Boehringer Ingelheim

Study Report: Comparative safety and effectiveness of oral anticoagulants: A cohort study

Document Number:	c02409758-01
BI Study Number:	1160.157
BI Investigational Product(s):	Dabigatran etexilate
Title:	Comparative effectiveness of oral anticoagulants: A cohort study (Phase 1 of the BI/BWH Pradaxa study program)
Version identifier of the final study report:	Version 1
Date of last version of the final study report:	20 February 2014
PASS:	Yes
EU PAS register number:	*# 3061*
Active substance:	Dabigatran etexilate
Medicinal product:	Pradaxa
Product reference:	N/A
Procedure number:	N/A
Marketing authorisation holder(s):	Boehringer Ingelheim GmbH
Joint PASS:	No
Research question and objectives:	Quantify associations between anticoagulant choice (warfarin and dabigatran) and the occurrence of selected outcomes including major thromboembolic events and major bleeds in patients with non-valvular atrial fibrillation at risk for stroke. The analyses will use data through June 2012.
Country(-ies) of study:	United States

Author:	[redacted], PharmD, DrPH [redacted]
Marketing authorisation holder(s):	Boehringer Ingelheim GmbH
MAH contact person:	[redacted]
Date:	14 Apr 2014

Page 1 of 195

TABLE OF CONTENTS

1. ABSTRACT

Name of company: Boehringer Ingelheim			
Name of finished medicinal product: PRADAXA			
Name of active ingredient: dabigatran etexilate			
Report date: 14 Apr 2014	**Study number:** 1160.157	**Version/Revision:** Version 1	**Version/Revision date:**
Title of study:	Comparative effectiveness of oral anticoagulants: A cohort study		
Keywords:	Dabigatran, warfarin, oral anticoagulants, thrombotic events, major bleeding		
Rationale and background:	A number of new oral anticoagulants have been developed and marketed. In Phase III studies, these drugs were found to be therapeutically advantageous or non-inferior over warfarin. In the coming years, as many as six new anticoagulants could be on the market with a lack of valid comparative evidence. This evaluation provides for the feasibility of a direct assessment of comparative effectiveness and safety across the anticoagulants. The current protocol addresses only the initial comparison between warfarin and dabigatran using data through June 2012.		
Research question and objectives:	The overall study objective is to quantify associations between anticoagulant use (warfarin and dabigatran) and the occurrence of selected outcomes, including major thromboembolic events, major bleeding events in patients with non-valvular atrial fibrillation (NVAF) at risk for stroke using a large US commercial health insurance database. The analyses described in this report are based on a cohort of patients identified between Jan 2009 and Jun 2012 in the UnitedHealth Research Database.		
Study design:	Observational cohort study		
Setting:	UnitedHealth July 2008 through June 2012		
Subjects and study size, including dropouts:	Patients ≥18 years with a diagnosis of NVAF at risk for stroke (CHA_2DS_2-VASc score ≥ 1) who initiated treatment with warfarin (N=7,724) or dabigatran (N=4,158) between October 2010 and June 2012. From this population, 2,991 patients initiating dabigatran were matched to 2,991 patients initiating warfarin, and these patients form the main study cohorts.		
Variables and data sources:	Exposure, outcomes, and baseline covariates were identified from UnitedHealth claims data from July 2008 through June 2012. Exposure is initiation of dabigatran or warfarin. Primary Outcomes are: stroke (hemorrhagic, ischemic, uncertain classification) and major bleeding.		

Name of company: Boehringer Ingelheim			
Name of finished medicinal product: PRADAXA			
Name of active ingredient: dabigatran etexilate			
Report date: 14 Apr 2014	**Study number:** 1160.157	**Version/Revision:** Version 1	**Version/Revision date:**
	Secondary outcomes include stroke or systemic embolism, systemic embolism, ischemic stroke, hemorrhagic stroke, stroke uncertain classification, major intracranial bleeding, major extracranial bleeding, major GI bleeding, major upper GI bleeding, major lower GI bleeding, major urogenital bleeding, major other bleeding, TIA, MI, VTE (DVT or PE), DVT, PE.. Covariates include demographic information and clinical risk factors for study outcomes.		
Results:	Following PS-matching, 2,991 dabigatran patients had a mean age of 63.86 ± 10.99 years with 31.53% being female and 2,991 warfarin patients had a mean age of 63.26 ± 11.04 years with 29.32% being female. Among the matched dabigatran patients (96% with the 150 mg dose) providing 1,237 person-years of follow-up and the warfarin patients providing 950 person-years of follow-up, there were 36 strokes (IR = 29.09, presented as events/1000 PY) among dabigatran users vs. 30 (IR = 31.59) among warfarin users (HR=1.05, 95% CI=0.64-1.70). With slightly different person-years follow-up (1,233 dabigatran vs. 944 warfarin), there were 74 major hemorrhages (IR = 60.00) among dabigatran users vs. 63 (IR = 66.72) in warfarin users (HR=0.97, 0.69-1.36). For the outcome of stroke or systemic embolism (1,233 person-years of dabigatran exposure vs. 944 person-years of warfarin exposure), there were 49 events (IR = 39.73) among dabigatran users vs. 57 events (IR = 60.38) among warfarin users (HR= 0.74, 95% CI=0.50-1.08, while for systemic embolism including PE, there were 1,242 person-years of dabigatran exposure and 949 person-years of warfarin exposure during which there were 14 events (IR = 11.27) vs. 29 events (IR = 30.55) (HR=0.40, 95% CI=0.21-0.76), with much of this effect due to PE. For ischemic stroke (1,240 person-years vs. 947 person-years) there were 34 events (IR = 27.40) vs. 31 events (IR = 32.71) (HR=0.95, 95% CI =0.58-1.55). There were only 8 hemorrhagic strokes in total (5 among dabigatran treated patients and 3 among warfarin treated patients). Hazard Ratios with corresponding 95% CI for hemorrhagic stroke and remaining secondary outcomes are listed below: Hemorrhagic stroke HR: 1.36; 95% CI: (0.32 - 5.72) Stroke Uncertain HR: 1.60; 95% CI: (0.83 - 3.08) TIA HR: 1.01; 95% CI: (0.47 - 2.20) MI HR: 1.24; 95% CI: (0.62 - 2.45) VTE HR: 0.29; 95% CI: (0.15 - 0.59) DVT HR: 0.31; 95% CI: (0.15 - 0.67) PE HR: 0.27; 95% CI: (0.09 - 0.84)		

Name of company: Boehringer Ingelheim			
Name of finished medicinal product: PRADAXA			
Name of active ingredient: dabigatran etexilate			
Report date: 14 Apr 2014	**Study number:** 1160.157	**Version/Revision:** Version 1	**Version/Revision date:**
	Major Intracranial bleeding HR: 1.45; 95% CI: (0.57 - 3.70) Major Extracranial bleeding HR: 0.95; 95% CI: (0.67 - 1.35) Major GI bleeding HR: 1.26; 95% CI: (0.79 - 2.01) Major Upper GI bleeding HR: 0.65; 95% CI: (0.21 - 1.98) Major Lower GI bleeding HR: 1.25; 95% CI: (0.78 - 2.01) Major Urogenital bleeding HR: NA; 95% CI: NA Other Major bleeding HR: 0.74; 95% CI: (0.48 - 1.14)		
Discussion:	Analyses suggested similarity of primary outcome occurrence between dabigatran and warfarin for stroke (HR = 1.05, 95% CI = 0.64-1.70) and major hemorrhage (HR = 0.97, 0.69-1.36) along with other effectiveness and safety outcomes, across a range of assumptions regarding exposure and within numerous subgroups and sensitivity analyses. Patients initiated on dabigatran appeared to be slightly more persistent with treatment, so this study reflects more person-time exposed to dabigatran than warfarin. These current feasibility results are limited by a small sample size, short follow up and few outcome events resulting in wide 95% confidence intervals. At this early stage, no comparative conclusions are possible. Future data in the context of the subsequent protocol (1160.207) will increase the number of patients and also expand to additional data sources.		
Marketing Authorisation Holder(s):	Boehringer Ingelheim GmbH		
Names and affiliations of principal investigators:	[redacted], PharmD, DrPH: [redacted] [redacted], MD, DSc: [redacted]		

3. LIST OF ABBREVIATIONS

ACE	Angiotensin Converting Enzyme
AF	Atrial Fibrillation
ARB	Angiotensin Receptor Blocker
$CHADS_2$	Congestive Heart Failure, Hypertension, Age>75, Diabetes Mellitus, Prior Stroke or Transient Ischemic Attack score
CHA_2DS_2-VASc	Congestive Heart Failure, Hypertension, Age>75, Diabetes Mellitus, Prior Stroke or Transient Ischemic Attack, Vascular Disease, Age 65-74, Sex Category score
CHF	Congestive Heart Failure
DRS	Disease Risk Score
DVT	Deep Venous Thrombosis
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
GI	Gastrointestinal
HAS-BLED	Hypertension, Abnormal Liver/Renal function, Stroke, Bleeding history or predisposition, Labile INR, Elderly (Age>65), Drugs-Alcohol usage
hdPS	high-dimensional Propensity Score
HR	Hazard Ratio
ICD-9	International Classification of Diseases, Ninth Revision
INR	International Normalized Ratio
LDL	Low Density Lipoprotein
MI	Myocardial Infarction
NOAC	New Oral Anticoagulants
NVAF	Non-Valvular Atrial Fibrillation
PE	Pulmonary Embolism
PGP	P-glycoprotein
PS	Propensity Score
PVD	Peripheral Vascular Disease
RE-LY	Randomized Evaluation of Long-Term Anticoagulant Therapy
SD	Standard Deviation
TIA	Transient Ischemic Attack
VTE	Venous Thromboembolism

4. INVESTIGATORS

Boehringer Ingelheim (Germany):

[redacted]

Brigham and Women's Hospital (United States)

[redacted]

5. OTHER RESPONSIBLE PARTIES

None

6. MILESTONES

Milestone	Planned date	Actual date	Comments
Start of data collection	March 2013	April 2013	Delay for ENCePP registration
End of data collection	June 2013	July 2013	
Registration in the EU PAS register	April 2013	April 2013	
Interim report	July 2013	July 2013	
Final report of study results	Sep 2013	Mar 2014	Delay for incorporation of comments and modification of report template

7. RATIONALE AND BACKGROUND

A number of new oral anticoagulants (NOACs) (dabigatran, rivaroxaban, and apixaban) are being developed and marketed. Unlike vitamin K antagonists, these new drugs do not require dose titration involving intensive therapeutic monitoring of anticoagulation level to achieve target anticoagulation within a narrow therapeutic range.

To date, these drugs have been studied in clinical trials for several indications, including but not limited to: (1) prevention of stroke and systemic embolism among patients with atrial fibrillation (AF); (2) prevention of deep venous thromboembolism (DVT) among patients undergoing hip or knee replacement therapy; and (3) treatment of venous thromboembolism (VTE). Phase III clinical trials comparing dabigatran, rivaroxaban, and apixaban to warfarin in patients with non-valvular atrial fibrillation (NVAF) have been completed. [P09-11669] [R11-4190] [R11-4223] There have been no direct comparisons of NOACs in randomized trials among patients with AF, but they have been compared to warfarin. In the Randomized Evaluation of Long-Term Anticoagulant Therapy (RE-LY) trial, patients with AF were randomly assigned to receive dabigatran (either 110 mg or 150 mg twice daily) or warfarin and followed for efficacy (stroke or systemic embolism prevention) and safety (major bleeding).[1] In this trial, dabigatran reduced the occurrence of stroke or systemic embolism relative to warfarin by an amount that varied by dose (RR = 0.91, 95% CI = 0.74-1.11 for 110 mg, and RR = 0.66, 95% CI = 0.53-0.82 for 150 mg). Major bleeding was also reduced with dabigatran relative to warfarin, and differently by dose (RR = 0.80, 95% CI = 0.69-0.93 for 110 mg, and RR = 0.93, 95% CI = 0.81-1.07 for 150 mg). A revised analysis based on some additional events identified subsequent to the original RE-LY study produced similar results: Stroke or systemic embolism RR = 0.90, 95% CI = 0.74-1.10 for 110 mg, and RR = 0.65, 95% CI = 0.52-0.81 for 150 mg; Major hemorrhage RR = 0.80, 95% CI = 0.70-0.93 for 110 mg, and RR = 0.93, 95% CI = 0.81-1.07 for 150 mg. [P10-12855]Since usage of either dabigatran or warfarin outside a clinical trial (i.e. in routine clinical practice) may differ from that within the clinical trial with respect to population treated, monitoring (especially for warfarin), and adherence to treatment, and differences in these aspects of use may lead to differences in effectiveness or safety, an observational study conducted in a routine practice setting will provide data that supplements results from clinical trials.

The observational study described in this report is part of a program designed to assess the comparative effectiveness and safety of warfarin versus dabigatran and additional NOACs as they become available. We sought to evaluate the comparative effectiveness and safety among oral anticoagulants in routine care. This report includes only the initial comparison between warfarin and dabigatran using routine care data through June 2012 to assess the feasibility of such an ongoing sequential cohort study program, which is still an innovative approach in post-marekting drug safety surveillance. The continued follow-up of these cohorts as well as the formation of additional cohorts and inclusion of an additional data source is planned at intervals defined in the subsequent protocol (1160.207).

8. RESEARCH QUESTION AND OBJECTIVES

The overall study objective is to quantify associations between anticoagulant use (warfarin and dabigatran – and other NOACs when available) and the occurrence of selected outcomes, including major thromboembolic events, major bleeding events, hepatotoxicity, and all-cause mortality in patients with non-valvular atrial fibrillation (NVAF) at risk for stroke using a large US commercial health insurance database.

The first analyses described in this report are based on a cohort of patients identified between Jan 2009 and Jun 2012 (the data available for this report) in the UnitedHealth Research Database.

Future analyses will extend the study size by the inclusion of additional cohorts identified in 6-month increments, along with continued follow-up of previously-identified cohorts. This sequential approach will extend initial analyses through accrual of greater numbers of people exposed to the study medications and by extending follow-up time, permitting evaluation of relative safety and effectiveness of longer follow-up, and also by inclusion of additional NOACs as they become available according to the protocol-defined report schedule outlined in the subsequent protocol (1160.207). [P12-00343]

9. AMENDMENTS AND UPDATES

None

10. RESEARCH METHODS

10.1 STUDY DESIGN

The study uses a cohort design with propensity score matching to address potential confounding. The primary study comparison was between new initiators of warfarin and new initiators of dabigatran (as dabigatran etexilate, but indicated as dabigatran throughout this report for brevity), and this comparison was conducted for each of the study outcomes. The primary study outcomes were stroke (hemorrhagic, ischemic, uncertain classification) and major bleeding. Secondary outcomes were: stroke or systemic embolism, systemic embolism (defined as an acute vascular occlusion of the extremities or any organ, such as kidneys, mesenteric arteries, spleen, retina or grafts), ischemic stroke, hemorrhagic stroke, stroke of uncertain classifications, major intracranial bleeding, major extracranial bleeding, major gastrointestinal (GI) bleeding, major upper GI bleeding, major lower GI bleeding, major urogenital bleeding, major other bleeding, transient ischemic attack (TIA), myocardial infarction (MI), venous thromboembolism (VTE), deep vein thrombosis (DVT), pulmonary embolism (PE). Further outcomes were hepatotoxicity and all-cause mortality. Effect measures were estimated using person-time based analyses. The primary analyses represented an 'as treated' approach, with follow-up starting the day after cohort entry and ending at the time of disenrollment, end of the observation period (available data), death, discontinuation of the index study exposure (+ 14 days), or switch to a different anticoagulant, whichever came first.

10.2 SETTING

Data for the patients included in this cohort study arose from a de-identified research database of United Healthcare members who had both medical and prescription benefits. The individuals covered by this health insurance are geographically diverse across the United States. The plan provides fully insured coverage minus applicable copayments for physician, hospital, and prescription drug services. The research database included claims for reimbursement of pharmacy dispensings, inpatient and outpatient services, and procedures including the associated diagnoses for an open cohort (i.e. one where the cohort members may enter or exit at varying times) of members with an average cohort residence time of 2.5 years (SD, 2.4) and a cross-sectional size around 14 million persons depending on the study year. This data source has an open formulary with tiered copayment structure so that use of medications is guided more by physician-patient interactions than by health insurer policies. This feature of the data source means that newly-marketed medications will be observed within it in relation to their use among the physicians who accept patients with United Healthcare insurance.

This study was approved by the institutional review board of the Brigham and Women's Hospital and a signed licensing agreement was in place. The study was registered at ENcEPP.org (# 3061) and ClinicalTrial.gov (#NCT01847547) before the initiation.

10.3 SUBJECTS

The study cohort consists of patients ≥18 years with a diagnosis of NVAF at risk for stroke who initiated treatment with warfarin or dabigatran between October 2010 and June 2012. Warfarin initiators between January 2009 and September 2010, before dabigatran became available, were used for estimation of disease risk scores. [R13-0524] The definition of initiation was a first dispensing with no prior dispensing of any oral anticoagulant in the 12 months preceding this initiation date. We therefore required continuous enrollment in the database for at least 12 months prior to treatment initiation, allowing for gaps of up to 32 days based on enrollment and disenrollment dates. [R14-1394] The diagnosis of non-valvular atrial fibrillation was defined by the presence of a diagnosis code for atrial fibrillation (ICD-9 427.31) at any time prior and including the date of the cohort entry (i.e., date of treatment initiation), provided there was no diagnosis or procedure codes indicating valvular

disease (see Table 1). Patients were considered at risk for stroke if they had a CHA_2DS_2-VASc score ≥ 1 (See Table 2). Patients' follow-up ended when there was a treatment gap ≥14 days beyond the computed end of days medication supplied, at treatment switch, at disenrollment, or 30 June 2012, whichever came first. Patients with ambiguous age or sex information were excluded. The patient flow diagram shows the development of the dabigatran and warfarin cohorts (Figure 1). These early phase numbers underscore the limited size of this analysis, and its value in providing insights into the feasibility of such a monitoring system rather than on providing precise results.

10.4 VARIABLES

See below in section 10.4.1, 10.4.2 and 10.4.3.

10.4.1 Exposures

The primary study comparison was between new initiators of warfarin and new initiators of dabigatran, and this comparison was conducted for each of the study outcomes (primary, secondary, and further). Effect measures were estimated using person-time based analyses. The primary analyses represented an 'as treated' approach, with follow-up starting the day after cohort entry and ending at the time of disenrollment, end of the observation period (available data), death, discontinuation of the index study exposure (+ 14 days), or switch to a different anticoagulant, whichever came first. These exposure rules were intended to focus the analysis on person-time during which patients were likely to be taking the medications.

Warfarin therapy involves dose titration following initiation before a stable therapeutic INR is obtained, and during this titration phase warfarin exposure will be less closely correlated with expected consumption (days of warfarin dispensed) than it will be later in a course of warfarin therapy (once a stable INR is achieved). To account for this uncertainty in duration of exposure following a dispensing, we used the same exposure definition as for dabigatran, which is days' supply from the pharmacy claim for all warfarin dispensings (+14 days), with exposure duration being updated according to each new warfarin dispensing. This exposure rule should both accurately reflect warfarin exposure during the stable phase of warfarin therapy where the days supplied by the pharmacy matches expected consumption and also accommodate the titration phase of warfarin therapy, since upward titration in warfarin dose increases warfarin consumption, prompting a new warfarin dispensing, which will lead to an update in warfarin exposure (days supplied with the new dose) under the exposure rule. Downward titration in warfarin dose leads to reduced warfarin consumption so that actual exposure to the medication occurs for longer than predicted based on days supplied, and the 14-day extension of exposure from the end of days supplied accounts for this. We also assessed different exposure assumptions in sensitivity analyses by reducing this grace period to 7 days and by increasing it to 30 days.

In contrast to the 14-day extension of days supplied before discontinuation of exposure, patients who switch anticoagulant therapy were censored in the as-treated approach on the day of the dispensing of the new anticoagulant medication. Where the switch in therapy was a response to an adverse effect of a drug, the adverse effect and the discontinuation would have preceded the dispensing of the new medication. Accordingly, clinical events recorded before the switch would be attributed to the index drug, and this is the rationale for not applying the 14-day extension in the case of a switch. In sensitivity analyses, we explored the effect of attributing events occurring in the first two days after the date of a switch to the index drug. These exposure rules are necessary due to the nature of the source database within which reasons and timing of switch are not explicitly recorded, but must be inferred from pharmacy dispensing data.

In secondary analyses, an intention to treat approach was taken that corresponds to a first exposure carried forward assumption, where patients are assumed to be exposed to the medication that defines

their cohort entry for all follow-up until 1 year (365 days), with censoring at the occurrence of a study outcome, disenrollment from the database, or the end of the study period (30 June 2012). This exposure rule accepts greater ambiguity with respect to use of the medication during follow-up, and consequently patients are less likely to be exposed to the medication at the time of an event than under the 'as treated' analyses. [R14-1391]

10.4.2 Outcome(s)

10.4.2.1 Primary outcome(s)

The primary study outcomes were stroke (hemorrhagic, ischemic, uncertain classification) and major bleeding. Please refer to Table 3 for a complete listing of the diagnostic and procedure codes used to define these outcomes and the secondary and further outcomes listed below.

10.4.2.2 Secondary outcome(s)

Secondary outcomes were: stroke or systemic embolism, systemic embolism (defined as an acute vascular occlusion of the extremities or any organ, such as kidneys, mesenteric arteries, spleen, retina or grafts), ischemic stroke, hemorrhagic stroke, stroke of uncertain classifications, major intracranial bleeding, major extracranial bleeding, major gastrointestinal (GI) bleeding, major upper GI bleeding, major lower GI bleeding, major urogenital bleeding, major other bleeding, transient ischemic attack (TIA), myocardial infarction (MI), venous thromboembolism (VTE), deep vein thrombosis (DVT), pulmonary embolism (PE).

10.4.2.3 Further outcome(s)

Further outcomes include hepatotoxicity and all-cause mortality.

This report involved UnitedHealth data without external linkage with the Social Security Death Master File, so mortality was not assessed.

10.4.3 Covariates

Numerous patient characteristics were considered, including demographics, calendar time, comorbidities, prior medications, and prior health care utilization (See Table 4 for definitions). Covariates were mostly ascertained from claims for medical services during the 6 months preceding (and including the date of) the index dispensing, with a few covariates (e.g. age, sex, and geographic location) being ascertained from enrollment information at the index date. Since some covariates may be incompletely captured within the 6 month timeframe, a sensitivity analysis was conducted with a 12 month covariate ascertainment period.

We balanced baseline covariates among patients initiating dabigatran and warfarin by matching on an exposure propensity score. [R14-1395] All covariates listed in Table 4 were included in a logistic regression propensity score (PS) model for dabigatran vs. warfarin initiators without further variable selection. Variables included in the PS were identified based on the a-priori expectation that they might be both associated with choice of anticoagulant and risk factors for study outcomes so that they represent likely confounding variables. As a secondary analysis, we employed an empirical identification of potential confounding variables using a high-dimensional propensity score (hdPS) algorithm. [R13-0525] This algorithm identifies confounding variables through an automated process

that involves estimating associations between hundreds of health insurance claims characteristics and both exposure (dabigatran or warfarin) and outcome. Since this study involves multiple outcomes, we applied two separate outcomes for the hdPS confounding assessment, one based on a composite of ischemic and embolic outcomes (labeled "composite ischemic outcome") and one based on the major hemorrhage outcome and conducted separate hdPS development and matching for each. The hdPS empirically identified confounders were combined with the investigator-identified covariates to potentially improve confounding adjustment. [R13-2767] Matching was conducted in a 1:1 fixed ratio using a nearest neighbor technique. [R14-1392] In addition, the patients were matched within calendar quarters (3 month periods).

10.4.4 Adverse events/adverse reactions

10.4.4.1 Definitions of adverse events

None

10.4.4.2 Adverse event and serious adverse event reporting

As this was a non-interventional study with secondary use of data from a US health insurance claims databases, adverse event reporting is not applicable.

10.5 DATA SOURCES AND MEASUREMENT

Exposure, outcomes, and covariates were defined using claims data. Definitions as well as information on the validity of outcome definitions are provided in tables 1- 6.

10.6 BIAS

We have compared distributions of socio-demographic, clinical and utilization characteristics between initiators of dabigatran and warfarin to assess the presence of differences that might result in confounding if not accounted for in the analysis. In adjusted analyses, we have used propensity score (PS) matching to balance potential confounders. We have tabulated individual covariates and compared these across matched cohorts. In addition, we have assessed balance across matched cohorts on the basis of empirical disease risk scores for each of the study outcomes.

In addition to using propensity scores developed on the basis of all covariates specified in this protocol, we have used high-dimensional propensity scores (hdPS) in confirmatory analyses. The hdPS algorithm evaluates thousands of diagnoses, procedures, and pharmacy claim codes to identify and prioritize those covariates that serve as proxies for unmeasured confounders [R13-0525]. These empirically identified confounders are combined with investigator-identified covariates to improve confounding adjustment. HdPS approaches have been shown to improve validity in longitudinal claims data studies, particularly when combined with pre-specified covariates.

10.7 STUDY SIZE

The study size depended on the number of dabigatran and warfarin initiators who met the study inclusion and exclusion criteria and were successfully matched into cohorts. Among patients with a minimum of 12 months baseline history and a diagnosis of non-valvular atrial fibrillation, there were 2,991 dabigatran incident users with no prior use of any oral anticoagulant matched to a similar number of warfarin incident users. These users provided 1,237 person-years (dabigatran) and 950 person-years (warfarin) of follow-up. The study size affects the precision with which this study will estimate effects (as a 95% confidence interval) so assessments of study size should focus on the width of the 95% confidence interval for each of the study outcomes.

10.8 DATA TRANSFORMATION

The source data for this study are the health insurance transactions between a large health insurer and providers of healthcare services. These transactional records are routinely conducted for reimbursement of healthcare services. The source data are transformed from raw insurance claims to a series of indicator variables (yes/no) for the presence or absence of patient characteristics (covariates) on a given date or over some period of time. During follow-up, the occurrence of a claim that is part of an outcome definition is assumed to represent the occurrence of the outcome on the date corresponding to the claim.

10.9 STATISTICAL METHODS

10.9.1 Main summary measures

Outcomes were identified during the follow-up of each cohort, and event rates (events divided by person-time) were estimated for each of the specified outcomes, and for each of the exposure rules (as treated vs. intention to treat). Relative risks (hazard ratios (HR)) and rate differences with corresponding 95% confidence intervals (CI) were estimated in the matched cohort. Kaplan-Meier curves were plotted for event-free survival as a function of the duration of use of the index anticoagulant in the matched cohorts.

10.9.2 Main statistical methods

Analyses involved description and comparison of socio-demographic, clinical and healthcare utilization characteristics among initiators of different anticoagulants. Individual covariates were tabulated before and after PS matching to assess the achieved balance. In addition, cohort balance was assessed on the basis of empirical disease risk scores (DRSs). Three DRSs were developed using data for warfarin patients from January 2009 through September 2010, one for stroke, one for the composite largely ischemic outcome (stroke, systemic embolism, TIA, MI or VTE) and one for the outcome of major hemorrhage. These DRSs were derived from predicted probabilities of outcomes estimated in logistic regression models with covariates. The comparison of mean DRS across matched cohorts is useful to assess balance with respect to a summary score rather than each individual variable, and was used for adjustment in secondary analyses. [R14-1471] Another summary assessment of covariate balance was made by determining the remaining discrimination (measured by c-statistic) between the 1:1 matched cohorts (the aim is a c-statistic close to 0.5, since a c-statistic exactly equal to 0.5 indicates perfect balance). [R14-1258]

Outcomes were identified during the follow-up of each cohort, and event rates (events divided by person-time) were estimated for each of the specified outcomes, and for each of the exposure rules (as treated vs. intention to treat). Relative risks (hazard ratios (HR)) and rate differences with corresponding 95% confidence intervals (CI) were estimated in the matched cohort. Kaplan-Meier curves were plotted for event-free survival as a function of the duration of use of the index anticoagulant in the matched cohorts.

Multiple Comparisons

In this study, with primary as well as secondary outcomes and main analyses along with secondary, exploratory and sensitivity analyses, there were many comparisons being made in the same study population. For example, there were 42 subgroup analyses for each of the two primary outcomes, 7 sensitivity analyses for all the outcomes and 3 analyses in patients with laboratory tests results available (overall for each lab, and further stratified by lab values tertiles). Conventional use of p-values in the context of such multiple comparisons may be associated with a substantially increased risk of type 1 error. [R14-1514] However, throughout this report we did not engage in formal hypothesis testing or multiplicity adjustment, and indeed no p-values are presented. Instead, this study focused on estimation of effects and quantified the precision of the effect estimates as 95% confidence intervals. Such confidence intervals can be interpreted as accurate estimates of precision even where multiple comparisons have been made. [R14-1393]

10.9.3 Missing values

Patient demographics (age and sex) may be missing in the source data, but since their presence was a criterion for study inclusion, so they were not missing for any of the study subjects.

Each of the other study variables (exposure, covariates, and outcomes) was considered present where represented by an appropriate code. Thus, lack of a code was assumed to represent non-presence of the variable rather than missing. This coding approach is appropriate for data sources that are comprehensive such as this health insurance database.

10.9.4 Sensitivity analyses

Stratified analyses. A number of stratified analyses were conducted, with stratification according to the following specified subgroups (see Table 7): 1) prior major bleed; 2) prior GI drug use; 3) duration of follow-up in 6 month blocks; 4) propensity score quintile; 5) antiplatelet use; 6) age at cohort entry (in four age categories); 7) sex; 8) dabigatran starting dose; 9) renal dysfunction; 10) hypertension; 11) diabetes; 12) atherosclerosis; 13) congestive heart failure; 14) $CHADS_2$ score (see Table 5); 15) CHA_2DS_2-VASc score; 16) HAS-BLED score (see Table 6). For each of the pre-defined subgroups, we assessed balance among the dabigatran and warfarin initiators within each stratum of the subgroup variable. This step provided an assessment of the balance achieved by propensity score matching even though this matching was not conducted within each of the subgroups.

Subgroup of patient with laboratory tests results. Since laboratory results were expected to be available for only approximately 30% of the cohorts, these laboratory results were not incorporated into the propensity score. However, the laboratory results were used to assess balance of lab test values for plausibly confounding laboratory measures among the compared cohorts and to form subgroups. [R14-1390] Specifically, we evaluated the extent to which patients who had laboratory tests results data differed from the overall study population. We tabulated patient characteristics

stratified by the exposure for the subsets of the matched cohorts that had available laboratory test result data, and stratified results (measures of association between anticoagulant use and outcomes) according to specified levels of the laboratory results to evaluate the potential for heterogeneity of effect. Matched cohort members with baseline low density lipoprotein (LDL) levels were divided into tertiles according to LDL (low, medium and high LDL), and effect measures (HR for dabigatran relative to warfarin for each outcome) overall and within each of these LDL strata were assessed. Similarly, members of the matched cohorts with baseline levels of glycosylated hemoglobin (Hb_{A1c}) formed a Hb_{A1c} subset and were divided into high, medium, and low levels according to tertiles of Hb_{A1c}, and members with baseline levels of creatinine (Cr) were divided into high, medium, and low levels according to tertiles of serum Cr.

Sensitivity analyses. Several sensitivity analyses were undertaken and included: 1) Extending the covariates assessment period from 6 months to 12 months in the primary analysis; 2) Relaxing the 'incident user' definition from 12 months without prior use of any anticoagulant to 6 months without prior use of any anticoagulant, with the same change applied to eligibility requirement (from 12 months to 6 months); 3) For patients who switch, extending the exposure risk window to 2 days after initiation of another anticoagulant; 4) For patients who discontinue the index drug, varying the exposure risk window by censoring exposure at the treatment gap ≥ 7 days or ≥ 30 days; 5) Excluding patients with a prior major GI bleed (defined as GI bleed associated with hospitalization) or intracerebral hemorrhage(ICH); 6) Censoring patients at the time of antiplatelet/another anticoagulant initiation. For each of these sensitivity analyses, comparison of dabigatran and warfarin initiators were made with results compared back to the main study results.

10.9.5 Amendments to the statistical analysis plan

10.10 QUALITY CONTROL

All aspects of the study conduct were consistent with the Good Pharmacoepidemiology Practice guidelines. [R09-0182] Programming for this project was conducted by a primary SAS analyst and validated by a separate analyst (validation analyst). For all data processing and analysis steps, the validation analyst reviewed the program along with input and output data sets, and for select steps of the project double programming techniques were employed to reduce the potential for programming errors. All analyses were conducted in SAS V9.2.

11. RESULTS

11.1 PARTICIPANTS

Dabigatran and Warfarin users

Between October 2010 and June 2012, a total of 43,768 patients had 12 months of enrollment and initiated treatment with either dabigatran or warfarin. Of these patients, there were 15,213 (34.8%) who had a diagnosis of atrial fibrillation prior to treatment initiation, and 12,219 (27.9%) who met our definition of non-valvular atrial fibrillation. A further restriction to patients with CHA_2DS_2-VASc ≥ 1 removed 337 patients for a final qualifying study size of 11,882 patients initiating treatment. Of these, 4,158 (35%) initiated with dabigatran and 7,724 (65%) initiated with warfarin (**Figure 1**). This set of eligible patients was used to develop and estimate the propensity score, and represented the pool of patients who could be included in the matching. This process resulted in 2,991 patients initiating dabigatran being matched to 2,991 patients initiating warfarin, and these patients form the main study cohorts. These small numbers underscore the emphasis on feasibility in this report.

11.2 DESCRIPTIVE DATA

Patient Characteristics

Prior to matching, the dabigatran initiators differed from the warfarin initiators on numerous variables. Dabigatran initiators at baseline were younger (average age of 62.4±10.8 years of age vs 64.7±11.5) and had generally fewer health conditions recorded (e.g., prior ischemic stroke 8.0% vs 11.3%), so that the disease risk scores tended to be lower for dabigatran initiators: mean CHA_2DS_2-VASc score was 2.5±1.5 vs 3.2±1.7, mean HAS-BLED score was 1.9±1.0 vs 2.2±1.1. (Table 8).

In contrast to the eligible patients, the matched dabigatran and warfarin initiators exhibited nearly identical characteristics to one another, such as average age (63.9 years vs. 63.3 years) and prevalence of prior ischemic stroke (8.9% vs. 8.6%), with no individual characteristic appearing out of balance.The similarities with respect to individual covariates meant that the groups were similar with respect to the clinical summary scores: both groups had a mean CHA_2DS_2-VASc score of 1.8 (±1.1) and a HAS-BLED score of 2.0 (±1.1) (Table 9). The c-statistic from the PS model was 0.805 before matching and 0.552 after matching reflecting the much improved covariate balance among the cohorts used for outcome analyses.

Other co-morbidities among the matched cohorts (dabigatran vs. warfarin) such as previous myocardial infarction (MI) (3.6% vs. 4.0%), congestive heart failure (CHF) (17.3% vs. 17.1%), diabetes (26.5% vs. 26.5%), and hypertension (95.2% vs 95.3%) were also similar.The low prevalence of aspirin use (.37% vs. .5%) in the prior 6 months reflects the fact that most aspirin use is over-the-counter and not dispensed by a pharmacy in response to a prescription (and thus not captured by this data source). Angiotensin receptor blockers (ARBs) were used by 18.4% vs. 19.5% of patients, angiotensin converting enzyme (ACE) inhibitors by 33.0% and 32.7%, beta-blockers by 67.5% and 67.9%, statins by 45.6% and 44.5%, proton pump inhibitors by 17.2% vs 16.5% and H_2 receptor antagonists by 3.7% in both groups (Table 9).

Matching with hdPS. The results from hdPS matching with respect to balance of cohorts were largely consistent with the protocol-guided propensity score matching (Table 10). There were 2,800 dabigatran initiators matched to 2,800 warfarin initiators using the hdPS algorithm that assessed potential confounding based on ischemic risk factors and 2,778 dabigatran initiators matched to 2,778 warfarin initiators using the hdPS algorithm that assessed potential confounding based on hemorrhagic risk factors. These compared to the 2,991 dabigatran and 2,991 warfarin initiators that matched without the hdPS algorithm. Covariates in the hdPS-matched cohorts were well-balanced,

with a c-statistic after matching of 0.536 with the ischemic risk factor hdPS implementation and 0.544 in the hemorrhagic risk factor hdPS implementation.

Disease Risk Scores (DRS). The development of the DRS among warfarin initiators before dabigatran became available (between January 2009 and September 2010) led to a DRS for stroke that had a c-statistic of 0.757 for identifying future stroke, while the DRS for the composite ischemic outcome had a c-statistic of 0.733 and for major hemorrhage the c-statistic was 0.691. Among the PS matched cohorts, the mean DRS for stroke was similar in the cohorts (0.23 vs. 0.18) as was the major hemorrhage DRS (0.59 vs. 0.55), and the DRS based on the composite ischemic outcome was identical (0.42 vs. 0.42) (Table 17). This balance on outcome DRSs within the matched cohorts reflects the high degree of similarity of the cohorts achieved through the propensity score matching process.

11.3 OUTCOME DATA

Follow-up. The 2,991 matched dabigatran patients (96% with the 150mg dose) had 1,237 person-years of follow-up (mean 0.41 years, median 0.28) while the 2,991 warfarin patients had 950 person-years of follow-up (mean 0.32 years, median 0.21).This difference in follow-up time within cohorts that are identical in size and where start of follow-up was balanced within calendar quarter was largely due to differences in the way the two treatments are used, with more frequent discontinuation of treatment among warfarin initiators leading to shorter average follow-up in the warfarin cohort.

Many patients end follow-up for administrative reasons (end of enrollment in United healthcare or end of June 2012), so that the average follow-up is less than half a year, which substantially limits comparative safety and effectiveness assessments. The large majority of patients ended follow-up due to treatment discontinuation (65% among dabigatran and 73% among warfarin) and many fewer patients (29% among dabigatran and 20% among warfarin) ended follow-up because they reached the end of the study (Table 12).

11.4 MAIN RESULTS

Outcomes

As a consequence of the small sample size and the limited follow-up in this feasibility stage of the project, outcomes were infrequent so that confidence intervals are consistent with a wide range of hypotheses. Incidence rates (IR) are provided as events per 1000 person-years.

As-treated analyses. Among the 2,991 matched dabigatran patients (96% with the 150mg dose) providing 1,237 person-years of follow-up and the 2,991 warfarin patients providing 950 person-years of follow-up, there were 36 strokes (IR = 29.09) among dabigatran users vs. 30 (IR = 31.59) among warfarin users (HR=1.05, 95% CI=0.64-1.70). With slightly different person-years follow-up (1,233 dabigatran vs. 944 warfarin), there were 74 major hemorrhages (IR = 60.00) among dabigatran users vs. 63 (IR = 66.72) in warfarin users (HR=0.97, 0.69-1.36). (Table 11) For the outcome of stroke or systemic embolism (same person-time as with major hemorrhage), there were 49 events (IR = 39.73) among dabigatran users vs. 57 events (IR = 60.38) among warfarin users (HR= 0.74, 95% CI=0.50-1.08, while for systemic embolism including PE, there were 1,242 person-years of dabigatran exposure and 949 person-years of warfarin exposure during which there were 14 events (IR = 11.27) vs. 29 events (IR = 30.55) (HR=0.40, 95% CI=0.21-0.76), with much of this effect due to PE. For ischemic stroke (1,240 person-years vs. 947 person-years) there were 34 events (IR = 27.40) vs. 31 events (IR = 32.71) (HR=0.95, 95% CI =0.58-1.55). There were only 8 hemorrhagic strokes in total (5 among dabigatran treated patients and 3 among warfarin treated patients). Other outcomes along with corresponding person-time and effect measures are listed in Table 11. The time to event curves for each of the outcomes are presented (Figures 2a-s). Visual inspection of some of these curves suggest

non-proportional hazards, but the small numbers of events limits formal assessment of this. However, estimates obtained from proportional hazards regression are informative even with violations in this assumption. The time-course of events also leads to the divergence of the hazard ratio from a straight-forward comparison of rates (as an incidence rate ratio). [R14-1400]

Intention to treat analyses. These analyses carried initial exposure forward for a period of up to 1 year and included more person-time and more events along with less certainty that the person-time and events were exposed to the medication (Table 14). The person-time in these analyses was more evenly balanced between the dabigatran and warfarin cohorts since treatment discontinuation or switching (more frequent among warfarin) is ignored. There were 52 strokes and 2,181 person-years among dabigatran initiators compared to 49 strokes and 2,163 person-years among the warfarin initiators (HR = 1.05, 95% CI = 0.71-1.55). There were 110 vs 121 major hemorrhages (HR = 0.90, 95% CI 0.70-1.17). For the outcome of stroke or systemic embolism, there were 75 events among dabigatran initiators vs 94 events among warfarin (HR = 0.79, 95% CI 0.58-1.07). For the outcome of systemic embolism alone (including PE), there were 27 vs 49 events (HR = 0.54, 95% CI 0.34-0.87), while for ischemic stroke, there were 50 vs 45 events (HR = 1.1, 95% CI 0.74-1.65). There were 15 hemorrhagic strokes (7 dabigatran and 8 warfarin) leading to wide confidence intervals (HR = 0.86, 95% CI 0.31-2.38).

HdPS-matched analyses. These results are largely consistent with the analyses based on the investigator-defined covariates (Tables 15 and 16). The hdPS matched cohorts had an association with stroke (HR = 1.06, 95% CI = 0.66-1.69) that was closely similar to the results from the protocol-driven propensity score matched cohorts (HR=1.05, 95% CI=0.64-1.70). For the outcome of major hemorrhage, the HdPS matched result (HR = 0.90, 95% CI 0.64-1.27) was also quite similar to the protocol-driven propensity score matched cohorts (HR=0.97, 0.69-1.36). This similarity of results provides reassurance that empiric identification of potential confounding variables does not substantially alter the results, reflecting the thoroughness of the original covariate identification.

DRS-adjusted analyses. Although there were some differences in the DRSs noted within the matched cohorts, adjustment for the DRS leads to only minor change in the study effect measures (Table 18). Accordingly, the small differences in DRS across compared groups does not represent or indicate the presence of meaningful confounding, a finding that is consistent with the observation of no substantial imbalance on either individual variables or the summary balance metrics (c-statistic).

11.5 OTHER ANALYSES

Sensitivity and Exploratory Analyses

Stratified analyses

These analyses were conducted for the primary outcomes of major hemorrhage and stroke. Baseline covariate balance was assessed in each subgroup (see Appendix Excel file) and based on this assessment, outcome models were further adjusted for CHA_2DS_2-VASc score and HAS-BLED score as these scores summarize the risk factors for both the ischemic events and hemorrhage. Kaplan-Meier curves were plotted for cumulative incidence of events in each subgroup (Figures 3-4) and results for these analyses were presented in tabular form and graphically (Tables 19-20 and Figures 5-6). The subgroup based on follow-up stratifies the follow-up into 6-month blocks of time. Results suggested little or no treatment effect heterogeneity for either stroke or major hemorrhage across any of the variables evaluated. Stratification by quintiles of propensity score provides the effect estimate

across people who are empirically more typical dabigatran initiators (quintile 5) and people who are less typical dabigatran initiators (quintile 1). The analysis stratifying follow-up time into 6-month blocks is a stratification based on a post-index characteristic rather than a baseline characteristic, and one that aids interpretation in the setting of non-proportional hazards. [R14-1400] It should be noted, however, that given the small size of the subgroups resulting in very imprecise estimates, these analyses had little capacity to identify even meaningful heterogeneity. However, these analyses will become more meaningful as more data accrue in the ongoing portion of this study.

Subgroup of patients with laboratory test results

Patients with laboratory test results were compared to patients who did not have laboratory test results available in both the unmatched and matched cohorts (Tables 21 and 22). Among the matched cohorts, patients with laboratory test results available were largely similar to those without laboratory test results, although those with labs tended to be younger (60 years vs. 65 years), more likely to be male (71% vs 69%), reside in the south (63% vs. 36%) and have a diagnosis of diabetes (33% vs 24%) or hyperlipidemia (63% vs. 53%). Other characteristics were quite similar between the full matched cohort and the subset of the matched cohort with laboratory test results available.

Among the matched cohorts, LDL lab test results were available for 18.3% of warfarin patients and 19.6% of dabigatran patients, Hb_{A1c} results were available for 10.1% of warfarin patients and 10.5% of dabigatran patients, and creatinine test results were available for 23.6% of warfarin patients and 25.8% of dabigatran patients. The mean values for the last laboratory test prior to treatment initiation were quite similar for warfarin and dabigatran patients in these matched cohorts: mean LDL of 91.4 mg/dl (median: 90, interquartile range: 68 – 137) for warfarin vs. 91.2 mg/dl (median: 88, interquartile range: 69 – 109) for dabigatran; mean Hb_{A1c} of 10.1% (median: 6.3%, interquartile range 5.8%-7.4%) for warfarin vs. 10.5 (median: 6.3%, interquartile range 5.8%-7.3%) for dabigatran; and mean serum creatinine of 1.0 mg/dl (median: 0.95, interquartile range: 0.81-1.14) for warfarin vs. 1.0 mg/dl (median: 0.98, interquartile range 0.84-1.14) for dabigatran (Table 23). These findings point to the thoroughness of the propensity score matching: even though laboratory values were not explicitly included in the variables that entered into the propensity score, the similarity of laboratory test results among warfarin and dabigatran initiators in the matched groups suggests that sufficient proxies for laboratory test results were represented among the variables in the propensity score to achieve balance on the test results. The similarity in these three laboratory tests achieved by propensity score matching indicates that residual confounding by, for example, severity of diabetes based on Hb_{A1c} or degree of hyperlipidemia is unlikely to be an important confounding factor in this study, and by implication, residual confounding by other patient characteristics that might not be explicitly captured by insurance claims data becomes less plausible.

Subgroup analyses examining the association between anticoagulant (dabigatran vs. warfarin) and primary outcomes of stroke and major hemorrhage within strata defined by laboratory results (tertiles) are presented in Tables 24 and 25. The results were largely consistent across strata, and although some strata exhibit numerically different associations, wide confidence intervals illustrate the potential of chance to account for the differences.

Sensitivity Analyses

In the framework of feasibility assessment that this analysis represents, many sensitivity analyses have been conducted.

Assessing the covariates over a period of 12 months. When we rebuild study cohorts after ascertaining covariates over a period of 12 months rather than the 6 month period used in the main analyses, there were 3,020 dabigatran initiators that matched to 3,020 warfarin initiators, and their characteristics were well balanced (Table 26). Within the cohorts matched on the basis of 12 month covariates (Table

27), there were 36 strokes among 1,248 dabigatran person-years and 32 strokes among 966 warfarin person-years (HR = 0.99, 95% CI = 0.61-1.60), a result that is quite similar to that based on covariates ascertained over 6 months (HR=1.05, 95% CI=0.64-1.70). For the other primary outcome, there were 73 major bleeds among 1,246 dabigatran person-years and 65 major bleeds among 964 warfarin person-years (HR = 0.93, 95% CI 0.66-1.30), also quite similar to the result based on covariates ascertained over 6 months (HR=0.97, 0.69-1.36). These results suggest that the 6 month covariate assessment period may not be improved by extending to 12 months, but this conclusion is limited by small numbers of outcomes.

Shorter baseline exposure assessment. When we relaxed the 'new user' definition to 6 months of no use prior to treatment initiation (as opposed to 12 months) and 6 months of prior enrollment in the United healthcare database, additional patients were eligible, so that following a similar process of building study cohorts (propensity score development and matching) produced larger matched cohorts comprised of 3,890 dabigatran and 3,890 warfarin initiators compared to 2,991 of each under the 12 month baseline criterion. These matched cohorts exhibited similar good balance with respect to patient characteristics (Table 28). For the outcome of stroke, there were 47 events among 1,597 dabigatran person-years and 39 events among 1,211 warfarin person years (HR = 1.06, 95% CI = 0.69-1.63), a result that is quite similar to the result based on cohorts with a 12 month baseline requirement (HR=1.05, 95% CI=0.64-1.70). For major hemorrhage, there were 92 events among 1,594 dabigatran person-years and 75 events among 1,205 warfarin person-years (HR = 1.02, 95% CI = 0.79-1.39), also quite similar to the result based on cohorts with a 12 month baseline (HR=0.97, 0.69-1.36). No meaningful differences were observed for any of the study outcomes between the base cohort and this expanded cohort (Table 29).However, the limited size of the cohorts and small numbers outcomes limit the ability to draw conclusions at this stage.

Extending the medication switch grace period by 2 days. The main primary analysis discontinues follow-up immediately upon dispensing of a different anticoagulant than the one that establishes cohort eligibility, and this sensitivity analysis extends the original medication exposure for 2 days after such a switch. Results for the primary outcomes were identical under this altered assumption, which suggests our main results are not affected by premature discontinuation of exposure at the time of a switch (Table 30).

Reducing the medication exposure grace period from 14 days to 7 days. The primary analysis extends dabigatran and warfarin exposure for 14 days beyond the end of days supplied, and this sensitivity analysis reduced this grace period to 7 days. The 2,991 matched dabigatran patients had 968 person-years of follow-up for the stroke outcome while the 2,991 warfarin patients had 761 person-years of follow-up. Results did change slightly under this 7 day grace period for stroke (HR = 1.17, 95% CI = 0.67-2.04) compared to (HR=1.05, 95% CI=0.64-1.70) under the 14 day grace period. Also, major hemorrhage changed to HR = 1.06 (95% CI = 0.74-1.54), compared to HR=0.97 (95% CI = 0.69-1.36) under the 14 day grace period (Table 31). However, these minor differences are consistent with chance variation and likely not suggestive of a systematic mis-attribution of exposure, supporting the primary exposure definition (14-day exposure grace period). Accumulation of larger numbers and longer follow-up will permit a more clear assessment.

Increasing the medication exposure grace period from 14 days to 30 days. The primary analysis extends dabigatran and warfarin exposure for 14 days beyond the end of days supplied, and this sensitivity analysis increases this grace period to 30 days. The 2,991 matched dabigatran patients had 1,536 person-years of follow-up for the stroke outcome while the 2,991 warfarin patients had 1,285 person-years of follow-up. Results only change slightly under this 30 day grace period for stroke (HR

= 1.06, 95% CI = 0.68-1.65) compared to (HR=1.05, 95% CI=0.64-1.70) under the 14 day grace period. Major hemorrhage only changed with respect to precision (HR = 0.97, 95% CI = 0.72-1.32), compared to (HR=0.97, 0.69-1.36) under the 14 day grace period (Table 32). As with the shorter grace period, these negligible differences under the longer grace period are consistent with chance variation and likely not suggestive of a systematic mis-attribution of exposure, providing further support of the primary exposure definition and the 14-day grace period.

Subset with no prior GI hemorrhage or ICH. These analyses apply an exclusion to patients with a baseline diagnosis of GI hemorrhage or ICH. After the exclusion, the matched cohort consisted of 2,938 dabigatran and 2,938 warfarin initiators (Table 33). Among the cohorts after this exclusion, the results do change slightly for some outcomes (Table 34). For the outcome of stroke, after the exclusion HR = 0.79 (95% CI = 0.49-1.26) compared to HR = 1.05 (95% CI = 0.64-1.70) without the exclusion. For the other primary outcome of major hemorrhage, the result after exclusion (HR = 0.91, 95% CI 0.65-1.28) is quite similar to the result (HR = 0.97, 0.69-1.36) without the exclusion. The potential for this criterion (no prior GI hemorrhage or ICH) to change results will continue to be evaluated as this project continues.

Censor at time of antiplatelet initiation. These analyses censor patients from either the dabigatran or matched cohorts when an antiplatelet medication is initiated during follow-up. The results are essentially unchanged with this additional censoring criterion (Table 35). For the outcome of stroke, after the censoring the HR was 1.06 (95% CI = 0.64-1.76) compared to HR = 1.05 (95% CI = 0.64-1.70) without the censoring. For the other primary outcome of major hemorrhage, the result after censoring (HR = 0.97, 95% CI 0.68-1.38) was quite similar to the result (HR = 0.97, 0.69-1.36) without the censoring. The potential for this censoring to alter results will continue to be evaluated as this project continues.

11.6 ADVERSE EVENTS/ADVERSE REACTIONS

Not Applicable

12. DISCUSSION

12.1 KEY RESULTS

The key results are described in Section 11

12.2 LIMITATIONS

As an observational study, there are inherent limitations with respect to potential for alternate explanations for any observed association. Exposure to dabigatran and warfarin is assumed based on dates of pharmacy dispensing and expected days of use rather than a direct measure of patient consumption. Study outcomes are based on claims for reimbursement of healthcare services, and even though they were defined on the basis of codes that are highly suggestive of the study outcomes and occurred within a comprehensive health insurance database that records any billable medical service, misclassification of outcomes is possible. Where available, we used validated outcome algorithms that have known performance characteristics (such as sensitivity, specificity, and PPV) in order to reduce the extent of misclassification. Further, the serious nature of the study outcomes reduces the potential for differential surveillance between cohorts to lead to spurious findings.

Some covariates or study entry criteria may have been incompletely captured such as duration of atrial fibrillation, valvular etiology, severity of comorbidities and other conditions (such as smoking, obesity). The incomplete capture of these characteristics could lead to confounding even within the propensity score matched cohorts, although the evaluation of balance with respect to laboratory tests suggests only limited potential for such confounding. Also reasons for discontinutation and switching are incompletely captured in the database and may lead to incorrect inferences. The linkage to electronic medical record data that is planned as part of this research program will provide in the future an assessment of the potential for these and other unmeasured patient characteristics to differ between dabigatran and warfarin patients and thereby confound the observed associations.

Although duration of atrial fibrillation may represent a risk factor for study outcomes, this covariate will be incompletely captured since the patient history in the dataset is relatively short (at least 6 months, and an average of approximately 2 years), and a first claim within the database may not represent atrial fibrillation onset since the condition is typically not diagnosed at its onset.

There are a number of covariates, such as the type of atrial fibrillation, cause-specific mortality, BMI, smoking, that are not directly assessed in a health insurer database. Further, some covariates that can be directly assessed through diagnosis and procedure codes (such as renal dysfunction) have uncertain sensitivity, specificity, and predictive value. The modified MDRD equation used for creatinine clearance estimation has limitations, and is only used for the assessment of balance between exposure cohorts.

Medication use in United Healthcare data is restricted to prescription drug medication. Consequently, the use of over-the-counter (OTC) medications (e.g., OTC aspirin) is not captured. In addition, exposure is assessed based on prescription pick-up at a pharmacy and might be misclassified in “as treated” analysis if patients do not take their medications as expected.

One of the main limitations of this study at this point is the relatively small study size and short follow-up period, which is a function of the data source and timeframe over which the study is conducted. Accordingly, circumspection in interpretation of results is appropriate. The continuation of this research project will lead to expansion of the data source by inclusion of an additional data resource and by extending the time over which cohorts are accrued.

12.3 INTERPRETATION

This initial cohort study includes matched cohorts of 2,991 patients with non-valvular atrial fibrillation who initiated dabigatran and 2,991 similar patients who initiated warfarin that were drawn from a pool of 11,882 eligible initiators (4,158 initiators of dabigatran and 7,724 initiators of warfarin).

This initial analysis illustrates the successful implementation and feasibility of a PS-matched new-user design to assess the safety of a newly marketed product and forms the foundation for a long-term ongoing follow-up post-marketing drug surveillance program.

Before matching, patients who initiated dabigatran were generally healthier than patients who initiated warfarin, but the propensity score matching achieved close balance between the matched cohorts including balance on selected lab test results. Analyses suggested similarity of primary outcome occurrence between dabigatran and warfarin for stroke (HR = 1.05, 95% CI = 0.64-1.70) and major hemorrhage (HR = 0.97, 0.69-1.36) along with other effectiveness and safety outcomes, across a range of assumptions regarding exposure and within numerous subgroups and sensitivity analyses. Patients initiated on dabigatran appeared to be slightly more persistent with treatment, so this study reflects more person-time exposed to dabigatran than warfarin. These current feasibility data are

limited by short follow up and few outcome events resulting in wide 95% confidence intervals. At this early stage, no comparative conclusions are possible. Future data in the context of the subsequent protocol (1160.207) will increase the number of patients and also expand to additional data sources.

Numerous analyses were conducted to support the main study results and the full set of analyses is summarized (Table 36). However, as described above those results across subgroups and exploratory analyses are also limited by the small numbers, short follow-up period and limited study power at this stage. The future data with increased number of patients will also further support subgroup analyses.

The consistency of study results over 6-month periods during follow-up suggests that the proportional hazards assumption is reasonable in this study and supports the use of proportional hazards regression for estimation of hazard ratios; however, this finding is based on limited numbers of patients and remains inconclusive.

12.4 GENERALISABILITY

Given the characteristics of the UnitedHealth data (commercial health insurer), patients older than 65 years are under-represented in the data. As NVAF predominately affects older patients, some of the target population is not available from the data source. This limitation can be mitigated by including an additional data source in the future analyses.

The overall findings in this analysis are consistent with analyses that focused on subgroups of patients defined with respect to numerous variables, suggesting little potential for treatment effect heterogeneity. This result argues for generalizability of the study findings even though the study population is derived from a commercially insured data source that may differ from the general population on several characteristics. For example, the commercial health insurer from which the data are derived are largely employed people, so there is an under-representation of patients older than 65 years when many people transfer their health coverage to Medicare. As a result, even though we observed that 40% of incident users were age 65 or older, a data source with an age distribution matching the general population would likely observe a higher percentage of elderly. However, unless there is treatment effect heterogeneity by age, a lower prevalence of older age people will lead to study results that are comparable to those that would be obtained from a study conducted in a general population data source. [R14-1389]

13. OTHER INFORMATION

None

14. CONCLUSION

Analyses suggested similarity of primary outcome occurrence between dabigatran and warfarin for stroke (HR = 1.05, 95% CI = 0.64-1.70) and major hemorrhage (HR = 0.97, 0.69-1.36) along with other effectiveness and safety outcomes, across a range of assumptions regarding exposure and within numerous subgroups and sensitivity analyses. Patients initiated on dabigatran appeared to be slightly more persistent with treatment, so this study reflects more person-time exposed to dabigatran than warfarin. These current feasibility data are limited by short follow up and few outcome events resulting in wide 95% confidence intervals. At this early stage, no comparative conclusions are possible. Future data in the context of the subsequent protocol (1160.207) will increase the number of patients and also expand to additional data sources.

15. REFERENCES

15.1 PUBLISHED REFERENCES

P08-02098	Myers RP, Leung Y, Shaheen AAM, Li B. Validation of ICD-9-CM/ICD-10 coding algorithms for the identification of patients with acetaminophen overdose and hepatotoxicity using administrative data. BMC Health Serv Res 7, 159 (2007)
P09-11669	Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. The New England journal of medicine 2009;361:1139-51.
P10-12855	Connolly SJ, Ezekowitz MD, Yusuf S, Reilly PA, Wallentin L, Randomized Evaluation of Long-Term Anticoagulation Therapy I. Newly identified events in the RE-LY trial. The New England journal of medicine 2010;363:1875-6.
P12-00343	Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. Clinical pharmacology and therapeutics 2011;90:777-90.
R05-0358	Cushman M, Tsai AW, White RH, Heckbert SR, Rosamond WD, Enright P, Folsom AR. Deep vein thrombosis and pulmonary embolism in two cohorts: the longitudinal investigation of thromboembolism etiology. Am J Med 117, 19 - 25 (2004)
R09-0182	ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Pharmacoepidemiology and drug safety 2008;17:200-8.
R11-2274	Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. Pharmacoepidemiol Drug Saf, (2011)
R11-4190	Granger CB, Alexander JH, McMurray JJ, et al. Apixaban versus warfarin in patients with atrial fibrillation. The New England journal of medicine 2011;365:981-92.
R11-4223	Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. The New England journal of medicine 2011;365:883-91.
R11-4316	Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avron J, Solomon DH. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. Am Heart J 148 (1), 99 - 104 (2004)

R11-4884	Wahl PM, Terrell DR, George JN, Rodgers KJ, Uhl L, Cataland S, Bohn RL. Validation of claims-based diagnostic codes for idiopathic thrombotic thrombocytopenic purpura in a commercially-insured population. Thromb Haemost 103, 1203 - 1209 (2010)
R12-3688	Andrade SE, Kahler KH, Frech F, Chan KA. Methods for evaluation of medication adherence and persistence using automated databases. Pharmacoepidemiol Drug Saf 15, 565 - 574 (2006)
R13-0519	Benesch C, Witter DM, Wilder AL, Duncan PW, Samsa GP, Matchar DB Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. Neurology 49 (3), 660 - 664 (1997)
R13-0520	Kokotailo RA, Hill MD Coding of stroke and stroke risk factors using International Classification of Diseases, revisions 9 and 10. Stroke 36 (8), 1776 - 1781 (2005)
R13-0522	Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. Pharmacoepidemiol Drug Saf 21 (Suppl 1), 154 - 162 (2012)
R13-0524	Hansen B. The prognostic analogue of the propensity score. Biometrika 2008;95:481-8.
R13-0525	Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 2009;20:512-22.
R13-2767	Rassen JA, Glynn RJ, Brookhart MA, Schneeweiss S. Covariate selection in high-dimensional propensity score analyses of treatment effects in small samples. American journal of epidemiology 2011;173:1404-13.
R14-1258	Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. Statistics in medicine 2013.
R14-1389	Weisberg HI, Hayden VC, Pontes VP. Selection criteria and generalizability within the counterfactual framework: explaining the paradox of antidepressant-induced suicidality? Clinical trials 2009;6:109-18.
R14-1390	Eng PM, Seeger JD, Loughlin J, Clifford CR, Mentor S, Walker AM. Supplementary data collection with case-cohort analysis to address potential confounding in a cohort study of thromboembolism in oral contraceptive initiators matched on claims-based propensity scores. Pharmacoepidemiology and drug safety 2008;17:297-305.
R14-1391	Hernan MA, Hernandez-Diaz S. Beyond the intention-to-treat in comparative effectiveness research. Clinical trials 2012;9:48-55.

R14-1392	Rassen JA, Shelat AA, Franklin JM, Glynn RJ, Solomon DH, Schneeweiss S. Matching by propensity score in cohort studies with three treatment groups. Epidemiology 2013;24:401-9.
R14-1393	Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1:43-6.
R14-1394	Seeger JDD, G.W. Commercial insurance databases. In: Strom BLK, S.E.; Hennessy, S., ed. Pharmacoepidemiology. Philadelphia: Wiley&Sons; 2012.
R14-1395	Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. Journal of clinical epidemiology 2006;59:437-47.
R14-1400	Hernan MA. The hazards of hazard ratios. Epidemiology 2010;21:13-5.
R14-1471	Gagne JJ, Bykov K, Willke RJ, Kahler KH, Subedi P, Schneeweiss S. Treatment dynamics of newly marketed drugs and implications for comparative effectiveness research. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2013;16:1054-62.
R14-1514	Bailar JCM, F. Medical Uses of Statistics. 2nd ed. Boston: NEJM Books; 1992.

15.2 UNPUBLISHED REFERENCES

Not applicable