-II cohort extended model$^c$ ($P$-value)</th>
<th>Meta-analysis extended model$^d$ ($P$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1998591</td>
<td>CYP2C18</td>
<td>96397968</td>
<td>$1.9 \times 10^{-9}$</td>
<td>1.0 $\times 10^{-6}$</td>
<td>$4.4 \times 10^{-4}$</td>
<td>$4.9 \times 10^{-12}$</td>
</tr>
<tr>
<td>10</td>
<td>2104543</td>
<td>CYP2C18</td>
<td>96419961</td>
<td>$2.5 \times 10^{-9}$</td>
<td>1.2 $\times 10^{-6}$</td>
<td>$4.1 \times 10^{-4}$</td>
<td>$6.5 \times 10^{-12}$</td>
</tr>
<tr>
<td>10</td>
<td>12772169</td>
<td>CYP2C18</td>
<td>96395319</td>
<td>$3.0 \times 10^{-9}$</td>
<td>1.4 $\times 10^{-6}$</td>
<td>$4.1 \times 10^{-4}$</td>
<td>$7.7 \times 10^{-12}$</td>
</tr>
<tr>
<td>19</td>
<td>2108622</td>
<td>CYP4F2</td>
<td>1585143</td>
<td>$2.0 \times 10^{-8}$</td>
<td>$9.8 \times 10^{-9}$</td>
<td>$3.0 \times 10^{-3}$</td>
<td>$2.5 \times 10^{-10}$</td>
</tr>
<tr>
<td>19</td>
<td>2074901</td>
<td>CYP4F2</td>
<td>15858422</td>
<td>1.9 $\times 10^{-7}$</td>
<td>9.3 $\times 10^{-8}$</td>
<td>1.2 $\times 10^{-2}$</td>
<td>$8.3 \times 10^{-9}$</td>
</tr>
<tr>
<td>3</td>
<td>10935268</td>
<td>CNTN4</td>
<td>138930175</td>
<td>7.7 $\times 10^{-7}$</td>
<td>7.9 $\times 10^{-7}$</td>
<td>0.42</td>
<td>8.18 $\times 10^{-7}$</td>
</tr>
<tr>
<td>3</td>
<td>9828150</td>
<td>CNTN4</td>
<td>2518302</td>
<td>9.7 $\times 10^{-7}$</td>
<td>4.5 $\times 10^{-7}$</td>
<td>0.56</td>
<td>1.94 $\times 10^{-6}$</td>
</tr>
<tr>
<td>3</td>
<td>4571230</td>
<td>CNTN4</td>
<td>2513769</td>
<td>1.3 $\times 10^{-6}$</td>
<td>9.9 $\times 10^{-7}$</td>
<td>0.56</td>
<td>9.1 $\times 10^{-6}$</td>
</tr>
<tr>
<td>10</td>
<td>1322179</td>
<td>CYP2C19</td>
<td>96565232</td>
<td>1.6 $\times 10^{-6}$</td>
<td>9.2 $\times 10^{-6}$</td>
<td>0.06</td>
<td>2.6 $\times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>12767583</td>
<td>CYP2C19</td>
<td>96639954</td>
<td>1.6 $\times 10^{-6}$</td>
<td>9.5 $\times 10^{-6}$</td>
<td>0.06</td>
<td>2.8 $\times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>1042194</td>
<td>CYP2C18</td>
<td>96597374</td>
<td>1.6 $\times 10^{-6}$</td>
<td>9.5 $\times 10^{-6}$</td>
<td>0.06</td>
<td>2.7 $\times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>1624312</td>
<td>CYP4F2</td>
<td>1585422</td>
<td>1.6 $\times 10^{-6}$</td>
<td>9.7 $\times 10^{-6}$</td>
<td>0.04</td>
<td>2.0 $\times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>1928614</td>
<td>CYP2C18</td>
<td>96467844</td>
<td>1.7 $\times 10^{-6}$</td>
<td>9.7 $\times 10^{-6}$</td>
<td>0.07</td>
<td>8.9 $\times 10^{-6}$</td>
</tr>
<tr>
<td>10</td>
<td>2111939</td>
<td>No gene</td>
<td>107497352</td>
<td>3.2 $\times 10^{-6}$</td>
<td>6.6 $\times 10^{-6}$</td>
<td>0.77</td>
<td>$8.3 \times 10^{-8}$</td>
</tr>
<tr>
<td>10</td>
<td>12610189</td>
<td>CYP4F2</td>
<td>1585422</td>
<td>1.6 $\times 10^{-6}$</td>
<td>9.2 $\times 10^{-6}$</td>
<td>0.06</td>
<td>2.6 $\times 10^{-7}$</td>
</tr>
<tr>
<td>10</td>
<td>1263789</td>
<td>No gene</td>
<td>107497352</td>
<td>3.2 $\times 10^{-6}$</td>
<td>6.6 $\times 10^{-6}$</td>
<td>0.06</td>
<td>2.7 $\times 10^{-7}$</td>
</tr>
</tbody>
</table>

Significant results are printed in bold. Abbreviations: CYP2C18 (cytochrome P450 2C18), CYP4F2 (cytochrome P450 4F2), CNTN4 (contactin 1), CYP2C19 (cytochrome P450 2C19).

$^a$The extended model contained beside the clinical factors age, gender, BMI and target INR within unstandardized residuals also rs10871454 (VKORC1) and rs4086116 (CYP2C9) as independent covariates.

$^b$The extended CYP2C model contained the clinical factors age, gender, BMI and target INR within unstandardized residuals and also rs10871454 (VKORC1), rs4086116 (CYP2C9) and rs1057910 (CYP2C9*3) as independent covariates.
In our GWAS of acenocoumarol dosage, rs10871454 had the lowest P-value. This SNP was located on chromosome 16, close to VKORC1. Cooper et al. (26) have previously found this SNP to be most strongly associated in a GWAS with warfarin dosage. Due to more subjects in our study, the P-value of this SNP in our study was \(2.0 \times 10^{-123}\) compared with \(P = 6.2 \times 10^{-13}\) in the earlier study. The clinical implications of this SNP are unknown. As it is situated within another gene, the Syntaxin 4 gene, it is not coding for VKORC1 and thus cannot directly alter the structure of this protein. However, such polymorphisms might lower the intra-hepatic mRNA expression for this enzyme, resulting in decreased enzyme activity which reduces the amount of drug target in the liver and leads to lower dosage requirement of coumarins. (25) The effect of rs10871454 may be through LD with a putative functional SNP within VKORC1. In our study population, rs10871454 was in complete LD \(\left(r^2 = 0.99\right)\) with rs9934438 \((VKORC1\ 1173C>T)\), one of the five single segregating SNPs that were able to explain all variation in warfarin dose caused by \(VKORC1\) (7,25). Earlier studies suggested rs9934438 or rs8050894 as the possible putative functional SNP responsible for the association between \(VKORC1\) and coumarin dosage, as within the five representative \(VKORC1\) SNPs, being in complete LD with each other, these two SNPs were the ones that remained conserved across species (5). However, clinical implications may be only small as rs10871454 contributed only 0.2% more to the explained variance in acenocoumarol maintenance dosage than rs9934439.

Besides chromosome 16, only SNPs on chromosome 10 reached genome-wide significance. This chromosome contains within a region of high LD a cluster of the following \(CYP2C\) genes: \(CYP2C8, CYP2C9, CYP2C18\) and \(CYP2C19\). Most strongly associated with acenocoumarol dosage was SNPs within the \(CYP2C9\) gene, 

\[
\text{rs4086116 reaching the lowest P-value with } P = 3.29 \times 10^{-24}.
\]

This SNP was in complete LD with SNP rs4917639 \(\left(r^2 = 1.0\right)\) which had been shown previously to significantly associated to warfarin dosage (12). This SNP was able to explain nearly as much of dosage variation as the combined \(CYP2C9^*2/3\) genotype (only 0.3% less). Besides \(CYP2C9\), only polymorphisms flanking \(CYP2C18\) showed associations with acenocoumarol maintenance dosage that could not be fully explained by \(CYP2C9\) polymorphisms. Possible associations of polymorphisms in the \(CYP2C19\) gene could be explained through complete LD of these SNPs with polymorphisms in \(CYP2C18\). For loci in the \(CYP2C8\) gene, no association with acenocoumarol dosage could be demonstrated. In an earlier study, associations of \(CYP2C18\) and \(CYP2C19\) with warfarin dosage were fully explained by \(CYP2C9^*2\) and/or \(^*3\) polymorphisms (12). The independence of \(CYP2C18\) from \(CYP2C9\) polymorphisms in our study may be due to the fact that pharmacokinetics by \(CYP2C\) enzymes differs between warfarin and acenocoumarol. Adding rs1998591 as an indicator for variance in \(CYP2C18\) activity to our regression model increased the \(r^2\)adj of acenocoumarol dosage variance by 1.2%. In comparison, adding the \(CYP2C9^*3\) SNP \((\text{rs1057910})\) to a model already including rs4086116 for \(CYP2C9\) polymorphisms increased the percentage of explained acenocoumarol dosage variation by only 0.1%.

The density and coverage of the polymorphism set used (HumanHap 550K) is substantial and represents \(~90\%\) of the common SNP variation in Caucasians as determined by the HapMap. Our measure of maintenance dosage was an incidental outcome as the dosage at the end of 6 weeks of the initiation period and might have been influenced by coincidental factors such as diet or concomitant use of substrates or inhibitors of metabolic enzymes such as \(CYP2C9\). However as diet and prescription of co-medication are independent from genotypes, this should lead to non-random misclassification.
and rather weaken than inflate associations found. Nevertheless, we repeated our analysis in the index population for outcomes adjusted for vitamin K intake and also for current use of CYP2C9 co-medication. Both subanalyses did not change our results. Analysis with the first INR after standard initial acenocoumarol dosage confirmed our results for the \( \text{VKORC1} \) polymorphisms as contributing most to interindividual variation in response to initial acenocoumarol treatment.

To our knowledge, this is the first GWAS for response to acenocoumarol dosage. We consider the chance of bias and confounding in this retrospective study as negligible. First, the Rotterdam study is a prospective cohort study and the regional anticoagulation clinic covered a complete area of more than one million inhabitants in the Rotterdam region. Consequently, everyone who is treated with a coumarin anticoagulant as an outpatient will be registered as such and selection bias is highly unlikely, confirmed by the fact that in our study population genotypes were within Hardy Weinberg equilibrium (HWE). Also allele frequencies of the variant alleles were comparable to results for the same SNPs in other Caucasian populations (13). Furthermore, acenocoumarol dosage and INR are well-defined phenotypes and our results were biologically plausible. We confirmed earlier associations with warfarin dosage for rs10871454 related to \( \text{VKORC1} \) (26) and rs2108622 within \( \text{CYP4F2} \) (13). One of the strengths of our study lies in the high numbers of patients included. Thus power to detect possible associations for new genes independently from \( \text{VKORC1} \) and \( \text{CYP2C9} \) was considerably larger than in earlier studies (26). However, we did not detect further independent associations with acenocoumarol dosage besides \( \text{VKORC1} \) and \( \text{CYP2C9} \) for other genes than \( \text{CYP4F2} \) and \( \text{CYP2C18} \).

**Conclusion**

Our findings confirmed that acenocoumarol maintenance dosage mainly depends on polymorphisms in the \( \text{VKORC1} \) and \( \text{CYP2C9} \) genes. Independent from these two genes, only

**Table 4.** Explained variance of acenocoumarol maintenance dosage in different models

<table>
<thead>
<tr>
<th>Basic model</th>
<th>Added variables in different models</th>
<th>( r^2 \text{adj}^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, gender, BMI, target INR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs9934438 (( \text{VKORC1} 1173C&gt;T ))</td>
<td></td>
<td>12.6%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs9934438 (( \text{VKORC1} 1173C&gt;T ), rs1799853 and rs1057910 (combined ( \text{CYP2C9}^*2/3 ) genotypes)</td>
<td></td>
<td>40.6%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results) combined rs1799853 and rs1057910 (combined ( \text{CYP2C9}^*2/3 ) genotypes)</td>
<td></td>
<td>46.4%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results)</td>
<td></td>
<td>46.6%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results), rs1057910 (( \text{CYP2C9}^*3 ))</td>
<td></td>
<td>46.7%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results), rs1057910 (( \text{CYP2C9}^*3 ), rs2108622 (top SNP for ( \text{CYP4F2} ) in our results))</td>
<td></td>
<td>48.0%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results), rs1057910 (( \text{CYP2C9}^*3 ), rs2108622 (top SNP for ( \text{CYP4F2} ) in our results), rs1998581 (top SNP for ( \text{CYP2C18} ))</td>
<td></td>
<td>48.9%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results), rs2108622 (top SNP for ( \text{CYP4F2} ) in our results), rs1998581 (top SNP for ( \text{CYP2C18} ))</td>
<td></td>
<td>48.8%</td>
</tr>
<tr>
<td>Age, gender, BMI, target INR rs10871454 (top SNP for ( \text{VKORC1} ) in our results), rs4086116 (top SNP for ( \text{CYP2C9} ) in our results), rs2108622 (top SNP for ( \text{CYP4F2} ) in our results), rs1998581 (top SNP for ( \text{CYP2C18} ))</td>
<td></td>
<td>48.8%</td>
</tr>
</tbody>
</table>

\( r^2 \text{adj}^a \) Adjusted for the variance of acenocoumarol dosage explained by the model.

**Figure 7.** Linkage disequilibrium of the within the \( \text{CYP2C} \) cluster on chromosome 10 within the RS-I cohort. The LD-plot with \( r^2 \)-values is shown for different polymorphisms within \( \text{CYP2C18}, \text{CYP2C19} \) and \( \text{CYP2C9} \).
DNA variants in the CYP4F2 gene had a small additional influence on acenocoumarol dosage. From the CYP2C gene cluster on chromosome 10, only polymorphisms in the CYP2C18 gene added some extra information besides CYP2C9 polymorphisms on acenocoumarol dosage variation.

MATERIALS AND METHODS

Setting and inclusion criteria

This study was carried out within the baseline cohort of the Rotterdam study (RS-I), consisting of 7983 (response 78%). Subsequently, findings were replicated in the first extended cohort of the Rotterdam study (RS-II) with 3011 (response 67%) participants. The rationale and design of the Rotterdam study have been described elsewhere (28,29). In brief, the Rotterdam study is a prospective population-based cohort study, designed to study neurological, cardiovascular, locomotor and ophthalmological diseases in a population of people of 55 years and older. The RS-I cohort had baseline examinations during 1990–1993 with completion of standardized questionnaires, sampling of blood and isolation of DNA. The RS-II cohort consisted of a second independent cohort formed in 1999 with baseline examinations between 2000 and 2001.

A regional anticoagulation clinic monitors all inhabitants of Ommoord with an indication for anticoagulant therapy. This clinic covers the patients of the RS-I cohort as well as those from the RS-II cohort. The physician who treats the patient makes the choice of anticoagulant. Up to now, all patients start with a standard dosing scheme of acenocoumarol (6–4–2 mg) during day 1 up to day 3. Prothrombin times are monitored every 1–6 weeks by reference to the INR, dependent on the stability of the anticoagulant level. Doses are adjusted on the basis of the INR of the patient by computerized dose calculations. Since 1984, all data on dosing, laboratory and clinical data, including data on bleeding complications are fully computerized.

From both cohorts, subjects were included in our study population if they were genotyped for GWAS, and if they had started with acenocoumarol in the study period between 30 October 1985 and 8 December 2006, and if they had an INR assessment within 4 days after initial standard dosage of their treatment. Patients who were switched to or from phenprocoumon during the first 6 weeks of acenocoumarol treatment were excluded.

Genotyping

RS-I cohort. Genomic DNA was extracted from whole blood samples using the salting out method (30). Micro array genotyping was performed in the whole RS-I cohort with proper quality DNA samples (n = 6449) using the Infinium II HumanHap 550K Genotyping BeadChip® version 3 (Illumina, San Diego, CA, USA). The Illumina 550K Bead Chip array was genotyped in all participants of the original Rotterdam study cohort with proper quality DNA samples (n = 6449). Intensity files were analyzed using the Bead Studio Genotyping Module software v.3.1.14. A no-call threshold of 0.15 was applied to a custom-generated cluster file derived from the Illumina-provided cluster file (based on the cluster definitions applied to the HapMap CEPH cohort). In the custom-cluster file, 2308 SNPs with GenCall scores less than 0.90 were visually checked by two observers and manually reclustered or zeroed accordingly. Poorly performing samples with low call rate and 10th percentile GenCall score were excluded prior to calling genotypes. Any samples with a call rate below
Genotyping of CYP2C9

97.5% (n = 209), excess autosomal heterozygosity >0.336 ~FDR <0.1% (n = 21), mismatch between called and phenotype gender (n = 36), or if there were outliers identified by the IBS clustering analysis (see below), clustering more than three standard deviations away from the population mean (n = 102) or if there were outliers identified by the IBS clustering analysis, clustering >3~97% (n = 129) were excluded from the analysis. In total, we disposed over 5974 analyzed samples having passed QC and inclusion criteria.

RS-II cohort. Genotyping was targeted in the whole RS-II cohort using the Infinium II HumanHap 550K Duo Genotyping BeadChipR version 3 (Illumina) as part of a large population-based project on genetics of complex traits and diseases, financed by the Dutch government through the Netherlands Scientific Organization—Large Investments (NWO Groot; 175.010.2005.011). Of the 3011 RS-II participants, 2613 (86.7%) gave consent for DNA analysis. Of these, 2548 (97.5%) were plated and included for the genome-wide genotyping. In the current study, we use the first set of RS-II samples made available on September 2008, which was genotyped successfully for 2020 samples of which 1895 remained for analyses after QC. We describe below the genotyping and QC processes for this first RS-II set. Genotyping procedures were followed according to Illumina manufacturer’s protocols. All participants of the RS-II cohort with proper quality DNA samples (n = 2611) were genotyped with the array. Intensity files were analyzed using the Beadstudio Genotyping Module software v.3.2.32. A no-call threshold of 0.15 was applied to a custom-generated cluster file derived from the Illumina-provided cluster file (based on the cluster definitions applied to the HapMap CEPH cohort). Poorly performing samples with very low call rate (90%) and 10th percentile GeneCall score were excluded prior to calling genotypes. Any samples with a call rate below 97.5% or pending to be processed (n = 528), excess of autosomal heterozygosity (F < 0.05; n = 5) mismatch between called and phenotypic gender (n = 14), or if there were genetic outliers identified by the IBS clustering analysis (clustering more than three standard deviations away from the RS-II population mean; n = 81) or with IBS probabilities more than 40% (n = 25) were excluded from the analysis; in total, 1895 samples were analyzed. Of the 1895 individuals with a mean age of 65.20 (SD 8.34) years, 1032 (54.45%) were women.

Genotyping of CYP2C9*2 and VKORC1 C1173C>T. In order to check LD with variant alleles of these two genes which were not included in the Illumina 550 array, we included the corresponding SNPs which had already been genotyped within both study cohorts and were used in earlier studies (18,24). Genotyping for CYP2C9*2 allele variants was performed using the polymerase chain reaction followed by restriction enzyme digestion analysis (PCR-RFLP), as previously described (31). For genotyping of the VKORC1 1173C>T SNP in intron 1, dbSNP: rs9934438, a TaqMan allelic discrimination assay was used as previously described (32).

Phenotype definition

The mean weekly acenocoumarol dosage at the end of the first 6 weeks of treatment was regarded as stable maintenance dosage. Steady state of a drug is usually achieved within 5–7 half-lives of drug elimination. The (S)-enantiomeric form of acenocoumarol has a 2–5-fold higher anticoagulant potency than the (R)-form. For acenocoumarol with a half-life of 2 h for the S-enantiomere, the initiation period of 6 weeks would be sufficient to reach steady state, even if its elimination was prolonged in patients with CYP2C9 variant alleles (33).

Statistical analysis

Basic model. The overall strategy involved linear regression analysis for the genotype–phenotype association of all SNPs within the Illumina 550 array in the index cohort and replication of the associations found in the second cohort. SNPs with a minor allele frequency <0.01, a genotype call rate more than 95% or an exact HWE P-value less than 0.0001 were excluded from the analysis. We a priori declared results significant at $P < 5 \times 10^{-8}$, based on estimation of the multiple testing burden for GWAS of nearly all common variants in the human genome of European ancestry individuals for a target genome-wide $\alpha = 0.05$ (34). For replication of the specific SNPs associated with acenocoumarol dosage within the index cohort, in the replication cohort, a $P$-value less than 0.05 was considered significant. This was done first with a basic model for dosage at the end of 6 weeks initiation period as outcome, adjusted for the clinical factors age, gender and BMI and the patient’s target INR. Target INR was included into our model in order to adjust for necessity of higher dosages due to intended therapeutic INR values. BMI was defined as (kg/m²) and missing values were imputed with a regression model consisting of maintenance dosage as outcome with age, gender and target INR as variables. In order to adjust acenocoumarol maintenance dosage by these covariates, the unstandardized residuals from an additive linear regression model were used for genotype–phenotype associations in GWAS.

Extended model and meta-analysis. In order to detect SNPs within other genes independently associated with acenocoumarol dosage, we repeated GWAS in an extended model further adjusting for influences of VKORC1 and CYP2C9 as done previously in the GWAS of warfarin (26). In the extended model, those SNPs were taken as independent co-variates that were found by the basic model to be the strongest associated SNPs from the VKORC1 and from the CYP2C9 locus. For the extended model, we performed a meta-analysis combining the results of both GWAS from RS-I and RS-II using the beta estimates with inverse variance weighting and genomic control to control for inflation.

Subanalyses. As acenocoumarol maintenance dosage is the result of dosage titration due to INR measurements, the first INR after standard dosage is an outcome more directly representing the reaction of a patient to the drug. Therefore, we repeated GWAS for the basic model in subanalyses with the first INR after the initial standard dosage scheme of
acenocoumarol at the fourth day of treatment as outcome. In order to adjust for influences of co-medication on the acenocoumarol maintenance dosage, we repeated the GWAS including also current use of substrates and inhibitors of the CYP2C9 enzyme besides the variables of the basic model in a second subanalysis. As CYP2C9 co-medication, the following drugs were considered: amiodarone, carbamazepine, chloramphenicol, cimetidine, diclofenac, fluconazole, fluvastatin, losartan, miconazole, phenylbutazone, phenytoin, sulphasafidine, sulphanethoxazole, sulphinpyrazone, tolbutamide, trimethoprim and zafirlukast (31). In a third subanalysis, we repeated GWAS of acenocoumarol maintenance dosage adjusting for vitamin K intake at baseline besides the variables of the basic model.

Regression model on dosage for clinical impact. In a multivariate linear regression model, the additional contribution of newly found genetic factors was assessed by the adjusted \( r^2 (r^2\text{adj}) \). The \( r^2\text{adj} \) statistic measures the proportion of the total variability explained by the model with adjustment for the number of parameters in the model.

**Software**

For the genome-wide association analysis, we used PLINK v.1.04 (59). The meta-analysis was performed using METAL. (http://www.sph.umich.edu/csg/abecasis/Metal/) For multivariate linear regression analysis, we used SPSS software (version 15.0).

**SUPPLEMENTARY MATERIAL**

Supplementary Material is available at HMG online.

**ACKNOWLEDGEMENTS**

We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, Marijn Verkerk and Sander Bervoets for their help in creating the GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

**Conflict of Interest statement.** None declared.

**FUNDING**

The generation and management of GWAS genotype data for the Rotterdam study is supported by the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2) and the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) project nr. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII) and the Municipality of Rotterdam.

**REFERENCES**


