Pharmacogenetics of warfarin project (submitted proposal 7/01/04)

II. DETAILS OF PROPOSED RESEARCH

1. **Purpose and outcomes of proposed research**

   Pharmacogenetics is the study of the genetic basis for the difference between individuals in response to drugs. Its promise to allow individualisation of medicines, and thereby maximise efficacy and minimise toxicity, is of great public health importance. Two examples where pharmacogenetics is already utilised in clinical practice for patient stratification and dose adjustment are the use of Herceptin in breast cancer and phenotyping for thiopurine methyltransferase in patients treated with either 6-mercaptopurine or azathioprine. These are however “niche” drugs, and the broader impact of pharmacogenetics needs to be demonstrated in a larger group of patients by developing a “proof of concept” model, which will then allow similar study designs with other drugs.

   Warfarin is an ideal drug with which to study the clinical utility of pharmacogenetics because (a) it is widely used, (b) it has a narrow therapeutic index, and (c) there is large variability in dose requirements and response, which is influenced by both genetic and environmental factors. However, the pharmacogenetic factors identified to date are not of sufficient predictive value to be of clinical use. The purpose of the study, using a patient-centred approach to treatment, is to define the genetic and environmental factors that determine variability in response to warfarin in a single comprehensive study. The proposed outcome would be the development of a clinically useful and practical algorithm (that takes into account the relevant genetic and environmental factors) that will help clinicians individualise anticoagulant therapy. The potential benefits of this would include:
   - Improved safety of warfarin with reduced morbidity and mortality;
   - Improvement in patient quality of life;
   - Improvement in the cost effectiveness of warfarin therapy;
   - Improved uptake of warfarin particularly for atrial fibrillation.

2. **Background to the project**

   Warfarin is the oral anticoagulant of choice in the UK for the treatment of venous thromboembolism and as prophylaxis in patients with mechanical heart valves and atrial fibrillation (AF). It has been estimated that of the 21000 strokes per year among patients with AF, 3000 are probably prevented by warfarin. Consequently, the use of warfarin has increased such that patient numbers have doubled in most anticoagulant clinics in the last 5 years, and the trend is set to continue. Currently, 1% of the whole UK population (600,000 patients) and 6% of those over 80 years (154,000 patients) are on warfarin (source: IMS Health). The major risk of warfarin treatment is haemorrhage; the incidence varies from 10-24 episodes per 100 patients for all bleeding complications and from 1.2-7.0 episodes per 100 patients for major bleeding complications. The risk of bleeding increases with the intensity of anticoagulation. A recent study in Mersey showed that warfarin was responsible in 1 in 10 of admissions due to adverse drug reactions (Pirmohamed, unpublished). Furthermore, there is between 10- to 50-fold inter-individual variability in the dosage requirements necessary to maintain the international normalized ratio (INR) within a target range (most commonly between 2-3). This is increasingly a primary care problem with a shift to oral anticoagulation management into the community. There remains no consensus as to starting regimes for community-based patients, with little in the way of clinical information to guide best practice.

   An option that has been considered to reduce the problems associated with warfarin therapy is to use alternative anticoagulants, in particular oral thrombin inhibitors (OTI). However, this seems unlikely in the near future because (a) there is an enormous amount of high quality RCT data which show the clinical effectiveness of warfarin, and to accrue this for other anticoagulants is going to take many years; (b) OTI are expensive; (c) OTI also lead to haemorrhage (at rates equivalent to warfarin); (d) OTI are also likely to be associated with inter-individual variability in response; (e) there is no effective method of anticoagulation monitoring; (f) there is no effective antidote for reversal of over-anticoagulation except the use of fresh frozen plasma which is expensive; and (g) they may have additional problems, for example, ximelagatran has been associated with a 7% frequency of liver function abnormalities, which may itself require monitoring. Furthermore, although most studies have shown the thrombin inhibitor ximelagatran to be equivalent to warfarin in preventing thrombosis, none has demonstrated superiority and one has suggested inferiority. Thus, it is important that strategies are put into place to improve the clinical effectiveness of warfarin. Moreover, an analysis of the genetic and environmental factors determining variability in warfarin sensitivity will provide a framework to individualise therapy with other anticoagulants.
2.1 Factors determining inter-individual variability in response to warfarin

**Genetic factors:** Warfarin is administered as a racemate, with S-warfarin being three times more potent than R-warfarin. S-warfarin is metabolised by the P450 isoform CYP2C9; allelic variants of this isoform with reduced catalytic activity (between 5-12% of the activity of wild-type alleles), and in some cases, altered substrate specificity have been identified. The variants CYP2C9*2 and CYP2C9*3 both show decreased clearance of warfarin *in vitro* and *in vivo* (compared with the wild type CYP2C9*1). In accordance with this, individuals with these allelic variants require low doses of warfarin to achieve anticoagulation (table) [ref 6, and references therein]. Control of warfarin therapy on commencement is also more difficult in these patients, and they are also more liable to bleed while on warfarin.

<table>
<thead>
<tr>
<th>CYP2C9 genotype</th>
<th>Number of patients</th>
<th>Aggregate mean dose (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2C9<em>1</em>1</td>
<td>639</td>
<td>5.5</td>
</tr>
<tr>
<td>CYP2C9<em>1</em>2</td>
<td>207</td>
<td>4.5</td>
</tr>
<tr>
<td>CYP2C9<em>1</em>3</td>
<td>109</td>
<td>3.4</td>
</tr>
<tr>
<td>CYP2C9<em>2</em>2</td>
<td>7</td>
<td>3.6</td>
</tr>
<tr>
<td>CYP2C9<em>2</em>3</td>
<td>11</td>
<td>2.7</td>
</tr>
<tr>
<td>CYP2C9<em>3</em>3</td>
<td>5</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Despite the consistent data on the effect of CYP2C9 allelic variants on warfarin dosage and the risk of haemorrhage, it may be premature to advise routine pre-prescription genotyping for three main reasons: first, a great deal of inter-individual variability exists even within the same genotype group such that the predictive accuracy of CYP2C9 genotyping is likely to be too low to make it clinically useful; second, polymorphisms in genes coding for other P450 isoforms and clotting factors may also be important determinants in dosage, and their inclusion may improve the predictability of dose requirement. The utility of this has recently been shown in a small Japanese study. Third, the interaction between these polymorphisms and environmental factors such as vitamin K intake has not been adequately defined.

**Environmental factors:** A review of the literature suggests that age, weight, liver disease, vitamin K intake, interacting medications and alcohol intake have been reported to affect warfarin dose requirements and susceptibility to bleeding complications. The elderly, the group with the highest usage of warfarin, are more sensitive to the effects of warfarin and when considered as a population, require lower dosages. Nevertheless, there is still a great deal of variability in dose requirements, and age by itself cannot be used as a predictive factor. The relationship between age and the risk of major bleeding is not clear-cut, with contradictory data between different studies. The mechanism for the age-related sensitivity is unclear, although it has been suggested that it is related to a reduction in hepatic mass. However, the interaction with genetic determinants remains largely un-investigated.

Warfarin is a vitamin K antagonist, and variation in dietary vitamin K intake may lead to changes in anticoagulant response, and may be responsible for intra-individual variation in INR during maintenance therapy. However, the relationship between INR and vitamin K has not been shown in all studies. This may be a consequence of the fact that (a) the vitamin K pool in the body is relatively stable, and deficiency is only likely to occur with prolonged starvation; and (b) triglyceride levels can act as a confounding factor but are not measured in all studies. It is also important to note that the relationship between vitamin K levels, anticoagulant response and any genetic determinants, apart from CYP2C9, has not been evaluated. Heavy alcohol intake may also affect vitamin K levels usually through prolonged poor nutrition, and also increases susceptibility to bleeding through the increased falls that inevitably occur in these patients. Warfarin metabolism can be perturbed by many other drugs which may either inhibit or induce metabolism, or affect vitamin K absorption. The role of these factors has not been considered in genetic studies, and indeed, patients on interacting drugs are actively excluded.

**Gene-environment and gene-gene interactions:** Ultimately, any response to warfarin is going to be the result of interplay between genetic and environmental factors. One study showed that both age and CYP2C9 genotype accounted for 12% and 10% of the variance in warfarin dose requirements, respectively. Although this study did not show any interaction between vitamin K status and warfarin sensitivity, another study suggested that plasma vitamin K and vitamin K epoxide levels accounted for 25% of the variance, while age accounted for another 25% of the variance in warfarin dosage. However, no study to date has looked at all the environmental and genetic factors in the
same patient cohort, and the interaction between them. Many studies have been relatively small, and most of the genetic studies have been cross-sectional rather than prospective.

2.2 Warfarin resistance

Although the above has largely concentrated on increased warfarin sensitivity, there are patients who demonstrate resistance to warfarin, and require extremely high doses to maintain the INR in the therapeutic range. Although these patients are not susceptible to bleeding, they are susceptible to thrombotic complications resulting from under anticoagulation, and they also consume a disproportionate amount of clinic time because they need more frequent visits. The basis of pharmacodynamic resistance in humans is poorly understood but has been better studied in rodents. In both rats and mice, it is believed to relate to an alteration in the target for warfarin, vitamin K 2,3-epoxide reductase (VKOR), the rat gene locus being termed \( Rw \). At present, it seems most likely that the human \( Rw \) homologue is located at chromosome 16p on the grounds that two of the most closely linked markers in the rat, \( IL4R \) and \( ITGAM \), both occur here in humans. In addition, in a recent genetic study of a family with a clotting defect apparently due to lack of VKOR, it was found that the trait mapped to chromosome 16p12 to q21 which would be consistent with the possible location for the \( Rw \) homologue.

3. Research plan including methodology proposed

The purpose of the study is to evaluate both genetic and environmental factors in a large sample of patients initiated and maintained on warfarin using a prospective study design. The ultimate aim of the study is to develop a simple algorithm, based on a limited subset of genes alongside other clinically relevant information, which will help in individualising anticoagulant therapy. The plan of the study is illustrated below. Each section has the name of the main applicants responsible for that aspect of the study.

3.1 Research team

In order to achieve the aims of the project, we have brought together a multidisciplinary team comprising clinical and basic science expertise in complementary areas including pharmacogenetics, clinical pharmacology, haematology, primary care, pharmacokinetics, genomics, modelling, social sciences and statistics. The team comprises researchers from The University of Liverpool (Hughes, Park, Pirmohamed, Toh, Walley, Williamson), The University of Newcastle (Daly, Kamali), The University of Birmingham (Fitzmaurice), the University of York (Lewis, Webster), The Royal Liverpool and Broadgreen University Hospital Trust (Martlew, Pirmohamed, Toh, Walley) and the Sanger Institute in Cambridge (Bentley, Deloukas, Rogers). Thus, the patient cohorts, the clinical expertise, laboratory expertise and equipment, and modelling and statistical expertise are in place to successfully accomplish this project, and provide the necessary evidence to allow the use of this technology in routine patient care. The overall project will be co-ordinated by Pirmohamed, and all participants will be in regular e-mail contact. We would also aim for 6-monthly meetings of the whole research team in order to ensure that the timelines were being met, and to overcome any difficulties. For the investigators in Liverpool, fortnightly meetings will also be held.

3.2 Patients (Pirmohamed, Toh, Martlew, Fitzmaurice)

The initial patient cohort will be recruited from the Royal Liverpool and Broadgreen University Hospitals Trust which has 100 referrals per month for initiation of warfarin. All patients
irrespective of indication will be recruited over a period of two years by two nurses (under the supervision of Pirmohamed); we will aim to recruit 2000 patients for the initial cohort, all of whom will be followed up for a minimum of six months. Patients will either be recruited from the wards (in-patients initiating warfarin for conditions such as venous thromboembolism) or from the anticoagulant clinic (usually patients with atrial fibrillation), and follow-up will be via the anticoagulant clinic where all patients will attend (run by Toh and Martlew). The only exclusion criterion will be inability or refusal to give informed consent. The study design is naturalistic in that patients will receive the usual clinical care with a standard loading regimen and maintenance dosages being determined in the anticoagulant clinic by a combination of clinical expertise and computer algorithm. The patient accrual rate will be carefully monitored particularly in the early stages, and any sign that the target accrual is not being met will be tackled by identifying any obstacles, and if necessary considering using other anticoagulant clinics in the vicinity, for example at the University Hospital Aintree and/or the Wirral Hospitals NHS Trust. The principal investigator has conducted numerous pharmacogenetic studies, and in his experience, very few patients (<1%) refuse to give informed consent. Based on the population of the locality, it is expected that approximately 95% of patients recruited will be Caucasians, with 5% belonging to other ethnic groups. The latter will be analysed as part of the whole cohort (taking into account any population stratification as determined by genotyping for unlinked markers – see below), as well as separately in order to identify any differences. Although numbers of non-Caucasian patients will be limited, we feel that recruiting the small sample will still provide important information that will allow (a) the development of evidence-based dosing guidelines, and whether these can be generalised to all populations, and (b) the possibility of mounting larger studies in different ethnic groups in the future.

A second patient cohort (n=400) will be recruited from a primary care setting in West Midlands. This will be used for validation of genetic and environmental factors identified in the first cohort and to test the possibility of extrapolating results obtained in a secondary care setting to a primary care setting, which in the future may be the site of anticoagulant monitoring clinics. Assuming 4 new patients per annum for an average 8,000-list size practice, 100 practices would be needed to recruit 400 patients within 12 months. The methodology for recruitment and follow-up will be essentially the same as for the secondary care study, with practice nurses undertaking baseline investigations. This will be co-ordinated by a non-clinical research fellow while a research associate will co-ordinate the fieldwork; both will be supervised by Fitzmaurice. An initial 6-month period will be required for practice set-up, while recruitment will take place over 1 year, and a final 6 months will be used for data verification. The researchers will be based in the Department of Primary Care and General Practice, the University of Birmingham, which has particular expertise in the design and conduct of large primary care based trials (within the Midlands Research Practices Consortium comprising a network of over 350 practices with representative populations for England and Wales, which receives the largest NHS Budget 1 funding for primary care research in the country) and health technology assessments, major epidemiological studies, and methodological research, notably modelling and health services research.

3.3 Patient clinical details and follow-up (Pirmohamed, Toh, Martlew, Fitzmaurice)

All patients will be seen at the initial visit (i.e. before starting warfarin) and at every follow-up visit in the anticoagulant clinic with a maximum follow-up of six months for every patient. The frequency of clinic visits will vary in different patients, being determined by the stability of INR control. All patients will also be seen at 2, 8, and 26 weeks (after warfarin initiation) for various measurements, as indicated below. At the index visit, we will record age, gender, height, weight, indication for warfarin therapy, additional medical problems, current medications, and whether there is use of herbal or other complementary medicines. The alcohol intake will also be recorded at the index visit using the AUDIT questionnaire (a validated tool developed by the WHO), while on the subsequent visits, we will record their approximate alcohol intake in the preceding week. Blood will be obtained for DNA at the index visit, while blood for other measurements will be obtained at the index visit as well as at follow-up appointments. For all patients, we will record the haemoglobin level, liver function, plasma albumin, and renal function at each visit. Some patients, for example those being treated for venous thromboembolism will initially have been treated with heparin; these data will be recorded and any effect on any parameters being measured determined. An existing Microsoft Access database in the Department of Pharmacology, which is (a) password protected and (b) undergoes back-up every day will be used for this study.
3.4 Outcome measures

**Co-primary**
- INR>4 during week 1
- Warfarin sensitivity (defined as a dose of 1.5 mg or less on 3 successive clinic visits)

**Secondary**
- Warfarin resistance (dose of >10mg/day on 3 successive clinic visits)
- Clinical events such as major (defined as bleeding that led to loss of 2 units of blood over a 7-day period or was otherwise life-threatening and minor (all other bleeds, including bruising) haemorrhagic complications
- Markers of lack of efficacy including recurrent thrombosis
- Markers of lack of efficacy including occurrence of stroke in patients with atrial fibrillation
- Time to therapeutic INR
- Time to stable warfarin dosing in days (defined as 3 consecutive INR measurements within therapeutic range for the same mean daily dose)
- Days spent in hospital to achieve target INR
- Time within target INR range during the 6 months of treatment

3.5 Genotyping strategy (Bentley, Deloukas, Rogers, Daly, Pirmohamed)

Blood samples will be sent directly from Liverpool and Birmingham to the Sanger Institute where DNA will be extracted. The Sanger Institute has developed quality-control procedures to ensure sample probity and measure misidentification rates. Automated robotic processing of samples has been established to reduce mishandling and cross contamination areas. An automated sample management system tracks and monitors all sample workflow within an Oracle database and allows for barcode reading, data transfer and tracking of individual samples. Pirmohamed and Bentley hold a Wellcome Trust grant in epilepsy, in which this DNA archiving procedure is used.

To date, genotyping for warfarin sensitivity has largely concentrated on CYP2C9. More recently, a Swedish study has suggested an association with MDR1 polymorphisms, while a study of 45 Japanese patients showed warfarin sensitivity to be related not only to CYP2C9, but also to polymorphisms in the factor II and VII genes, accounting for 50% of the variance. For this study, we therefore intend to look at the genes identified so far in a more systematic manner, but also to expand the gene list to include those coding for proteins involved in the absorption, distribution, metabolism and excretion of both enantiomers of warfarin, and the target proteins which modulate the anticoagulant effects of warfarin (these include genes coding for clotting factors, and those involved in the vitamin K cycle). Using these criteria, we have identified 25 candidate genes (appendix 1). We hypothesise that by using a limited subset of SNPs derived from our candidate gene list, together with the environmental factors, we will be able to develop a predictive algorithm that will be clinically relevant with an acceptable threshold for clinical use. Furthermore, the high quality data obtained for genes involved in the clotting cascade will also serve as templates for investigation of variability in response to other anticoagulants.

In preliminary work performed at the Sanger Institute, all 25 genes have been manually annotated in order to accurately define exons, introns, exon-intron boundaries and 5'- and 3'- untranslated regions. Initial bioinformatic analysis has led to the identification of 700 SNPs in these 25 genes. In order to determine which SNPs to test in our patient cohorts, we will use a two-pronged approach. First, common variants (≥10%), identified through the HapMap project and re-sequencing, will be detected by using tagging SNPs (tSNPs). We estimate between 4 and 6 htSNPs will account for ≥95% of haplotype diversity at each candidate gene. Second, in order to detect less common variants (3-10%), which are unlikely to be in linkage disequilibrium with tSNPs, we will undertake resequencing of all exons, intron-exon boundaries and untranslated regions in 50 unrelated patients on warfarin. Rare variants whose presence is not captured by the tSNPs will be identified, and a subset that is predicted to have a functional effect, for example, because of an amino acid substitution, will be incorporated into the genotyping strategy. The combination of tSNPs and selected rare variants will be related to the clinical outcome measures above, and incorporated into the model. Using this strategy, we estimate that we will need to genotype for an average of 8 SNPs per gene, or a total of 200 SNPs in each patient (for the 25 genes). This strategy can be illustrated with respect to CYP2C9, which has been widely studied. Three common haplotypes associated with variation in enzyme activity have been identified for the CYP2C9 gene. These can be identified by genotyping for the two common non-synonymous SNPs associated with the Arg144Cys and Ile359Leu polymorphisms
(CYP2C9*2 and CYP2C9*3 alleles). A number of other functionally significant SNPs are also known to occur but at lower frequencies. However, since a significant role for CYP2C9 in determining warfarin metabolism has already been established, these polymorphisms are functionally significant and show ethnic variation, we propose to also genotype the study DNAs for the single SNPs which represent the CYP2C9*4, CYP2C9*5, CYP2C9*6 and CYP2C9*11 alleles. In particular, CYP2C9*5, CYP2C9*6 and CYP2C9*11 have been reported to occur at frequencies between 1 and 5% in Africans and African-Americans while CYP2C9*4 has been detected in some Asians. In addition to the above strategy, every patient will also be genotyped for 20 unlinked SNPs to control for population stratification.

Genotyping is currently carried out at the Sanger Institute using the Mass Extend method of Sequenom. Individual SNP loci are amplified (in quadruplex format) using the polymerase chain reaction, which provides a template for allele-specific primer extension. Allele-specific products are detected using MALDI-TOF mass spectrometry. Two such systems are currently in operation enabling determination of approximately 60,000 individual genotypes per day and this very high throughput permits a major economy of scale.

3.6 Pharmacokinetic analysis (Hughes, Kamali, Park)

Warfarin is administered as a racemate (R and S warfarin). Unbound levels of both enantiomers will be measured in all patients at each clinic visit using chiral HPLC according to a previously described method17. The time of the blood samples will be noted and related to the time of dose. A population pharmacokinetic-pharmacodynamic analysis, based on a one-stage approach that allows the calculation of population parameters from sparse, imbalanced observations, obtained under routine clinical conditions will be performed18,19. The statistical model applied will be a hierarchical non-linear mixed effects model that includes relationships between covariates and model parameters. This will allow inter-individual variability in pharmacokinetic and pharmacodynamic parameters, and residual variability (e.g. drug-assay error, intra-individual variability and deviations caused by the incorrect specification of the model) to be quantified. Pharmacokinetic models of (S)- and (R)-warfarin will be developed based on previous studies that suggest one-compartment models describe the data adequately. An indirect response model that incorporates a sigmoidal inhibitory I_max function, and which accounts for the relative competitive inhibition by (R)- and (S)- warfarin, will be used to describe the rate of change of clotting factor activity20,21. The influence of covariates (described in the different sections) on pharmacokinetic (clearance, volume of distribution and elimination rate constant) and pharmacodynamic modelling parameters (rate of production and dissipation of clotting factor activity, I_max and IC_50 values for (S)- and (R)-warfarin and sigmoidicity of the concentration-response relationships) will be incorporated in the mixed effects model. A stepwise generalized additive model, based on the parameter estimates from the basic population model as dependent variables, will be used to select the most important covariates and select the functional relationship between covariates and the parameter. To evaluate the significance of covariate effects, the difference in minimum value of the objective effects, between a model with and without a specific covariate relationship will be compared with χ^2 distributions. This will be performed by a PhD student, supervised by Hughes, with the assistance of a technician.

Compliance is an important factor to assess in order to reduce the error rates (false positives) in mis-classifying patients as being warfarin resistant. Analysis of plasma warfarin concentrations will serve as a measure of compliance, as will a tablet count, assessed at each clinic visit and a validated structured questionnaire on dosing history22. Taken together, these analyses will allow us to relate genotype through pharmacokinetic phenotype to haematological phenotype and clinical outcome.

3.7 Determination of vitamin K levels (Kamali, Park)

Vitamin K is an essential cofactor for the synthesis of the clotting factors II, VII, IX, X, rendering them functionally active through carboxylation of glutamic acid residues. Plasma vitamin K levels exhibit inter-individual variability. Vitamin K1 2,3-epoxide, the inactive form of vitamin K, accumulates in plasma during warfarin therapy, and thus the ratio epoxide:vitamin K provides an indication of degree of inhibition of epoxide reductase by warfarin. Recent studies have suggested that vitamin K status is associated with both early warfarin sensitivity, i.e. at initiation, and during maintenance, and may account for 25% of the variance in warfarin sensitivity14. However, this has not been replicated in other studies11. All these studies have been small. We therefore feel it is essential that vitamin K levels are monitored in our patients to assess the impact on warfarin
sensitivity, and the interaction with other clinical, haematological and genetic markers, in order to make an evidence-based assessment of whether including or excluding vitamin K levels has an impact on the predictive ability of the proposed clinical algorithm. We will measure vitamin K1 and K2,3-epoxide levels in all patients before starting warfarin and 2 and 8 weeks after warfarin using a HPLC method previously reported in the literature - this will be a fasting sample since vitamin K levels show a temporal variation, and are affected by food. Every vitamin K assessment will also be accompanied by measurement of triglycerides, which correlate with plasma clotting factor activity. The PhD student and technician will both be involved in measurement of vitamin K levels; this will be undertaken in labs in Newcastle (supervised by Kamali) and in Liverpool (supervised by Park).

3.8 Determination of clotting factor levels (Toh, Martlew)
Vitamin K antagonism by warfarin results in a reduction of both pro-coagulant factors II, VII, IX and X as well as the anticoagulant proteins, C, S and Z. The rate of change in these proteins on initiation of warfarin will be dependent on their half-lives. Thus, factors with a short half-life and high synthesis rate (e.g. factor VII) influence the INR earlier in therapy, and therefore may be particularly important in determining the variability in initiation of warfarin. By contrast, factors with a longer half-life and lower synthesis rate (e.g., prothrombin) influence the INR later, and may therefore be more important in determining variability in maintenance doses. There is some evidence that clotting factor levels can predispose to thrombosis. For example, high levels of factors II and IX are associated with thrombosis as well as reduced proteins C, S and Z. However, what is unclear is the rate and extent of falls in these individually and in combination with regard to bleeding. Factor IX has been implicated with two missense mutations at Ala-10 of the propeptide causing factor IX levels to drop to 3% or less upon warfarinisation and causing severe bleeding despite therapeutic INRs. Similar variants may be inconsequential normally, but significant on the use of warfarin, with the possibility of inducing bleeding. All of these clotting factors have not been measured in one large cohort previously – this will provide valuable information regarding how variation in their levels leads to variability in INR, warfarin dose requirements and bleeding complications. Important, by determining the SNP haplotype structure of these factors, we will also be able to determine the phenotype-genotype relationships for all the clotting factors. We will measure the vitamin K-dependent proteins before warfarin commencement and in the first two subsequent visits (2 and 8 weeks). These will be measured in the Department of Haematology, The Royal Liverpool Hospital, where this is a routine service.

3.9 Modelling and cost-effectiveness analysis (Walley, Hughes)
Health care resource utilisation data relating to anticoagulation clinic service, genotyping, bleeds, thromboembolic events, drug therapy will be collected prospectively. A costing exercise will be performed alongside the clinical study in order to enable the calculation of the cost-effectiveness of pharmacogenetic identification of warfarin dose both in hospital and in primary care. This will form part of the PhD project for the student funded on the grant. Unit costs will be taken from appropriate sources (e.g. hospital finance department, Netten et al., Drug Tariff). For patients at the highest risk of bleeding with warfarin, we will also use economic modelling to assess the cost-effectiveness of treating them with oral thrombin inhibitors rather than warfarin. An outline model to address the cost effectiveness of ximelagatran is currently under development on behalf of the National Institute for Clinical Excellence by Walley and others, and can be adapted for the purposes of this study. Patients will be asked to complete EuroQol 5D quality of life questionnaires at the beginning and upon completion of the study to derive health state utilities. The perspective of the NHS will be adopted, and the time horizon of analysis will be six months. Modelling to longer periods will also be possible. The primary economic outcome for the analysis will be cost per quality-adjusted life year. A secondary endpoint of cost per serious adverse event avoided will also be calculated. The economic modelling will be based on the multivariate regression model for identifying risk according to patient demographic and genetic traits. An economic model based on the population pharmacokinetic-pharmacodynamic model as proposed previously, will also be evaluated. These will provide an opportunity to (a) explore the impact of pharmacogenetic variability on cost- as well as clinical-effectiveness, and (b) assess the cost-effectiveness of targeting patients according to risk of complications or treatment resistance. Uncertainty in the analysis will be addressed by use of a probabilistic analysis that conserves the variance and co-variances in the risk model and assigns distributions to cost and quality of life data, according to standard methodology. Data relating to
alternative anticoagulation treatments and additional data on warfarin will be identified through literature searches.

3.10 Statistical considerations and development of algorithm at a clinical level (Williamson, Hughes)

Univariate analysis of the secondary care cohort data will be undertaken to identify factors affecting each outcome and correlations between putative factors explored. Regression models will be developed to assess both genetic and environmental factors and their interactions, and their effect on warfarin dose requirements. These will include survival and logistic regression models, depending on the nature of the outcome assessed. Recently developed methods for analysing time-dependent covariates will be employed\textsuperscript{35}. Consistency between risk factors for different outcomes will be examined. The final model for each outcome will be applied to the primary care cohort and the predictive ability of the model assessed.

Missing outcome data will be minimized as far as possible but where unavoidable, information on the reason will be recorded and sensitivity analyses undertaken to assess the robustness of the conclusions to this problem. Various approaches to the handling of missing baseline data will be investigated but will not result in the exclusion of a patient from the analysis as this may lead to bias (Dr Ian White, MRC Biostatistics Unit – personal communication). A statistical analysis plan will be developed prior to any data analysis.

The study aims to recruit 2000 patients, the largest study to date. Consider first the rate of INR>4 during week one in the group with a CYPC29*2 variant compared to no variant. The prevalence of CYPC29*2 is taken to be 20%\textsuperscript{7,8}. Since there is little information in the literature concerning the rate of INR>4 in the ‘no variant’ group, we have varied the estimate from 5-90% and calculated the minimum odds ratio (OR) that could be detected at the 5% significance level with at least 80% power. All ORs were less than 1.85. Assuming the rate of INR>4 in the ‘no variant’ group is 5-10% (figure B, ref 36), ORs of between 1.6 and 1.85 can be detected with 80% power. Since the estimate of the prevalence of the CYPC29*3 variant is also 20%\textsuperscript{7}, similar ORs can be detected. For exposure defined as either CYPC29*2 or CYPC29*3, smaller ORs can be detected since the estimated prevalence is higher (Aithal\textsuperscript{38}-38%, Higashi\textsuperscript{39}-31%). There is no information from previous studies on the rate of warfarin sensitivity, defined as a dose of 1.5 mg or less on 3 successive clinic visits, in the ‘no variant’ group. However, as above, if this rate is between 5 and 90%, ORs of 1.85 or less can be detected with 80% power. For rarer variants in this and other genes to be clinically important, their effect size (measured here by the allelic odds ratio, OR) must be large. For a rare variant (allelic frequency 5%), and an estimate of the rate of INR>4 in the ‘no variant’ group of 20%, our study would have 99% power to detect an OR of 3 or more. A lower significance level has not been adopted in this study since (a) the independently recruited validation cohort will help to reduce false positives arising from the initial analyses; and (b) concomitant functional analyses are being undertaken to assess the effect of genotypic variants. It is also important to note that these effect sizes are for a single causal variant, and we expect to realise much larger overall effect sizes via combinations of causal variants.

It is our intention to develop a clinically relevant and robust dosing algorithm based on a limited number of genes that have a significant impact on response to warfarin, but which also takes into account readily evaluable clinical/environmental factors. This will be achieved using a two-stage procedure. First, the statistical regression modelling approach will be used to identify a limited subset of genetic and environmental factors, and their interactions, that predict variation in response to warfarin, and the identification of patients who may be unduly sensitive or resistant to warfarin. The second step will involve population pharmacodynamic modelling, which will incorporate the identified risk factors, to develop the dosing algorithm.

3.11 Development of the algorithm at a patient/organizational level (Lewis, Webster)

An important part of developing the algorithm will be an examination of clinician and patient acceptance of pharmacogenetics\textsuperscript{37}. A detailed sociological examination by the York team of the problems and challenges thrown up during the proposed study would provide an exemplar for introduction of pharmacogenetics prescribing more generally\textsuperscript{38}, deriving broad lessons from this study that have utility for (a) other hospital settings/therapeutic areas, (b) GPs and other health professionals in primary care; and if so, to what extent. In this regard, the proposed study will be complemented by concurrent research at the Science and Technology Studies Unit (SATSU) on pharmacogenetics, funded under the ESRC Science in Society programme, that focuses on clinician acceptance as well as
international and regulatory aspects of pharmacogenetics. Another focus of SATSU involvement would be patient care. Despite its promise, many aspects of patient care remain unexplored in the context of pharmacogenetics, including patient perspectives on pharmacogenetics-based therapy, and the protocols and regimes employed; patient compliance on leaving the hospital setting; and the role of carers. All of these can be expected to impact on the overall success or otherwise of pharmacogenetics, and the extent to which the technology can, and will, be introduced into the clinical setting, and the time frame within which this is likely to occur.

The objectives of the SATSU component of the study will include:
- An analysis of the impact on monitoring practice, and changes in same, of the introduction of pharmacogenetics, from the perspective of clinicians, nursing staff, and patients.
- Identification of problems and challenges specific to pharmacogenetics that emerge during the study around patient experience and problems with organisational delivery.

The methodology will include a series of open-ended semi-structured interviews with (a) study participants (clinical research staff, research nurses, hospital managers) and other clinicians in therapeutic areas where pharmacogenetics is likely to be introduced; (b) patients, focusing on factors influencing clinician and patient acceptance. The latter will comprise three groups of purposively selected patients: 30 hospital-based patients (Patient Group 1 - Liverpool); a warfarin-treated but non-genotyped comparison group of 15 patients (Patient Group 2 - Liverpool); and 15 warfarin-treated primary care patients (Patient Group 3 – Birmingham).

For this element of the overall project, the emphasis is principally qualitative, and as such complements the quantitative pharmacological elements in the study. In doing this, it also directly relates to the oft-expressed need for qualitative methods to form an integral part of clinical trials and patient care more generally. This component will be undertaken by a Research Fellow supervised by Lewis and Webster.

### 3.12 Added value
The project will have the following added value:
- The setting up of a carefully clinically characterised DNA archive, which will be available for analysis by other researchers. Access to this DNA archive will be granted and monitored by a steering committee that will include some of the applicants, as well as independent scientists (one of whom will act as the Chair) and lay members.
- The availability of a cohort of patients with non-valvular atrial fibrillation.
- The availability of a cohort of patients with venous thromboembolism.
- The availability of a qualitative database on patient and clinician acceptance.

### 3.13 Timelines

<table>
<thead>
<tr>
<th>Months</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>15</th>
<th>18</th>
<th>21</th>
<th>24</th>
<th>27</th>
<th>30</th>
<th>33</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient recruitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Database development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Database entry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re-sequencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warfarin levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin K levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clotting factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PK-PD model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost-effectiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SATSU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data analysis and write up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 4. Methods for disseminating and implementing research
The project will be registered on the NRR. Given the multi-disciplinary nature of the project, the results will be presented at medical, pharmacological, haematological and primary care conferences in order to disseminate widely. Specific interdisciplinary dissemination activity already
identified includes the Wellcome Trust pharmacogenetics network; annual meetings of the ESRC Innovative Health Technologies programme (of which Webster is director) and the ESRC Science in Society programme (of which Lewis is a member); and the bi-annual Cold Spring Harbor Laboratory/Wellcome Trust meeting (of which Pirmohamed and Bentley are organisers). Papers (estimated 5-10) from the project will be published in leading international clinical, pharmacogenetics and social science journals. The development of an algorithm may also have IP value, with future return of any investment back to the NHS.

The study as proposed involves the use of very specialist pharmacogenetic resources not routinely available in the NHS. However, the technology is not difficult and if shown to be of practical clinical use, could be made available throughout the NHS. The cost effectiveness of such an approach will depend on the costs of such a test and this will be explored in sensitivity analyses within the current study. This would encourage manufacturers to develop simple kits at a cost consistent with broadly acceptable cost effectiveness thresholds within the NHS, and the NHS to adopt such a genetic testing approach. This study is essential before thrombin inhibitors become established practice within the NHS, when a less expensive and more appropriate alternative approach has not been fully explored.

5. Funding collaborations

Funding is requested for staff and consumables to perform the project (see appendix B for full justification); no funds are required for the applicants, who are already funded via their own institutions. In addition, no costs are required to cover the costs of office accommodation for the staff employed on the project. As an indicator of its support for the grant, The University of Liverpool will provide funding for a research technician for 3 years to work on the project if the grant is awarded (see letter). A major added value of this project is the ability to utilise the expertise and infrastructure of the Sanger Institute, which is covered by Core funding.

6. References