Protocol for Genetics-InFormatics Trial (GIFT) of Warfarin to Prevent DVT
September 2017

Washington University in St. Louis HRPO	09-10101
NIH R01	HL097036
FDA IDE	G100317
ClinicalTrial.gov ID	NCT01006733

GIFT Investigators

Aims and Hypothesis

The overall objective of the *Genetics-InFormatics Trial (GIFT) of Warfarin to Prevent DVT* is to elucidate novel strategies to improve the safety and effectiveness of warfarin therapy. With this study we directly respond to Health and Human Services (HHS) priorities to advance the field of personalized medicine and to prevent venous thromboembolic disease. Recently, the Honorable Mike Leavitt, Secretary of HHS, announced the Personalized Health Care Initiative and wrote that a key goal was, "… to use our personal genetic information to tailor treatments more effectively to each patient."[1] On September 15, 2008, the Acting Surgeon General (Dr. Steven K. Galson, MD, MPH) issued a *Call to Action* to reduce the number of cases of deep vein thrombosis and pulmonary embolism in the United States.[2] To facilitate the dosing strategies for the trial proposed herein and for the public at large, we have made publically available a non-profit, decision-support web application, www.WarfarinDosing.org.

Aim 1: To determine how pharmacogenetic-based warfarin therapy affects the safety and effectiveness of warfarin therapy. The intensity of anticoagulant therapy is measured by the International Normalized Ratio (INR). During initiation, the INR often falls outside the therapeutic range. INRs that are too low predispose patients to thromboembolism [3] while supratherapeutic INR values increase risk of bleeding. [4,5] In August 2007, the FDA approved the label change of warfarin/Coumadin™ to recommend considering lower initial doses in patients known to have certain polymorphisms in genes affecting warfarin metabolism and sensitivity.[6] However, whether this strategy improves the safety and effectiveness of warfarin therapy in general is unknown. In particular, how this strategy affects subgroups with and without the genetic variants of interest is also unknown. To test the resulting joint hypothesis while preserving an Aim-wide Type I error rate of $\leq$ 0.05 we will partition our expected error rate as described in the methods section below.

Primary Joint hypothesis: Pharmacogenetic therapy decreases the composite risk of a non-fatal VTE, non-fatal major hemorrhage, death, or INR$\geq$4.0 in all patients, and in the subgroup of patients whose pharmacogenetic and clinical predicted therapeutic maintenance doses differ by ≥ 1.0 mg/day (**Appendix 4**). Based on our meta-analysis of prior trials[7-10] (Sections B.3 and B.7 of grant proposal) and our pilot studies (Section C of grant proposal), we anticipate 80% power to simultaneously detect a reduction in the composite outcome, as measured by a chi-square test in both populations.

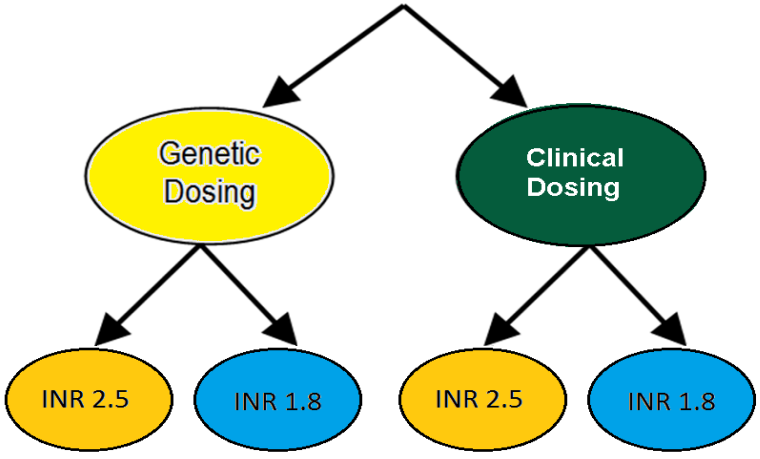
Aim 2: To determine whether warfarin therapy with a target INR of 1.8 is non-inferior to therapy with a target INR of 2.5 at preventing VTE or death in orthopedic patients. One randomized trial (PREVENT) found that a target INR value of 1.5–2.0 prevented 64% of VTE *recurrence*.[11] Although that trial excluded orthopedic patients, such an approach has been endorsed by the American Academy of Orthopedic Surgeons (AAOS) and by many academic orthopedists (Table C.4). On page 15 of the latest AAOS guidelines (American Academy of Orthopaedic Surgeons, 2007) they offer the following recommendation for VTE prophylaxis around the time of joint replacement: "Warfarin, with an INR goal of ≤2.0, starting either the night before or the night after surgery, for 2-6 weeks." However, the AAOS grade the overall evidence for VTE prophylaxis in this population as low (level III) because no randomized trials have answered key clinical questions in this area—what is the optimal target INR value and whether pharmacogenetic therapy can improve clinical outcomes. The AAOS guidelines conflict with the American College of Chest Physician (ACCP) guidelines, [12] which recommend, as one of their (Grade 1A) options (page 338S), using an "…adjusted-dose vitamin K antagonist (INR target, 2.5; range 2.0 to 3.0)." Because lower target INR values may reduce the risk of hemorrhage and simplify warfarin management [11] we propose to test the following:

Hypothesis 2: For prevention of non-fatal VTE or death, a target INR of 1.8 will be non-inferior to a higher target INR (2.5). Using a non-inferiority margin of 3%, we will have 83% power to detect the non-inferiority of a target INR of 1.8 in 1600 patients.

1. Trial Overview

Over 4 years, *GIFT of Warfarin* will enroll 1600 orthopedic patients from Washington University in St. Louis, Intermountain Health Care, University of Utah Hospital, the Hospital for Special Surgery (Weill-Cornell, NYC), University of Miami, Rush University, and University of Texas Southwestern. Participants will be aged 65 years or older and scheduled for 4-6 weeks of warfarin therapy for venous thromboembolism VTE prophylaxis after elective hip or knee arthroplasty. After informed consent and genotyping, patients will be randomized to: (Aim 1) pharmacogenetic vs. clinical dosing of warfarin; and (Aim 2) a target INR of 2.5 vs. 1.8 (Figure D.1)

Figure 1 Overview of Randomization

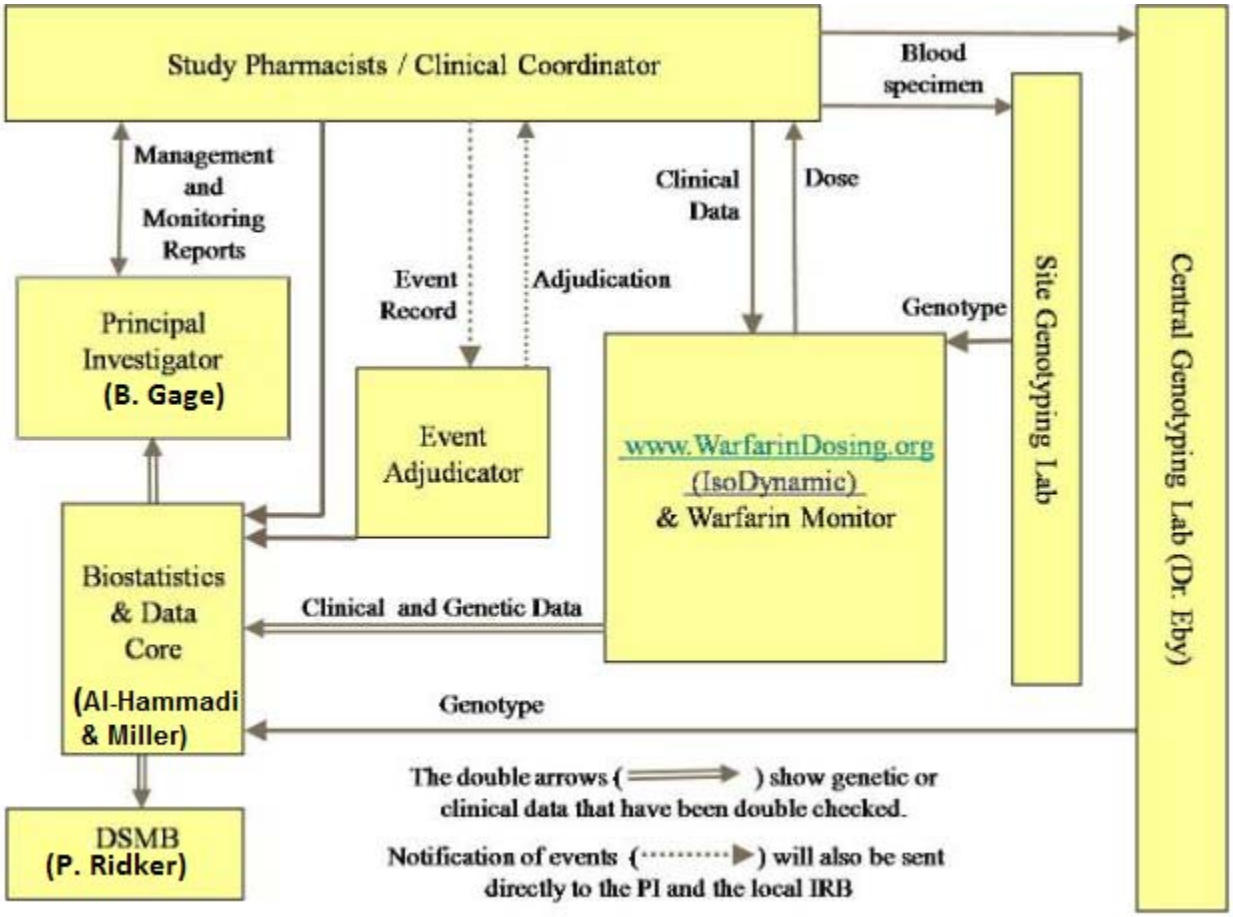


Randomization will be performed by the study website, WarfarinDosing.org and be stratified by study site, race, and type of joint replacement (hip vs. knee). WarfarinDosing.org will collect baseline information and display the therapeutic warfarin dose predicted by either a pharmacogenetic or clinical dosing algorithm, depending on treatment arm. WarfarinDosing.org has also been configured to prompt for regular INR follow up and adherence to the dosing protocol.

This protocol is organized by the order in which information will be collected and/or transmitted (figure 2. illustrates this information flow). Orthopedic patients will be referred by their orthopedist to the Anticoagulation Service for warfarin initiation. The research coordinator will then contact the patient to explain the study and screen for eligibility. After informed consent, the coordinator will collect initial clinical data and three blood samples. He or she will enter the clinical data into the Warfarin Monitor database and WarfarinDosing.org. The coordinator will then send the 3 deidentified blood samples to the site's genotyping laboratory. After the laboratory technician has performed genotyping, he or she will enter the genotypes at www.WarfarinDosing.org. GIFT sites that do not have genotyping available locally will FedEx the blood samples directly to the GIFT Central Laboratory for genotyping. After the genotype has been entered, WarfarinDosing.org will email a dose (according to the study arm randomly assigned to the patient) back to the study pharmacist and the Study PI. This email will blind the researchers to genotype/Aim 1 study arm, but not to Target INR/Aim 2 study arm. A fourth blood specimen will be drawn on post-op day 2 and frozen for later study. The extra deidentified samples will be sent via FedEx on a regular basis to the central genotyping laboratory for genotype quality assurance, pre-operative warfarin-related analytic substudies, and long-term storage.

Inpatients will be monitored for adverse events daily by members of the clinical research team (section 8). During the telephone follow-up of INR values after discharge, adverse events will be collected systematically. If a symptomatic VTE has not been documented, patients will have a Doppler ultrasound at the time of their 3-7 week follow-up visit. Whenever possible, the Doppler US will be done within 6 weeks of arthroplasty. If a major hemorrhage, death, or symptomatic VTE has occurred (as determined locally to prevent delay in clinical follow-up, if necessary), the clinical coordinator will notify the participant's primary care physician, their local IRB, and the PI at Washington University. After local reading, the research team members will send the necessary documents and/or US images to the centralized event adjudicator, who will report back to the site clinical coordinator in case of discrepancy. Adjudication will be by a physician experienced in the interpretation of ultrasound images. While blinded to the original reading, the physician-adjudicator will interpret a sample of positive Doppler ultrasound reports and an equal number of negative ultrasound reports. After central adjudication and the resolution of any discrepant reports, the Biostatistics and Data Core will generate reports for the study PIs, and the IRB, regarding event rates of adjudicated outcomes (not stratified by arm). While the study PIs, IRB, clinicians and patients will remain blinded, the Data Safety and Monitoring Board (DSMB), will receive event rates stratified by trial arm. Updates on the trial status and trial observations will be made available through various means (Appendix 3). Trial data will be released to the public via PharmGKB (or similar password-protected public data base) 12 months after trial completion.

Figure 2 Summary of the organization and flow of information and DNA for the proposed trial. DSMB=Data Safety and Monitoring Board; PI= Principal Investigator.



2 Anticoagulation Services

The Anticoagulation Services are staffed by clinical pharmacists and physicians with expertise in warfarin therapy. WarfarinDosing.org will email these pharmacists with the predicted therapeutic doses, but genotype will be masked in this email.

3 Participant Eligibility

Inclusion Criteria

Patients will be eligible to participate if they are aged 65 years or older, and anticipate taking warfarin therapy for 4-6 weeks for VTE prophylaxis after elective hip or knee arthroplasty. They must be able to give written, informed consent, have reliable telephone access, and be willing/able to follow-up in 3-7 weeks with a Doppler Ultrasound. Participants must have venous access, a life expectancy > 6 months, and plans to have regular INR monitoring.

Exclusion Criteria

We will exclude patients, who know their genotype or their therapeutic warfarin dose from prior therapy, or who refuse to give written consent. They must not have an absolute contraindication or allergy to warfarin therapy or plan to receive any anticoagulant besides warfarin (except heparin flushes). However, if LMWH, fondaparinux, or subcutaneous heparin is deemed necessary by the clinician after enrollment, such patients will remain in the study. Participants must not be incarcerated or institutionalized at the time of enrollment, or unlikely to be compliant (e.g. due to history of non-compliance or alcoholism), but nursing home residents are eligible to participate. Use of antiplatelet therapy (e.g. aspirin, dipyridamole, ticlopidine, clopidogrel, or prasugrel) will be allowed, as clinically indicated. As in the pilot studies, use of intermittent compression devices will be allowed for inpatients (but not used after hospital discharge). Patients with known thrombophilia, a bleeding disorder, a serious bleed in the past 2 years unless caused by trauma, or a baseline INR $\geq$ 1.35 will be excluded because it would be unethical to randomize them to the 2 target INR ranges. Patients will not be recruited if their clinicians are of the opinion that warfarin dosing needs to be adjusted for any reason not accounted for by the dosing algorithms.

4 Participant Recruitment and Involvement (see also Protection of Human Subjects)

As in the pilot studies, we will recruit patients who plan to undergo knee or hip replacement surgery. Sites will include Washington University in St. Louis, Intermountain Health Care, University of Utah Hospital, the Hospital for Special Surgery (Weill-Cornell, NYC), Rush University, and University of Texas Southwestern. Orthopedic patients will be recruited upon initial evaluation for total hip or knee arthroplasty, at the group education meeting that they attend pre-operatively, or in the preadmission testing pre-operatively. After written consent, an extra 11 ml of blood for research purposes will be drawn with their pre-operative labs. If blood is not available, then saliva or a buccal swab may be used for DNA recovery.

Recruiters will be provided with a checklist of inclusion and exclusion criteria, as well as a sample recruitment script. The consent process will begin with a verbal description of the study, the risks and benefits of participating in the study, confidentiality of health information, and the right to refuse or withdraw participation without consequence at any time. Informed consent will be obtained by study pharmacists or other personnel certified by the IRB (coordinators, research assistants) to recruit patients. Because we are not asking sensitive questions, nor testing for paternity, HIV, or illicit drug use, we will not provide a certificate of confidentiality. Recruiters will be trained to answer questions about the protocol and consent forms.

Folders for potentially recruited patients will be prepared in advance and will contain blank consent forms, intake questionnaires, eligibility checklist, a recruitment script, and INR/dose monitoring forms (a paper-based copy of the daily information collected in warfarin monitor) affixed with a unique study identifier. Concomitant anticoagulant and antiplatelet agents will be recorded in case report forms. In addition, the timing of the warfarin dose will be standardized and recorded and the time of each blood draw will be noted on each specimen. Folders for each consented participant will be maintained and data from each folder will be entered into the Warfarin Monitor database and www.WarfarinDosing.org. The recruiters (or hospital phlebotomist in the presence of the recruiter) will collect the blood specimens at the time of recruitment, label the specimens with the study ID, and send them to the local genotyping lab for processing. They will then enter clinical data into Warfarin Monitor

and www.WarfarinDosing.org. The genetic laboratory assistants will enter genotype separately onto the website; they will be trained to withhold communication about the genotype of particular patients from the pharmacists and research assistants. When feasible, a fourth 3 mL blood specimen will be drawn on post-op day 2 and also sent to the local genotyping laboratory and frozen for later study (of S- and R- warfarin levels, high-sensitivity C-reactive protein [hs-CRP], IL-6, high-sensitivity cardiac troponin I (hs-cTnI), and clotting factors). The deidentified blood specimens will be sent to the Central Genotyping Laboratory via FedEx for genotyping, analyte substudies, and DNA archiving.

Inpatients will be monitored for adverse events daily by members of the clinical research team (section 8). During the telephone follow-up of INR values after discharge, adverse events will be collected systematically. If a major hemorrhage or symptomatic VTE has not been documented clinically, patients will have a Doppler ultrasound at the time of their 3-7 week follow-up visit.

Together, the participating sites initiate warfarin therapy on at least 13,000 orthopedic patients per year. Our refusal rate averages 6-7%, and the drop-out rate (due mainly to change in orthopedic procedure after recruitment) is < 10%. Thus, we plan to recruit 1600 participants.

5 Randomization and Blinding

Randomization

Using a 2 x 2 factorial design, we will randomize patients to each of the following:
(1) Pharmacogenetic (50%) vs. clinical dosing of warfarin (50%); and
(2) a standard target INR (2.5) (50%) vs. a lower target INR (1.8) (50%)

This randomization scheme allows us to answer two essential questions about VTE prevention in this high-risk population (Aims 1-2). To allow for randomization stratified by site, race, and type of arthroplasty (knee vs. hip), www.WarfarinDosing.org will randomize patients after these data have been entered by clinical coordinator. Lists for block randomization will be prepared in advance by the trial statistician, and monitored to ensure proper function.

Blinding

To maintain double-blinding to genotype, www.WarfarinDosing.org will randomize participants to pharmacogenetic or clinical dosing and email the recommended warfarin dose to the study pharmacist and PI. Thus, these clinicians will receive an email with the estimated dose tailored to genotype and/or clinical factors, but the genotype will be masked. We acknowledge that rare genotypes will require unusually small doses, potentially leading to unblinding. However, because so many factors affect warfarin dose besides genotype, inadvertent unmasking will be rare. Randomization to standard vs. lower target INR value will not be double blinded because clinicians must know patients' target INR to monitor their INR properly. Bleeding and VTE events will be assessed locally (blinded to genotype), so that any clinical action deemed necessary may proceed without delay and appropriate IRBs can be notified. Event reports and/or deidentified portions of the medical record will then be sent to the central adjudicators, who will be blinded to both genotype and target INR, for confirmation. The central adjudicators' adjudications will follow standardized guidelines and be considered final for analysis purposes (see section10 for details). Collaborating orthopedic surgeons will be blinded to the genotype and study arm (pharmacogenetic or not) of participants, but not target INR.

Rules for Unblinding

We will not include genotyping results in the patient's medical record nor provide them to study participants nor their physicians. All procedures during the unblinding process will be reported to the study PI and DSMB, where they will be documented. The DSMB will have access to unblinded data, when requested, for evaluating and/or reporting on patient safety during the conduct of the trial. The consent form will contain language that makes our intent clear to participants.

6 Algorithm-Based Dosing of Warfarin

The genetic and clinical dosing algorithms have been published[13-17] or are in the manuscript writing stage and have been programmed into WarfarinDosing.org. The algorithms use the following polymorphisms to estimate the therapeutic warfarin dose among participants randomized to genetic dosing: *VKORC1*2 (-1639 G>A*, dbSNP rs9923231), *CYP2C9*2 (430C>T*, dbSNP rs1799853), *CYP2C9*3 (1075A>C*, dbSNP rs1057910*)*, and *CYP4F2*3* (V433M, 1297G>A, dbSNP rs2108622).[18-20] The anticoagulation service or managing physician will enter warfarin doses and INRs for the initial 11 days of therapy for each participant using a unique patient ID. The website will then calculate a dose refinement based on this information as well as the study arm and genotype (when applicable). The genotype and study arm will be electronically stored in the background and not visible by the pharmacist seeking a dose-refinement.

7 Standardization of INR Monitoring

Table D.1. Protocol for INR Monitoring, Stratified by Duration of Warfarin Therapy

Week of Rx	INR Monitoring*
1	Daily while an inpatient; biweekly (usually Mon & Thur) for outpatients
2-4	At least 1X/week*

*INRs should also be drawn when clinically indicated

The effect of warfarin will be assessed by the INR. For this measure, a 3- or 4.5-ml blood sample will be collected by venipuncture into a vacutainer containing 0.5 ml of 3.2% sodium citrate and centrifuged to obtain plasma for analysis. Prothrombin times will be performed on automated coagulation instruments (e.g. STAR™) using a thromboplastin with an international sensitivity index $\leq$ 1.5. Point-of-care (POC) INR testing will be allowed. The frequency of INR monitoring will be per protocol (Table D.1). Deviations from this schedule will be made if deemed clinically necessary as determined by protocol (i.e., for bleeds, after dose adjustments, or following INR values > 3.5) and when a national holiday falls on a Monday or Thursday. Inpatients will have daily INR measurements, as is standard of care post joint replacement.

8 Standardization of Dosing Protocol & Event Monitoring

After randomization, but prior to surgery, WarfarinDosing.org will be used to predict a therapeutic warfarin dose (either based on the clinical or the pharmacogenetic algorithm [15]) for each patient. The initial dose will be administered to participants starting the day before or the night of surgery (depending on local practice). CYP2C9 variants will be ignored during the 1st two days of

therapy. This convention, which is based on pharmacokinetic modeling and our prior work,[15] prevents under-dosing slow metabolizers. It also allows clinicians to remain blinded to study arm, even in patients who have usual CYP2C9 genotypes.

After two warfarin doses, the research coordinator will enter the warfarin doses and INR (called "INR_2") into www.WarfarinDosing.org. The website then estimates a refined estimate of the therapeutic warfarin dose that incorporates INR_2 and genotype (for patients in the pharmacogenetic arm). Then, each day that an INR is available, the researcher will enter INR value and get a refined dose estimate up to day 11 of warfarin therapy. On days when the INR is not drawn, patients will continue to receive their estimated maintenance dose. WarfarinDosing.org also indicates whether the dose of warfarin given on the day of INR testing should differ from the estimated therapeutic dose. This feature (called "Today recommendations") allows WarfarinDosing.org to compensate for missed doses, large doses, or other dosing errors.

Clinicians will be allowed to round up or down to the nearest 1 mg (for doses > 3 mg/day) or 0.5 mg (for doses $\leq$ 3.0 mg/day) at their clinical discretion. If the clinician overrides this protocol-defined dose, he or she will receive prompting by the website to adhere to the protocol (unless a deviation is clinically indicated). When necessary (e.g., an INR > 5), low-dose oral vitamin K will be administered to reverse the effects of warfarin, as clinically indicated.

Inpatients will be monitored by the study coordinators for warfarin-related adverse events daily until hospital discharge. During the hospital stay, recommendations for warfarin doses will also be made daily by the study pharmacist or managing physician. Thereafter, the coordinator will monitor for adverse events (including VTE and, bleeding) whenever an INR is checked (per schedule in Table D.1). Patients who stop their warfarin prematurely should be called weekly through day 30 to assess for study outcomes and adverse events. All patients also should be called after 30 days and 90 days of follow-up to assess for study outcomes and adverse events. Serious adverse events will be immediately reported to the IRB and DSMB (if the DSMB requests this information). Patients will be given a list of potential adverse effects and telephone numbers to report adverse events to the study pharmacist or managing physician. Patients will be given warfarin for 4-6 weeks after hospital discharge and scheduled to return for Doppler screening of leg veins 3-7 weeks after the date of surgery (when they have their routine follow up).

9 Event Definitions

Major Bleeding

As per the Control of Anticoagulation Subcommittee of the International Society on Thrombosis and Haemostasis (ISTH) definition,[21] all major bleeds must be symptomatic or clinically-overt according to 3 criteria:

1. Fatal bleeding defined as bleeding that has been adjudicated as the cause of patient death by a panel of experts blinded to treatment and study arm, and/or
2. Overt bleeding in a critical area or organ, such as intracranial, intraspinal, intraocular, retroperitoneal, joint or soft tissue hematoma requiring return to the operating room, intraarticular or pericardial, or intramuscular with compartment syndrome, and/or

3. Overt bleeding causing a fall in hemoglobin* level of 20 g/L (i.e., 2 g/dL or 1.24 mmol/L) or more, or leading to transfusion of 2 or more units of whole blood or red cells.

*we will use post-op day 1 Hgb as "baseline" because this value accounts for intra-operative blood loss and volume resuscitation in the OR.

Because the threshold for blood transfusion has evolved[22 23] since the ISTH guideline was formulated, GIFT also will classify as "major" any bleed that meet this fourth criterion:

4. An overt bleed causing hemodynamics changes and leading to transfusion of 1 or more units of blood. The number of additional major bleeds identified by criteria #4, will be reported.

Minor Bleeding

We will report minor bleeding, defined as bleeding that is neither major nor occult using the following categories:

1. Wound hematoma

2. Gross Hematuria, excluding hematuria caused by traumatic insertion of a Foley catheter

3. Gastrointestinal bleeding

4. Intra-operative bleeding that is unexpected for the type of surgical procedure and that requires blood transfusion of ≤ 2 units packed red blood cells, even if the transfusion is administered post-operatively

5. Other site

6. Hemovac drainage (total) will be collected for all participants

Minor bleeding that is clinically relevant, according to the criteria in Amadeus[24,25] (see their Table 1, below, © New Engl J Med) shall be so noted.

Table 1. Definition of Major and Clinically Relevant Bleeding.*

Major bleeding
Bleeding associated with a fall in hemoglobin of 2 g per deciliter or more
Bleeding that led to a transfusion of 2 or more units of packed red cells or whole blood†
Bleeding that involved a critical organ (intracranial, intraocular, intraspinal, retroperitoneal, or pericardial)
Bleeding that contributed to death
Clinically relevant bleeding
Any bleeding compromising hemodynamics
Any bleeding leading to hospitalization
Subcutaneous hematoma larger than 25 cm^2, or 100 cm^2 if there was a traumatic cause
Intramuscular hematoma documented by ultrasonography
Epistaxis that lasted for more than 5 minutes, was repetitive (i.e., two or more episodes of bleeding more extensive than spots on a handkerchief within 24 hours), or led to an intervention (e.g., packing or electrocoagulation)
Gingival bleeding occurring spontaneously (i.e., unrelated to eating or tooth brushing) or lasting for more than 5 minutes
Hematuria that was macroscopic and was spontaneous or lasted for more than 24 hours after instrumentation (e.g., catheter placement or surgery) of the urogenital tract
Macroscopic gastrointestinal hemorrhage, including at least one episode of melena or hematemesis, if clinically apparent with positive results on a fecal occult-blood test
Rectal blood loss, if more than a few spots on toilet paper
Hemoptysis, if more than a few speckles in the sputum and not occurring within the context of pulmonary embolism
Any other bleeding type considered to have clinical consequences for a patient — such as medical intervention, the need for unscheduled contact (visit or telephone call) with a physician, or temporary cessation of a study drug — or associated with pain or impairment of activities of daily life

* Any one or more of the criteria met the definition of either major or clinically relevant bleeding.
† A red-cell unit was defined as the quantity of red cells obtained from or corresponding to approximately 500 ml of whole blood.

Venous Thromboembolic Event (VTE)

VTE includes any DVT or PE that has been objectively confirmed by a Doppler US, venography, pulmonary-perfusion scan, spiral CT scan, MRI, or pulmonary angiogram; an elevated D-Dimer test will not be sufficient for a diagnosis of VTE. We will screen for asymptomatic DVT by Doppler US 3-7 weeks post-operatively. Using the classic method,[26] DVT examination will consist of comprehensive venous compression, color flow imaging, and pulse wave evaluation with

augmentation of the common femoral, superficial femoral, popliteal veins, and calf veins at 2-cm intervals in the transverse plane. A positive Doppler is defined by the detection of any noncompressible intraluminal venous thrombus; adjunct measures of a positive Doppler will include lack of color flow and diminished augmentation.

Myocardial Infarction (MI)

An event will be considered an MI when "there is evidence of myocardial necrosis in a clinical setting consistent with myocardial ischemia" [27]. Under these conditions any one of the following criteria meets the diagnosis for myocardial infarction:

Detection of rise and or fall of cardiac biomarkers (per the ACC/AHA guidelines, preferred cardiac biomarker is troponin) with at least one value above the 99th percentile of the upper reference limit together with myocardial ischemia with at least one of the following:

a. Symptoms of ischemia
b. ECG changes indicative of new ischemic (new ST-T changes or new left bundle branch block)
c. Development of pathological Q waves in the ECG
d. Imaging evidence of new loss of viable myocardium or new regional wall motion abnormality.

The 99% percentile of the upper reference limit will be assay specific. For example, using the current reagents at Barnes-Jewish Hospital, the upper limit will be $\geq$ 0.07 ng/ml.

10 Adjudication of Outcomes

First, major bleeds or VTEs will be assessed by the local research coordinator and site PI, both of who will be masked to genotype and to PGx vs. clinical treatment assignments. To provide double data entry of the primary outcomes, www.WarfarinDosing.org will also solicit INR, warfarin dose, and clinical outcomes if not entered at 30 days of follow up. Next, deidentified event reports, including the relevant portions of the medical record or of ultrasound images (if available for review) will then be sent to the centralized event adjudicators for validation. Central adjudicators will be blinded to genotype, Target INR, study arm, and warfarin doses.

The safety outcomes will be major bleeding (as defined above, Section 9), and the occurrence of INR≥4.0 within 30 days of arthroplasty. The effectiveness outcomes will be nonfatal VTE (including asymptomatic DVTs found by Doppler ultrasound within 60 days of arthroplasty) or death within 30 days of arthroplasty. Patients who develop a nonfatal efficacy or safety outcome will continue follow-up until the end-of-study assessment except that a Doppler US will not be done in patients who already have had an incident VTE documented.

11 Primary Study Endpoints and Hypotheses

11.1 Timeframe for Study Endpoints

The study is a 2 x 2 factorial design. The primary endpoint for Aim 1 is the composite of nonfatal VTE (DVT or PE), major hemorrhage, INR>4.0, or death occurring with 30 days of arthroplasty. Because the date of onset of VTE is ambiguous in the post-operative setting (and depends on the date of the Doppler US screening), VTEs diagnosed within 60 days of arthroplasty will be included in the primary endpoint.

The primary endpoint for Aim 2 is the composite of nonfatal VTE or death. We opted to include all-cause death in the endpoint (as opposed to *vascular death*) because cause of death may not be discernable. Primary analyses for both aims will be on an intent-to-treat basis, but we also will report an on-treatment analysis.

Table 2. Summary of Aims, Endpoints, Hypotheses, and Statistical Tests

Aim	Factor	Endpoints, Primary [Secondary]	Hypothesis, Primary	Statistical Test
1.	Genetic vs. Clinical Dosing	VTE, major hemorrhage, death, or INR $\geq$ 4.0 [INR control]	Fewer events with genetic dosing in whole population, and in subpopulation whose clinical and genetic predicted doses* differ by $\geq$1 mg/day.	Chi-square test (partitioning alpha as described below)
2.	Target INR 2.5 vs. 1.8	VTE or death [INR control, bleeding]	The event rate is non-inferior in the lower target INR arm	Chi-square test

VTE = Venous Thromboembolic Event; INR International Normalized Ratio
*Predictions based on dose initiation models [15].

11.2 Effect of Censoring on the Analysis of the Primary Outcomes

Withdrawal or loss to follow-up may occur either due to circumstances unrelated or related to the trial. We will plot time to study discontinuation in the study arms and compare them using the log-rank test. The frequency of INR testing and dose-adjustment will be captured and analyzed in all patients (including study dropouts). In pilot studies at Washington University, which were observational, the rate of withdrawal was 3.4%, but we anticipate a higher drop out in this randomized trial. Subjects who withdraw from the study after randomization will be included in the analyses on an intent-to-treat basis. Because all aims will benefit from completion of the trial, we do not plan an interim analysis, but the DSMB has authority to stop the trial in the event of unexpectedly high number of serious adverse events.

12 Power and Statistical Analyses of Primary Endpoints

Aim 1. Primary endpoint for clinical vs. pharmacogenetic warfarin dosing

For Aim 1, our composite endpoint is VTE (DVT or PE), death, major bleed, and an INR$\geq$4.0. We will analyze the primary endpoint in the whole population and in the subpopulation whose clinical and genetic predicted doses differ by $\geq$ 1.0 mg/day (50% of the population) using a two-sided chi-square test. For the primary analysis, we opted not to weigh clinical events due to complications relating to the arbitrariness of weights or non-linear trends in the frequency of endpoint severity (see

secondary analyses). To preserve the type I error rate of this primary endpoint, we will partition our alpha for the tests in the whole group and the subgroup, as described below.

We hypothesize that the rate of VTE in the subpopulation whose clinical and genetic predicted doses differ by $\geq$ 1.0 mg/day will be 1.6 times as high as that of the remaining population. The 1.6-fold increase accounts for a greater rate of adverse events in patients who have genetic variants, especially CYP2C9*2 and/or CYP2C9*3.

We originally estimated VTE rates of 18% in patients randomized to clinical dosing and 15% in patients randomized to genetic dosing based on older data.[28-31] In the clinical arm, we originally anticipated a rate of major bleeding as 2.4% and rate of death as 1.0%, for a total of 3.4%. In the pharmacogenetic arm, we anticipated the rate of major bleeding or death as 2.3%, a 32% relative risk reduction based on a meta-analysis of clinical trials[32] and observational studies.[16,33] We had estimated the rate of INRs $\geq$ 4.0 in clinical and pharmacogenetic arms from previous research[16] as 12.3% and 7.4% respectively. We anticipate that half of the bleeding events will be associated with INRs $\geq$ 4.0, and account for this correlation in our original and updated power calculations.

When GIFT was planned, the expected rate of the composite endpoint (non-fatal VTE, non-fatal major hemorrhage, death, or INR$\geq$4.0) in Aim 1 was estimated as 27.3%. This rate would have provided a power of 99% for a sample size of 1600 participants. Midway through the trial (when data from 775 GIFT participants were available), the composite endpoint in Aim 1 was observed to be 13.15%, which provides for a power of 80%.

The 80% power was calculated using a two-sided alpha of 0.05 for a test of proportions, a drop-out rate of 2%, a modest (5%) correction for continuity, and assuming the 32% relative risk reduction (RRR) in adverse events from our original meta-analysis,[32] which yielded estimated rates of 10.7% in the pharmacogenetic arm and 15.7% in the clinical arm. The 80% power includes use of the partitioned alpha with 0.044 allocated to the whole population and 0.01 to the high-risk subgroup. Because of correlation between these two subgroups, using these alphas preserves an overall type 1 error rate of 0.05.

Alpha partitioning

To preserve a Type 1 error rate of 5% for Aim 1, we will partition the alpha between the whole group analysis and the subgroup analysis. The subgroup consists of patients for whom PGx and clinical doses differ by 1+ mg/day. Due to correlation between outcomes in main study and subgroup, Bonferroni splitting would be overly conservative. Due to the lack of a closed equation, we simulated this correlation to allocate the alpha optimally. In the table (below) 'α_{whole}' is the alpha spent by the test of the whole group (and should be ~0.05). 'α_{sub}' is the alpha spent by the test of the subgroup. 'α_{total}' is the total alpha (=sum minus intersection of the probability spaces) of these correlated tests. Partitioning the alpha in this manner maximizes power for Aim 1 while limiting the overall type 1 error rate to 0.05.

α_{whole}	α_{sub}	α_{total}
0.04	0.01	**0.049**
0.025	0.025	**0.044**
0.027	0.027	**0.047**
0.028	0.028	**0.051**
0.044	0.01	**0.050**
0.04	0.014	**0.051**
0.04	0.013	**0.051**

We elected to partition the alpha according to the highlighted row above, as it maximizes the power for the test in the whole group, without jeopardizing the power in the subgroup.

Aim 2. Primary endpoint for low vs. high target INR

We hypothesize that orthopedic patients randomized to a target INR of 1.8 will have a rate of VTE or vascular death that is non-inferior to a target INR of 2.5. Using the average of our estimates above, we had expected the rate of VTE (including fatal events) with warfarin therapy and Doppler US screening to be 16.5%, which yielded the original power calculation (see appendix). Based on the aggregate (blinded) analysis done half-way through GIFT, we observed that the VTE rate averaged only 5.56%, which increased the power for this non-inferiority analysis. Therefore, we were able to *decrease* the non-inferiority margin from the original value (5%), to a more stringent one (3%), while increasing our power from 80% to 83%. For the updated power calculation, we used a sample size of 1600 participants, drop-out rate of 2%, and non-inferiority margin of 3%. For this calculation, we used a one-sided test (because we are testing for non-inferiority) and a minimum absolute difference of $\underline{3}$% in VTEs detectable by Doppler US.

Differences of VTEs of $\underline{\leq}$ 3% (as detected by screening) seem unlikely to motivate orthopedic surgeons and other physicians to use a therapy with a higher risk of hemorrhage. For example, LMW heparins have a 3% lower absolute rate of VTEs (on venography), yet warfarin remains more popular in the US, because of its lower rate of hemorrhage, oral administration, and low cost.

Contingency plan for statistical analyses.

For both aims, if the expected counts in any cell are less than 5, we will use a Fisher's exact test. If by chance randomization were to result in an unbalanced distribution of any clinical variable associated with VTE [i.e., age, body mass index, hormonal replacement therapy, or male gender [34]], we would adjust for the unbalanced variables using logistic regression. Analyses will be performed in SAS, version 9.1.3 or greater or in *R*.

13 Secondary Outcomes and their Statistical Analyses

13. 1 Percentage of Time in Therapeutic Range (PTTR)

We will compare percentage time spent in therapeutic range (PTTR) during days 4-28 of warfarin therapy for pharmacogenetic vs. clinical dosing in a regression model that uses linear interpolation, as recommended.[35] If there is no INR measured on day 28, but there is an INR measured later (e.g. day 29), then the latter INR will be used so that the analysis can be completed for

days 4-28 of therapy. Our approach to missing data (Table D.3) is based on that used by the COAG investigators (Table 3):

Table 3 Approaches to dealing with missing INRs

Missing Data Status	**PTTR Computation**
No INR on or after day 4	Missing, PTTR = "."
Only 1 INR on or after day 4	If INR: < 2, PTTR = 0 2-3, PTTR = 0.5 > 3, PTTR = 0
Any temporary discontinuation after day 4	If $\leq$ 5 days: Compute PTTR with all available INRs If >5 days: Compute PTTR with all available INRs for 5 days after hold, then use all INRS after restart and concatenate.
Permanent discontinuations after day 4/5	Compute PTTR with available INRs up through 5 days after discontinuation.

A "restart" is defined as starting warfarin after it had been held for at least 1 day. For patients who have the drug held for 5 days or fewer, all available INRs will be used in the calculation of the PTTR. For those who have the drug held for more than 5 days, any INRs measured in the 5 days after the drug was held will be used in the calculation of PTTR. Once the drug is restarted, the first INR drawn will then be used to calculate PTTR from that point on. The overall PTTR will be concatenated between the courses of warfarin therapy; that is, for both short and longer term holds, a single PTTR will be calculated for a patient using all INRs available during the time on warfarin. For patients who have their warfarin permanently discontinued, the PTTR will be calculated using all INRs through 5 days after discontinuation.

We will also conduct a separate analysis of time supratherapeutic (i.e. >0.5 units more than the target INR) during the first 28 days of warfarin. In that analysis we will test for an interaction between the study arm in Aim 1 and CYP2C9 genotype. Specifically, we will code 2C9*1*1 as 0, 2C9*1*2 as 1, 2C9*1*3 or 2C9*2*2 as 2, 2C9*2*3 as 3, and 2C9*3*3 as 4 because the effect of the 2C9*3 allele on warfarin metabolism is approximately twice the 2C9*2 effect.

13.2 INR Variability

We will report INR variability, defined as the standard deviation of transformed INR values from days 4-28 of therapy, calculated according to the method of Lind et al.[36]

13.3 Time to First Event

We will compare time to first laboratory event (number of days until INR > target INR + 1.5) graphically (using Kaplan–Meier curves) and statistically (using the log-rank test, Wilcoxon test, or Cox-proportional hazard model), as appropriate. We will censor participants at the time of withdraw, loss to follow-up, death from an unrelated event, or 90-days of follow-up (whichever come first).

Likewise, we will compare time to the first major or non-major, clinically relevant bleed that is clinically relevant (safety endpoint) within 90-days of follow up.

13.4 Secondary Statistical Analysis of Primary Endpoint in Aim 1

As a secondary outcome, we will analyze the rank of events and test the hypothesis that genetic dosing decreased the rank of adverse events vs. clinical dosing in the whole cohort. We will use the following tiers, in hierarchical order, from worst to best: (1) death; (2) PE; (3) Major bleed; (4) symptomatic DVT; (5) INR $\geq$ 4 with minor bleed; (6) asymptomatic DVT; (7) INR $\geq$ 4 (w/out major/minor bleed); (8) PTTR. For the clinical outcomes (1-7) the duration of follow-up will be 60 days, events that happen earliest receive the lowest (worst) score. For PTTR, lower time in the target INR range is worse. This approach, similar to that used in the RELAX trial (Redfield et al. 2013) weighs outcomes according to their clinical relevance. Ranks are compared using a standard non-parametric test (Mann-Whitney 1947) to determine if one arm improves outcomes. The Steering Committee favored this approach, rather than using weighted outcomes, because it avoids assigning *ad hoc* weights to these adverse events.

13.5 Secondary Analyses for Aim 2

We also will report the secondary analyses above for the two arms in Aim 2. Furthermore, we will compare the two arms in Aim 2 using the same composite outcome from Aim 1: VTE (within 60 days), or any of the following within 30 days: major hemorrhage, death, or INR $\geq$ 4.0.

14 Potential Problems and Their Resolution

Incorporating New Variants into Pharmacogenetic Arm

We recognize that we or others may discover additional relevant SNPs during the course of the trial and will establish criteria for whether these SNPs are clinically relevant. The primary metric for this decision will be whether this SNP has been validated to warrant scientific agreement of its effect as well as the ability of the new and validated SNP to decrease the prediction error of the pharmacogenetic algorithm in affected individuals by some minimum threshold. We will also need to establish the frequency of such variants in the subgroups studied in the trial, and determine whether a minimum allele frequency should also be established, so that the overall effect (as measured by the R^2) improves significantly in a clinical and statistical manner (e.g. 1%).

If a variant improves the predictive accuracy significantly, the pharmacogenetic algorithm would be modified. This adjustment would be easily incorporated by the website, since all predictive variables translate to multipliers in the dosing algorithm. For example, if we find a variant whose effect in carriers is to increase the necessary dose by 20%, we would program WarfarinDosing.org to apply the 1.2 multiplier where appropriate. For common SNPs, a complementary dose adjustment would be made to the non-carriers. Our positive working relationship with Osmetech's eSensor, Pyrosequencing, and Autogenomics' INFINITI™ will facilitate the implementation of a rapid genotyping platform for any new SNP warranting pharmacogenetic implementation as well.

15 Timetable for the Study.

Table 3 Three phases of the proposed trial

Phase	**Timeframe**	**Goal**
Protocol Revision	18 months	Revise and finalize protocol, obtain IRB approval, configure website for trial
Randomization & Follow up	5 years	Enroll and follow patients; monitor data for cleanliness and adverse events
Complete Data Analysis	2 months	Analyze and promulgate results

16 Data Sharing and the Dissemination of Results

As detailed in Appendix 3, we will disseminate the results of the trial using the Internet, traditional medical publications, professional societies, lay press, and collaboration with the FDA. We will register the trial following instructions at http://prsinfo.clinicaltrials.gov/registering.pdf, and regularly update information pertaining to its status. One year after the completion of the trial and data analysis, we will release the trial data to the public via the PharmGKB site maintained at Stanford University or equivalent site. We will submit the methods of the clinical trial for publication.

17 Archiving of DNA and Blood for Subsequent Analyses

DNA and plasma from the proposed study will be archived at the central genotyping laboratory for subsequent studies of SNPs or biomarkers that may affect bleeding or VTE, and SNPs or biomarkers that may further influence warfarin dosing. These future studies may include whole genome association studies, or candidate gene studies about warfarin dosing, thrombosis, or hemorrhage. We also will archive blood for future pharmacokinetic modeling, and (possible) future proteomic and metabolomic studies about hemorrhage and VTE.

Appendix 1: Justification for Protocol Modifications Made April 7, 2015

p. 1, 13. **Power for Aim 1 has been changed to 80%**
Previous: ">99% power"
Change: "80% power"
Rationale: When the study was planned, the expected rate of the composite endpoint (non-fatal VTE, non-fatal major hemorrhage, death, or INR≥4.0) in Aim 1 was estimated as 27.3%. However, improvements in standard of care (early mobilization, shorter lengths of stay, and avoidance of general endotracheal anesthesia) and other factors have resulted in fewer outcomes than anticipated. Midway through the trial, when data from 775 GIFT participants were available, the composite endpoint in Aim 1 was observed to be 13.15%, which provides for a power of 80%.

The 80% power was calculated using a two-sided alpha of 0.05 for a test of proportions, a drop-out rate of 2%, a modest (5%) correction for continuity, and assuming a 32% relative risk reduction (RRR) in adverse events from our original meta-analysis,[32] which yielded estimated rates of 10.7% in the pharmacogenetic arm and 15.7% in the clinical arm. The 80% power includes use of the partitioned alpha with 0.044 allocated to the whole population and 0.01 to the high-risk subgroup. Because of correlation between these two subgroups, using these alphas preserves an overall type 1 error rate of 0.05.

p. 3, 6. **Expanded use of the Central Laboratory**
Previous: Initially, each GIFT site did its own genotyping locally.
Change: "GIFT sites that do not have genotyping available locally will FedEx the blood samples directly to the GIFT Central Laboratory for genotyping."
Rationale: It is more efficient to have GIFT specimens FedEx'd to the Central Laboratory, who runs batched genotyping at least weekly.

p. 3. **Adjudication of Doppler US images clarified**
Previous: Adjudication of Doppler US was not stated explicitly in the protocol.
Change: "Adjudication will be by a physician experienced in the interpretation of ultrasound images. While blinded to the original reading, the physician-adjudicator will interpret a sample of positive Doppler ultrasound reports and an equal number of negative ultrasound reports."
Rationale: The adjudication will allow for an estimate of inter-observed agreement when interpreting Doppler (Duplex) US images.

p. 3 **Release of data 12 months after GIFT completion**
Previous: **"**Trial data will be released to the public via PharmGKB six months after the last recruited patient has finished participation."
Change: **"**Trial data will be released to the public via PharmGKB (or similar password-protected public data base) **12** months after trial completion."
Rationale: The GIFT investigators are committed to making GIFT data publicly available 12 months after publishing the main findings. However, to prevent duplicate analyses, GIFT data will not be made publicly available until 12 months of trial completion. However, collaborators are encouraged to pursue ancillary studies at any time, by submitting an application to the GIFT Ancillary Study.

p. 3 and 14. **Non-inferiority Margin has been changed to 3%**
Previous: **"Hypothesis 2:** For prevention of non-fatal VTE or death, a target INR of 1.8 will be non-inferior to a higher target INR (2.5). Using a non-inferiority margin of 5%, we will have 80% power to detect the non-inferiority of a target INR of 1.8 in 1600 patients."
Change: **"Hypothesis 2:** For prevention of non-fatal VTE or death, a target INR of 1.8 will be non-inferior to a higher target INR (2.5). Using a non-inferiority margin of **3%**, we will have **83**% power to detect the non-inferiority of a target INR of 1.8 in 1600 patients."
Rationale: Previously, the statisticians pointed out that given the lower rate of VTE and death in GIFT, that a non-inferiority margin of 5% (in the absolute event rate) was too high. Using a 3% non-inferiority margin, we estimated a power of 83% with 1600 enrollees. We used a one-sided alpha of 0.05 for a non-inferiority test of proportions, composite rates of 5.56%, and a drop-out rate of 2%.

p. 5 (Figure D.2) **Updated head of the GIFT Biostatistics and Data Core**
Previous: Juan Li, MPH was the head of the Biostatistics and Data Core.
Change: J. Phil Miller now heads the Biostatistics and Data Core and Noor Al-Hammadi is the statistical data analyst.
Rationale: Juan Li has moved to China,

p. Multiple. **Timing of follow-up visit for US made more flexible**
Previous: **"**If a symptomatic VTE has not been documented during the first 4-6 weeks of therapy, patients will have a Doppler ultrasound at the time of their 4-6 week follow-up visit."
Change: "If a symptomatic VTE has not been documented, patients will have a Doppler ultrasound at the time of their **3-7** week follow-up visit."
Rationale: Sometimes the orthopedic follow-up visit occurs after 6 weeks, so the ultrasound (US) is late. We now make explicit that DVTs diagnosed on or before 60 days post-operatively will be included in the primary outcomes for both Aims. This delay is the reality of getting an US test in elderly patients who cannot drive in the post-op period.

p. 5. **Patients with a prior bleed caused by trauma are no longer excluded from GIFT**
Previous: "Patients with known thrombophilia, a bleeding disorder or serious bleed in the past 2 years, or a baseline INR ≥ 1.35 will be excluded because it would be unethical to randomize them to the 2 target INR ranges."
Change: "Patients with known thrombophilia, a bleeding disorder, a serious bleed in the past 2 years unless caused by trauma, or a baseline INR ≥ 1.35 will be excluded because it would be unethical to randomize them to the 2 target INR ranges."
Rationale: Patients with **non**-traumatic bleed will continue to be ineligible for GIFT because there is not clinical equipoise re: which target INR they should use—many clinicians prefer a lower target INR for this population. This preference does not extend to patients whose only major bleed in the past 2 years was due to trauma. Therefore, the latter patients are still eligible to participate in GIFT.

p. 5 **Addition of new GIFT sites**
Previous: "Washington University Medical Center, Intermountain Health Care, University of Utah Hospital, or the Hospital for Special Surgery (Weill-Cornell, NYC)."
Change: "Sites will be Washington University in St. Louis, Intermountain Health Care, University of Utah Hospital, the Hospital for Special Surgery (Weill-Cornell, NYC), Rush University, and University of Texas Southwestern."
Rationale: To meet our targeted enrollment and to increase enrollment of Hispanic patients, Rush University and University of Texas Southwestern will now participate in GIFT.

p. 6 **Reiterate target enrollment of 1600 patients, but not necessarily in 4 years.**
Previous: "Thus, we should have no difficulty recruiting 400 participants per year, or 1600 participants in 4 years at the 4 participating sites."
Change: "Thus, we plan to recruit 1600 participants."
Rationale: We now anticipate that enrollment will complete in March 2016.

p. 9. **Definition of *Major Bleed* has been expanded slightly**
Previous: GIFT used only the 3 ISTH criteria to define a major bleed
Change: "Because the threshold for blood transfusion has evolved[22 23] since the ISTH guideline was formulated, GIFT also will classify as "major" any bleed that meet this fourth criterion:

4. An overt bleed causing hemodynamics changes and leading to transfusion of 1 or more units of blood. The number of additional major bleeds identified by criteria #4 (estimated to be 2), will be reported."

Rationale: Because the threshold for blood transfusion has evolved since the ISTH guideline was first published,[21] GIFT also will classify as "major" any bleed that meets the fourth criterion. At the time that the ISTH guidelines were written (2005), patients with overt bleeding causing hemodynamics changes typically would have been transfused at least 2 units of packed RBCs. Now, these patients often receive only a single unit of blood.[22 23] The fourth criteria allows these bleeds to be capture as *major*, even if only 1 unit is transfused.

p. 10 **Minor bleeding that is clinically relevant shall be so noted.**
Previous: Minor bleeding was not further characterized
Change: "Minor bleeding that is clinically relevant, according to the criteria in Amadeus (see their Table 1 [reproduced, above]) shall be so noted."
Rationale: Many "minor bleeds" have clinically consequences. Providing this additional information may inform clinical decision making.

p. 12. **The maximum timeframe for VTE diagnosis**
Previous: No upper limit for date of VTE was provided.
Change: "The primary endpoint for Aim 1 is the composite of nonfatal VTE (DVT or PE), major hemorrhage, INR≥4.0, or death occurring with 30 days of arthroplasty. Because the date of onset of

VTE is ambiguous in the post-operative setting (and detection depends on the date of the Doppler US screening), VTEs diagnosed within 60 days of arthroplasty will be included in the primary endpoint."
Rationale: A reasonable delay in US should not prompt GIFT to ignore US results that were obtained within 60 days of surgery. The modified protocol makes the longer time frame explicit. The rationale for the 60-day time limit is that DVTs detected up to this date mostly likely developed during the 30-day intervention period.

p. 14 **The power analysis for Aim 2 was revised, based on the lower-than-expected VTE rates observed half-way through the trial**

Previous: "If we recruit 1600 participants (800 each for greater and lesser INR targets), and assume an 18% drop-out rate, we will have 1312 participants left for analysis. Using these figures, we will have 80% power to reject the null hypothesis of a difference greater than 5% (the non-inferiority margin) in the two arms."
Change: "Using the average of our estimates above, we had expected the rate of VTE (including fatal events) with warfarin therapy and Doppler US screening to be 16.5%, which yielded the original power calculation (see appendix). Based on the aggregate (blinded) analysis done half-way through GIFT, we observed that the VTE rate averaged only 5.56%, which increased the power for this non-inferiority analysis. Therefore, we were able to *decrease* the non-inferiority margin from the original value (5%), to a more stringent one (3%), while increasing our power from 80% to 83%. For the updated power calculation, we used a sample size of 1600 participants, drop-out rate of 2%, and non-inferiority margin of 3%."
Rationale: Given the observed VTE rate mid-way through GIFT, a non-inferiority margin of 5% would not have been sufficiently stringent.

p. 14-15, Section 13, **Secondary Outcomes and their Statistical Analyses**

Previous: "We will compare time spent in therapeutic range (PTTR) during the first 30 days of warfarin for pharmacogenetic vs. clinical dosing in a regression model using linear interpolation, as recommended.[35]"
Change: "We will compare percentage time spent in therapeutic range (PTTR) during days 4-28 of warfarin therapy for pharmacogenetic vs. clinical dosing in a regression model that uses linear interpolation, as recommended.[35] If there is no INR measured on day 28, but there is an INR measured later (e.g. day 29), then the latter INR will be used so that the analysis can be completed for days 4-28 of therapy. Our approach to missing data (Table D.3) is based on that used by the COAG investigators (Table 3):

A "restart" is defined as starting warfarin after it had been held for at least 1 day. For patients who have the drug held for 5 days or fewer, all available INRs will be used in the calculation of the PTTR. For those who have the drug held for more than 5 days, any INRs measured in the 5 days after the drug was held will be used in the calculation of PTTR. Once the drug is restarted, the first INR drawn will then be used to calculate PTTR from that point on. The overall PTTR will be concatenated between the courses of warfarin therapy; that is, for both short and longer term holds, a single PTTR will be calculated for a patient using all INRs available during the time on warfarin. For patients who have their warfarin permanently discontinued, the PTTR will be calculated using all INRs through 5 days after discontinuation."

Rationale: The above approach maximizes use of INR data while mitigating the effect on PTTR when warfarin is held > 5 days.

p. 15, Section 13.2 **INR Variability will be reported**

Previous: This metric was not included in the original version of the protocol.
Change: "We will report INR variability, defined as the standard deviation of transformed INR values from days 3-28 of therapy, calculated according to the method of Lind et al.[36]"
Rationale: Using administrative data from 19,180 patients with atrial fibrillation, M. Lind et al. found the INR variability (defined as the standard deviation of transformed INR) was a stronger predictor than TTR for both stroke and of bleeding.[36]

p. 16, Section 13.4 **The secondary statistical analysis in Aim 1 has been detailed**
Previous: **"**As a secondary outcome, we will rank events as INR ≥ 4, asymptomatic DVT, symptomatic DVT, major bleed or PE, death, and analyze with ordinal logistic regression."
Change: "As a secondary outcome, we will analyze the rank of events and test the hypothesis that genetic dosing decreased the rank of adverse events vs. clinical dosing in the whole cohort. We will use the following tiers, in hierarchical order, from worst to best: (1) death; (2) stroke; (3) MI; (4) PE; (5) symptomatic DVT; (6) INR > 4, (7) DVT detected on screening ultrasound; (8) PTTR. For the clinical outcomes (1-7), events that happen earliest receive the lowest (worst) score. For PTTR, lower time in the target INR range is worse. This approach, used in the RELAX trial[37] weighs outcomes according to their clinical relevance. Ranks are compared using a standard non-parametric test (Mann-Whitney 1947) to determine if one arm improves outcomes. The Steering Committee favored this approach, rather than using weighted outcomes, because it avoids assigning *ad hoc* weights to these adverse events."
Rationale: This non-parametric approach[38] avoids the assumptions necessary for ordinal logistic regression to be valid.

p. 16 Section 13.5 **Secondary analyses for Aim 2 have been clarified.**

Previous: Secondary analyses for Aim 2 were not detailed
Change: "We also will report the secondary analyses above for the two arms in Aim 2. Furthermore, we will compare the two arms in Aim 2 using the same composite outcome from Aim 1: VTE, major hemorrhage, death, or INR > 4.0."
Rationale: The latter analysis will allows us to quantify any difference in the four-part composite outcome in the two target INR groups.

p. 17 Section 15 **Timetable for the Study**

Previous:

"Phase	**Timeframe**	**Goal**
Protocol Revision	12 months	Revise and finalize protocol, obtain IRB approval, configure website for trial
Randomization & Follow up	4 years	Enroll and follow patients; monitor data for cleanliness and adverse events
Data Analysis	6 months	Analyze and promulgate results"

Change:

"Phase	**Timeframe**	**Goal**
Protocol Revision	18 months	Revise and finalize protocol, obtain IRB approval, configure website for trial
Randomization & Follow up	5 years	Enroll and follow patients; monitor data for cleanliness and adverse events
Complete Data Analysis	2 months	Analyze and promulgate results"

Rationale: The new time table is more realistic.

Appendix 2: Justification for Protocol Modifications Made July 15, 2015

p. 8 **Timing of follow-up has been made explicit.**

Previous: The study sites were told to have the follow-up intervals described above.

Change: "Patients who stop their warfarin prematurely should be called weekly through day 30 to assess for study outcomes and adverse events. All patients also should be called after 30 days and 90 days of follow-up to assess for study outcomes and adverse events."

Rationale: To prevent ascertainment bias, patients who stop their warfarin early should be followed with equal frequency. The 90-day outcomes has been part of GIFT since the 1st patient was recruited, but the protocol was not explicit.

Section 12 Power and Statistical Analyses of Primary Endpoints, p. 13

Original: *Aim 1. Primary endpoint for clinical vs. pharmacogenetic warfarin dosing*

… Based on a variety of data, we estimate VTE rates of 18% in patients randomized to clinical dosing and 15% in patients randomized to genetic dosing. Historically, DVT rates with warfarin therapy after joint arthroplasty have been variable [28] often with rates around 25%. However, because seminal studies (e.g. [29,30]) screened for DVT using a more sensitive test, venography, we estimate that an 18% VTE rate will be detected by Doppler US screening in *GIFT*. The rate of VTE with pharmacogenetic-dosed warfarin also is uncertain. In our pilot study,[31] the rate of DVT was 11.5%, but the 95% confidence interval was large. For *GIFT*, we estimated a 15% VTE rate in participants dosed pharmacogenetically.

We suspect that the rate of VTE in the subpopulation whose clinical and genetic predicted doses differ by ≥ 1.0 mg/day will be 1.6 times as high as that of the remaining population. The 1.6-fold increase accounts for a greater rate of adverse events in patients who have genetic variants, especially CYP2C9*2 and/or CYP2C9*3.

Compared to VTEs, major bleeds and deaths will be uncommon. In the clinical arm, we anticipate that the rate of major bleeding will be 2.4% and the rate of death will be 1.0%, for a total of 3.4%. In the pharmacogenetic arm, we anticipate the rate of major bleeding or death will be 2.3%, a 32% relative risk reduction based on a meta-analysis of clinical trials[32] and observational research.[16,33]

We used one of these observational studies[16] to estimate the reduction in supra-therapeutic INR values. We estimated the rate of INRs ≥ 4.0 in clinical and pharmacogenetic arms from previous research in this area as 12.3% and 7.4% respectively. We anticipate that half of the bleeding events will be associated with INRs ≥ 4.0, and account for this correlation in our power calculations.

If we recruit 1600 participants (800 each for pharmacogenetic and for clinical arms), and assume an 18% drop-out rate, we will have 1312 participants left for analysis. Using these figures and partitioning our alpha to preserve an overall Type I error rate of 0.05 for the primary endpoints in Aim 1, we calculate 99% power to detect a difference in the rate of the composite endpoint between clinical and pharmacogenetic arms in the whole population or the subgroup.

Change: *Aim 1. Primary endpoint for clinical vs. pharmacogenetic warfarin dosing*

…We hypothesize that the rate of VTE in the subpopulation whose clinical and genetic predicted doses differ by > 1.0 mg/day will be 1.6 times as high as that of the remaining

population. The 1.6-fold increase accounts for a greater rate of adverse events in patients who have genetic variants, especially CYP2C9*2 and/or CYP2C9*3.

We originally estimated VTE rates of 18% in patients randomized to clinical dosing and 15% in patients randomized to genetic dosing based on older data [28-31]. In the clinical arm, we originally anticipated a rate of major bleeding as 2.4% and rate of death as 1.0%, for a total of 3.4%. In the pharmacogenetic arm, we anticipated the rate of major bleeding or death as 2.3%, a 32% relative risk reduction based on a meta-analysis of clinical trials [39] and observational studies[16]. We had estimated the rate of INRs > 4.0 in clinical and pharmacogenetic arms from previous research[16] as 12.3% and 7.4% respectively. We anticipate that half of the bleeding events will be associated with INRs > 4.0, and account for this correlation in our original and updated power calculations.

When GIFT was planned, the expected rate of the composite endpoint (non-fatal VTE, non-fatal major hemorrhage, death, or INR>4.0) in Aim 1 was estimated as 27.3%. This rate would have provided a power of 99% for a sample size of 1600 participants. Midway through the trial (when data from 775 GIFT participants were available), the composite endpoint in Aim 1 was observed to be 13.15%, which provides for a power of 80%.

The 80% power was calculated using a two-sided alpha of 0.05 for a test of proportions, a drop-out rate of 2%, a modest (5%) correction for continuity, and assuming the 32% relative risk reduction (RRR) in adverse events from our original meta-analysis,[32] which yielded estimated rates of 10.7% in the pharmacogenetic arm and 15.7% in the clinical arm. The 80% power includes use of the partitioned alpha with 0.044 allocated to the whole population and 0.01 to the high-risk subgroup. Because of correlation between these two subgroups, using these alphas preserves an overall type 1 error rate of 0.05.

Rationale:

The rationale for the above changes is to make explicit the original and updated power for Aim 1. The primary outcomes, sample size, effect size (32% relative risk reduction), and proposed analytic approach have not changed.

Appendix 3: Justification for Protocol Modifications Made April 14, 2017

p. 6 Added hs-cTnI to the list of post-operative labs:

Previous: When feasible, a fourth 3 mL blood specimen will be drawn on post-op day 2 and also sent to the local genotyping laboratory and frozen for later study (of S- and R- warfarin levels, high-sensitivity C-reactive protein [hs-CRP], IL-6, and clotting factors).

Change: When feasible, a fourth 3 mL blood specimen will be drawn on post-op day 2 and also sent to the local genotyping laboratory and frozen for later study (of S- and R- warfarin levels, high-sensitivity C-reactive protein [hs-CRP], IL-6, high-sensitivity cardiac troponin I (hs-cTnI), and clotting factors)

Rationale: The hs-cTnI will allow us to determine whether there was occult myocardial damage and the risk factors for a rise in hs-cTnI.

p. 15 Correction of time frame for calculation of INR Variability:

Previous: We will report INR variability, defined as the standard deviation of transformed INR values from days 3-28 of therapy, calculated according to the method of Lind et al.

Change: We will report INR variability, defined as the standard deviation of transformed INR values from days 4-28 of therapy, calculated according to the method of Lind et al.

Rationale: For consistency with the PTTR analysis, we will analyze days **4**-28.

p. 16 Correction of "*Secondary Statistical Analysis of Primary Endpoint in Aim 1"*

Previous: We will use the following tiers, in hierarchical order, from worst to best: (1) death; (2) stroke; (3) MI; (4) PE; (5) symptomatic DVT; (6) INR ≥ 4, (7) DVT detected on screening ultrasound; (8) PTTR.

Change: We will use the following tiers, in hierarchical order, from worst to best: (1) death; (2) PE; (3) Major bleed; (4) ~~PE; (5)~~ symptomatic DVT; (5) INR ≥ 4 with minor bleed; (6) asymptomatic DVT; (7) ~~DVT detected on screening ultrasound~~;INR > 4 (w/out major/minor bleed); (8) PTTR.

Rationale: Stroke and MI were not primary endpoints in GIFT, were not centrally adjudicated, and were very rare. In contrast, major bleeds were primary endpoints in GIFT, were centrally adjudicated, and were more common than stroke or MIs. Therefore, we will include major bleeds, but not stroke and MI in this secondary analysis. We moved asymptomatic DVTs above supratherapeutic INR values (provided that the high INR was not accompanied by bleeding) because any DVT predisposes to PE and post-phlebitic syndrome.

Appendix 4: Identification of the High-Risk Subgroup

If $|\text{Clinical\_Dose}_i - \text{Genetic\_Dose}_i| \geq 1$ mg/day, then patient i was in the high-risk group, where

| | is the absolute value function,

Clinical_Dose$_i$ is the estimated clinical dose for patient i in mg/d as calculated below and rounded to the nearest 0.1 mg/d.

Genetic_Dose$_i$ is the estimated pharmacogenetic dose for patient i in mg/d as calculated below and rounded to the nearest 0.1 mg/d.

The clinical dose estimate (on day 1) is from on Gage et al.[15]:
Clinical_Dose$_i$ (mg/day) = exp [0.613 − (0.0075 × Age) − (0.257 × Amiodarone) + (0.425 × BSA) + (0.156 × RaceAA) + (0.108 × Smokes) + (0.216 × Target_INR) + (0.0784 ×VTE)] × Sulfa_Factor × Azole_Factor

where

exp is the exponential function.

Age is in years.

Amiodarone = 1 if prescribed; 0 otherwise.

BSA is body surface area is in m^2 calculated from Dubois and Dubois[40]:
BSA (m^2) = $0.20247 \times \text{Height(m)}^{0.725} \times \text{Weight(kg)}^{0.425}$.

RaceAA = 1 if patient self identifies as black or African-American; 0 otherwise;

Smokes = 1 if the patient uses a tobacco product; 0 otherwise;

Target_INR = the target INR, which was either 1.8 or 2.5 in GIFT.

VTE = 1 if the indication for warfarin was venous thromboembolism treatment; 0 otherwise. In GIFT, the indication was either hip or knee arthroplasty so **VTE** was always 0.

Sulfa_Factor = 0.86 if sulfamethoxazole is prescribed[41]; 1 otherwise.

Azole_Factor = 0.5 if an azole (fluconazole, itraconazole, ketoconazole, miconazole, posaconazole, voriconazole) is prescribed[42,43]; 1 otherwise.

The genetic dose (on day 1) also was adapted from Gage et al.[15]:
Genetic_Dose$_i$ (mg/day) = exp[0.97505 − (0.00745 × Age) – (0.2538 × Amiodarone)+ (0.43172 × BSA) – (0.09007 × RaceAA) + (0.09215 × Smokes) + (0.20291 × TargetINR) + (0.0664 × VTE) – (0.32376 × VKOR_1639G>A) − (0.40075 × CYP2C9*3) − (0.20658 × CYP2C9*2)] × Sulfa_Factor × Azole_Factor × CYP4F2_Factor

where the SNPs (VKOR_1639G>A, CYP2C9*3, CYP2C9*2, CYP4F2_Factor) are coded 0 if absent, 1 if heterozygous, and 2 if homozygous. CYP4F2_Factor accounts for the V433M polymorphism[20] with a coefficient that is slightly greater among patients who self-identify as black or African-American:

	CYP4F2_Factor	
	Value in AA	**Value in Other Races**
CC (wildtype)	0.979	0.956
CT (heterozygous)	1.058	1.033
TT (homozygous mutant)	1.142	1.115

Note to reader: WarfarinDosing.org ignores *CYP2C9* variants (*CYP2C9*2* and *CYP2C9*3*) during the 1st two days of therapy. This convention, which is based on pharmacokinetic modeling and our prior

work,[15] prevents under-dosing slow metabolizers. It also allows clinicians to remain blinded to study arm, even in patients who have usual CYP2C9 genotypes.

References

1. Remarks Prepared for the Honorable Mike Leavitt, Secretary of Health and Human Services, on Personalized Medicine Coalition. U.S. Department of Health and Human Services, 2007. (Accessed October, 2, 2008, at http://www.hhs.gov/news/speech/2007/sp20070919a.html.)
2. Acting Surgeon General Issues 'Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism'. U.S. Department of Health and Human Services, Office of the Surgeon General, News Room, News Releases, 2008. (Accessed October 2, 2008, at http://www.surgeongeneral.gov/news/pressreleases/pr20080915.html.)
3. Kearon C, Ginsberg JS, Kovacs MJ, et al. Comparison of low-intensity warfarin therapy with conventional-intensity warfarin therapy for long-term prevention of recurrent venous thromboembolism. N Engl J Med 2003;349:631-9.
4. Hylek EM, Singer D. Risk factors for intracranial hemorrhage in outpatients taking warfarin. Annals of internal medicine 1994;120:897-902.
5. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115:2689-96.
6. Wood S. New Warfarin Labeling Reminds Physicians About Genetic Tests to Help Guide Initial Warfarin Dosing. theHeart.com; 2007:1.
7. Hillman MA, Wilke RA, Yale SH, et al. A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. Clin Med Res 2005;3:137-45.
8. Anderson JL, Horne BD, Stevens SM, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation 2007;116:2563-70.
9. Caraco Y, Blotnick S, Muszkat M. CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: a prospective randomized controlled study. Clinical pharmacology and therapeutics 2008;83:460-70.
10. Huang SW, Chen HS, Wang XQ, et al. Validation of VKORC1 and CYP2C9 genotypes on interindividual warfarin maintenance dose: a prospective study in Chinese patients. Pharmacogenetics and genomics 2009;19:226-34.
11. Ridker PM, Goldhaber SZ, Danielson E, et al. Long-term, low-intensity warfarin therapy for the prevention of recurrent venous thromboembolism. N Engl J Med 2003;348:1425-34.
12. Geerts WH, Pineo GF, Heit JA, et al. Prevention of Venous Thromboembolism: The Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126:338S-400.
13. Lenzini P, Grice G, Milligan P, et al. Optimal Dose Adjustment in Orthopaedic Patients Beginning Warfarin Therapy. Ann Pharmacother 2007;41 1798-804.
14. Millican E, Lenzini P, Milligan P, et al. Genetic-based dosing in orthopaedic patients beginning warfarin therapy. Blood 2007;110:1511-5.
15. Gage BF, Eby C, Johnson JA, et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clinical pharmacology and therapeutics 2008;84:326-31.
16. Lenzini PA, Grice GR, Milligan PE, et al. Laboratory and clinical outcomes of pharmacogenetic vs. clinical protocols for warfarin initiation in orthopedic patients. J Thromb Haemost 2008;6:1655-62.
17. Mant J, Hobbs FD, Fletcher K, et al. Warfarin versus aspirin for stroke prevention in an elderly community population with atrial fibrillation (the Birmingham Atrial Fibrillation Treatment of the Aged Study, BAFTA): a randomised controlled trial. Lancet 2007;370:493-503.

18. Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 2005;352:2285-93.
19. Linder MW, Bon Homme M, Reynolds KK, et al. Interactive modeling for ongoing utility of pharmacogenetic diagnostic testing: application for warfarin therapy. Clin Chem 2009;55:1861-8.
20. Caldwell MD, Awad T, Johnson JA, et al. CYP4F2 genetic variant alters required warfarin dose. Blood 2008;111:4106-12.
21. Schulman S, Kearon C, Subcommittee on Control of Anticoagulation of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. J Thromb Haemost 2005;3:692-4.
22. Carson JL, Grossman BJ, Kleinman S, et al. Red blood cell transfusion: a clinical practice guideline from the AABB. Annals of internal medicine 2012;157:49-58.
23. Hebert PC, Carson JL. Transfusion threshold of 7 g per deciliter--the new normal. N Engl J Med 2014;371:1459-61.
24. Lane DA, Kamphuisen PW, Minini P, Buller HR, Lip GY. Bleeding risk in patients with atrial fibrillation: the AMADEUS study. CHEST 2011;140:146-55.
25. Amadeus I, Bousser MG, Bouthier J, et al. Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomised, open-label, non-inferiority trial. Lancet 2008;371:315-21.
26. Cronan JJ. Venous thromboembolic disease: the role of US. Radiology 1993;186:619-30.
27. Lane DA, Lip GY. Female gender is a risk factor for stroke and thromboembolism in atrial fibrillation patients. Thrombosis and haemostasis 2009;101:802-5.
28. Geerts W, Bergqvist D, Pineo G, et al. Prevention of venous thromboembolism. CHEST 2008;133:381-453.
29. Francis CW, Davidson BL, Berkowitz SD, et al. Ximelagatran versus warfarin for the prevention of venous thromboembolism after total knee arthroplasty. A randomized, double-blind trial. Annals of internal medicine 2002;137:648-55.
30. Francis CW, Berkowitz SD, Comp PC, et al. Comparison of ximelagatran with warfarin for the prevention of venous thromboembolism after total knee replacement. N Engl J Med 2003;349:1703-12.
31. Voora D, Eby C, Linder MW, et al. Prospective dosing of warfarin based on cytochrome P-450 2C9 genotype. Thrombosis and haemostasis 2005;93:700-5.
32. Eckman MH, Rosand J, Greenberg SM, Gage BF. Cost-effectiveness of using pharmacogenetic information in warfarin dosing for patients with nonvalvular atrial fibrillation. Annals of internal medicine 2009;150:73-83.
33. Epstein RS, Moyer TP, Aubert RE, et al. Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo Warfarin Effectiveness study). Journal of the American College of Cardiology 2010;55:2804-12.
34. Davidson HC, Mazzu D, Gage BF, Jeffrey RB. Screening for deep venous thrombosis in asymptomatic postoperative orthopedic patients using color Doppler sonography: analysis of prevalence and risk factors. AJR Am J Roentgenol 1996;166:659-62.
35. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. Thrombosis and haemostasis 1993;69:236-9.
36. Lind M1 FM, Kosiborod M, Eliasson B, Odén A. Variability of INR and its relationship with mortality, stroke, bleeding and hospitalisations in patients with atrial fibrillation. Thrombosis research 2012;129:32-5.

37. Redfield MM, Chen HH, Borlaug BA, et al. Effect of Phosphodiesterase-5 Inhibition on Exercise Capacity and Clinical Status in Heart Failure With Preserved Ejection Fraction: A Randomized Clinical Trial. Jama 2013:1-10.
38. Packer M. Proposal for a new clinical end point to evaluate the efficacy of drugs and devices in the treatment of chronic heart failure. J Card Fail 2001;7:176-82.
39. Eckman M, Rosand J, Greenberg S, Gage B. Cost-Effectiveness of Pharmacogenetic-Based Dosing of Warfarin in Patients with Atrial Fibrillation. Annals of internal medicine 2009.
40. DuBois D, DuBois E. Clinical Calorimetry; a formula to estimate the approximate surface area if height and weight be known. Arch Int med 1916;17:863-71.
41. Ahmed A, Stephens JC, Kaus CA, Fay WP. Impact of preemptive warfarin dose reduction on anticoagulation after initiation of trimethoprim-sulfamethoxazole or levofloxacin. Journal of thrombosis and thrombolysis 2007.
42. Kunze KL, Wienkers LC, Thummel KE, Trager WF. Warfarin-fluconazole. I. Inhibition of the human cytochrome P450-dependent metabolism of warfarin by fluconazole: in vitro studies. Drug Metab Dispos 1996;24:414-21.
43. Purkins L, Wood N, Kleinermans D, Nichols D. Voriconazole potentiates warfarin-induced prothrombin time prolongation. British journal of clinical pharmacology 2003;56 Suppl 1:24-9.