CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

206316Orig1Orig2s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

Secondary review

NDA/BLA #:	NDA 206316
Serial #:	0000
Drug Name:	Edoxaban
Indication(s):	Deep vein thrombosis (DVT) and pulmonary embolism (PE)
Applicant:	Daiichi Sankyo Pharma Development
Date(s):	Submission Date: 8 January, 2014 PDUFA due Date: 8 January 2015 Review Completion Date: 6 January 2015
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V
Primary Statistical Reviewer:	Yun Wang, PhD
Secondary Reviewer	Lei Nie, PhD.
Concurring Reviewer:	Rajeshwari Sridhara, PhD, Division Director
Medical Division:	Office of Hematology and Oncology Product
Clinical Team:	Saleh Ayache, MD Kathy Robie Suh, MD, PhD, Team Leader
Project Manager:	Janet Higgins, RPM

Keywords: Deep Vein Thrombosis, Pulmonary Embolism, Venous thrombosis embolism, Non-inferiority, Subgroup analysis, Dose selection.

In this New Drug Application submission, the applicant seeks full approval of Edoxaban for the following two indications:

1) To reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation (AF)
2) For the treatment of deep vein thrombosis (DVT) and pulmonary embolism (PE) following 5-10 days of initial therapy with a parenteral anticoagulant

The support of Edoxaban for the second indication, treatment of DVT and PE, was based on trial Hokusai VTE, a Phase III, randomized, multi-center, double-blind, parallel-group study in subjects with acute symptomatic VTE comparing Edoxaban with Warfarin. The primary efficacy objective of Hokusai VTE was to evaluate whether Edoxaban is non-inferior to Warfarin. This trial randomized 8292 patients, 4143 to Edoxaban arm and 4149 to Warfarin arm respectively. Non-inferiority was demonstrated in the primary efficacy endpoint, time to symptomatic recurrent VTE or VTE-related death, with the estimated hazard ratio of 0.89 and a 95% confidence interval of (0.70 , 1.13). The upper 95% confidence limit of 1.13 demonstrated that treatment with Edoxaban retained more than 90% treatment effect of Warfarin. The protocol specified primary analysis is based on non-inferiority margin of 1.5. The estimated HR for the pre-specified primary safety endpoint, time to major or clinically relevant non-major bleeding (CRNM), was 0.81 (95% confidence interval: 0.71-0.94) based on 772 major/CRNM events. For further details regarding the design, data analyses, and results of trial Hokusai VTE, please refer to the statistical review by Dr. Yun Wang. This team leader concurs with Dr. Wang's recommendations and conclusions.

In this secondary review, two additional issues are addressed.

First, the sponsor has proposed (b) (4)

(b) (4)

Second, while patients with creatinine clearance (CrCL) 30 to 50 mL/min or body weight ≤60 kg or who used certain P-glycoprotein (P-gp) inhibitors received Edoxaban 30 mg in both pivotal trials: Hokusai VTE for DVT and PE and trial DU176B-C-U301 (also called ENGAGE AF-TIMI 48) for AF, Edoxaban 60 mg is under consideration for all AF patients with CrCL ≤ 95 mL/min. This review documents our rationale for recommending Edoxaban 30 mg for patients

with CrCL 30 to 50 mL/min or body weight ≤60 kg or who use P-gp inhibitors for patients with DVT or PE, and 60 mg per day for all other patients with DVT or PE.

According to the sponsor, the Edoxaban dosage regimen adjustments in Phase 3 trials were pre-determined by the pharmacometric analysis of extrinsic and intrinsic factors affecting the exposure of Edoxaban relative to bleeding. P-gp inhibitors and renal impairment (CrCL≤ 50 mL/min) were identified as factors which resulted in clinically relevant increased exposure, and low body weight (≤ 60 kg) was associated with a higher incidence of bleeding. Therefore, subjects with these risk factors were to receive half of the randomized Edoxaban dosage regimen.

In addition, an early Japanese study in subjects with AF also showed that the incidence of overall bleeding in the Edoxaban 30 mg QD group was similar to that of the Warfarin control group, and the incidence of overall bleeding in the Edoxaban 60 mg QD group was higher than that of the Warfarin control group. Edoxaban-treated Japanese subjects with body weight ≤ 60 kg had a significantly higher incidence of bleeding than those with body weight > 60 kg. Subjects with body weight > 60 kg in the 60 mg QD group showed a similar incidence of bleeding to those in the Warfarin group.

The results from the trial Hokusai VTE with the pre-specified dose regimen demonstrated non-inferiority in the primary efficacy endpoint and superiority in the primary safety endpoint. Our exploratory subgroup analyses also indicate that Edoxaban's efficacy is similar to Warfarin in patients with CrCL 30 to 50 mL/min or body weight ≤ 60 kg and Edoxaban's safety is superior to Warfarin in this subgroup. See Tables 1 and 2, as well as the statistical review by Dr. Yun Wang for more details.

Table 1: Subgroup analysis of primary efficacy endpoint by Edoxaban/Placebo dose

Edoxaban/Placebo dose level	Edoxaban N=4118	Warfarin N=4122	HR (95% CI)
30mg, n/N (%)	22/733 (3.0)	30/719 (4.2)	0.73 (0.42, 1.26)
60mg, n/N (%)	108/3385 (3.2)	116/3403 (3.4)	0.93 (0.72, 1.21)

- Primary efficacy endpoint: time to VTE or VTE-related death.

Table 2: Subgroup analysis of primary safety endpoint by Edoxaban/Placebo dose

Edoxaban/Placebo dose level	Edoxaban N=4118	Warfarin N=4122	HR (95% CI)
30mg, n/N (%)	58/733 (7.9)	92/719 (12.8)	0.62 (0.44, 0.86)
60mg, n/N (%)	291/3385 (8.6)	331/3403 (9.7)	0.87 (0.74, 1.02)

- Primary safety endpoint: Time to major bleeding or clinically relevant non-major bleeding.

We also performed the following exploratory analyses (Tables 3 and 4) to assess the efficacy and safety of Edoxaban 30 mg in subjects who are ≤ 60 kg or CrCL ≤ 50 mL/min. The analyses indicate that subjects with weight ≤ 60 kg or CrCL ≤ 50 mL/min and received Edoxaban dose of 30 mg daily had similar efficacy and safety as those patients who are > 60 kg and CrCL > 50 mL/min and received Edoxaban dose of 60 mg daily. Subgroup analysis by concomitant use of P-gp inhibitors was not preformed because only 51 (0.6%) of patients received concomitant P-gp inhibitor.

Table 3: Primary Efficacy Analyses by subgroups

Subgroups	Edoxaban	Warfarin	HR (95% CI)
CrCL Level, n/N (%)			
30 – 50 mL/min	8/268 (3.0)	16/273 (5.9)	0.50 (0.21, 1.17)
> 50 ml/min	122/3850 (3.2)	130/3849 (3.4)	0.94 (0.73, 1.20)
Weight, n/N (%)			
≤ 60 kg	15/524 (2.9)	18/519 (3.5)	0.84 (0.43, 1.68)
> 60 kg	115/3594 (3.2)	128/3603 (3.6)	0.90 (0.70, 1.16)

- Primary efficacy endpoint: time to VTE or VTE-related death.

Table 4: Primary Safety Analyses by subgroups

Subgroups	Edoxaban	Warfarin	HR (95% CI)
CrCL Level, n/N (%)			
30 – 50 mL/min	28/268 (10.5)	39/273 (14.3)	0.71 (0.44, 1.15)
> 50 ml/min	321/3850 (8.3)	384/3849 (10.0)	0.82 (0.71, 0.96)
Weight, n/N (%)			
≤ 60 kg	39/524 (7.4)	64/519 (12.3)	0.60 (0.40, 0.89)
> 60 kg	310/3594 (8.6)	359/3603 (10.0)	0.85 (0.73, 0.99)

- Primary safety endpoint: Time to major bleeding or clinically relevant non-major bleeding.

The additional exploratory analyses (Tables 5 and 6) suggest that subjects with weight ≤ 60 kg and CrCL > 50 mL/min, who received Edoxaban dose of 30 mg daily primarily due to low weight, had similar or comparable efficacy and safety as patients in other categories. The results support that weight alone is a factor of determining Edoxaban dose of 30 mg or 60 mg and subjects with body weight ≤ 60 kg should receive Edoxaban 30mg regardless their CrCL level.

Table 5: Subgroup analysis of primary efficacy endpoint by patients' weight and CrCL at baseline

Subgroups	**Edoxaban N=4118**	**Warfarin N=4122**	**HR (95% CI)**
Weight ≤ 60 kg, CrCL ≤ 50 mL/min, n/N (%)	2/81 (2.5)	4/91(4.4)	0.58 (0.11, 3.16)
Weight ≤ 60 kg, CrCL > 50 mL/min, n/N (%)	13/443 (2.9)	14/428(3.3)	0.92 (0.43, 1.95)
Weight > 60 kg, CrCL ≤ 50 mL/min, n/N (%)	6/187 (3.2)	12/182 (6.6)	0.47 (0.18, 1.26)
Weight > 60 kg, CrCL > 50 mL/min, n/N (%)	109/3407 (3.2)	116/3421 (3.4)	0.94 (0.73, 1.22)

- Primary efficacy endpoint: time to VTE or VTE-related death.

Table 6: Subgroup analysis of primary safety endpoint by patients' weight and CrCL at baseline

Subgroups	**Edoxaban N=4118**	**Warfarin N=4122**	**HR (95% CI)**
Weight ≤ 60 kg, CrCL ≤ 50 mL/min, n/N (%)	10/81 (12.35)	13/91(14.3)	0.78 (0.34, 1.77)
Weight ≤ 60 kg, CrCL > 50 mL/min, n/N (%)	29/443 (6.5)	51/428(11.9)	0.56 (0.35, 0.88)
Weight > 60 kg, CrCL ≤ 50 mL/min, n/N (%)	18/187 (9.6)	26/182 (14.3)	0.68 (0.37, 1.24)
Weight > 60 kg, CrCL > 50 mL/min, n/N (%)	292/3407 (8.6)	333/3421 (9.7)	0.86 (0.74, 1.01)

- Primary safety endpoint: Time to major bleeding or clinically relevant non-major bleeding.

Finally, we note that no clinical trials were conducted to evaluate Edoxaban 60 mg in patients with DVT or PE who had CrCL 30 to 50 mL/min or body weight ≤60 kg or who use P-gp inhibitors at baseline.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

LEI NIE
01/07/2015

RAJESHWARI SRIDHARA
01/07/2015

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

DIVISION DIRECTOR MEMO

NDA /Serial Number:	206316/N_000
Drug Name:	Edoxaban
Applicant:	Daiichi Sankyo Pharma Development
Indication(s):	Deep vein thrombosis (DVT) and pulmonary embolism (PE)
Date(s):	Submission Date: January 8, 2014 PDUFA Date: January 8, 2015
Review Priority:	Standard Review
Biometrics Division:	Division of Biometrics V
Statistical Team :	Rajeshwari Sridhara, Ph.D., Division Director Lei Nie, Ph.D., Team Leader Yun Wang, Ph.D. Primary Reviewer
Medical Division:	Division of Hematology Products
Clinical Team:	Saleh Ayache, M.D. Kathy Robie Suh, M.D., Ph.D.
Project Manager:	Janet Higgins
Keywords:	Time to event, Non-inferiority, Subgroup analyses, DVT, PE, VTE

Introduction

This is the Division Director's memo in addition to the primary statistical review by Dr. Yun Wang and secondary team leader memo by Dr. Lei Nie. I concur with the conclusions and recommendations of both Drs. Wang and Nie. Additional comments are presented here to clarify our approach in labeling the product for the treatment of deep vein thrombosis (DVT) and pulmonary embolism (PE) following 5-10 days of initial therapy with a parenteral anticoagulant.

Overview

In this application, the applicant has submitted results from two double-blind, randomized non-inferiority studies: the ENGAGE AF-TIMI 48 study and the Hokusai VTE study. The applicant is seeking two indications for edoxaban: (1) to reduce the risk of stroke and systemic embolism (SE) in patients with non-valvular atrial fibrillation (NVAF) based on results from the AF-TIMI 48 study and (2) as a treatment of DVT and PE based on results from VTE study.

In the AF-TIMI 48 study, a total of 21,105 patients with non-valvular atrial fibrillation were randomized to receive either edoxaban (one of two doses: 30 mg or 60 mg) or warfarin. The dose of edoxaban was halved per protocol if CrCL $\leq$ 50 mL/min, weight $\leq$ 60 kg, or concomitant use of specific P-gp inhibitors occurred. Approximately 25% of patients in all treatment groups received a reduced dose at baseline, and an additional 7% during the study. This part of the application for the indication of NVAF was reviewed by the Division of Biometrics I (Primary reviewer: Dr. John Lawrence, Team Leader: Dr. James Hung), and results of this study were discussed at the cardio-renal drugs advisory committee (CRDAC) meeting on October 30, 2014. The committee was asked to vote on the following question:

Should edoxaban be approved to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation?
If you recommend approval, please discuss the following options:
a) Approval of the 60-mg dose for patients with normal or mildly impaired renal function.
b) Approval of a dose higher than 60 mg for patients with normal renal function.
c) Approval only for patients with mild and moderate renal impairment.

The committee voted 9 to 1 favoring approval. Out of those 9 who voted for approval, the majority said that in their opinion, 60 mg was the dose they recommended for patients in the normal renal function subgroup and only 1 (out of 9) said that a 60 mg dose was not acceptable for that subgroup. Please refer to the minutes of the advisory committee meeting and the statistical review by Dr. Lawrence for further details on this part of the application.

In support of the claim for treatment of DVT and PE, the applicant submitted results from one randomized clinical trial comparing edoxaban 60 mg orally once daily versus

warfarin titrated to INR 2.0-3.0 in patients with acute symptomatic venous thromboembolism (VTE; VTE defined as DVT or PE with or without DVT), in which a total of 8292 patients were randomized. Patients randomized to edoxaban received a lower dose of 30 mg if they met one or more of the following criteria: CrCL ≥ 30 and ≤ 50 mL/min, body weight ≤ 60 kg, or concomitant use of specific P-gp inhibitors. Approximately 17.5% of the patients received a 30 mg per day dose based on their baseline characteristics (Table 1) and an additional 2% had reduction in dose from 60 mg to 30 mg during the course of the study. Please refer to the review by the primary reviewer, Dr. Wang, for a detailed description of the results of this trial. The primary efficacy endpoint was time to recurrent VTE or VTE-related death during the 12-month study period, and the primary objective was to demonstrate that edoxaban was non-inferior to warfarin. The non-inferiority margin for the hazard ratio was set at 1.5, as documented in the statistical analysis plan. The primary safety endpoint was time to occurrence of major bleeding or clinically relevant non-major bleeding. The VTE study demonstrated that edoxaban was non-inferior to warfarin with respect to both efficacy and safety as presented in Table 2.

Table 1: Baseline CrCL, Weight and Use of P-gp Inhibitors in the VTE Study

	Edoxaban (N=4118) n (%)	**Warfarin (N=4122) n (%)**
Edoxaban 30 mg (low dose) at randomization		
Yes	733 (17.8)	719 (17.4)
No	3385 (82.2)	3403 (82.6)
Weight at randomization (kg)		
≤ 60	524 (12.7)	519 (12.6)
> 60	3594 (87.3)	3603 (87.4)
Creatinine clearance at randomization (mL/min)		
≥ 30 to ≤ 50	268 (6.5)	273 (6.6)
> 50	3850 (93.5)	3849 (93.4)
Verapamil or quinidine use at randomization		
Yes	26 (0.6)	25 (0.6)
No	4092 (99.4)	4097 (99.4)

Table 2: Primary Efficacy and Safety Analyses in the VTE Study

Primary Endpoint	Edoxaban N = 4118	Warfarin N = 4122	HR (95% CI)
Efficacy: Time to VTE or VTE related death	130 (3.2%) events	146 (3.5%) events	0.89 (0.70, 1.13)
Safety: Time to major bleeding or clinically relevant non-major bleeding	349 (8.5%) events	423 (10.3%) events	0.81 (0.71, 0.94)

HR < 1 favors edoxaban

Edoxaban dose consideration for the treatment of VTE

We conducted further several exploratory subgroup analyses (Tables 3-6) in order to evaluate if a higher dose is warranted for patients with renal impairment. Because restriction to patients with CrCL ≤ 95 mL/min is being considered for the treatment of non-valvular atrial fibrillation, we also examined the safety and efficacy in the subgroup of patients with CrCL ≤ 95 mL/min versus those with CrCL > 95 mL/min at baseline (Table 7). All these subgroup analyses are exploratory and should be viewed as such. All the subgroup results were consistent with the overall population results presented in Table 2. Based on the conduct of the VTE study and the exploratory subgroup analyses, our recommendation is that patients with impaired renal function (CrCL between 30-50 mL/min), or ≤ 60 kg body weight, or are using specific P-gp inhibitors, should receive 30 mg edoxaban and all others (including patients with CrCL > 95 mL/min) receive 60 mg per day as studied.

Table 3: Exploratory efficacy and safety analyses in subgroup of patients who received 30 mg of Edoxaban

Primary Endpoint	Edoxaban N = 733	Warfarin N = 719	HR (95% CI)
Efficacy: Time to VTE or VTE related death	22 (3.0%) events	30 (4.2%) events	0.73 (0.42, 1.26)
Safety: Time to major bleeding or clinically relevant non-major bleeding	58 (7.9%) events	92 (12.8%) events	0.62 (0.44, 0.86)

HR < 1 favors edoxaban

Table 4: Exploratory efficacy and safety analyses in subgroup of patients who received 60 mg of Edoxaban

Primary Endpoint	Edoxaban N = 3385	Warfarin N = 3403	HR (95% CI)
Efficacy: Time to VTE or VTE related death	108 (3.2%) events	116 (3.4%) events	0.93 (0.72, 1.21)
Safety: Time to major bleeding or clinically relevant non-major bleeding	291 (8.6%) events	331 (9.7%) events	0.87 (0.74, 1.02)

HR < 1 favors edoxaban

Table 5: Exploratory Efficacy Analyses by subgroups

Subgroups	Edoxaban		Warfarin		HR (95% CI)
	N	Events (%)	N	Events (%)	
CrCL Level					
30 – 50 mL/min	268	8 (3.0)	273	16 (5.9)	0.50 (0.21, 1.17)
> 50 ml/min	3850	122 (3.2)	3849	130 (3.4)	0.94 (0.73, 1.20)
Weight					
<= 60 kg	524	15 (2.9)	519	18 (3.5)	0.84 (0.43, 1.68)
> 60 kg	3594	115 (3.2)	3603	128 (3.6)	0.90 (0.70, 1.16)

HR < 1 favors edoxaban; Primary efficacy endpoint: time to VTE or VTE-related death; Subgroup analysis by concomitant use of P-gp inhibitors was not done due to small number of patients in the subgroup receiving concomitant use of P-gp inhibitors.

Table 6: Exploratory Safety Analyses by subgroups

Subgroups	Edoxaban		Warfarin		HR (95% CI)
	N	Events (%)	N	Events (%)	
CrCL Level					
30 – 50 mL/min	268	28 (10.5)	273	39 (14.3)	0.71 (0.44, 1.15)
> 50 ml/min	3850	321 (8.3)	3849	384 (10.0)	0.82 (0.71, 0.96)
Weight					
<= 60 kg	524	39 (7.4)	519	64 (12.3)	0.60 (0.40, 0.89)
> 60 kg	3594	310 (8.6)	3603	359 (10.0)	0.85 (0.73, 0.99)

HR < 1 favors edoxaban; Primary safety endpoint: time to major bleeding or clinically relevant non-major bleeding; Subgroup analysis by concomitant use of P-gp inhibitors was not done due to small number of patients in the subgroup receiving concomitant use of P-gp inhibitors.

Table 7: Exploratory efficacy and safety analyses by CrCL level (≤ 95 mL/min vs. > 95 mL/min)

	Edoxaban		Warfarin		HR (95% CI)
	N	Events (%)	N	Events (%)	
Efficacy: Time to VTE or VTE related death					
≤ 95 mL/min	1935	60 (3.1)	1960	83 (4.2)	0.73 (0.53, 1.02)
> 95 ml/min	2183	70 (3.2)	2162	63 (2.9)	1.10 (0.78, 1.55)
Safety: Time to major bleeding or clinically relevant non-major bleeding					
≤ 95 mL/min	1935	173 (8.9)	1960	243 (12.4)	0.71 (0.58, 0.86)
> 95 ml/min	2183	176 (8.1)	2162	180 (8.3)	0.96 (0.78, 1.18)

HR < 1 favors edoxaban

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

RAJESHWARI SRIDHARA
01/07/2015

LISA M LAVANGE
01/07/2015

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 206-316
Supplement #:	
Drug Name:	Edoxaban
Indication(s):	(b) (4) stroke or systemic embolic event in patients with Atrial Fibrillation
Applicant:	Daiichi Sankyo
Date(s):	1/8/2014
Review Priority:	Standard
Biometrics Division:	DBI
Statistical Reviewer:	John Lawrence, Ph D
Concurring Reviewers:	Jim Hung
Medical Division:	Cardiorenal.
Clinical Team:	Melanie Blank MD, Tzu-Yun McDowell MD, Martin Rose MD
Project Manager:	Alison Blaus
Keywords:	active control/non-inferiority, Cox regression

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

There was only one phase 3 trial for this indication in the submission. Two dose regimens were studied. Both regimens were safe and effective.

INTRODUCTION

1.1 Overview

Table 1 List of all studies included in analysis

	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
Study DU176B-C-U301	*Phase 3*	*2.5 years*	*2.8 years*	*7002 (low dose edoxaban), 7012 (high dose), 7012 (warfarin)*	*subjects with documented AF within the preceding 12 months and in whom anticoagulant therapy was indicated.*

Source: Study Report.

1.2 Data Sources

Electronic datasets and Study Reports:

\\cdsesub1\evsprod\NDA206316\\206316.enx

\\cdsesub1\evsprod\NDA206316\0009\m5\datasets\du176b-c-u301\analysis\legacy\datasets

STATISTICAL EVALUATION

1.3 Data and Analysis Quality

Some of the datasets were too large to be used. The laboratory dataset was 8.6 Gb and I could not copy it to my hard drive or open it in any software on my computer. Other datasets that I could open were also very large (on the order of 0.5 Gb). Files of this size are difficult to work with. It takes a long time to copy them from one place to another and takes a long time to open and do any analysis. The structure of the datasets was complicated and made it difficult to understand how to do simple analyses such as counting the number of primary endpoint events in

each group or how much time of exposure in each group. I needed to communicate with the sponsor several times to understand how to do things that should have been simple if the datasets had been designed in a better way. In defense of the sponsor, this is a common and recurring problem across many applications. I would judge the data quality as fair.

The analysis was complicated because the sponsor used an on-treatment approach. In my opinion, the intention to treat analysis should be used for the primary analysis, even in an active control, non-inferiority trial. Some people have concerns that low compliance in all treatment arms could bias the results toward showing no difference between the two arms, which could increase the chances that an inferior drug could be demonstrated as non-inferior. That is something to be concerned about, but using the on treatment analysis approach is not the way to fix the problem. Instead, steps should be taken to make sure every subject stays on their randomized treatment. The intention to treat analysis should be the primary analysis. The on treatment analysis should be a sensitivity analysis. Studies with a large amount of non-compliance to either treatment or a large amount of loss to follow-up are not interpretable. When the amount of non-compliance is small, the two approaches should give the same results. To the extent that they differ, the reasons for the difference should be explored. Although I disagree with the approach used, the data analysis, given the decision was made to use that approach, was excellent.

1.4 Evaluation of Efficacy

1.4.1 Study Design and Endpoints

Study DU176B-C-U301, also called the ENGAGE AF-TIMI 48 trial, was a phase 3, randomized, double-blind, double-dummy, parallel-group, multi-center, multi-national study for evaluation of efficacy and safety of du-176b (edoxaban) versus warfarin in subjects with atrial fibrillation (AF). Subjects needed to be at least 21 years old, with a history of AF documented by any electrical tracing within the prior 12 months and for which anticoagulation therapy was indicated and planned for the duration of the study. Subjects with or without previous vitamin K antagonist experience (abbreviated VKA, warfarin is one such VKA) were allowed; it was anticipated that approximately 40% of subjects would be VKA-naive). Subjects needed a $CHADS_2$ index score ≥ 2. The $CHADS_2$ scoring was performed by assigning 1 point each for a history of congestive heart failure (CHF), hypertension, age ≥ 75 years, or diabetes mellitus; and by assigning 2 points for history of stroke or transient ischemic attack (TIA).

Eligible subjects were stratified by $CHADS_2$ risk score at randomization (2 or 3 vs. 4, 5, or 6). Within each $CHADS_2$ stratum, subjects were further stratified based on whether or not a

subject required edoxaban dose reduction for factors such as low estimated creatinine clearance using the Cockroft-Gault equation (eCrCL less than 50 mL/min), low body weight (less than 60 kg), or a need for concomitant treatment with P-glycoprotein (P-gp) inhibitors such as quinidine and/or verapamil. Randomization was stratified by these two stratification factors.

Warfarin was the active control used in this study. Warfarin was titrated within each subject to achieve a target INR between 2.0 and 3.0. There were two experimental treatment arms. The usual dose in the high dose edoxaban arm was 60 mg qd. The usual dose in the low dose edoxaban arm was 30 mg qd. Within each treatment arm, subjects who required a dose adjustment (for low eCrCL, low body weight, or concomitant treatment with P-gp inhibitors) was cut in half of the usual dose, i.e. 30 mg in the high dose arm and 15 mg in the low dose arm. In order to maintain the study blind, a double dummy strategy was used and sham INR values were reported for subject given warfarin placebo. It is difficult to conduct a double-blind trial with warfarin. The sponsor is commended for making the effort to do this because a double-blind trial is more credible than an open label trial.

The primary endpoint was time to first stroke (of any kind) or systemic embolic event (SEE) while on treatment. Subjects were considered on treatment for 3 days after their last dose. If a study drug interruption occurred and the subject returned to study drug later, they were considered on treatment during the first 3 days with no treatment, not on treatment the remaining days with no treatment, then back on treatment when they continued treatment. Any events that occurred during the off-treatment period did not count in the primary analysis.

This was an event-driven study. The statistical considerations and plan for the study required approximately 672 primary endpoints overall, with 448 on-treatment primary endpoint (composite of stroke and SEE) events for each of the following 2 pairwise comparisons: i) edoxaban 30 mg group versus warfarin ii) edoxaban 60 mg group versus warfarin.

There were three secondary endpoints: i) composite of stroke, SEE, and CV mortality; ii) MACE, which is the composite of non-fatal MI, non-fatal stroke, non-fatal SEE, and death due to CV cause or bleeding; iii) composite of stroke, SEE, and all-cause mortality.

1.4.2 Statistical Methodologies

The analysis of the primary endpoint used the modified intent-to-treat (mITT) population, which was defined as subjects who were randomized and received at least one dose of study drug. The analysis approach used the on-treatment period to count events. Subjects were excluded from the at risk set during periods of treatment interruptions. Subjects were considered on treatment for 3

days after last dose and back on treatment if the treatment recommenced. The primary analysis used the counting process approach to include only subjects at risk in each time interval in the Cox proportional hazards regression model analysis. The Cox proportional hazards model included treatments and the following 2 stratification factors as covariates:
1. The dichotomized CHADS2 score (1 if CHADS2 ≥ 4; 0 otherwise)
2. The dichotomized dose adjustment variable (1 if eCrCL ≤ 50 mL/min, or body weight ≤ 60 kg, or taking verapamil or quinidine; 0 otherwise)

Because there were two treatment arms being compared simultaneously to the active control group, a Bonferroni type approach was used to control the overall error rate. The two-sided 97.5% confidence interval (CI) for the hazard ratios (each edoxaban treatment group versus warfarin) was estimated using the proportional hazards model. If the upper limit of this CI of the hazard ratio was below 1.38, then non-inferiority to warfarin was considered established for the corresponding edoxaban treatment group. The margin of 1.38 was appropriate and has been used in many other studies for this indication. It was derived from the estimated effect of warfarin compared to placebo in historical trials.

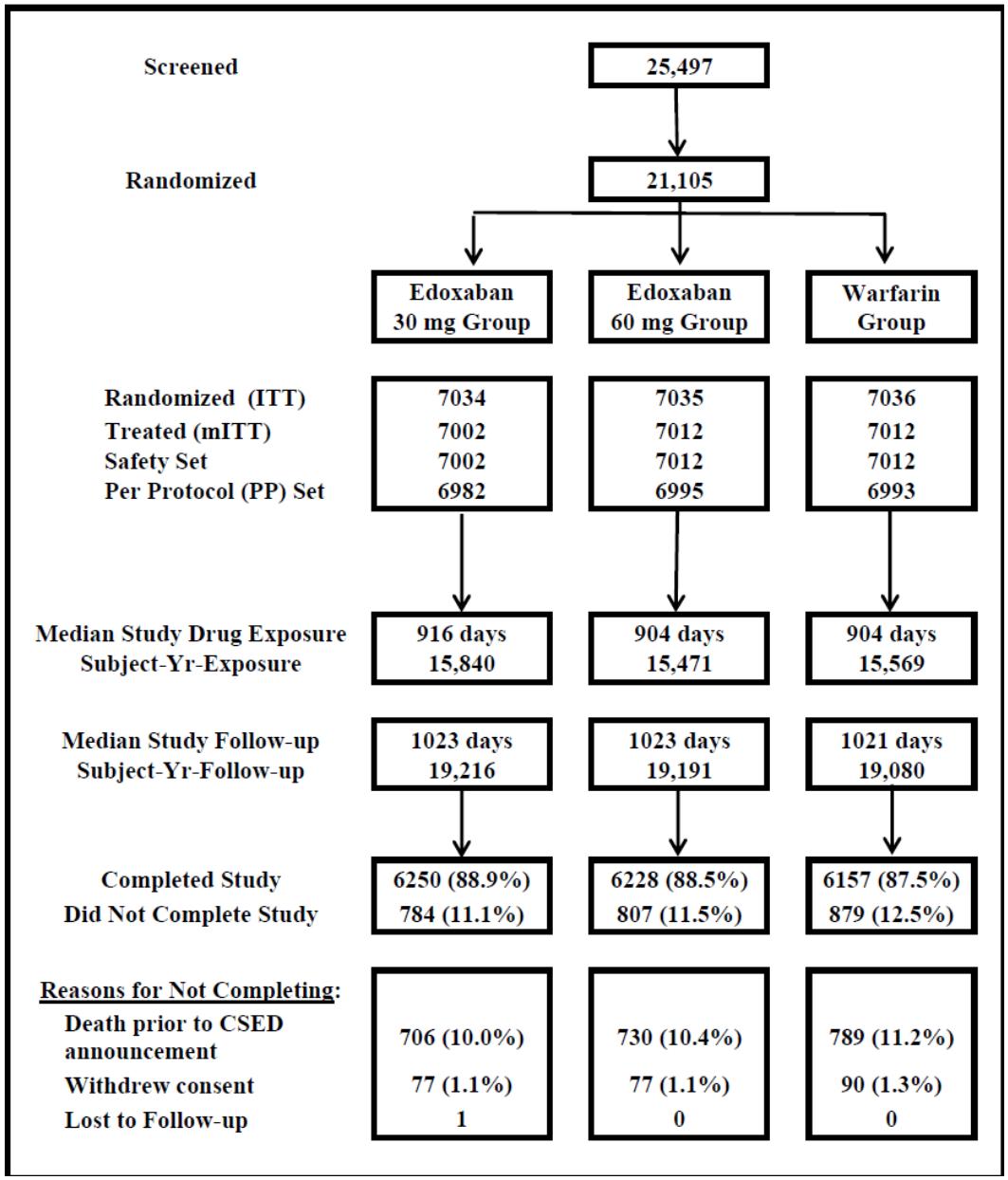
If the high dose arm was noninferior to warfarin, then the high dose would be compared to warfarin for superiority on the primary endpoint using two-sided error rate of α=0.01. For the test of superiority, the ITT analysis and overall study period would be used (i.e. all events would be counted, not just the on-treatment events). If superiority was concluded, then the three secondary endpoints would be tested in order using the same approach (ITT analysis with α=0.01). No testing for superiority of the primary or any secondary endpoints was planned for the low dose arm.

1.4.3 Patient Disposition, Demographic and Baseline Characteristics

The patient disposition is shown in Figure 1 and the baseline and demographic characteristics are shown in Table 2. The figure shows that very few subjects were lost to follow-up or did not complete the study except for those who died. However, nearly ¼ of the follow-up time was not included in the primary analysis because it only included on-treatment periods. A strength of the trial is that there is very little loss to follow-up and subjects were followed after stopping study drug. However, the amount of information that is intentionally censored (during the off-treatment period) could be a problem. The duration off-treatment in the three arms seems to be roughly equal. If that had not been the case, then it would be even more concerning. Still, that does not prove that the censored information can be ignored. In consideration of that, we should look carefully at the consistency between the analysis of the on-treatment period events with the analysis of the overall study period events. In Table 2, it is seen that the average age was about 70

years, 60% were male, ¾ of the subjects fell in the usual dose category, 80% were Caucasian. There were no significant differences in the demographics between groups. In addition to what is shown in Table 2, other useful demographic information includes: approximately 18% of the subjects were from the US region and about half of the subjects were VKA naive.

Figure 1 Patient disposition



Source: Figure 10-1 of Study Report.

Table 2 Patient demographic and baseline characteristics

	Edoxaban 30 mg (15mg DosAdj) (N=7002)	Edoxaban 60 mg (30mg DosAdj) (N=7012)	Warfarin (N=7012)
Age (years), n	7002	7012	7012
Mean	70.6	70.6	70.5
SD	9.31	9.51	9.44
Median	72.0	72.0	72.0
Minimum	27	25	27
Maximum	95	96	95
>= 65 years n(%)	5218 (74.5)	5182 (73.9)	5143 (73.3)
>= 75 years n(%)	2789 (39.8)	2838 (40.5)	2805 (40.0)
>= 80 years n(%)	1197 (17.1)	1177 (16.8)	1195 (17.0)
Gender, n (%)	7002	7012	7012
Male	4284 (61.2)	4353 (62.1)	4383 (62.5)
Female	2718 (38.8)	2659 (37.9)	2629 (37.5)
Race, n (%)[a]	7001	7012	7012
Caucasian	5650 (80.7)	5679 (81.0)	5679 (81.0)
Black	94 (1.3)	96 (1.4)	88 (1.3)
Asian	975 (13.9)	956 (13.6)	963 (13.7)
Other	282 (4.0)	281 (4.0)	282 (4.0)
Edoxaban/Placebo Dose Adjusted at Randomization, n (%)	7002	7012	7012
Yes	1774 (25.3)	1776 (25.3)	1780 (25.4)
No	5228 (74.7)	5236 (74.7)	5232 (74.6)
CrCL (mL/min), n (%)[b]	6961	6954	6973
< 30	42 (0.6)	70 (1.0)	51 (0.7)
30 - <= 50	1274 (18.2)	1287 (18.4)	1297 (18.5)
> 50 - < 80	3034 (43.3)	2985 (42.6)	3030 (43.2)
>= 80	2611 (37.3)	2612 (37.3)	2595 (37.0)
Weight (kg), n (%)[c]	6996	7007	7007
<= 50	148 (2.1)	158 (2.3)	172 (2.5)
<= 60	692 (9.9)	681 (9.7)	697 (9.9)
> 60	6304 (90.0)	6326 (90.2)	6310 (90.0)
Mean (SD)	83.9 (20.11)	84.2 (20.40)	83.7 (20.09)

Source: Table 10.4 of Study Report.

1.4.4 Results and Conclusions

Both doses were non-inferior to warfarin on the primary endpoint. The results are shown in Table 3. The primary analysis is in the first row of the table (mITT Analysis Set, On Treatment Period). Both doses were non-inferior to warfarin because the upper limit of both two-sided 97.5% confidence intervals were less than the margin of 1.38. The primary analysis did not include events that happened during treatment interruptions. The second row is the sensitivity analysis which includes all of the events. There were about 50% more subjects with events in each group when these off-treatment events were counted. As stated before (Section 1.3 of this review), I believe that it would have been best to make this the primary analysis. In any case, both methods of analysis showed that both doses were non-inferior to warfarin.

Table 3 Adjudicated Primary Endpoint (Stroke or SEE), mITT Analysis Set - On-Treatment and Overall Study Period (Non-Inferiority)

Primary Endpoint	Edoxaban 30 mg (15mg DosAdj) (N=7002)		Edoxaban 60 mg (30mg DosAdj) (N=7012)		Warfarin (N=7012)		Edoxaban 30 mg (15mg DosAdj) vs Warfarin		Edoxaban 60 mg (30mg DosAdj) vs Warfarin	
First Stroke or SEE	# of events	Event Rate (%/yr) [a]	# of events	Event Rate (%/yr) [a]	# of events	Event Rate (%/yr) [a]	HR (97.5% CI)	p-value[b]	HR (97.5% CI)	p-value[b]
mITT Analysis Set On Treatment Period	253	1.61	182	1.18	232	1.50	1.07 (0.874, 1.314)	0.0055	0.79 (0.632, 0.985)	<0.0001
mITT Analysis Set Overall Study Period	382	2.04	292	1.55	336	1.80	1.13 (0.955, 1.336)	0.0074	0.86 (0.719, 1.029)	<0.0001

Abbreviations: DosAdj = Dose Adjusted, HR = Hazard Ratio versus Warfarin, CI = Confidence Interval, mITT = Modified Intent-to-Treat, SEE = Systemic Embolic Event, yr = year.
[a]: The event rate (%/yr) is calculated as # of events/subject-year exposure.
[b]: The two-sided p-value is based on the non-inferiority margin of 1.38
Source: Table 11.2 of Study Report and confirmed by FDA.

In any active control study, we want to make sure that the active control was used appropriately. Table 4 shows the time in therapeutic range (INR 2-3) and also the percent time in other ranges. The TTR was about 65%, which is very good. 23% of the time, the subjects had INR<2 and 12% of the time, the subjects had INR>3. In the range INR<2, the warfarin dose is too low, leading to subjects having a greater risk of ischemic stroke. In the range INR>3, the warfarin dose is too high, leading to greater risk of bleeding. Overall, this is about as good as can be expected in a clinical trial and I do not think that this causes any concern with respect to the interpretation of the non-inferiority of the two edoxaban doses.

Table 4 Time in various INR ranges for subjects randomized to warfarin, safety analysis set – on treatment period, excluding initial 7 days

	Percent Time in INR Range[a]								
	<1.5	**1.5-2.0**	**<2**	**2-3 (TTR)**	**>3**	**>=4**	**>5**	**>=8**	**1.8-3.2**
Overall (N=6897)									
Mean (SD)	6.10(13.8)	22.70(13.3)	22.80(18.9)	64.90(18.7)	12.40(10.3)	1.80(4.5)	0.30(2.3)	0.00(0.8)	78.40(18.1)
Median	1.90	21.00	17.70	68.40	10.80	0.40	0.00	0.00	83.10

Abbreviations: INR = International Normalized Ratio, SD = Standard Deviation, TTR = Time in Therapeutic Range.
[a]: Percent Time in INR range is defined by the percentage of days the subjects have been within the specified range. Percent Time in Therapeutic Range (TTR) is calculated as the mean percentage in the range 2-3.
Note: N = Number of subjects with at least 1 INR recorded beyond Day 7.
Note: All INRs taken while on-treatment, excluding the initial 7 days of study medication are considered.
Note: Analyses of INR use a linear interpolation method to impute INR for study days that do not have an actual INR value.
Source: Table 10.10 of Study Report.

The components of the primary endpoint and the results for different types of strokes are shown in Table 5. There were nearly equal numbers of ischemic strokes in the high dose and warfarin groups, but more in the low dose group. The low dose group had the fewest number of hemorrhagic strokes and the warfarin group had the most. All groups had about the same number of fatal strokes, but the low dose group had very few fatal hemorrhagic strokes. The low dose had the most disabling strokes.

Table 5 Components of the primary endpoint and different types of strokes (mITT analysis set, on-treatment period)

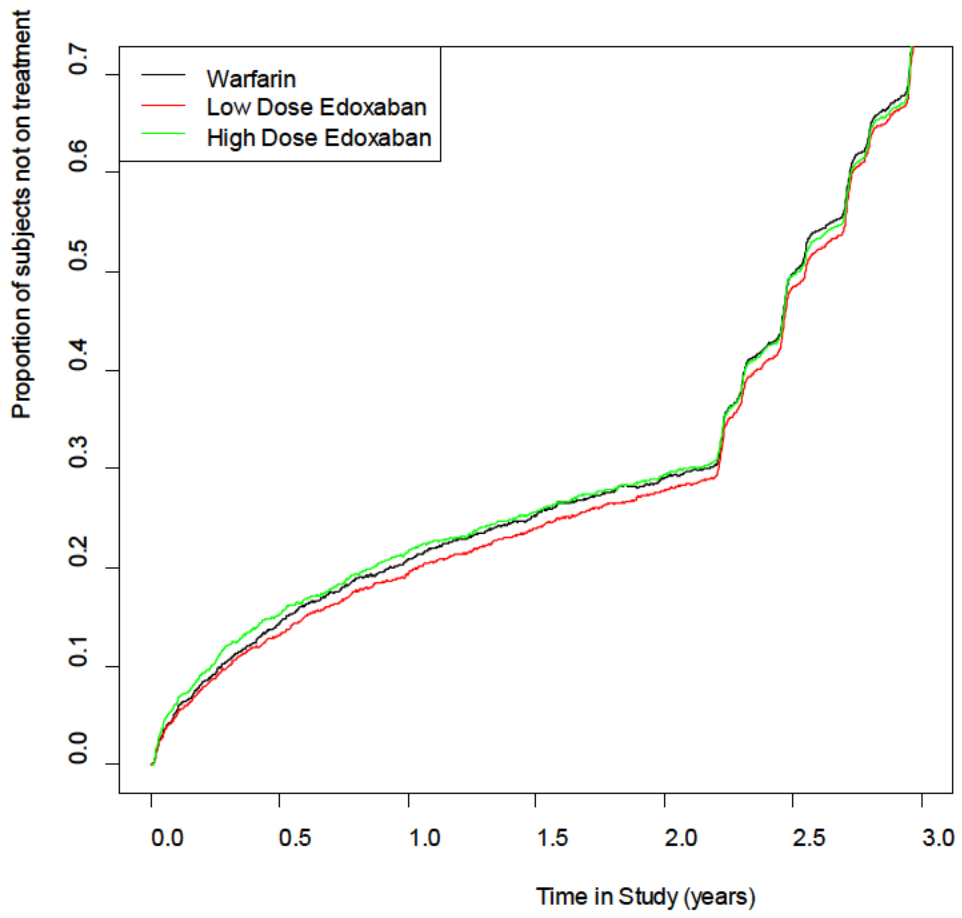
Endpoint	Edoxaban low dose	Edoxaban high dose	Warfarin
Stroke/SEE	253	182	232
First Stroke	244	174	219
First Ischemic Stroke	226	135	144
First Hemorrhagic Stroke	18	40	76
Fatal Stroke	40	45	43
Fatal Ischemic Stroke	35	22	13
Fatal Hemorrhagic Stroke	5	23	30
First Disabling Stroke	57	35	41
First SEE	11	8	13

Source: Tables 14.2.1.10 and 14.2.1.15 of Study Report.

In a normal time to event analysis, subjects can be right censored and I would draw a figure that shows the percent of subject still remaining at risk at each time point in the trial to compare the

dropout rates between groups. In this trial, the rate of loss to followup (not including death) was almost 0 in all groups. But the percent of subjects at risk over time changes because people go in and out of the on-treatment period. For this trial, I made a figure that shows the percent of subject days missing from the analysis at each time point. For example, at 1 year, the percentage of people not on treatment was about 19.3% (low dose), 21.5% (high dose), and 20.7% (warfarin). That is a high percentage of people being censored in the analysis.

Figure 2 Proportion of subjects not included in primary analysis over time in the study.

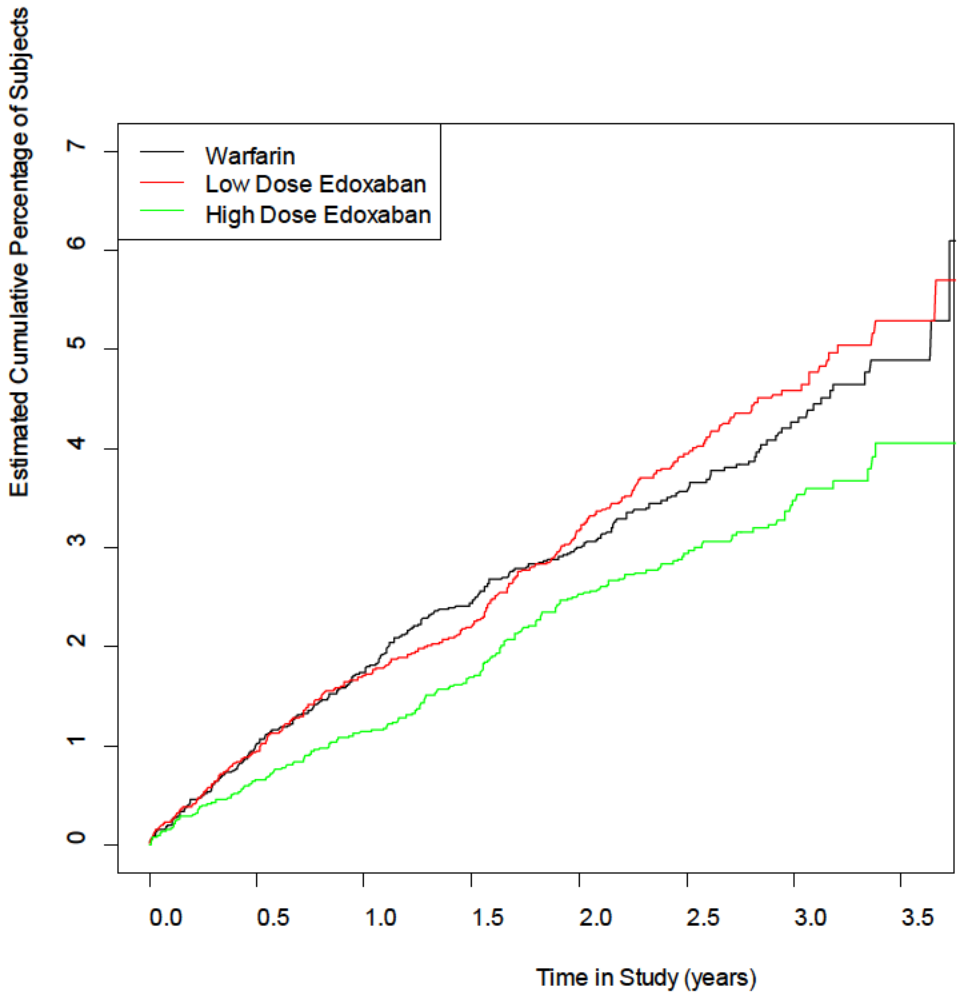


Source: FDA analysis.

Next, in Figure 3,the Kaplan-Meier estimates for the primary endpoint are shown. This shows the estimate of the proportion of people without a first on-treatment stroke/SEE. In this figure, no proportional hazards assumption is made and no adjustment is made for any factors. Figure 4 is similar to Figure 3, but transforms both the x and y-axis. In Figure 4, log(-log(estimated survival)) is on the y-axis and log(time) in the x-axis. The reason for drawing this latter figure is to check

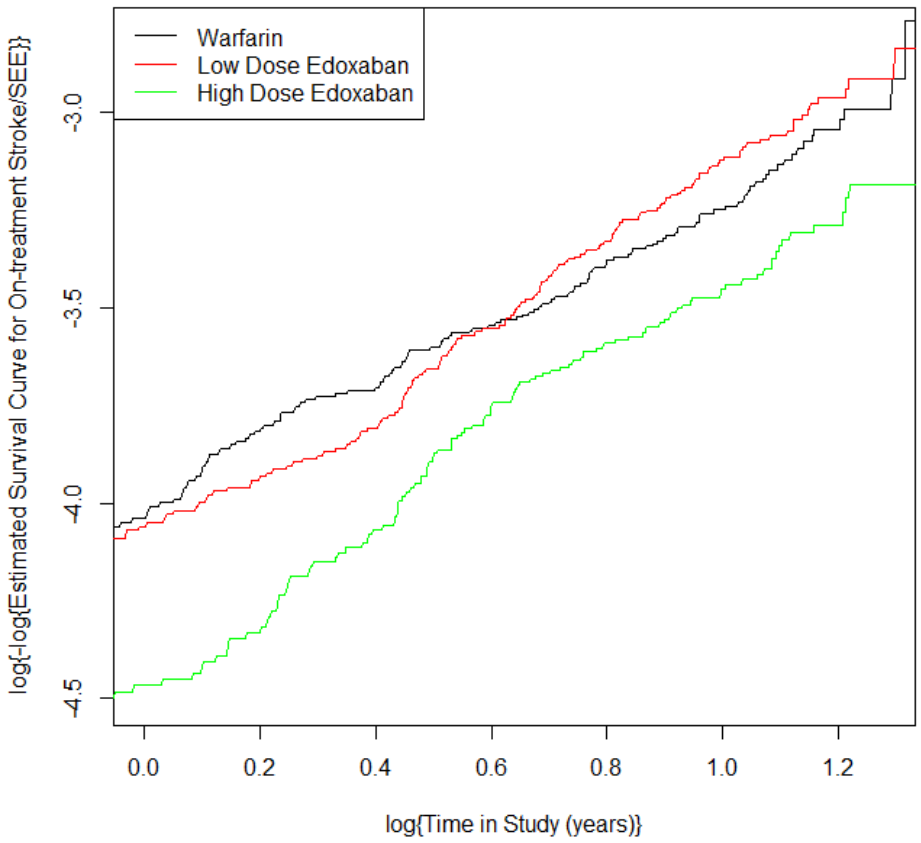
whether the proportional hazards assumption is correct. In addition, we can use this figure to check whether the distributions can be approximated by a Weibull distribution. Since the curves cross each other (red and black), the curves are not quite parallel and the hazards are not proportional. The curves are not exactly straight lines, but Weibull distributions may be a reasonable approximation to the true survival curves depending on the purpose.

Figure 3 Kaplan-Meier estimates of event rates over time (mITT, On treatment period).



Source: FDA analysis

Figure 4 Log{-log(survival)) plot (mITT, On treatment period).

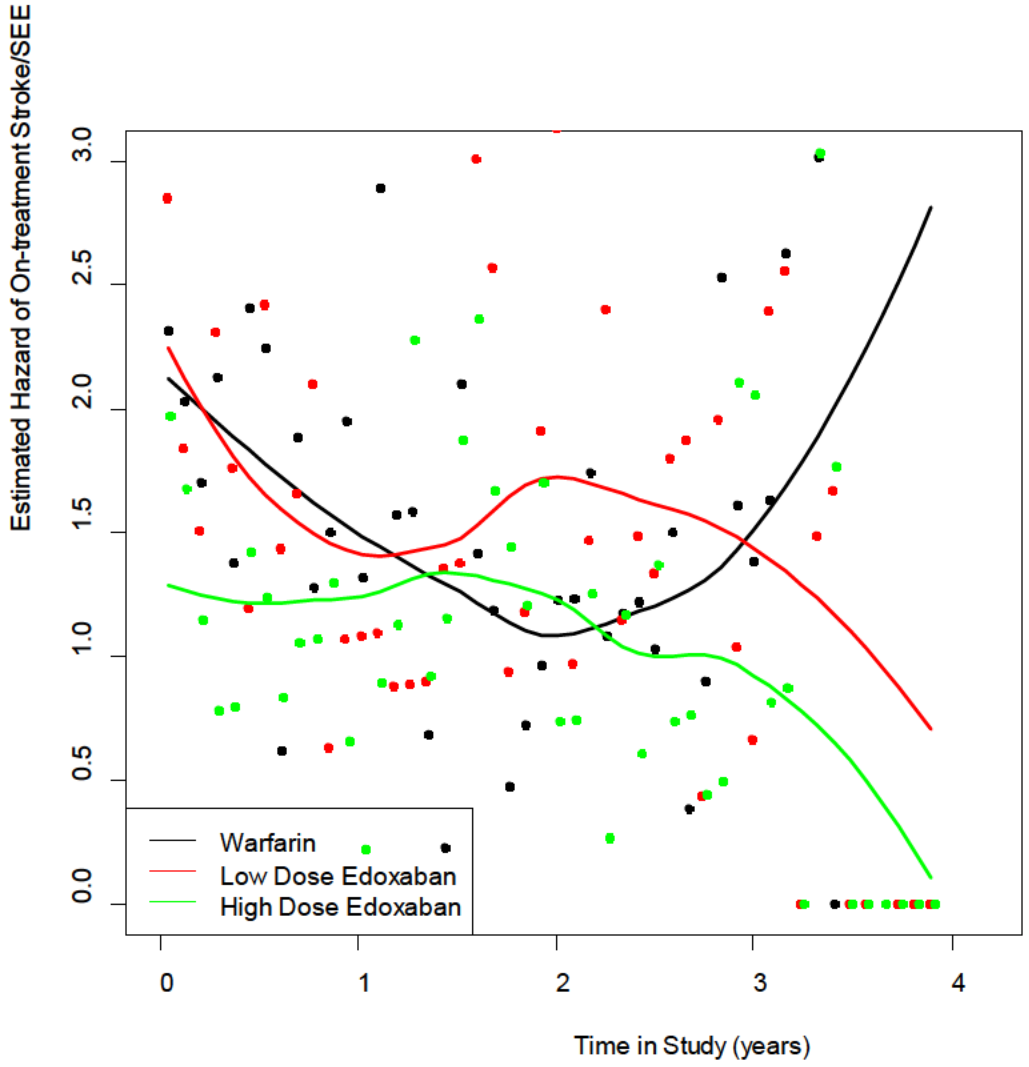


Source: FDA analysis

In the next figure, Figure 5, the estimated hazard functions are shown for each group. As in Figure 3, each hazard function is estimated independently and no adjustments are made for any covariates. I found these estimated hazard functions by first finding the number of events in each 30 day period and dividing by the amount of subject time on treatment during those time intervals. There should be 12 dots of each color within each year interval. If that many cannot be seen, then some are hidden behind dots of other color; 1 or 2 dots have a y-coordinate greater than 3 and are not shown. I then used a locally weighted regression (weighting each point by the duration of exposure in the denominator) to draw smooth curves through these points. Of note, referring back to Figure 2, it can be seen that at any given time after 3 years, only about 30% or less of the subjects are on treatment. Therefore, the curves are not very reliable beyond 3 years. The green curve starts out below the other two and stays fairly constant for the first 3 years. The red and black curves start out higher than the green curve, but decline over time. One possible explanation for the decline in the black curve in particular, is that warfarin is difficult to titrate initially. Therefore, I separated each group into the subjects who were VKA naive 30 days before

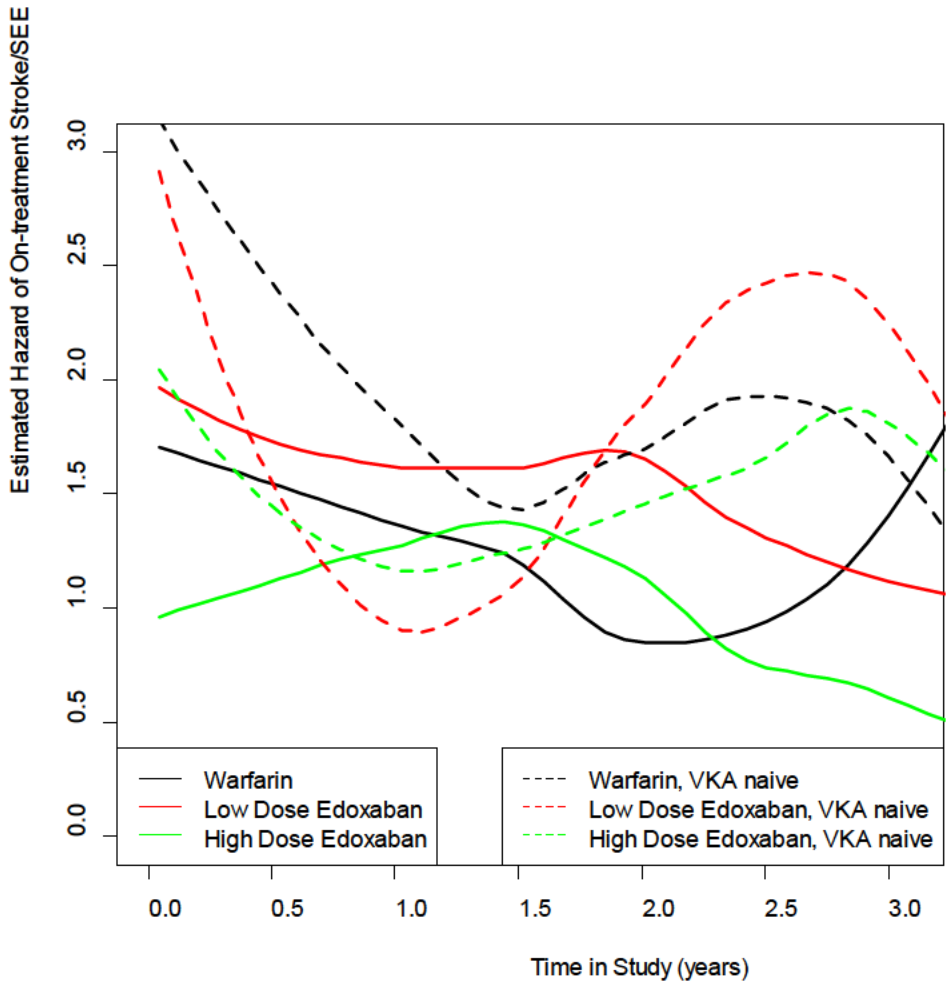
randomization and those who were not. I drew similar hazard function curves for the 6 different groups. Within each treatment group, not only in the warfarin assigned group, the subjects who were VKA naive started out with a higher hazard rate. In the warfarin group, the dashed curve (VKA naive) stayed higher throughout the first 3 years, so it may not just be an issue of initial warfarin titration.

Figure 5 Estimated hazard functions over time (mITT, On treatment period).



Source: FDA analysis

Figure 6 Estimated hazard functions over time (mITT, On treatment period) by VKA naive status (solid curves are not VKA naive at randomization, dashed curves are VKA naive).



Source: FDA analysis

I tried to fit different parametric survival distributions for the primary endpoint. As can be seen from Figure 5, a constant hazard function model (exponential distribution) would not fit the data very well. In addition, the hazard functions do not appear to be proportional, nor do they appear to satisfy the assumptions for an accelerated failure time model. I tried the Weibull family, which includes the Exponential family as a special case. I did not include any covariates or adjust for other factors. Rather, I fit one set of parameters for each treatment arm separately. See the Appendix for more details.

For a discussion of results in special subgroups, including subgroups defined by renal function, see Section 1.8.

The superiority test for the primary endpoint (high dose vs. warfarin) was not significant at the pre-specified two-sided 0.01 level. Thus, according to the analysis plan, none of the secondary endpoints would be tested for superiority. Although none of the mortality results would be statistically significant using the analysis plan, they are shown in Table 6. The low dose had substantially fewer deaths than the high dose group, which in turn had substantially fewer deaths than the warfarin group.

Table 6 Mortality results (ITT, overall study period).

	Statistic	Edoxaban 30mg (15mg DosAdj) (N=7034)	Edoxaban 60mg (30mg DosAdj) (N=7035)	Warfarin (N=7036)
All Cause Mortality	# of Events	737	773	839
	Subj Yr Expo	19414.02	19355.51	19286.20
	Event Rate (%/yr)	3.80	3.99	4.35
	HR (95% CI)	0.87 (0.788, 0.960)	0.92 (0.831, 1.011)	
	p-value	0.0058	0.0816	
Cardiovascular Mortality	# of Events	527	530	611
	Subj Yr Expo	19414.02	19355.51	19286.20
	Event Rate (%/yr)	2.71	2.74	3.17
	HR (95% CI)	0.85 (0.760, 0.960)	0.86 (0.768, 0.970)	
	p-value	0.0080	0.0133	

Source: Table 14.2.2.6 of Study Report.

1.5 Evaluation of Safety

See clinical review.

1.6 Benefit-Risk Assessment (Optional)

See clinical review.

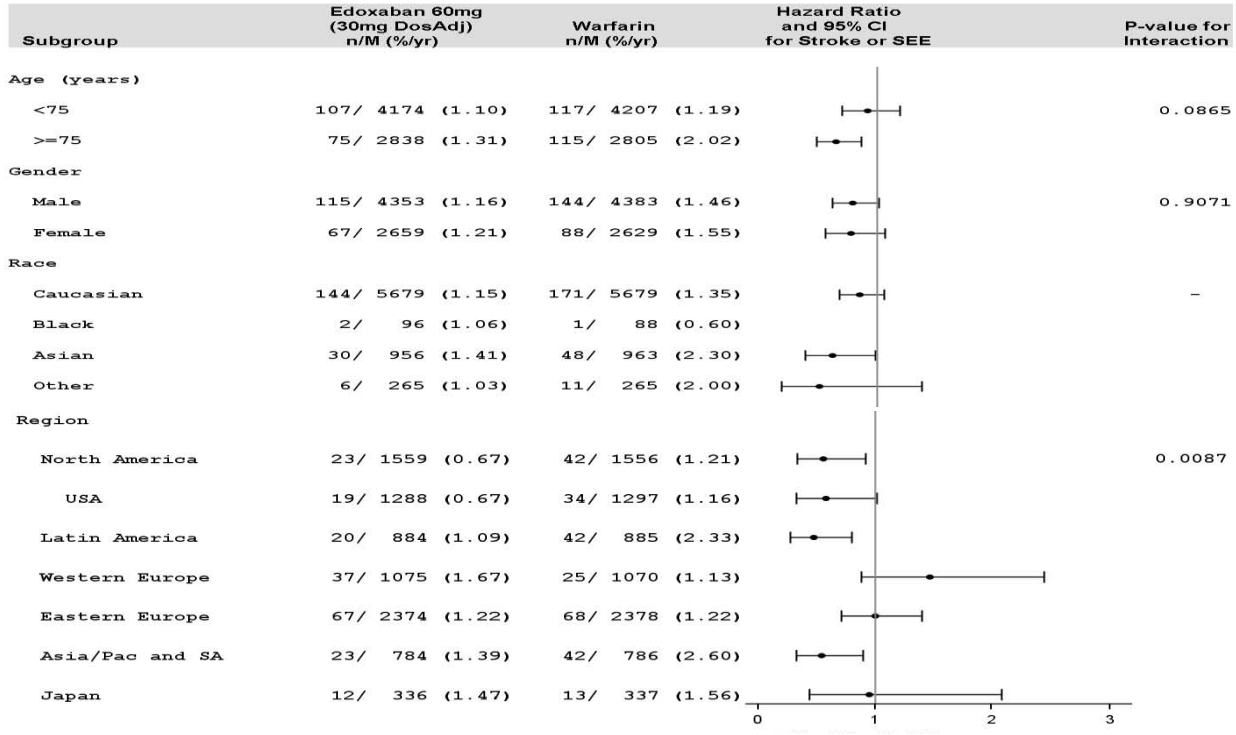
FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

1.7 Gender, Race, Age, and Geographic Region

The results comparing the high dose to warfarin for these subgroups for the primary endpoint are shown in Figure 7. The only significant interaction was by geographic region. In Eastern Europe, Western Europe, and Japan, warfarin tended to be better than edoxaban. In the remaining

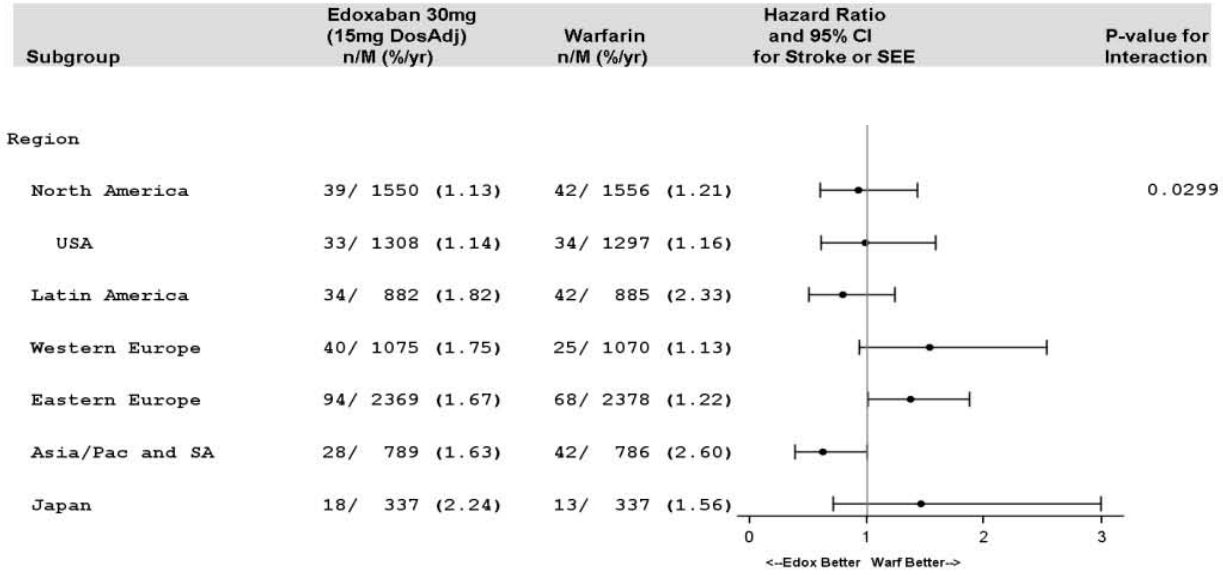
regions, edoxaban tended to be better than warfarin. This same trend was also suggested when comparing the low dose regimen to warfarin (Figure 8). One possible explanation is that this observation has something to do with how warfarin is performing across those regions.

Figure 7 High dose/warfarin results for subgroups defined by gender, race, age, and geographic region.



Source: Figure 11-3 of Study Report.

Figure 8 Low dose/warfarin results for subgroups defined by gender, race, age, and geographic region.

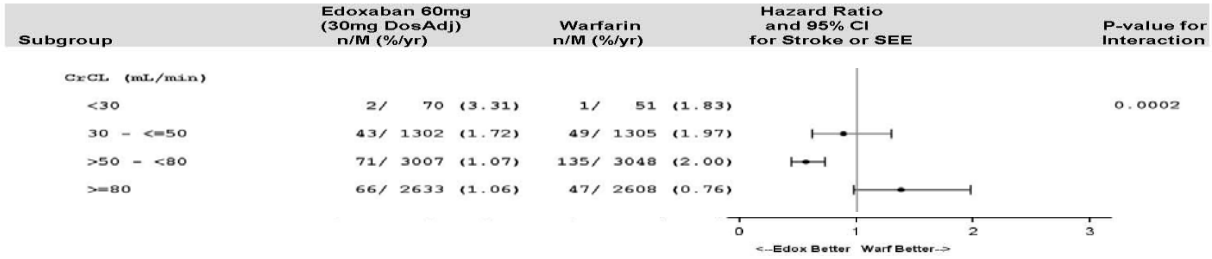


Source: Figure 11-3 of Study Report.

1.8 Other Special/Subgroup Populations

The results for the high dose compared to warfarin in subgroups defined by baseline eCrCL are shown in Figure 9. Similar results were observed when comparing the low dose regimen to warfarin (not shown).

Figure 9 High dose/warfarin results (primary endpoint) for subgroups defined by baseline eCrCL.



Source: Figure 11-3 of Study Report.

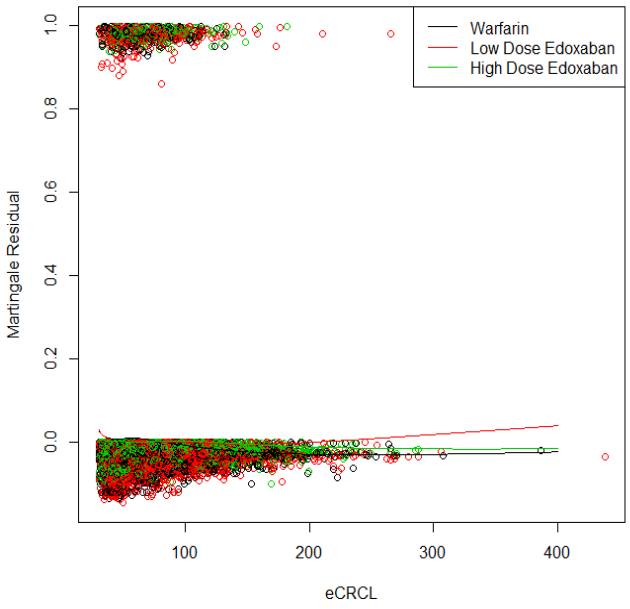
The first question I will look at is whether there is a qualitative interaction. A qualitative interaction, roughly, is when one confidence interval is completely below 1 and the other confidence interval is completely above 1. I will look only at the two subgroups eCrCL≥80 and 50≤eCrCL<80 because the patients in the subgroup with eCrCL<50 were supposed to receive a

dose adjustment by half. If we look at the treatment effect in those two subgroups (high dose vs. warfarin only) adjusting only for chads score variable and treatment, the estimated hazard ratios are 0.53 (50≤eCrCL<80 subgroup) and 1.41 (eCrCL≥80 subgroup). The corresponding estimated log-hazard ratios and standard errors are -0.641 (s.e. 0.146) and 0.346 (s.e. 0.191). Using the likelihood ratio test for a qualitative interaction relative to a test for superiority (Gail, M., and R. Simon. "Testing for qualitative interactions between treatment effects and patient subsets." *Biometrics* (1985): 361-372.), the p-value is $\Phi\left(\frac{0.346}{0.191}\right) = 0.035$ where $\Phi(x)$ is the standard normal distribution function. However, I would argue this is the wrong test in this situation. The study was not designed primarily to show superiority. A qualitative interaction with respect to the noninferiority margin of 1.38 would happen if one confidence interval was below 1.38 (edoxaban was effective) and the other was completely above 1.38 (edoxaban was not effective). The corresponding p-value for testing for a qualitative interaction with regard to noninferiority is $\Phi\left(\frac{0.346-\log(1.38)}{0.191}\right) = 0.45$. In the Appendix, there is a discussion about how likely it is to see a qualititative interaction for superiority in a noninferiority trial.

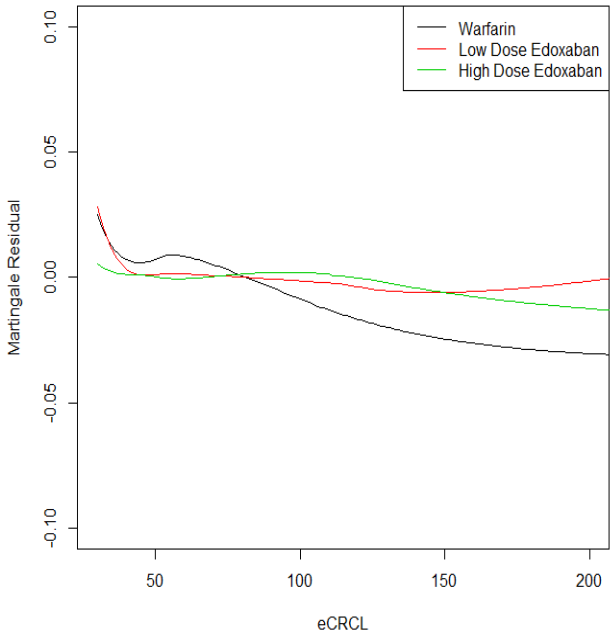
Next, I looked at the martingale residuals to see if there was any additional predictive value of eCrCL and what functional form to use in the model. The martingale residuals are the difference between the expected number of events and the observed number of events for each individual conditional on their exposure times. The residuals and the loess curves fit to them are shown in Figure 10. In panel (b), I zoomed into the part of the graph where the curves are and removed the points. Since the red and green curves are nearly constant, there is no predictive ability in the low and high dose groups, but the black curve shows there may be some added predictive ability for the warfarin group. In Figure 11 the estimated hazard ratios are shown as a function of eCrCL using splines (cubic splines with boundary knots at 60 and 100). For this figure, I started with the model used in the primary analysis and included the cubic spline function of eCrCL together with the two way interactions between dose and eCrCL.

Figure 10 Martingale residuals (primary endpoint) as a function of eCrCL.

(a)

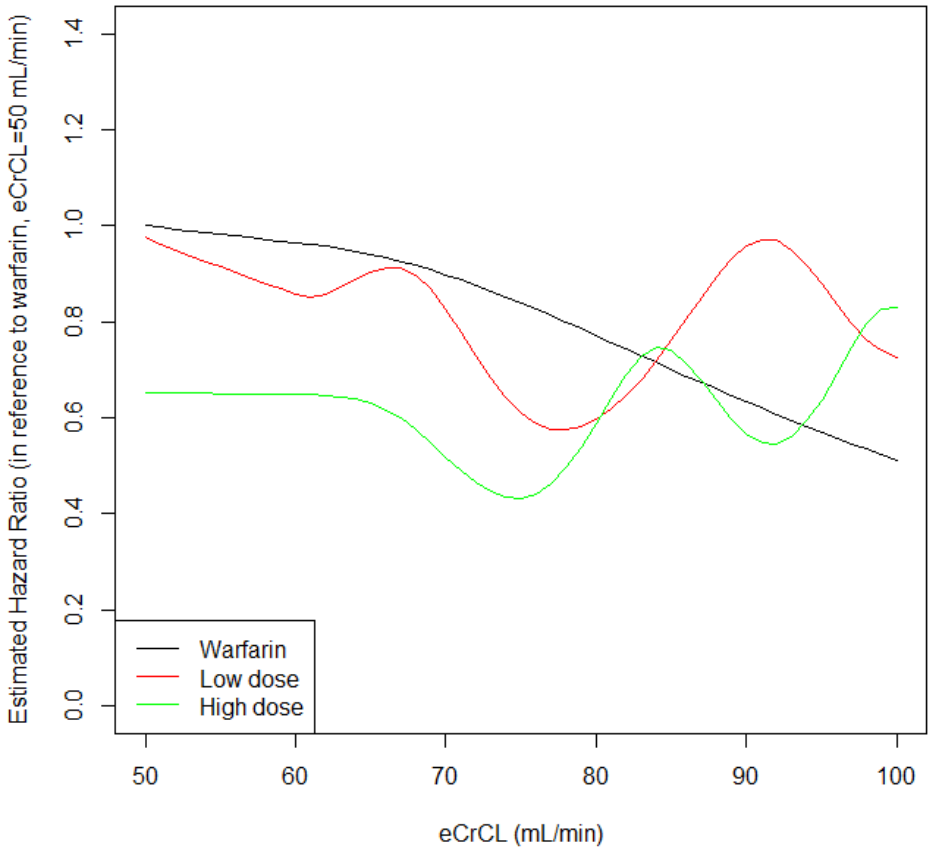


(b)



Source:FDA analysis.

Figure 11 Hazard ratios (primary endpoint) as a function of eCrCL.

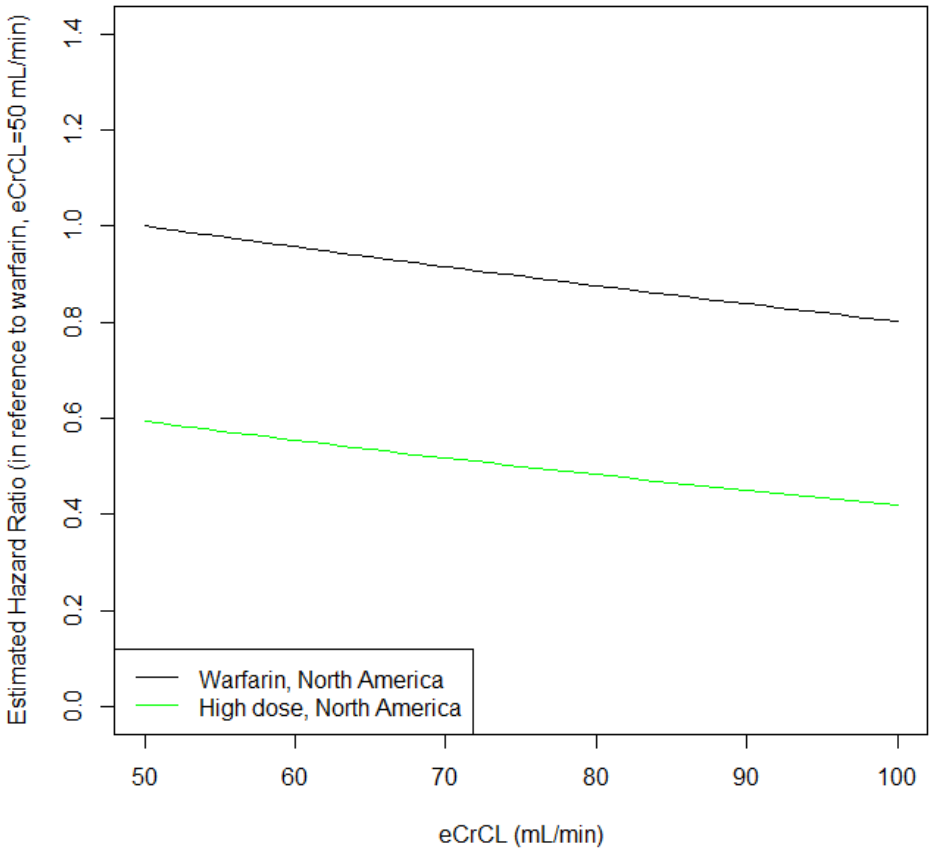


Source:FDA analysis.

Since we noticed that the effect of edoxaban varied across regions (Figure 7 and Figure 8), I wanted to look at the relative efficacy as a function of eCrCL across the regions. In particular, one region of interest to me was the North American region (mostly US subjects). To do this, I started with the proportional hazards model used in the primary analysis and added terms for eCrCL as a continuous variable, the 6 regions, and all the two-way interactions between region, dose and eCrCL. The estimated hazard ratio in the North American region from this model as a function of eCrCL is shown in Figure 12. This figure suggests that in North America, the high dose is consistently better than warfarin across the entire range of eCrCL. Furthermore, I did a similar analysis where I pooled the 3 regions where warfarin appeared to be better than edoxaban (Eastern Europe, Western Europe, and Japan) and did the same analysis. The hazard ratios for those regions as a function of eCrCL are shown in Figure 13. This figure suggests that in the

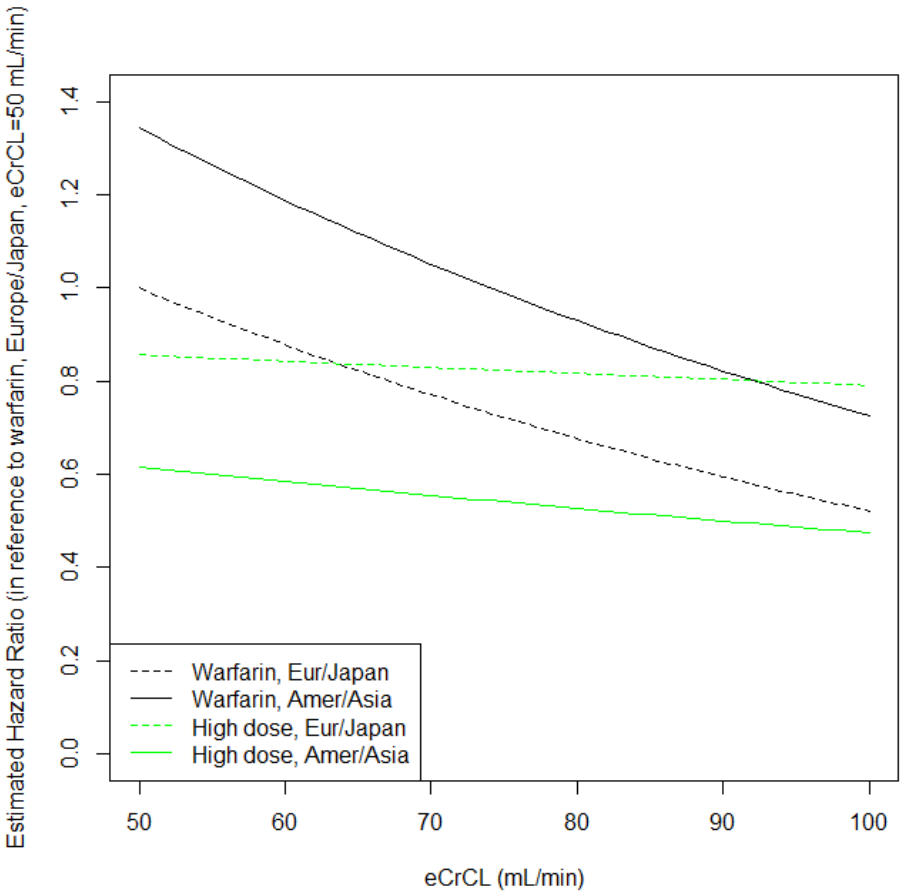
regions where warfarin performed well (dashed curves), warfarin did not do so well in the low eCrCL range, but improved with higher eCrCL (>65 mL/min). However, in the regions where warfarin did not do well (North America, Latin America, Asia/Pac and SA), edoxaban was consistently better across the entire range of eCrCL.

Figure 12 Estimated hazard ratios (primary endpoint) as a function of eCrCL in high dose/warfarin groups in North American region.



Source:FDA analysis.

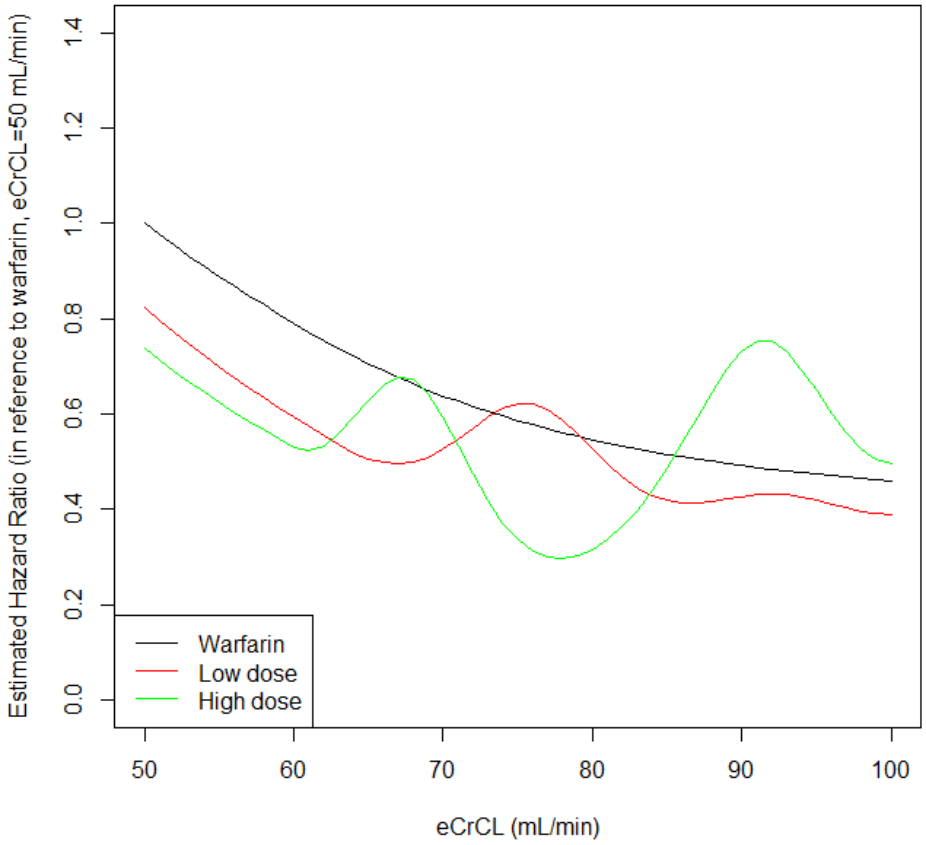
Figure 13 Estimated hazard ratios (primary endpoint) as a function of eCrCL in high dose/warfarin groups in different regions.



Source:FDA analysis.

In Figure 14, the hazard ratios for the endpoint of CV death are shown. I used the same model as I used to make Figure 11. Here, the low dose (red curve) seems to be consistently better than warfarin for the entire range of eCrCL. The high dose seems to be less effective for higher eCrCL. However, the evidence such as it is, suggests that an even higher dose would not decrease the rate of CV death for patients with higher eCrCL; a higher dose may even increase the risk of CV death.

Figure 14 Estimated hazard ratios (overall CV death) as a function of eCrCL.



Source:FDA analysis.

The cause of death in the subgroup with eCrCL>80 mL/min is shown in **Table 7**. This table is for the subgroup with eCrCL>80 mL/min and who had no dose adjustment indicated (i.e. weight >60 kg etc.). The table is divided into two sections. The top section includes people who had a history of stroke or TIA. The bottom section is for those with no history of stroke/TIA. The information in the table seems to suggest that there might be a tradeoff between different causes of death between the low dose and the high dose in this subgroup. However, it does not suggest that an even higher dose than the high dose used in the study (60 mg) could provide any further benefit on CV mortality and may do harm.

Table 7 Cause of death in subgroup with eCrCL>80 mL/min with no dose adjustment indicated.

A. eCrCL >80, no dose adjustment, history of stroke/TIA

	low (N=612)	high (N=594)	warfarin (N=614)
SUDDEN OR UNWITNESSED DEATH	14	17	17
CONGESTIVE HEART FAILURE/CARDIOGENIC SHOCK	2	7	4
OTHER	3	3	2
ISCHEMIC STROKE	7	4	3
INTRACRANIAL HEMORRHAGE	1	3	3
DYSRHYTHMIA	1	0	0
ATHEROSCLEROTIC VASCULAR DISEASE (EXCLUDING CORONARY)	0	0	1
DIRECTLY RELATED TO REVASCULARIZATION (CABG OR PCI)	0	0	0
NON-INTRACRANIAL HEMORRHAGE	1	0	0
PULMONARY EMBOLISM	1	0	1
SYSTEMIC ARTERIAL EMBOLIC EVENT	0	0	0

B. eCrCL>80, no dose adjustment, no history of stroke/TIA

	low (N=1879)	high (N=1894)	warfarin (N=1868)
SUDDEN OR UNWITNESSED DEATH	46	60	51
CONGESTIVE HEART FAILURE/CARDIOGENIC SHOCK	13	18	19
OTHER	7	9	5
ISCHEMIC STROKE	5	4	8
INTRACRANIAL HEMORRHAGE	0	9	6
DYSRHYTHMIA	7	2	3
ATHEROSCLEROTIC VASCULAR DISEASE (EXCLUDING CORONARY)	1	0	1
DIRECTLY RELATED TO REVASCULARIZATION (CABG OR PCI)	0	2	0
NON-INTRACRANIAL HEMORRHAGE	2	3	4
PULMONARY EMBOLISM	2	0	0
SYSTEMIC ARTERIAL EMBOLIC EVENT	0	0	0

Putting all of the analyses in this section together, the relative efficacy for the primary endpoint of a fixed dose of edoxaban (with dose adjustment as done in the study) to warfarin changes as a function of eCrCL. One possible explanation is that the effect of warfarin improves with higher eCrCL. Another possible theory is that the concentration of edoxaban is lower for people with higher eCrCL and therefore, they need a higher dose of edoxaban to achieve the same efficacy as their counterparts with lower eCrCL. It has not been proven that there exists any higher dose that

would be safe and effective. Secondly, even if there were, we do not know what dose would be high enough to achieve greater efficacy nor when to stop to maintain safety. Third, if we were to accept that some dose would be more safe and effective, there could be dosing errors caused by people using the wrong estimating equation (e.g. MDRD or CKD-EPI instead of Cockroft-Gault). Of note, in the ENGAGE trial, there were people who should have gotten a dose adjustment according to the protocol because their eCrCL was less than 50 mL/min but did not (and vice versa).

SUMMARY AND CONCLUSIONS

1.9 Statistical Issues

The major statistical issues are the margin used to test for noninferiority and the use of the on-treatment period for the analysis. For Atrial Fibrillation trials using warfarin as the active control and the endpoint of Stroke/SEE, the issue of the margin has been more or less settled. The margin of 1.38 has been used in other trials and the FDA has accepted this margin. The on-treatment analysis can be problematic because a subject can have a period of study drug interruption for several weeks and have an event during that period, which would not count in the analysis. I prefer the intent-to-treat approach even in a non-inferiority trial. If the results are significantly different (which could only happen if there was a substantial amount of time when people were interrupting study drug), I think it would be difficult to interpret. In this study, the results for the overall treatment period were nearly identical to the on-treatment analysis even though there was a substantial number of events that happened off treatment. So, there was no problem interpreting this trial as positive.

1.10 Collective Evidence

There was only one phase 3 trial for this indication in the submission. Two dose regimens were studied. Both regimens were safe and effective.

1.11 Conclusions and Recommendations

The 3 doses used in the study (15 mg, 30 mg, 60 mg) were safe and effective. This range of doses should be sufficient to provide doses for individual treatment needs and I think all should be approved. Most people should take the 60 mg dose with dose adjustment based on renal

function and other factors such as body weight and concomitant medications. The strategy for adjusting dose for individuals that was used in the trial may not be the best one.

1.12 Labeling Recommendations (as applicable)

NA.

APPENDICES

Tests for qualitative interaction with respect to superiority in a noninferiority study

Suppose a study is designed to show a test drug is noninferior to warfarin using a margin of 1.38 and there are two subgroups of interest. This discussion is about a hypothetical scenario and not necessarily anything specific to the ENGAGE-AF trial. The true log hazard ratios in the two subgroups are λ_1 and λ_2; the estimated log hazard ratios in the two subgroups are denoted by $\hat{\lambda}_1$ and $\hat{\lambda}_2$ and the standard errors are s_1 and s_2. The overall estimated treatment effect is $\hat{\lambda} = \frac{\frac{\hat{\lambda}_1}{{s_1}^2} + \frac{\hat{\lambda}_2}{{s_2}^2}}{\frac{1}{{s_1}^2} + \frac{1}{{s_2}^2}}$ and its standard error is $s = \left(\frac{1}{{s_1}^2} + \frac{1}{{s_2}^2}\right)^{-1/2}$ If the study has 90% power to show noninferiority, then the test statistic $\frac{\hat{\lambda} - \log(1.38)}{s}$ has expected value $\Phi^{-1}(0.1) + \Phi^{-1}(0.025) = -3.24$. Assuming that $\mathrm{E}\hat{\lambda} = 0$, this implies $\lambda_2 = \frac{-\lambda_1 {s_2}^2}{{s_1}^2}$ and $s_2 = \frac{(1-\log(1.38))s_1}{\sqrt{(3.24s_1)^2 - (1-\log(1.38))^2}}$. Now, a statistically significant qualitative interaction for superiority will happen when the test statistics in both subgroups are statistically significant at one-sided level 0.05 and have opposite signs. This means either $\left\{\frac{\hat{\lambda}_1}{s_1} < -z_{0.05} \text{ and } \frac{\hat{\lambda}_2}{s_2} > z_{0.05}\right\}$ or $\left\{\frac{\hat{\lambda}_1}{s_1} > z_{0.05} \text{ and } \frac{\hat{\lambda}_2}{s_2} < -z_{0.05}\right\}$ where $z_{0.05} \approx 1.645$ is the upper 0.05 quantile of the standard normal distribution. The unconditional probability of the events can be found because the statistics are independent and the events are disjoint,

$$\Phi\left(-z_{0.05} - \frac{\lambda_1}{s_1}\right)\left(1 - \Phi\left(z_{0.05} - \frac{\lambda_2}{s_2}\right)\right) + \left(1 - \Phi\left(z_{0.05} - \frac{\lambda_1}{s_1}\right)\right)\Phi\left(-z_{0.05} - \frac{\lambda_2}{s_2}\right).$$

Moreover, conditional on an overall positive result for noninferiority, the conditional probability of a statistically significant qualitative interaction will be

$$\frac{P\left[\left\{\left\{\frac{\hat{\lambda}_1}{s_1} < -z_{0.05} \,\&\, \frac{\hat{\lambda}_2}{s_2} > z_{0.05}\right\} \text{ or } \left\{\frac{\hat{\lambda}_1}{s_1} < -z_{0.05} \,\&\, \frac{\hat{\lambda}_2}{s_2} > z_{0.05}\right\}\right\} \,\&\, \frac{\hat{\lambda} - \log(1.38)}{s} < -z_{0.025}\right]}{P\left[\frac{\hat{\lambda} - \log(1.38)}{s} < -z_{0.025}\right]}$$

Since we are assuming the study has 90% power, the denominator is 0.9. The numerator can be expressed as the sum of two terms. The calculations below show how to find the first term; the other term in the numerator is similar to the first term with the indices 1 and 2 interchanged.

Let $x_1 = \frac{\hat{\lambda}_1}{s_1}$ and $x_2 = \frac{\hat{\lambda}_2}{s_2}$. The first term is

$$P\left[\frac{\hat{\lambda}_1}{s_1} < -z_{0.05} \,\&\, \frac{\hat{\lambda}_2}{s_2} > z_{0.05} \,\&\, \frac{\hat{\lambda} - \log(1.38)}{s} < -z_{0.025}\right]$$

$$= P\left[x_1 < -z_{0.05}\ \&\ x_2 > z_{0.05}\ \&\ \frac{\frac{x_1}{s_1}+\frac{x_2}{s_2}}{\frac{1}{{s_1}^2}+\frac{1}{{s_2}^2}} - \log(1.38) < -z_{0.025}\right]$$

$$= \int_{z_{0.05}}^{\infty} \Phi\left(min\left\{-z_{0.05}, \frac{s_1(\log(1.38)\, s_2 - s(sx_2 + s_2 z_{0.025}))}{s^2 s_2}\right\} - \frac{\lambda_1}{s_1}\right)\phi\left(x_2 - \frac{\lambda_2}{s_2}\right)dx_2$$

Now, we have to consider two cases:

Case 1) $s_2\left(\frac{\log(1.38)}{s^2} - \frac{z_{0.025}}{s} + \frac{z_{0.05}}{s_1}\right) < z_{0.05}$. In this case, over the range of integration $min\left\{-z_{0.05}, \frac{s_1(\log(1.38)s_2 - s(sx_2+s_2z_{0.025}))}{s^2 s_2}\right\} = \frac{s_1(\log(1.38)s_2 - s(sx_2+s_2z_{0.025}))}{s^2 s_2}$ and the integral is

$$\int_{z_{0.05}}^{\infty} \Phi\left(\frac{s_1(\log(1.38)\, s_2 - s(sx_2 + s_2 z_{0.025}))}{s^2 s_2} - \frac{\lambda_1}{s_1}\right)\phi\left(x_2 - \frac{\lambda_2}{s_2}\right)dx_2$$

Case 2) $s_2\left(\frac{\log(1.38)}{s^2} - \frac{z_{0.025}}{s} + \frac{z_{0.05}}{s_1}\right) > z_{0.05}$. In this case, the integral is

$$\int_{s_2\left(\frac{\log(1.38)}{s^2} - \frac{z_{0.025}}{s} + \frac{z_{0.05}}{s_1}\right)}^{\infty} \Phi\left(\frac{s_1(\log(1.38)\, s_2 - s(sx_2 + s_2 z_{0.025}))}{s^2 s_2} - \frac{\lambda_1}{s_1}\right)\phi\left(x_2 - \frac{\lambda_2}{s_2}\right)dx_2$$

$$+\Phi\left(-z_{0.05} - \frac{\lambda_1}{s_1}\right)\left(\Phi\left(s_2\left(\frac{\log(1.38)}{s^2} - \frac{z_{0.025}}{s} + \frac{z_{0.05}}{s_1}\right) - \frac{\lambda_2}{s_2}\right) - \Phi\left(z_{0.05} - \frac{\lambda_2}{s_2}\right)\right)$$

In the example shown in the R program below, the hazard ratios in the two subgroups are 0.77 and 1.3. This means the drug is truly noninferior in both subgroups (with respect to the margin 1.38). But, there is about a 37% chance of finding a qualitative interaction with respect to superiority conditional on an overall positive result for the study. The program calculates this two different ways; by simulation and also using the exact formulas derived above.

```
#R program to estimate conditional probability of qualitative interaction.

nsim=10000000
z025=qnorm(0.975)
z05=qnorm(0.95)
l1=-0.26
s1=0.14
s2=log(1.38)*s1/sqrt((3.24*s1)^2-(log(1.38))^2)
l2=-l1*s2^2/s1^2
s2
exp(l1)
exp(l2)
s=1/sqrt(1/s1^2+1/s2^2)
lam1=rnorm(nsim,l1,s1)
lam2=rnorm(nsim,l2,s2)
x1=lam1/s1
x2=lam2/s2
lamhat=((x1/s1+x2/s2)/(1/s1^2+1/s2^2))
mean(lamhat)                          #check the mean is 0
mean((lamhat-log(1.38))/s<(-z025)) #90% power for noninferiority

#unconditional prob of qual inter. using simulation and using exact formula
mean((x1>z05 & x2<(-z05)) | (x2>z05 & x1<(-z05)))
pnorm(-z05-l1/s1)*(1-pnorm(z05-l2/s2))+(1-pnorm(z05-l1/s1))*pnorm(-z05-l2/s2)


#this section calculates the two integrals in the numerator
f1=function(x, l1,l2,s1,s2,s,delta,z025) {
pnorm((s1*(log(1.38)*s2 - s*(s*x + s2*z025)))/(s^2*s2)-l1/s1)*dnorm(x-l2/s2)}

if (s2*(log(1.38)/s^2-z025/s+z05/s1)<z05) i1= integrate(f1,lower= z05,upper=Inf,
l1=l1,l2=l2,s1=s1,s2=s2,s=s,delta=log(1.38),z025=z025)$val else i1=pnorm(-z05-l1/s1)*
(pnorm(s2*(log(1.38)/s^2 - z025/s + z05/s1)-l2/s2)-pnorm(z05-l2/s2))+
integrate(f1,lower= s2*(log(1.38)/s^2 - z025/s + z05/s1),upper=Inf,
l1=l1,l2=l2,s1=s1,s2=s2,s=s,delta=log(1.38),z025=z025)$val

f2=function(x, l1,l2,s1,s2,s,delta,z025) {
pnorm((s2*(log(1.38)*s1 - s*(s*x + s1*z025)))/(s^2*s1)-l2/s2)*dnorm(x-l1/s1)}

if (s1*(log(1.38)/s^2-z025/s+z05/s2)<z05) i2= integrate(f2,lower= z05,upper=Inf,
```

```
l2=l2,l1=l1,s2=s2,s1=s1,s=s,delta=log(1.38),z025=z025)$val else i2=pnorm(-z05-l2/s2)*
(pnorm(s1*(log(1.38)/s^2 - z025/s + z05/s2)-l1/s1)-pnorm(z05-l1/s1))+
integrate(f2,lower= s1*(log(1.38)/s^2 - z025/s + z05/s2),upper=Inf,
l2=l2,l1=l1,s2=s2,s1=s1,s=s,delta=log(1.38),z025=z025)$val

#conditional prob. of qual. int. using simulation and exact formula
mean(((x1>z05 & x2<(-z05)) | (x2>z05 & x1<(-z05))) & (lamhat-log(1.38))/s<(-z025))/0.9
(i1+i2)/0.9
```

Fitting parametric distributions

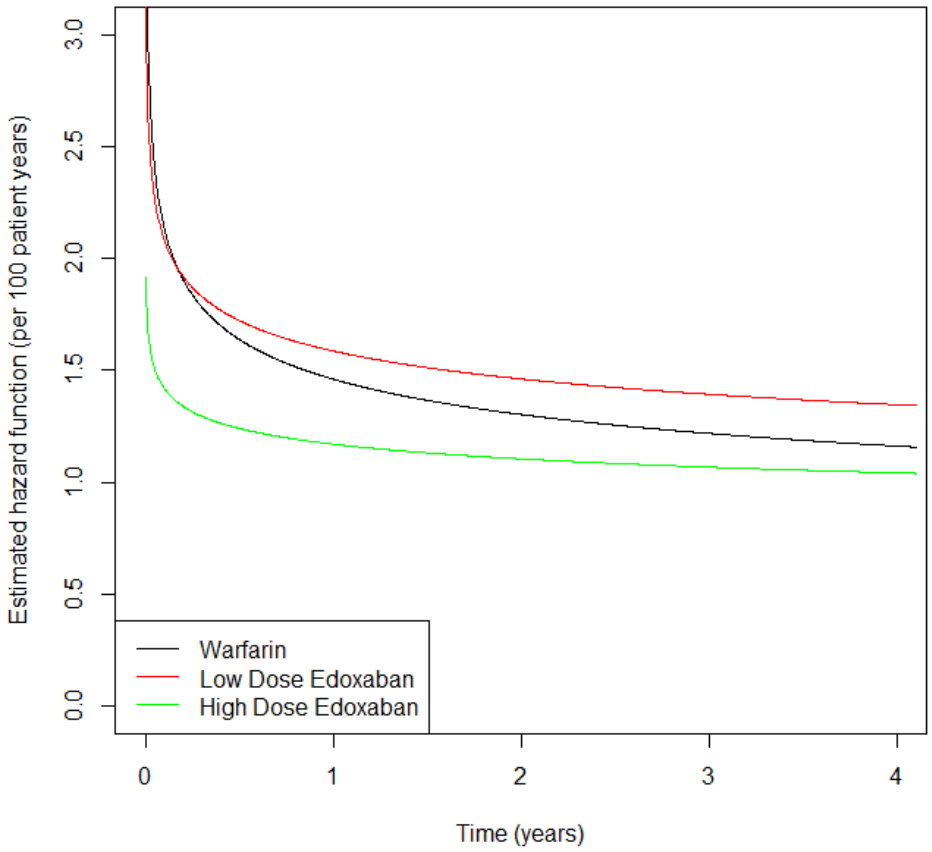
The cumulative distribution function is denoted by $F(x)$, the density is $f(x)$, the hazard function is $h(x) = \frac{f(x)}{1-F(x)}$, and cumulative hazard function, $H(x) = -log\{1 - F(x)\}$. Suppose subject *i* has exposure at time intervals $(s_{i,1}, t_{i,1}), (s_{i,2}, t_{i,2}), \ldots, (s_{i,m_i}, t_{i,m_i})$ where $0 = s_{i,1} \leq t_{i,1} < s_{i,2} \leq t_{i,2} < \cdots < s_{i,m_i} \leq t_{i,m_i}$, the number of subjects is *n* and they have been ordered so that the first *D* subjects have events and the remainder do not. Then the log-likelihood is

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\{H(s_{i,j}) - H(t_{i,j})\} + \sum_{i=1}^{D} log\{h(t_{i,m_i})\}$$

For the Weibull family, the density is $f(x) = \frac{a}{b}\left[\frac{x}{b}\right]^{a-1} e^{-(x/b)^a}$. The best fitting parameters for each group are shown below and the estimated hazard functions are in Figure A1 that follows.

Group	*a*	*b*
warfarin:	0.835	127
low:	0.882	95.3
high:	0.916	116

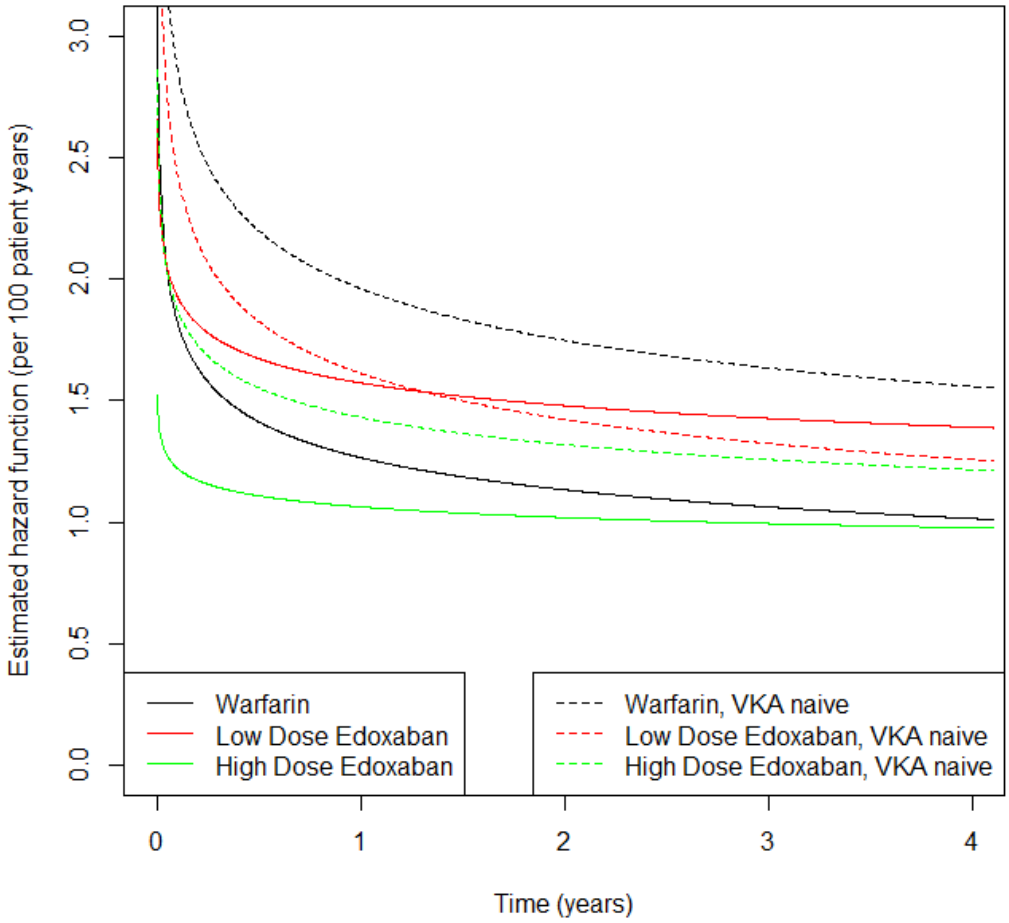
Figure A1. Estimated hazard functions using the Weibull distribution (primary endpoint, mITT, on treatment).



Source: FDA analysis

The next figure (A2) shows the estimated hazard functions in each group separated by VKA use 30 within 30 days before randomization. As in Figure 6, this figure suggests that the subjects who were VKA naive had a higher hazard rate, particularly in the warfarin group.

Figure A2. Estimated hazard functions using the Weibull distribution by VKA status (primary endpoint, mITT, on treatment). Solid curves are for the subgroups of patients who were not VKA naive, dashed curves are for VKA naive subgroups.



Source: FDA analysis

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOHN P LAWRENCE
09/22/2014

HSIEN MING J HUNG
09/22/2014

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 206316
Supplement #:	S-0000
Drug Name:	Edoxaban
Indication(s):	Deep vein thrombosis (DVT) and pulmonary embolism (PE)
Applicant:	Daiichi Sankyo Pharma Development
Date(s):	Submission Date: 8 January, 2014
	PDUFA due Date: 8 January 2015
	Review Completion Date: 8 September 2014
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V
Statistical Reviewer:	Yun Wang, PhD
Concurring Reviewers:	Lei Nie, PhD, Team Leader
	Rajeshwari Sridhara, PhD, Division Director
Medical Division:	Office of Hematology and Oncology Product
Clinical Team:	Saleh Ayache, MD
	Kathy Robie Suh, MD, PhD, Team Leader
Project Manager:	Janet Higgins, RPM

Keywords: Deep Vein Thrombosis, Pulmonary Embolism, venous thrombosis embolism, non-inferiority.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Deep vein thrombosis (DVT) and pulmonary embolism (PE), often collectively referred to as venous thromboembolism (VTE), are associated with a high morbidity and may progress to fatal outcome if left untreated. Per Applicant, Edoxaban (also referred to as Savaysa®) is a Factor Xa (FXa) inhibitor, and inhibition of FXa in the coagulation cascade is expected to reduce thrombin formation and prolongs clotting time.

In this New Drug Application (NDA) submission, the applicant seeks full approval of Edoxaban for the treatment of DVT and PE (b) (4) and for reduction of the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation. This review provides statistical evaluation of Edoxaban for the treatment of DVT and PE only. Please refer to Dr. John Lawrence's review for statistical evaluation of Edoxaban for reduction of the risk of stroke and systemic embolism.

The support of Edoxaban for the treatment of DVT and PE was based on one pivotal trial, Study Hokusai VTE (DU176b-D-U305), which was a Phase III, randomized, multi-center, double-blind, double-dummy, and parallel-group study with two parallel treatment groups: (low molecular weight [LMW]) Heparin/Edoxaban and [LMW] Heparin/Warfarin. The primary efficacy objective of the Study Hokusai VTE was to evaluate whether initial [LMW] Heparin followed by Edoxaban([LMW] Heparin/Edoxaban) was non-inferior to initial [LMW] Heparin overlapping with Warfarin, followed by Warfarin ([LMW] Heparin/Warfarin) in the treatment of subjects with acute symptomatic VTE.

Study Hokusai VTE randomized 8292 patients, 4143 to Heparin/Edoxaban arm and 4149 to Heparin/Warfarin arm respectively. Primary efficacy analysis was based on modified intent-to-treat (mITT) population, which consisted of 8240 patients who received at least one dose of study treatment. Non-inferiority was demonstrated in the primary efficacy endpoint, time to symptomatic recurrent VTE or VTE-related death, for patients treated with Heparin/Edoxaban versus Heparin/Warfarin. The estimated hazard ratio (HR) for time to symptomatic recurrent VTE or VTE-related death was 0.89 (95% confidence interval: 0.70 – 1.13) for the Heparin/Edoxaban arm versus Heparin/Warfarin arm based on 276 recurrent VTE or VTE-related death. The upper 95% confidence limit of 1.13 demonstrated that treatment with Heparin/Edoxaban retained at least 91% treatment effect of Heparin/Warfarin. The median time to symptomatic recurrent VTE or VTE-related death was not reached in either treatment arm.

The estimated HR for primary safety endpoint, time to major or clinically relevant non-major bleeding (CRNM), was 0.81 (95% confidence interval: 0.71-0.94) based on 772 major/CRNM events, median time to major/CRNM bleeding was not reached for either treatment arm.

This statistical reviewer believes the efficacy and Safety data from Study Hokusai VTE support the claim of non-inferiority of Edoxaban compared to Warfarin for the treatment of recurrent VTE (b) (4)

2 INTRODUCTION

2.1 Overview

The classical management of subjects with VTE consists of an initial treatment with bodyweight-adjusted subcutaneous (SC) low molecular weight (LMW) Heparin, adjusted-dose intravenous (IV) or fixed dose SC unfractionated Heparin (UFH),or bodyweight-adjusted subcutaneous fondaparinux, followed by long-term treatment with an oral vitamin K antagonist (VKA). However, the use of VKAs is complicated by several inherent problems, including a delayed onset of anticoagulant action; a narrow therapeutic index that requires close laboratory monitoring using the INR; an unpredictable and variable pharmacological response; and food and drug interactions requiring frequent dosage adjustment. Warfarin is the most commonly used VKA.

According to the sponsor, Edoxaban is an orally active, selective, reversible FXa inhibitor that has more predictable pharmacokinetics (PK), does not require monitoring for blood tests, and has fewer drug/food interactions than Warfarin.

One of the proposed indications submitted in the NDA application and reviewed in this document is for the treatment of DVT and PE (b) (4) Symptoms of DVT include erythema, warmth, pain, swelling, tenderness, and pain upon dorsiflexion of the foot. Symptoms of PE include sudden onset dyspnea, tachypnea, tachycardia, syncope, hypotension, and hypoxemia.

Study Hokusai VTE was a Phase III, randomized, multi-center, double-blind, double-dummy, study with two parallel treatment groups: Heparin/Edoxaban and Heparin/Warfarin. The primary efficacy endpoint was time to the first occurrence of symptomatic recurrent VTE or VTE-related death. All VTE events were adjudicated by clinical events committee (CEC). The secondary efficacy endpoint was time to first occurrence of recurrent VTE and all-cause mortality.

A total of 8292 patients with acute symptomatic VTE were enrolled between 28 January 2010 and 05 October 2012 from 439 sites in 37 countries. The data cut-off date was 01 September 2013. Among the enrolled 8292 patients, 8240 actually received at least one dose of study treatment, which were considered as the mITT population and safety analysis set. The efficacy and safety analyses were based on mITT population and safety analysis set respectively.

The original protocol for Study Hokusai was dated 24 August 2009, and the latest version, Amendment 5, was dated 16 April 2012.

Throughout this review, patients received [LMW] Heparin/Edoxaban or [LMW] Heparin/Warfarin, are referred as “Edoxaban” arm or “Warfarin” arm respectively in the text and the tables/figures.

TABLE 1: LIST OF ALL STUDIES INCLUDED IN ANALYSIS

Study	Phase and Design	Treatment Period	Follow-up Period	# of Subjects	Enrollment period Geographic region
Hokusai VTE	Phase 3, randomized, multi-center, double-blind, double-dummy, and parallel-group study with two parallel treatment groups: [LMW] Heparin/Edoxaban and [LMW] Heparin/Warfarin for patients with acute symptomatic VTE	Treatment until completion of planned treatment duration of 12 months or any other reason listed in the protocol for mandatory withdrawal.	All subjects were expected to be followed for 12 months after randomization even if study drug has been temporarily interrupted or permanently discontinued.	N=8292	28 January 2010 – 05 October 2012 439 sites in 37 countries

2.2 Data Sources

Analysis datasets, SDTM tabulations, and software codes are located on network with network path: \\CDSESUB1\EVSPROD\NDA206316\206316.enx.

3 STATISTICAL EVALUATION

This statistical evaluation is based on data from the pivotal Study Hokusai VTE.

3.1 Data and Analysis Quality

Data from the pivotal Study Hokusai VTE were provided electronically with standard formats. The primary efficacy and safety endpoints were derived and saved in analysis datasets "ADJEFF" and "ADJSAF" respectively. Documentations on datasets and programming were included with sufficient details for this reviewer to reproduce the applicant's key efficacy and safety results.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

3.2.1.1 Study Design

The pivotal trial Hokusai VTE was a Phase III, randomized, multi-center, double-blind, double-dummy, study with two parallel treatment groups: Edoxaban and Warfarin. Approximately 7500 patients were planned to be randomized 1:1 to the 2 treatment arms via an interactive voice/web response system (IXRS). Eligible subjects were stratified by presenting diagnosis: PE with or without DVT vs. DVT only. Within each diagnostic stratum, eligible subjects were further

stratified by baseline risk factors (a. temporary risk factors only [such as trauma, surgery, immobilization, estrogen therapy, etc.] vs. b. all others), and need for adjustment (body weight ≤ 60 Kg; creatinine clearance [CrCL] between 30 and 50 mL/min inclusive, and concomitant use of the P-gp inhibitors verapamil or quinidine).

The primary objective of the Study Hokusai VTE was to evaluate whether Edoxaban was non-inferior to Warfarin in the treatment of subjects with acute symptomatic VTE.

The study was designed to accumulate approximately 220 symptomatic recurrent VTE events in the mITT analysis set. This design would have a power of 85% and type I error of 0.05 to demonstrate that Edoxaban is non-inferior to Warfarin, with a non-inferiority margin for the hazard ratio of 1.5. Assuming an incidence rate of 3.0% for symptomatic recurrent VTE during the study period of 12 months, 7500 subjects were expected to be randomized.

Non-inferiority margin was derived based on indirect confidence interval comparison method. This method focused on identifying the maximally acceptable loss of active treatment benefit. Active treatment benefit was defined as the difference in treatment effect between available "more effective" treatment and "less effective" treatment, such as placebo or no treatment. Based on 14 historical studies, the odds ratio for available "more effective" treatment in comparison to "less effective" treatment was 0.18 (95% CI: 0.14 to 0.25). Considering the upper 95% confidence limit of 0.25 as the active treatment benefit, non-inferiority margin would be (1/0.25)^(1-0.7) = 1.5 to retain at least 70% of available treatment benefit and (1/0.25)^(1-0.9) = 1.15 to retain at least 90% of available treatment benefit.

3.2.1.2 Efficacy Endpoints

The primary efficacy endpoint was time to first symptomatic recurrent VTE and VTE-related death (i.e., the composite of DVT, non-fatal PE, and fatal PE), which was defined as time from the day of randomization to the first symptomatic recurrent VTE and VTE-related death experienced by a subject during the 12-month study period. Subjects who did not have a primary efficacy outcome during the 12-month study period would be censored at Day 365 or the last day the subject had a complete assessment for the study outcome, whichever came first. All events were adjudicated by CEC.

The secondary efficacy endpoint was time to composite of symptomatic recurrent DVT, non-fatal symptomatic recurrent PE, and all-cause mortality. Similar definition and censoring rules for the primary efficacy endpoint were applied to the secondary efficacy endpoint.

3.2.2 Statistical Methodologies

The primary efficacy endpoint was analyzed using un-stratified Cox proportional hazard model. The hazard ratio and corresponding 95% confidence interval (CI) were estimated.

The sponsor performed two sensitivity analyses of primary efficacy endpoint:

- An analysis based on the per-protocol population with the "on-treatment" approach. Per-protocol population was defined as randomized subjects with a CEC confirmed baseline

diagnosis of VTE and received at least one dose of randomized study treatment. On treatment was defined as the time period the subject was taking study drug up to 3 days after interruption or stopping study treatment.

- An analysis based on the per-protocol population with the "treatment + 30 days" approach. Treatment + 30 days were defined as the time period from randomization to up to 30 days after last dose of study drug.

The secondary efficacy endpoint was compared between Edoxaban and Warfarin using un-stratified log-rank test. The hazard ratio and corresponding 95% confidence interval (CI) were estimated for secondary efficacy endpoint using un-stratified Cox proportional hazard model.

The significance level (α) for the non-inferiority test for the primary efficacy analysis was 0.05. If non-inferiority was established for the primary efficacy endpoint, the secondary efficacy endpoint would be tested for superiority using a significance level of 0.01.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Analysis Population

Intent-to-treat (ITT) population was defined as all randomized subjects. ITT population was used for descriptions of patient disposition. Modified ITT population (mITT) was defined as all randomized subjects who received at least one dose of study treatment. mITT population was the primary analysis population for all efficacy analyses, and was used for descriptions of demographics, and baseline disease characteristics.

Per-protocol (PP) population was defined as randomized subjects with a CEC confirmed baseline diagnosis of VTE and received at least one dose of randomized study treatment.

The safety population was defined the same as mITT population. All safety analyses were based on the safety population.

Study Hokusai VTE randomized 8292 subjects with acute symptomatic VTE, 4143 to Edoxaban arm and 4149 to Warfarin arm respectively, from 439 sites in 37 countries. Twenty-five subjects in Edoxaban arm and 27 subjects in Warfarin arm did not receive study treatment. Therefore, mITT and safety population consisted of 8240 subjects, 4118 in Edoxaban arm and 4122 in Warfarin arm.

Subject Disposition

At the time of study cutoff of September 01, 2013, 181 (4.4%) treated subjects in Edoxaban arm and 167 (4.1%) treated subjects in Warfarin arm did not complete study treatment. The most common reason for not completing study treatment was death in both treatment arms (3.3% in Edoxaban arm and 3.1% in Warfarin arm, respectively).

TABLE 2: SUBJECT DISPOSITION, ITT POPULATION

	Edoxaban (N=4143) n (%)	**Warfarin (N=4149) n (%)**
All randomized	4143 (100)	4149 (100)
Never Treated	25 (0.6)	27 (0.7)
Treated (mITT, Safety population)	4118 (99.4)	4122 (99.3)
Completed study	3937 (95.6)	3955 (95.9)
Full 12 month follow-up	3058 (74.3)	3074 (74.6)
<12 month follow-up due to end of study	879 (21.3)	881 (21.4)
Did not complete study	181 (4.4)	167 (4.1)
Death	136 (3.3)	127 (3.1)
Withdrew consent	32 (0.8)	33 (0.8)
Lost to follow-up	7 (0.2)	4 (0.1)
Other	6 (0.1)	3 (<0.1)

[Source: Study Hokusai VTE CSR Pages 81 Table 10.1]

Subject Demographics and Baseline Disease Characteristics

Subject demographics appeared to be balanced between Edoxaban and Warfarin arms (Table 3).

TABLE 3: DEMOGRAPHICS, MITT POPULATION

	Edoxaban (N=4118)	Warfarin (N=4122)	Total (N=8240)
Age (years)			
Mean (SD)	55.7 (16.3)	55.9 (16.2)	55.8 (16.2)
Median	57.0	57.0	57.0
Range	(18.0, 106.0)	(18.0, 95.0)	(18.0, 106.0)
Category, n (%)			
< 65	2784 (67.6)	2752 (66.8)	5536 (67.2)
≥ 65	1334 (32.4)	1370 (33.2)	2704 (32.8)
Sex, n (%)			
Male	2360 (57.3)	2356 (57.2)	4716 (57.2)
Female	1758 (42.7)	1766 (42.8)	3524 (42.8)
Race, n (%)			
White	2867 (69.6)	2895 (70.2)	5762 (69.9)
Black	156 (3.8)	144 (3.5)	300 (3.6)
Asian	866 (21.0)	861 (20.9)	1727 (21.0)
Other	229 (5.6)	222 (5.4)	451 (5.5)
Region, n (%)			
North American	416 (10.1)	420 (10.2)	836 (10.2)
Western Europe	1396 (33.9)	1394 (33.8)	2790 (33.9)
Eastern Europe	911 (22.1)	913 (22.2)	1824 (22.1)
Asian	850 (20.6)	847 (20.6)	1697 (20.6)
Other	545 (13.2)	548 (13.3)	1093 (13.3)

SD: standard deviation;

[Source: Study Hokusai VTE CSR Page 83 Table 10.2 and statistical reviewer's analysis]

Reviewer's note:

- The region used by the applicant was not consistent with the common definition of regions in clinical trials. The statistical reviewer derived the region according to commonly used definition.
- Race information was missing for 20 subjects (9 in Edoxaban arm and 11 in Warfarin arm), these subjects were included in race "Other" group.

Baseline disease characteristics are summarized in Table 4. There were 4890 (59.3%) subjects with diagnosis of DVT only and 3350 (40.7%) subjects with diagnosis of PE at baseline based on IVXS. There were 2272 (27.6%) subjects with temporary risk factors, such as trauma, surgery, immobilization, estrogen therapy etc. Edoxaban dose was adjusted to 30mg at randomization for 1452 (17.6%) subjects due to body weight ≤ 60 Kg, creatinine clearance between 30 and 50 mL/min inclusive, or concomitant use of the P-gp inhibitors verapamil or quinidine.

There were 3319 subjects with an index PE (with or without DVT) and 4921 patients with an index DVT only confirmed by CEC adjudication or by the investigator if CEC could not adjudicate. For PE subjects, PE severity was assessed by protocol-specified assessments of baseline anatomic extent, baseline serum NT-proBNP, and baseline right ventricular (RV) dysfunction. Of 3319 subjects with an index PE, 3133 (94.4%) had anatomic extent of the PE assessed, 2989 (90.1%) had NT-proBNP at baseline assessed, and 1002 (30.2%) had RV dysfunction at baseline assessed. Using these three techniques the proportion of subjects identified as having more severe PE for both Edoxaban and Warfarin subjects was 47.9% and 49.1% by extensive anatomic extent, 30.6% and 32.2% by NT-proBNP ≥ 500 pg/ml, and 34.5% and 35.5% by RV dysfunction present at baseline, respectively.

TABLE 4: BASELINE DISEASE CHARACTERISTICS, MITT POPULATION

	Edoxaban (N=4118) n (%)	**Warfarin (N=4122) n (%)**	**Total (N=8240) n (%)**
Presenting diagnosis (IVXS)			
Pulmonary Embolism	1671 (40.6)	1679 (40.7)	3350 (40.7)
With DVT	611 (14.8)	560 (13.6)	1171 (14.2)
Without DVT	1060 (25.7)	1119 (27.1)	2179 (26.4)
DVT only	2447 (59.4)	2443 (59.3)	4890 (59.3)
Risk factors			
Temporary	1132 (27.5)	1140 (27.7)	2272 (27.6)
Other	2986 (72.5)	2982 (72.3)	5968 (72.4)
Intended treatment duration*			
3 months	221 (5.4)	245 (6.0)	466 (5.7)
6 months	1555 (37.8)	1502 (36.5)	3057 (37.1)
12 months	2339 (56.8)	2371 (57.6)	4710 (57.2)
Edoxaban 30 mg at randomization**			
Yes	733 (17.8)	719 (17.4)	1452 (17.6)
No	3385 (82.2)	3403 ()82.6)	6788 (82.4)

	Edoxaban (N=4118) n (%)	Warfarin (N=4122) n (%)	Total (N=8240) n (%)
Weight at randomization (kg)			
<= 60	524 (12.7)	519 (12.6)	1043 (12.7)
> 60	3594 (87.3)	3603 (87.4)	7197 (87.3)
Creatinine clearance at randomization (mL/min)			
>= 30 to <=50	268 (6.5)	273 (6.6)	541 (6.6)
> 50	3850 (93.5)	3849 (93.4)	7699 (93.4)
Verapamil or quinidine use at randomization			
Yes	26 (0.6)	25 (0.6)	51 (0.6)
No	4092 (99.4)	4097 (99.4)	8189 (99.4)
Index diagnosis			
PE	1650 (40.1)	1669 (40.5)	3319 (40.3)
DVT	2468 (59.9)	2453 (59.5)	4921 (59.7)
Anatomic extent of the PE	1550	1583	3133
Limited	128 (8.3)	123 (7.8)	251 (8.0)
Intermediate	679 (43.8)	682 (43.1)	1361 (43.4)
Extensive	743 (47.9)	778 (49.1)	1521 (48.6)
NT-ProBNP at baseline	1484	1505	2989
< 500 pg/mL	1030 (69.4)	1021 (67.8)	2051 (68.6)
>= 500 pg/mL	454 (30.6)	484 (32.2)	938 (31.4)
RV dysfunction at baseline	498	504	1002
Yes	172 (34.5)	179 (35.5)	351 (65.0)
No	326 (65.5)	325 (64.5)	651 (35.0)

IXRS: interactive voice/web response system; DVT: deep vein thrombosis; PE: pulmonary embolism; NT-proBNP: N-terminal pro brain naturetic Peptide; RV: right ventricular.

*: Mitigating factors related to the subject's clinical status could influence the total duration of treatment a given subject actually received with intended treatment duration of 3, 6, and 12 months as determined by the Investigator.

**: At randomization, subjects with low body weight (≤ 60kg), and moderate renal impairment (CrCL 30 to 50 ml/min), or taking pre-specified concomitant medications (e.g. verapamil, quinidine) in the Edoxaban group received active Edoxaban 30 mg (and placebo Warfarin) while subjects in the Warfarin group with the same low body weight, moderate renal impairment, or pre-specified concomitant medications received placebo Edoxaban 30 mg

[Source: Study Hokusai VTE CSR pages 83, 84 Table 10.2 and statistical reviewer's analysis]

Reviewer's note:

- Based on the submitted data set, the numbers of subjects with < 500 or >=500 pg/mL N-terminal pro brain naturetic Peptide (NT-ProBNP) at baseline derived by the statistical reviewer in Table 4 were slightly different from what were summarized by the applicant in Table 10.2.

Treatment with Warfarin and Heparin Prior to Randomization

There were 701 (8.6%) subjects who had at least 1 dose of Warfarin and 6729 (81.7%) subjects used Heparin within 2 days prior to randomization (Table 5).

TABLE 5: PRIOR ANTICOAGULANT THERAPIES, MITT POPULATION

	Edoxaban (N=4118) n (%)	**Warfarin (N=4122) n (%)**	**Total (N=8240) n (%)**
Warfarin use within 2 days prior to randomization			
No dose taken	3794 (92.1)	3745 (90.9)	7539 (91.5)
1 dose taken	279 (6.8)	319 (7.7)	598 (7.3)
> 1 Dose taken	45 (1.1)	58 (1.4)	103 (1.3)
(LMW) Heparin use within 2 days prior to randomization			
None	755 (18.3)	756 (18.3)	1511 (18.3)
<= 2 days duration	3260 (79.2)	3262 (79.1)	6522 (79.2)
> 2 days duration	103 (2.5)	104 (2.5)	207 (2.5)

[Source: Study Hokusai VTE CSR Page 92 Table 10.7 and statistical reviewer's analysis]

Protocol Deviation

Major protocol deviations were defined as follows in the statistical analysis plan (SAP):

- Subjects having thrombectomy, insertion of a cava filter, or the use of a fibrinolytic agent to treat the current [index] episode of DVT and/or PE.
- Subjects having an indication for Warfarin other than DVT and/or PE (i.e., subjects receiving non-study Warfarin for an indication other than DVT and/or PE).
- Subjects who had more than 48 hours pre-treatment with therapeutic dosages of anticoagulant treatment, or more than a single dose of VKA prior to randomization to treat the current [index] episode.
- Subjects having a treatment misallocation at any time during the study.

Other notable protocol deviations included:

- Subjects who received disallowed concomitant medication that impacted the evaluation of primary endpoints for efficacy and safety.
- Violation of inclusion/exclusion criteria, such as CrCL < 30 mg/
- Subjects at any site for which subject data authenticity was suspect and cannot be confirmed.

A total of 1896 subjects (23.0%) (953 [23.1%] in Edoxaban arm and 943 [22.9%] in Warfarin arm) had at least one notable protocol deviation, among that ~5% were major protocol deviations in both treatment arms (Table 6).

TABLE 6: SUBJECTS WITH NOTABLE PROTOCOL VIOLATIONS, MITT POPULATION

	Edoxaban (N=4118) n (%)	Warfarin (N=4122) n (%)	Total (N=8240) n (%)
Subjects with at least 1 notable protocol violation	953 (23.1)	943 (22.9)	1896 (23.0)
Subjects having a treatment misallocation at any time during the study	2 (<0.1)	0 (0.0)	2 (<0.1)
Subjects having thrombectomy, insertion of a caval filter, or the use of a fibrinolytic agent to treat the index episode of DVT and/or PE	16 (0.4)	27 (0.7)	43 (0.5)
Subjects receiving non-study Warfarin for an indication other than DVT and/or PE	28 (0.7)	21 (0.5)	49 (0.6)
Subjects who had more than 48 hours pre-treatment with therapeutic dosages of anticoagulant treatment or more than a single dose of VKA prior to randomization to treat the index episode	162 (3.9)	168 (4.1)	330 (4.0)
Subjects who received disallowed concomitant medications that impacts the evaluation of primary endpoints for efficacy or safety:			
NSAIDs	651 (15.8)	640 (15.5)	1291 (15.7)
Aspirin use > 100mg qd	39 (0.9)	29 (0.7)	68 (0.8)
Dual antiplatelet therapy	30 (0.7)	20 (0.5)	50 (0.6)
Any other prohibited medications during the study	75 (1.8)	65 (1.6)	140 (1.7)
Any other prohibited medications at randomization	68 (1.7)	76 (1.8)	144 (1.7)
CrCL <30 mL/min at randomization	10 (0.2)	10 (0.2)	20 (0.2)
Subjects at sites for which subject data authenticity is suspect and cannot be confirmed	16 (0.4)	8 (0.2)	24 (0.3)

DVT: deep vein thrombosis; PE: pulmonary embolism; VKA: vitamin K antagonist; NSAIDs: non-aspirin non-steroidal anti-inflammatory drugs; CrCL: creatinine clearance.

[Source: Study Holusai VTE CSR Table 14.1.2.1]

3.2.4 Results and Conclusions

3.2.4.1 Primary efficacy analysis results

The primary efficacy analysis was based on 276 recurrent VTE or VTE-related death observed by the study cutoff date. The primary efficacy analysis results are summarized in Table 7 and Figure 1. The estimated hazard ratio (HR) for time to symptomatic recurrent VTE or VTE-related death was 0.89 (95% confidence interval: 0.70 – 1.13) for the Edoxaban arm versus Warfarin arm. The upper 95% confidence limit of 1.13 demonstrated, with a high confidence level, that treatment with Edoxaban retained at least 91% treatment effect of Warfarin. Therefore, non-inferiority was demonstrated in the primary efficacy endpoint for patients treated with Edoxaban versus Warfarin. The median time to symptomatic recurrent VTE or VTE-related death was not reached in either treatment arm.

TABLE 7: PRIMARY EFFICACY ANALYSIS RESULTS, MITT POPULATION

Primary efficacy endpoint	**Edoxaban (N=4118)**	**Warfarin (N=4122)**
Subjects with recurrent VTE or VTE-related death, n (%)	130 (3.2)	146 (3.5)
PE with/without DVT, n (%)	73 (1.8)	83 (2.0)
Fatal PE, n (%)	24 (0.6)	24 (0.6)
DVT only, n (%)	57 (1.4)	63 (1.5)
Un-stratified Hazard Ratio (95% CI)	0.89 (0.70, 1.13)	
Nominal P value for non-inferiority	< 0.0001	

- CI: confidence interval;

- P value from asymptotic normal test.

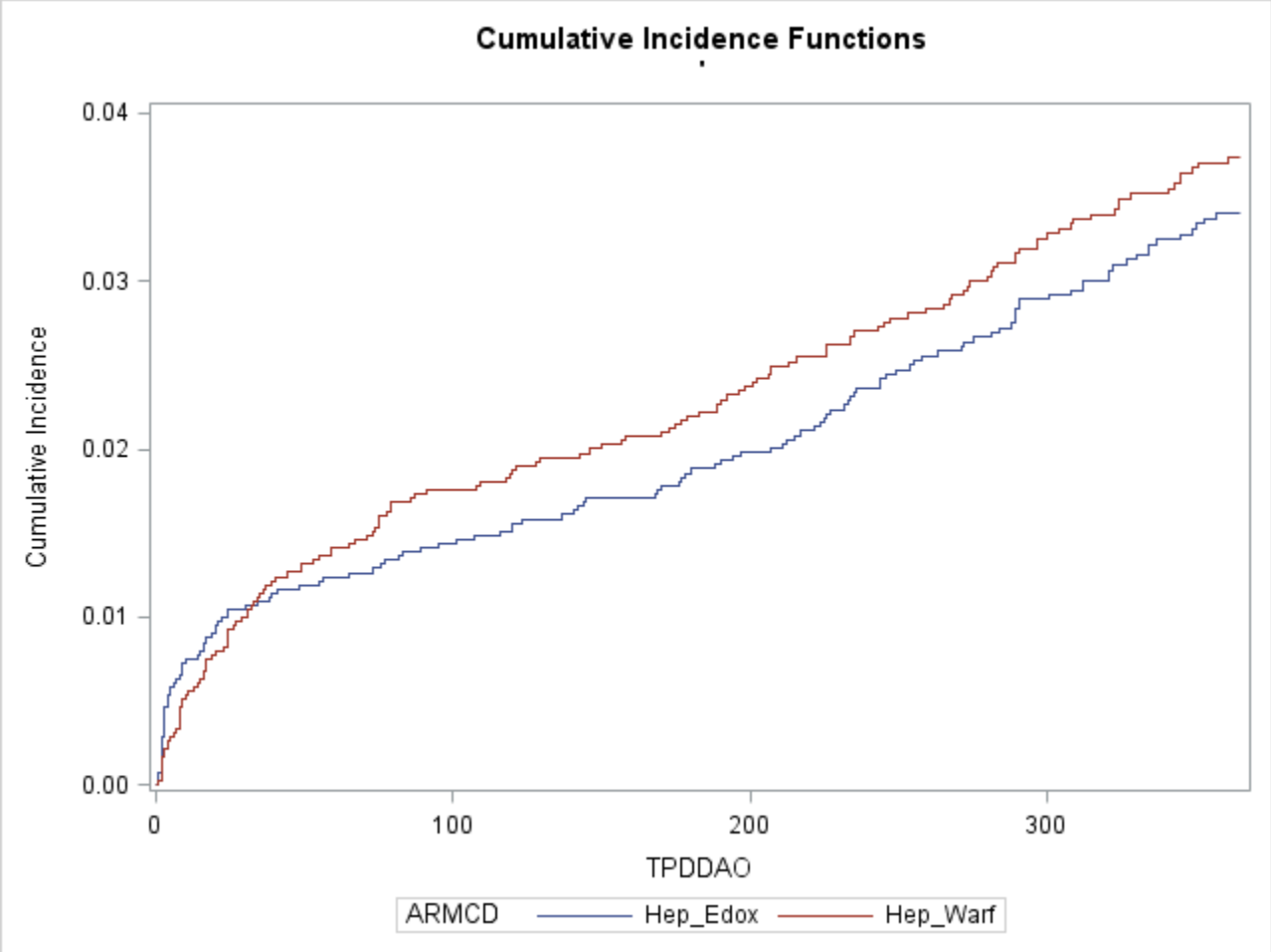
- Hazard ratio is from un-stratified proportional hazard model. Hazard ratio < 1 favors Edoxaban arm.

[Source: Study Hokusai CSR Page 99 Table 11.2 and statistical reviewer's analysis]

Reviewer's note:

- The nominal P value of <0.0001 for non-inferiority testing was based a non-inferiority margin of 1.5 for upper 95% confidence limit for HR. The Agency and the Applicant did not reach agreement with this non-inferiority margin, which only retained about 70% of Warfarin treatment effect. The Agency recommended greater percentage (85-90%) retention of Warfarin effect.
- Nominal P value for testing superiority in primary efficacy endpoint was 0.34. Therefore, Edoxaban arm was not superior to Warfarin arm.

FIGURE 1: CUMULATIVE INCIDENCE RATE FOR VTE OR VTE-RELATED DEATH, MITT POPULATION



TPDDAO: days since randomization.
[Source: Statistical reviewer's analysis.]

3.2.4.2 Primary efficacy analysis by Edoxaban dose adjustment at randomization

Edoxaban dose was adjusted to 30mg at randomization for 1452 (17.6%) subjects due to body weight ≤ 60 Kg, creatinine clearance between 30 and 50 mL/min inclusive, or concomitant use of the P-gp inhibitors verapamil or quinidine. Table 8 summarizes primary efficacy analysis results by dose level 30mg and 60 mg.

TABLE 8: PRIMARY EFFICACY ANALYSIS BY EDOXABAN DOSE ADJUSTMENT AT BASELINE, MITT POPULATION

	Dose of 30mg		Dose of 60 mg	
Primary efficacy endpoint	**Edoxaban (N=733) n (%)**	**Placebo* (N=719) n (%)**	**Edoxaban (N=3385) n (%)**	**Placebo* (N=3403) n (%)**
Subjects with recurrent VTE or VTE-related death, n (%)	22 (3.0)	30 (4.2)	108 (3.2)	116 (3.4)
PE with/without DVT, n (%)	14 (1.9)	19 (2.6)	59 (1.7)	64 (1.9)
Fatal PE, n (%)	7 (1.0)	10 (1.4)	17 (0.5)	14 (0.4)
DVT only, n (%)	8 (1.1)	11 (1.5)	49 (1.4)	52 (1.5)
Un-stratified Hazard Ratio (95% CI)	0.73 (0.42, 1.26)		0.93 (0.72, 1.21)	

* In active Warfarin arm, Edoxaban placebo dose was adjusted based on the specified risk factors.
- CI: confidence interval;
- Hazard ratio is from un-stratified proportional hazard model. Hazard ratio < 1 favors Edoxaban arm.
[Source: Statistical reviewer's analysis]

Reviewer's note: It seems the differences in treatment effect on primary efficacy endpoint between Edoxaban arm and Warfarin arm were mainly observed in subjects received dose of 30mg.

3.2.4.3 Sensitivity analysis of primary efficacy endpoint

The sponsor performed two sensitivity analyses of primary efficacy endpoint as discussed in Section 3.2.2. The sensitivity analysis results (Table 9) were consistent with those from the primary analysis.

TABLE 9: SENSITIVITY ANALYSIS RESULTS OF PRIMARY EFFICACY ENDPOINT, PP POPULATION

Sensitivity analyses	**Edoxaban N=4057 n (%)**	**Warfarin N=4078 n (%)**	**HR and 95% CI**
On treatment study period	65 (1.6)	80 (2.0)	0.81 (0.59, 1.13)
Treatment + 30 days study period	87 (2.1)	106 (2.5)	0.85 (0.64, 1.14)

- HR: Hazard ratio; CI: confidence interval;
[Source: Study Hokusai VTE CSR Page 103 Table 11.4 and statistical reviewer's analysis]

Reviewer's note: The statistical reviewer detected one more recurrent VTE or VTE-related death in the Edoxaban arm than what was reported by the applicant for sensitivity analysis on treatment study period. The sensitivity analysis results listed in Table 9 were based on statistical reviewer's analyses.

3.2.4.4 Secondary efficacy endpoint analyses results

The analysis results of secondary efficacy endpoint are summarized in Table 10. The p value for testing superiority in secondary efficacy endpoint, time to recurrent VTE or all-cause mortality, was 0.99, for Edoxaban compared to Warfarin in patients with acute symptomatic VTE.

TABLE 10: SECONDARY EFFICACY ENDPOINT ANALYSIS RESULTS, MITT POPULATION

	Edoxaban (N=4118)	**Warfarin (N=4122)**
Subjects with recurrent VTE or all-cause mortality, n (%)	228 (5.5)	228 (5.5)
Recurrent non-fatal VTE	106 (2.6)	122 (3.0)
All-cause mortality	122 (3.0)	106 (2.5)
VTE-related death	24 (0.6)	24 (0.6)
Infectious disease related death	25 (0.6)	12 (0.3)
Other death	73 (1.8)	76 (1.8)
Hazard Ratio (95% CI)	1.00 (0.83, 1.20)	
P value	0.99	

-CI: confidence interval;
- P value from un-stratified log-rank test.
- Hazard ratio is from un-stratified proportional hazard model. Hazard ratio < 1 favors Edoxaban arm.

[Source: Study Hokusai CSR Page 115 Table 11.10 and statistical reviewer's analysis]

Reviewer's note:

- Compared to Warfarin, Edoxaban did not provide more treatment benefit on recurrent VTE or all-cause mortality.
- All-cause mortality was numerically higher in the Edoxaban arm compared to Warfarin arm.

3.2.4.5 Conclusions for efficacy

The pivotal Study Hokusai VTE demonstrated non-inferiority in primary efficacy endpoint, time to recurrent VTE or VTE-related death, for Edoxaban compared to Warfarin in subjects with acute symptomatic VTE. Sensitivity analyses support the non-inferiority in primary efficacy endpoint. However, superiority was not established for neither primary nor secondary efficacy endpoints. Numerically higher incidence of all-cause mortality was observed in Edoxaban arm compared to Warfarin arm.

Based on exploratory analyses, the differences in treatment effect on primary efficacy endpoint between Edoxaban arm and Warfarin arm were mainly observed in subjects received dose of 30mg.

3.3 Evaluation of Safety

The primary safety endpoint was time to major or clinically relevant non-major (CRNM) bleeding. Study Hokusai VTE was adequately powered to test superiority in primary safety endpoint for Edoxaban compared to Warfarin. Table 11 summarizes the primary safety analysis results. Edoxaban was superior to Warfarin in reducing major or CRNM bleeding (p value = 0.004).

TABLE 11: PRIMARY SAFETY ENDPOINT ANALYSIS RESULTS, SAFETY POPULATION

	Edoxaban (N=4118)	**Warfarin (N=4122)**
Subjects with major or CRNM bleeding, n (%)	349 (8.5)	423 (10.3)
Major bleeding	56 (1.4)	66 (1.6)
CRNM bleeding	298 (7.2)	368 (8.9)
Hazard Ratio (95% CI)	0.81 (0.71, 0.94)	
P value	0.004	

CRNM: clinically relevant non-major; CI: confidence interval;

- P value from un-stratified log-rank test.

- Hazard ratio is from un-stratified proportional hazard model. Hazard ratio < 1 favors Edoxaban arm.

[Source: Study Hokusai CSR Page 130 Table 12.6]

Please refer to clinical review of this application for additional safety analyses results and conclusions for safety.

3.4 Benefit-Risk Assessment

Pivotal Study Hokusai VTE in this NDA application provided evidence for the non-inferiority of Edoxaban compared to Warfarin for the treatment of patients with acute symptomatic VTE, and convincing evidence for less risk of Edoxaban in primary safety endpoint, time to major or CRNM bleeding. The statistical reviewer believes the submission demonstrated a favorable benefit-risk profile on Edoxaban. The final judgment on overall benefit-risk is deferred to the clinical team reviewing this submission.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Subgroup Analyses by Gender, Age, Race and Region

Table 12 summarizes the subgroup analyses of primary efficacy endpoint by gender, age, race and region for the Study Hokusai VTE.

TABLE 12: SUBGROUP ANALYSES OF PRIMARY EFFICACY ENDPOINT BY GENDER, AGE, RACE AND REGION

Subgroup	Edoxaban Event/N (%)	Warfarin Event/N (%)	HR (95% CI)
Gender			
Male	82/2360 (3.5)	87/2356 (3.7)	0.94 (0.69, 1.27)
Female	48/1758 (2.7)	59/1766 (3.3)	0.82 (0.56, 1.20)
Age			
< 65 yrs	84/2784 (3.0)	83/2752 (3.0)	1.00 (0.74, 1.36)
≥ 65 yrs	46/1334 (3.5)	63/1370 (4.6)	0.75 (0.51, 1.10)
Race			
White	91/2867 (3.2)	98/2895 (3.4)	0.94 (0.70, 1.25)
Asian	27/866 (3.1)	34/861 (4.0)	0.79 (0.48, 1.31)
Other	12/385 (3.1)	14/366 (3.8)	0.83 (0.38, 1.79)
Region			
North America	15/416 (3.6)	18/420 (4.3)	0.85 (0.43, 1.68)
West Europe	45/1396 (3.2)	55/1394 (4.0)	0.81 (0.55, 1.20)
East Europe	16/911 (1.8)	19/913 (2.1)	0.84 (0.43, 1.63)
Asia	27/850 (3.2)	32/847 (3.8)	0.84 (0.50, 1.40)
Other	27/545 (5.0)	22/548 (4.0)	1.23 (0.72, 2.22)

HR: hazard ratio; CI: confidence interval.

[Source: Statistical reviewer's analysis.]

Reviewer's comment:

- Most patients were White and Asian in the Study Hokusai VTE. Therefore, patients with race other than White and Asian were combined together as a subgroup of "Other".
- Subgroup analyses demonstrated similar trend of treatment effect as the primary analyses, except for race "Other" group, which had reverse treatment benefit for Edoxaban arm vs. Warfarin arm.
- The analyses results for subgroups had high variation due to small sample sizes and few numbers of events.

5 SUMMARY AND CONCLUSIONS

Edoxaban demonstrated its non-inferiority compared to Warfarin in the Study Hokusai VTE. In Sections 5.1 and 5.2, we discussed some statistical issues in this application.

5.1 Statistical Issues

Here are the statistical issues we identified in this submission:

- The Applicant seeks approval for (b) (4) indications:

(b) (4)

(b) (4) indication claims were based on one primary endpoint analysis in one randomized study. (b) (4)

(b) (4)

5.2 Collective evidence

Table 13 summarized exploratory analyses results of time to first PE or VTE-related death and time to first DVT only respectively. The separate analyses results for time to first PE or VTE-related death and time to first DVT only were similar to the primary efficacy analysis results for time to combined first PE/DVT or VTE-related death. The review team recommends first 2 indications proposed by the Applicant should be combined.

TABLE 13: EXPLORATORY ANALYSIS RESULTS OF TIME TO PE/DEATH AND DVT RESPECTIVELY, MITT POPULATION

	Edoxaban N=4118 n (%)	Warfarin N=4122 n (%)	HR and 95% CI
Subjects with PE or VTE-related death	76 (1.9)	85 (2.1)	0.90 (0.66, 1.22)
Subjects with DVT only	62 (1.5)	71 (1.7)	0.88 (0.62, 1.23)

- HR: Hazard ratio; CI: confidence interval;

[Source: Statistical reviewer's analysis]

Exploratory analyses of primary efficacy endpoint by actual treatment duration were performed to check whether patients received longer treatment (> 3 months) experienced more treatment benefit than patients received shorter treatment (≤ 3 months).

TABLE 14: EXPLORATORY ANALYSIS RESULTS OF PRIMARY EFFICACY ENDPOINT, MITT POPULATION

Primary efficacy endpoint	**Edoxaban N=4118**	**Warfarin N=4122**	**HR and 95% CI**
Subject treated <= 3 months, n/N (%)	42/485 (8.7)	56/528 (10.6)	0.83 (0.55, 1.23)
Subject treated > 3 months, n/N (%)	86/3633 (2.4)	90/3594 (2.5)	0.97 (0.72, 1.30)

- HR: Hazard ratio; CI: confidence interval;

[Source: Statistical reviewer's analysis]

Reviewer's note:

- It seems the differences in treatment effect on primary efficacy endpoint between Edoxaban arm and Warfarin arm were mainly observed in subjects received <= 3 months treatment.
- Longer than 3-month treatment continued to benefit subjects in both treatment arms similarly.

(b) (4)

Diagnosis for the index DVT event at baseline required one of the following:

- A noncompressible vein on ultrasonography,
- An intraluminal filling defect on venography,
- An intraluminal filling defect on spiral/contrast computed tomography (CT) of the legs.

Diagnosis for the index PE event at baseline required one of the following:

- An intraluminal filling defect on spiral CT or pulmonary angiography,
- Cutoff of contrast material in a vessel more than 2.5 mm in diameter on pulmonary angiography.

In Study Hokusai VTE, subjects were identified as having more severe PE if they had extensive anatomic extent, NT-proBNP ≥ 500 pg/ml, or RV dysfunction present at baseline. Table 15 summarized subgroup analysis of primary efficacy endpoint by index PE/DVT at baseline and for subjects with severe PE at baseline.

TABLE 15: SUBGROUP ANALYSIS RESULTS OF PRIMARY EFFICACY ENDPOINT BY INDEX PE/DVT, AND FOR PATIENTS WITH SEVERE PE

Primary efficacy endpoint	Edoxaban N=4118	Warfarin N=4122	HR and 95% CI
Recurrent VTE or VTE-related death by index DVT at baseline, n/N (%)	83/2468 (3.4)	81/2453 (3.3)	1.02 (0.75, 1.38)
Recurrent VTE or VTE-related death by index PE at baseline, n/N (%)	47/1650 (2.9)	65/1669 (3.9)	0.73 (0.50, 1.06)
Recurrent VTE or VTE-related death by PE severity at baseline, n/N (%)			
Anatomic extent = extensive	24/743 (3.2)	30/778 (3.9)	0.84 (0.49, 1.44)
NT-proBNP >=500 pg/mL	15/454 (3.3)	30/484 (6.2)	0.54 (0.29, 1.00)
RV dysfunction =Yes	5/172 (2.9)	12/179 (6.7)	0.42 (0.15, 1.20)

- VTE: venous thrombosis embolism; DVT: deep vein thrombosis; PE: pulmonary embolism; NT-proBNP: N-terminal pro brain naturetic Peptide; RV: right ventricular; HR: Hazard ratio; CI: confidence interval;
[Source: Statistical reviewer's analysis]

Reviewer's note:

- It seems the differences in treatment effect on primary efficacy endpoint between Edoxaban arm and Warfarin arm were mainly observed in subjects with index PE at baseline.
- Definition of severe PE based on NT-proBNP ≥ 500pg/mL is not widely accepted.
- Of the 3319 subjects who presented with an index PE (with or without DVT), PE severity based on NT-proBNP level was missing for 330 (~10%) patients.
- The subgroup analyses by PE severity were pre-specified. However, multiplicity adjustment was not planned for testing efficacy in these subgroups.
- The observed treatment benefit for Edoxaban vs. Warfarin in subjects with severe PE may be not reliable due to small sample sizes and few numbers of events in these subgroups.

(b) (4)

5.3 Conclusions and Recommendations

This NDA application was based on a multicenter Phase III randomized trial (Hokusai VTE) comparing Edoxaban versus Warfarin for the treatment of patients with acute symptomatic VTE. The Study Hokusai VTE demonstrated non-inferiority of Edoxaban compared to Warfarin for treatment of patients with acute symptomatic VTE in time to recurrent VTE or VTE-related death, the primary efficacy endpoint, and convincing evidence for less risk of Edoxaban in primary safety endpoint, time to major or CRNM bleeding.

The statistical reviewer believes the efficacy and Safety data from Study Hokusai VTE support the claim of non-inferiority of Edoxaban compared to Warfarin for the treatment of DVT and PE (b) (4)

5.4 Labeling

The Applicant seeks the full approval of Edoxaban for (b) (4) indications: treatment of DVT, treatment of PE, (b) (4) The review team recommends first 2 indications proposed by the Applicant should be combined as one indication: treatment of deep vein thrombosis (DVT) and pulmonary embolism (PE) in patients who have been treated with a parenteral anticoagulant for 5 -10 days. (b) (4)

(b) (4)

(b) (4)

(b) (4)

In Figure 14.3 in the labeling, (b) (4) which made the figure confusing to read. The figure should be revised to show cumulative incidence rate for recurrent VTE or VTE-related death based on mITT population for overall study period only.

(b) (4)

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YUN WANG
09/08/2014

LEI NIE
09/08/2014

RAJESHWARI SRIDHARA
09/09/2014

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CARCINOGENICITY STUDIES

IND/NDA Number:	NDA 206-316/IND 77,254
Drug Name:	DU-176b (Savaysa)
Indication(s):	104 Week Rat and Mouse Carcinogenicity Studies
Applicant:	**Sponsor:** Daiichi Sankyo Co., LTD. 717 Horikoshi, Fukuroi, Shizuoka 437-0065, Japan **Testing Facility:** (b) (4)
Documents Reviewed:	Electronic submission submitted on January 8, 2014 Electronic data submitted on January 8, 2014
Review Priority:	Standard
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Mohammad Atiar Rahman, Ph.D.
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	Division of Cardiovascular and Renal Products
Reviewing Pharmacologist:	Baichun Yang, Ph.D.
Project Manager:	Alison Blaus
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of DU-176 (Savaysa) when administered orally daily through gavage at appropriate drug levels for 105 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Yang.

In this review the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as the dose increases.

2. Rat Study

Two separate experiments were conducted, one in male and one in female rats. In each of these two experiments there were three treated groups and one control group. Two hundred and sixty Crl:CD(SD) rats of each sex were assigned randomly to the treated and control groups in equal size of 65 rats per group. The dose levels for treated groups were 60, 200, or 600/400 mg/kg/day for male rats, and 50, 100, or 200 mg/kg/day for female rats. In this review these dose groups would be referred to as the low, medium, and high dose groups, respectively. The rats in the control group received the control article [aqueous solution of 0.5% methylcellulose].

Due to high mortality the dose level of male high dose group was lowered to 400 mg/kg/day on Week 44 (Day 302). Also due to high mortality, dosing of the male high dose group was discontinued during Week 80 and all of the remaining male rats in this group were sacrificed during Week 88 when the survival of male rats in this group was 25%. The remaining male rats were sacrificed during Week 90 when the survival for control, low, and medium dose were 29%, 45%, and 38%, respectively. All female rats were sacrificed during Week 105.

During the administration period all rats were observed twice daily morbidity and mortality. A detailed clinical examination was performed once before the start of treatment, once weekly thereafter and on the day of scheduled sacrifice. The rats were palpated regularly during the clinical observations for the appearance of masses.

Individual body weights were recorded prior to treatment, before dosing on Day 1, and weekly from Weeks 1-14 and every 4 weeks thereafter.

2.1. Sponsor's analyses

2.1.1. Survival analysis

The sponsor estimated the survival rates of rats in all treatment groups using the Kaplan-Meier product-limit estimation method and graphically presented the Kaplan-Meier curves by sex. The sponsor analyzed the survival data using the Cox-Tarone binary regression and Gehan-Breslow nonparametric tests for dose response relationship across the treated groups and pairwise comparisons of treated groups with the control.

Sponsor's findings: The sponsor reported that there were a total of 361 rats found dead, sacrificed as moribund, or died accidentally. The sponsor mentioned that following the FDA guidance for carcinogenicity studies and data analysis, when the survival in either sex of the control group reached 20 rats, all surviving rats of that sex were sacrificed, and when the survival in either sex of any treated group reached 15 rats, all surviving rats in that group/sex were sacrificed. In addition, when survival of the high-dose group males reached 20, dosing for this group was discontinued and when survival of this male group reached 15, all remaining surviving rats were sacrificed. Consequently, dosing for the 600/400 mg/kg/day male group was discontinued during Week 80 with all remaining rats in this group sacrificed during Week 88. All male rats in the control, low, and medium dose groups were sacrificed during Week 90. All female rats were sacrificed during Week 105.

The sponsor's count showed 46, 36, 40, and 49 deaths in male rats and 43, 48, 47 and 49 deaths in female rats in control, low, medium, and high dose groups, respectively. The sponsor's analysis showed a statistically significant positive dose response relationship in mortality for male rats through Week 88 (Cox-Tarone test: $p = 0.0074$; Gehan-Breslow test: $p = 0.0004$). The pairwise comparison showed statistically significantly higher mortality in male rats high dose group compared with their control (Cox-Tarone test: $p = 0.0039$; Gehan-Breslow test: $p = 0.0002$). The low and medium dose groups showed similar mortality compared to their control through Week 88 and to the end of study (Week 90). The pairwise comparison also showed a marginal but statistically significant increase mortality in the female rats high dose group compared to their control through Week 105 (Gehan-Breslow $p = 0.0500$).

2.1.2. Tumor data analysis

The sponsor analyzed the incidental[1] tumors using the logistic regression, and analyzed the rapidly lethal[2] tumors and palpable tumors using the same methodologies as they used for survival data analysis, with the day of death for lethal tumors and the day of first palpation for palpable tumors as the tumor onset time. In the cases where the study pathologist can assign particular occult[3] neoplastic lesions as the cause of death in the animals, IARC [International Agency for Cancer Research, 1980] type analysis was used by incorporating such information.

In the cases of sparse tables, exact form of survival adjusted method of tumor analysis will be used.

Adjustment for multiple testing: The sponsor followed the methodologies given in the FDA guidance for carcinogenicity studies and data analysis for the adjustment for multiple testing in tumor data analysis. It may be mentioned that for dose response relationship tests, the guidance suggests the use of test levels of $\alpha=0.005$ for common tumors and $\alpha=0.025$ for rare tumors for a submission with two species, and a significance level $\alpha=0.01$ for common tumors and $\alpha=0.05$ for rare tumors for a submission with one species. Also for multiple pairwise comparisons of treated group with control the guidance suggests the use of test levels of $\alpha=0.01$ for common tumors and $\alpha=0.05$

[1] Tumors considered as not the cause of death and is detected after the animals die.
[2] Tumors kill the animals very quickly. The onset time may be considered as the time of death.
[3] The animals die due to tumor but tumor may not be the direct cause of death (non-lethal).

for rare tumors for both submissions with two or one species.

Sponsor's findings: The sponsor's analyses did not show statistically significant dose response relationship among the treatment groups in any of the observed tumor types in either sex. Pairwise comparisons also did not show increased incidence in any of the observed tumor types in either sex.

2.2. Reviewer's analyses

To verify the sponsor's analyses and to perform additional analyses suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically.

2.2.1. Survival analysis

This reviewer estimated the survival distributions of rats in all four treatment groups using the Kaplan-Meier product limit method. The dose response relationship was tested using the likelihood ratio test and the homogeneity of survival distributions was tested using the log-rank test. The Kaplan-Meier curves for survival rates are given in Figures 1A and 1B in the appendix for male and female rats, respectively. The intercurrent mortality data are given in Tables 1A and 1B in the appendix for male and female rats, respectively. Results of the tests for dose response relationships and homogeneity of survivals, are given in Tables 2A and 2B in the appendix for male and female rats, respectively.

Reviewer's findings: This reviewer's analysis showed 46, 36, 40, and 49 number of deaths in male rats and 43, 48, 47, and 49 number of deaths in female rats in control, low, medium, and high dose groups, respectively. The tests showed statistically significant dose response relationship in mortality across control and treated groups in male rats. The pairwise comparisons showed statistically significant increased mortality in the male rat high dose group compared to their control.

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pairwise comparisons of each of the treated groups with control group. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). In this method an animal that lives through the full study period ($w_{\max}$) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of $s_h = 1$. An animal that dies at week w_h without a tumor before the end of the study gets a score of $s_h = \left(\frac{w_h}{w_{\max}}\right)^k < 1$.

The adjusted group size is then defined as Σs_h. As an interpretation, an animal with score $s_h = 1$ can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size Σs_h is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops at least one tumor, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose

response relationship (or the pairwise) tests using the Cochran-Armitage test. One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used. For this analysis the actual dose levels for control, low and medium dose groups were used as score. For high dose group a weighted dose level [(600*44+400*36+8*0)/88=464] was used. The tumor rates and the p-values of the tested tumor types are listed in Tables 3A and 3B in the appendix for male and female rats, respectively.

Multiple testing adjustment: For the adjustment of multiple testing this reviewer used the methodologies suggested in the FDA guidance for statistical aspects of the design, analysis, and interpretation of chronic rodent carcinogenicity studies of pharmaceuticals. For dose response relationship tests, the guidance suggests the use of test levels of α=0.005 for common tumors and α=0.025 for rare tumors for a submission with two species, and a significance level α=0.01 for common tumors and α=0.05 for rare tumors for a submission with one species in order to keep the false-positive rate at the nominal level of approximately 10%. A rare tumor is defined as one for which the published spontaneous tumor rate is less than 1%, the tumor is termed as common otherwise. For multiple pairwise comparisons of treated group with control the guidance suggests the use of test levels of α=0.01 for common tumors and α=0.05 for rare tumors, in order to keep the false-positive rate at the nominal level of approximately 10% for both submissions with two or one species.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Lin and Rahman (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

Reviewer's findings: Based on the criteria of adjustment for multiple testing discussed above, none of the observed tumors was considered to have statistically significant dose response relationship in either sex. The pairwise comparison also did not show statistically significant increased incidence in any observed tumor type in any treated group in either sex compared to their respective control.

3. Mouse Study

Two separate experiments were conducted, one in male and one in female mice. In each of these two experiments there were three treated groups and one control group. Two hundred and sixty Crl:CD1(ICR) mice of each sex were assigned randomly to the treated and control groups in equal size of 65 mice per group. The dose levels for treated groups were 50, 150 or 500 mg/kg/day. In this review these dose groups were referred to as the low, medium, and high dose groups, respectively. The mice in the control group received the control article [aqueous solution of 0.5% methylcellulose].

Due to high mortality, all of the remaining mice in the female medium dose group were sacrificed during Week 96. The surviving females in the control, low, and high dose groups were sacrificed during Week 103. All male mice were sacrificed during Week 105. At the time of the Week 96

sacrifice, survival in the female medium dose group was 23%, while at the Week 103 scheduled sacrifice, the survivals in the control, low, and high dose groups were 35%, 28%, and 22%, respectively.

During the administration period all mice were observed twice daily morbidity and mortality. A detailed clinical examination was performed once before the start of treatment, once weekly thereafter and on the day of scheduled sacrifice. The mice were palpated regularly for the appearance of masses during the clinical observations.

Individual body weights were recorded prior to treatment, before dosing on Day 1, and weekly from Weeks 1-14 and every 4 weeks thereafter.

3.1. Sponsor's analyses

3.1.1. Survival analysis

The sponsor used similar methodologies to analyze the mouse survival data as those used to analyze the rat survival data.

Sponsor's findings: The sponsor reported that there were a total of 356 mice found dead, sacrificed as moribund, or died accidentally. The sponsor mentioned that all of the remaining mice in the female medium dose group were sacrificed at Week 96. The surviving females in the control, low, and high dose groups were sacrificed during Week 103. All male mice were sacrificed at Week 105. At the time of the Week 96 sacrifice, survival in the female mice medium dose group was 23%, while at the time of Week 103 scheduled sacrifice, the survival in the female mice control, low, and high dose groups were 35%, 28%, and 22%, respectively.

The sponsor's count showed 38, 42, 40, and 46 deaths in male mice and 42, 44, 49 and 49 deaths in female mice in control, low, medium, and high dose groups, respectively. The sponsor's analysis showed a statistically significant positive dose response relationship (Gehan-Breslow text: $p = 0.0281$) in mortality in male mice. The pairwise comparison showed a statistically significant increased mortality in male high dose group and female medium dose group compared to their respective control.

3.1.2. Tumor data analysis

The sponsor used similar methodologies to analyze the mouse tumor data as those used to analyze the rat tumor data.

Adjustment for multiple testing: The sponsor followed the methodology given in the FDA guidance for carcinogenicity studies and data analysis for the adjustment for multiple testing in tumor data analysis.

Sponsor's findings: The sponsor's analyses did not show statistically significant positive dose response relationship among the treatment groups in any of the observed tumor types in either sex.

The pairwise comparison showed statistically significant increased incidence of subcapsular cell adenoma in the adrenal cortex in male low dose group (p = 0.0317), hemangiosarcoma in body, whole/cavity for female low and medium dose groups compared to their controls (p = 0.0464 and 0.0230, respectively), and lymphosarcoma in body, whole/cavity for females medium dose group (p = 0.0416) compared to their respective controls.

3.2. Reviewer's analyses

Similar to the rat study, to verify sponsor's analyses and to perform additional analyses suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses of mouse data. Data used in this reviewer's analyses were provided by the sponsor electronically.

For the analysis of both the survival data and the tumor data this reviewer used similar methodologies as he used for the analyses of the rat survival and tumor data.

3.2.1. Survival analysis

The Kaplan-Meier curves for survival rates of all treatment groups are given in Figures 2A and 2B in the appendix for male and female mice, respectively. The intercurrent mortality data of all treatment groups are given in Tables 4A and 4B in the appendix for male and female mice, respectively. Results of the tests for dose response relationships and homogeneity of survivals for control, low, medium, and high dose groups are given in Tables 5A and 5B in the appendix for male and female mice, respectively.

Reviewer's findings: This reviewer's analysis showed 38, 42, 40, and 46 number of deaths in male mice, and 42, 44, 49, and 49 number of deaths in female mice in control, low, medium, and high dose groups, respectively. The tests showed statistically significant dose response relationship in mortality across the treatment groups in male mice. The pairwise comparison showed statistically significant increased mortality in male mice high dose group compared to their control. The pairwise comparison also showed statistically significant increased mortality in female mice medium dose group compared to their control.

3.2.2. Tumor data analysis

The tumor rates and the p-values of the tested tumor types are given in Tables 6A and Table 6B in the appendix, for male and female mice respectively.

Reviewer's findings: Following tumor type showed p-values less than or equal to 0.05 either for dose response relationships or pairwise comparisons of treated groups and control.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pairwise Comparisons of Treated Groups and Control in Mice

			Cont	Low	Med	High	...P_Value			
Sex	Organ Name	Tumor Name	N=65	N=65	N=65	N=65	Dose Resp	C vs L	C vs M	C vs H
Male	Adrenal, Cortex	B-Adenoma, Subcapsular Cell	0	4	2	1	0.5368	0.0500*	0.2471	0.4430 Female
Body, Whole/Cav	M-Hemangiosarcoma		0	4	4	0	0.8518	0.0503	0.0287*	.

Based on the criteria of adjustment for multiple testing discussed in the rat data analysis section, the incidence of none of the observed tumor types was considered to have statistically significant dose response relationship in either sex. The pairwise comparisons showed statistically significant increased incidence of adrenal cortex B-adenoma in subcapsular cell in low dose male mice and whole body cavities M-hemangiosarcoma in medium dose female mice compared to their respective control.

4. Evaluation of the validity of design of rat and mouse studies

As has been noted, except for the pairwise comparisons of low dose group vs. control for adrenal cortex subcapsular cell B-adenoma in male mice, and medium dose group vs. control for whole body cavity M-hemangiosarcoma in female mice, no other tumor types showed statistically significant dose response relationship or increased incidences in the treated groups compared to their respective control. However, before drawing any conclusion regarding the carcinogenic or non-carcinogenic potential of the study drug in rats and mice, it is important to look into the following two issues, as have been pointed out in the paper by Haseman (1984).

(i) Were enough animals exposed, for a sustained amount of time, to the risk of late developing tumors?
(ii) Were dose levels high enough to pose a reasonable tumor challenge to the animals?

There is no consensus among experts regarding the number of animals and length of time at risk, although most carcinogenicity studies are designed to run for two years with about fifty to sixty animals per treatment group. The following are some rules of thumb regarding these two issues as suggested by experts in this field.

Haseman (1985) has done an investigation on the first issue. He gathered data from 21 studies using Fischer 344 rats and B6C3Fl mice conducted at the National Toxicology Program (NTP). It was found that, on the average, approximately 50% of the animals in the high dose group survived the two-year study period. Also, in a personal communication with Dr. Karl Lin of Division of Biometrics-6, Haseman suggested that, as a rule of thumb, a 50% survival of 50 initial animals or 20 to 30 animals still alive in the high dose group, between weeks 80-90, would be consider as a sufficient number and adequate exposure. In addition Chu, Cueto and Ward (1981), suggested that "to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one-year."

It appears, from these three sources that the proportions of survival at 52 weeks, 80-90 weeks, and two years are of interest in determining the adequacy of exposure and number of animals at risk.

Regarding the question of adequate dose levels, it is generally accepted that the high dose should be close to the maximum tolerated dose (MTD). In the paper of Chu, Cueto and Ward (1981), the following criteria are mentioned for dose adequacy. A high dose is considered as close to MTD if any of the criteria is met.

(i) "A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls."

(ii) "The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical."

(iii) "In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls."

We will now investigate the validity of the DU-176 rat and mouse carcinogenicity study, in the light of the above guidelines.

4.1. Rat Study

The following is the summary of survival data of rats in the high dose groups:

Percentage of Survival in the High Dose Group at the End of Weeks 52, 78, and 91 in Rats

	Percentage of survival		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Male	63%	35%	23%*
Female	89%	63%	38%

***At the end of 88 weeks**

Based on the survival criterion Haseman proposed, it may be concluded that not enough male rats were exposed to the high dose for a sufficient amount of time, however enough female rats were exposed to the high dose for a sufficient amount of time.

The following table shows the percent difference in mean body weight gain in rats from the concurrent control, defined as

$$\text{Percent difference} = \frac{(\text{Final BW} - \text{Baseline BW})_{\text{Treated}} - (\text{Final BW} - \text{Baseline BW})_{\text{Control}}}{(\text{Final BW} - \text{Baseline BW})_{\text{Control}}} \times 100$$

Percent Difference in Mean body Weight Gain from Controls in Rats

Male			Female		
Low	Medium	High	Low	Medium	High
-2.17	-5.24	1.99	4.15	6.42	8.68

Source: Table 3 of sponsor's submission

Therefore, relative to the control the male rats in high dose group had about 2% and the female rats had about 9% increased body weight gains. Thus, the body weight data indicate that the used high dose might be under MTD.

The mortality rates at the end of the experiment were as follows:

Mortality Rates at the End of the Experiment in Rats

	Control	**Low**	**Medium**	**High**
Male	71%	55%	61%	77%
Female	66%	74%	72%	75%

The morality rates in male rat high dose group was 6% higher, and in female rats high dose group was 9% higher than their respective control.

Thus, from the mortality data it can be concluded that the used high dose level might have reached the MTD in both sexes. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

4.2. Mouse Study

The following is the summary of survival data of mice in the high dose groups:

Percentage of Survival in the High Dose Group at the End of Weeks 52, 78, and 91 in Mice

	_____Percentage of survival_____		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Male	80%	54%	40%
Female	91%	74%	59%

Based on the survival criterion Haseman proposed, it may be concluded that not enough male mice were exposed to the high dose for a sufficient amount of time, however enough female mice were exposed to the high dose for a sufficient amount of time.

The following table shows the percent difference in mean body weight gain in mice from the concurrent control,

Percent Difference in Mean body Weight Gain from Controls

Male			Female		
Low	Medium	High	Low	Medium	High
-17.54	-16.67	-28.95	21.71	-8.53	-17.83

Source: Table 3 of sponsor's submission

Therefore, relative to control the high dose male mice had about 29% and the female mice had about 18% decreases body weight gain.

The mortality rates at the end of the experiment were as follows:

Mortality Rates End of the Experiment

	Control	**Low**	**Medium**	**High**
Male	58%	65%	61%	71%
Female	65%	68%	75%	75%

The morality rate was 13% higher in the male mice high dose group, and 10% higher in the female mice high dose group than their respective control.

Thus, from the body weight gain and mortality data it can be concluded that the used high dose level might have reached or exceeded the MTD in both sexes of mice. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

5. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of DU-176 when administered orally daily through gavage at appropriate drug levels for 105 weeks.

In this review the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as the dose increases.

Rat Study: Two separate experiments were conducted, one in male and one in female rats. In each of these two experiments there were three treated groups and one control group. Two hundred and sixty Crl:CD(SD) rats of each sex were assigned randomly to the treated and control groups in equal size of 65 rats per group. The dose levels for treated groups were 60, 200, or 600/400 mg/kg/day for male rats, and 50, 100, or 200 mg/kg/day for female rats. The rats in the control group received the control article [aqueous solution of 0.5% methylcellulose].

During the administration period all rats were observed twice daily morbidity and mortality. A detailed clinical examination was performed once before the start of treatment, once weekly thereafter and on the day of scheduled sacrifice. The rats were palpated regularly for the appearance of masses during the clinical observations.

Due to high mortality the dose level of male high dose group was lowered to 400 mg/kg/day on Week 44 (Day 302). Also due to high mortality, dosing of the male high dose group was discontinued during Week 80 and all of the remaining male rats in this group were sacrificed during Week 88 when the survival of male rats in this group was 25%. The remaining male rats were sacrificed during Week 90 when the survival for control, low, and medium dose were 29%, 45%, and 38%, respectively. All female rats were sacrificed during Week 105. Body weights were recorded prior to treatment, before dosing on Day 1, and weekly from Weeks 1-14 and every 4 weeks thereafter.

The tests showed statistically significant dose response relationship in mortality across control and treated groups in male rats. The pairwise comparisons showed statistically significant increased mortality in the male rat high dose group compared to their control.

The tests did not show statistically significant dose response relationship in any observed tumor type in either sex. The pairwise comparison also did not show statistically significant increased incidence in any observed tumor type in any treated group in either sex compared to their respective control.

The mortality data indicate that the used high dose level might have reached the MTD in both sexes of rats. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

Mouse Study: Two separate experiments were conducted, one in male and one in female mice. In each of these two experiments there were three treated groups and one control group. Two hundred and sixty Crl:CD1(ICR) mice of each sex were assigned randomly to the treated and control groups in equal size of 65 mice per group. The dose levels for treated groups were 50, 150 or 500 mg/kg/day. The mice in the control group received the control article [aqueous solution of 0.5% methylcellulose].

Due to high mortality, all of the remaining mice in the female medium dose group were sacrificed during Week 96. The surviving females in the control, low, and high dose groups were sacrificed during Week 103. All male mice were sacrificed during Week 105. At the time of the Week 96 sacrifice, survival in the female medium dose group was 23%, while at the Week 103 scheduled sacrifice, the survivals in the control, low, and high dose groups were 35%, 28%, and 22%, respectively.

During the administration period all mice were observed twice daily morbidity and mortality. A detailed clinical examination was performed once before the start of treatment, once weekly thereafter and on the day of scheduled sacrifice. The mice were palpated regularly for the appearance of masses during the clinical observations. Body weights were recorded prior to treatment, before dosing on Day 1, and weekly from Weeks 1-14 and every 4 weeks thereafter.

The tests showed statistically significant dose response relationship in mortality across the treatment groups in male mice. The pairwise comparison showed statistically significant increased mortality in male mice high dose group compared to their control.

The tests did not show statistically significant dose response relationship in any observed tumor type in either sex. The pairwise comparisons showed statistically significant increased incidence of adrenal cortex B-adenoma in subcapsular cell in low dose male mice and whole body cavities M-hemangiosarcoma in medium dose female mice compared to their respective control.

From the body weight gain and mortality data it can be concluded that the used high dose level might have reached or exceeded the MTD in both sexes of mice. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

Mohammad Atiar Rahman, Ph.D.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, Biometrics-6

cc:
Archival NDA 206-316
Dr. Yang
Ms. Blaus
Dr. Tsong
Dr. Lin
Dr. Rahman
Ms. Patrician

6. Appendix

Table 1A: Intercurrent Mortality Rate
Male Rats

Week	0 mg|kg|day No. of Death	Cum. %	60 mg|kg|day No. of Death	Cum. %	200 mg|kg|day No. of Death	Cum. %	600/400 mg|kg|day No. of Death	Cum. %
0 - 52	7	10.77	6	9.23	11	16.92	24	36.92
53 - 78	18	38.46	20	40.00	16	41.54	18	64.62
79 - 89	21	70.77	10	55.38	13	61.54	7	76.92
Ter. Sac.	19	29.23	29	44.62	25	38.46	16	23.08
Total	N=65		N=65		N=65		N=65	

The terminal sacrifice week for male rats in the high dose group was Week 88, and that for male rats in other treatment groups was Week 90.

Table 1B: Intercurrent Mortality Rate
Female Rats

Week	0 mg|kg|day No. of Death	Cum. %	50 mg|kg|day No. of Death	Cum. %	100 mg|kg|day No. of Death	Cum. %	200 mg|kg|day No. of Death	Cum. %
0 - 52	3	4.62	2	3.08	1	1.54	7	10.77
53 - 78	17	30.77	21	35.38	14	23.08	17	36.92
79 - 91	14	52.31	13	55.38	21	55.38	16	61.54
92 - 104	9	66.15	12	73.85	11	72.31	9	75.38
Ter. Sac.	22	33.85	17	26.15	18	27.69	16	24.62
Total	N=65		N=65		N=65		N=65	

The terminal sacrifice week for female rats in all treated groups was Week 105.

Table 2A: Intercurrent Mortality Comparison
Male Rats

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.0016
Homogeneity	Log-Rank	0.0013

Table 2B: Intercurrent Mortality Comparison
Female Rats

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.2658
Homogeneity	Log-Rank	0.5717

Table 3A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons
Male Rats

Organ Name	Tumor Name	0 mg Control N=65	60 mg Low N=65	200 mg Med N=65	600/400 mg High N=65	P_Value Dose Resp	P_Value L vs Com C	P_Value M vs Com C	P_Value H vs Com C
Adipose Tissue	B-Lipoma	0	0	1	1	0.1271	.	0.5055	0.3919
Adrenal, Cortex	B-Adenoma	3	0	0	1	0.6074	0.8791	0.8791	0.5036
	M-Carcinoma	1	2	0	0	0.8565	0.5165	0.5055	0.3919
Adrenal, Medull	B-Pheochromocytoma	6	7	2	1	0.9754	0.5000	0.8666	0.8470
	M-Malignant Pheochromocyt	1	4	0	2	0.3699	0.1944	0.5055	0.3382
Body, Whole/Cav	M-Hemangiosarcoma	1	1	1	0	0.6497	0.2527	0.2527	0.3919
	M-Histiocytic Sarcoma	2	1	1	0	0.8171	0.5082	0.5000	0.6270
Brain	B-Astrocytoma	0	0	1	0	0.4518	.	0.5055	.
	B-Granular Cell Tumor	1	0	0	0	0.7289	0.5055	0.5055	0.3919
	B-Oligodendroglioma	1	0	0	0	0.7246	0.5000	0.5000	0.3867
	M-Malignant Astrocytoma	2	1	0	2	0.2657	0.5000	0.7527	0.5317
Cavity, Abdomin	M-Liposarcoma	0	0	1	0	0.4518	.	0.5055	.
Cavity, Thoraci	B-Hibernoma	0	1	0	1	0.2299	0.5109	.	0.4000
Heart	M-Endocardial Schwannoma	1	0	0	0	0.7289	0.5055	0.5055	0.3919
Kidney	B-Adenoma, Tubule Cell	1	0	0	0	0.7246	0.5000	0.5000	0.3867
	M-Liposarcoma	0	0	2	0	0.3953	.	0.2527	.
	M-Malignant Renal Mesench	0	1	0	0	0.4491	0.5109	.	.
	M-Nephroblastoma	0	0	1	0	0.4551	.	0.5109	.
Liver	B-Adenoma, Hepatocellular	1	1	0	0	0.7766	0.2527	0.5055	0.3919
	M-Carcinoma, Hepatocellul	1	0	3	2	0.0856	0.5000	0.3083	0.3424
	M-Cholangiocarcinoma	0	1	0	0	0.4518	0.5055	.	.
Muscle, Bi Fem	M-Osteosarcoma	1	0	0	0	0.7289	0.5055	0.5055	0.3919
Pancreas	B-Adenoma, Acinar Cell	1	2	0	1	0.4718	0.5000	0.5000	0.6368
	B-Adenoma, Islet Cell	1	1	0	1	0.4220	0.2527	0.5055	0.6335
	M-Carcinoma, Islet Cell	0	1	0	1	0.2305	0.5055	.	0.4000
Parathyroid	B-Adenoma	4	2	1	2	0.5484	0.6718	0.8195	0.4270
Pituitary	B-Adenoma	37	26	28	19	0.7436	0.9718	0.8875	0.8545
Skin/Subcutis	B-Fibroadenoma, Mammary G	0	0	0	1	0.1796	.	.	0.4000
	B-Fibroma	1	0	3	1	0.2308	0.5055	0.3166	0.6335
	B-Keratoacanthoma	1	0	2	0	0.5239	0.5055	0.5083	0.3919
	B-Lipoma	0	0	1	0	0.4518	.	0.5055	.
	B-Papilloma, Squamous Cel	2	1	0	1	0.5289	0.5000	0.7527	0.3307
	B-Schwannoma	1	0	0	0	0.7246	0.5000	0.5000	0.3867
	B-Trichoepithelioma	0	0	0	1	0.1796	.	.	0.4000
	M-Carcinoma, Basal Cell	0	1	0	0	0.4518	0.5055	.	.
	M-Carcinoma, Sebaceous Gl	0	1	0	0	0.4518	0.5055	.	.

Table 3A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Male Rats

Organ Name	Tumor Name	0 mg Control N=65	60 mg Low N=65	200 mg Med N=65	600/400 mg High N=65	P_Value Dose Resp	P_Value L vs Com C	P_Value M vs Com C	P_Value H vs Com C
Skin/Subcutis	M-Carcinoma, Squamous Cel	0	1	0	0	0.4518	0.5055	.	.
	M-Malignant Schwannoma	1	0	0	0	0.7246	0.5000	0.5000	0.3867
	M-Schwannoma	1	0	0	0	0.7289	0.5055	0.5055	0.3919
Spinal Cord	M-Malignant Astrocytoma	1	0	0	0	0.7246	0.5000	0.5000	0.3867
Stomach, Nongl	B-Papilloma, Squamous Cel	0	1	0	0	0.4491	0.5109	.	.
	M-Leiomyosarcoma	1	0	1	0	0.5498	0.5055	0.2527	0.3919
Testis	B-Interstitial Cell Tumor	0	1	1	0	0.3953	0.5055	0.5055	.
	B-Mesothelioma	1	0	0	0	0.7289	0.5055	0.5055	0.3919
Thyroid	B-Adenoma, Follicular Cel	2	2	0	2	0.3428	0.3166	0.7527	0.5179
	B-C-Cell Adenoma	7	7	8	9	0.0568	0.4203	0.5000	0.1352
	M-Carcinoma, C-cell	1	3	2	2	0.2642	0.3250	0.5083	0.3382
	M-Carcinoma, Follicular C	2	0	1	0	0.7728	0.7527	0.5000	0.6270

Table 3B: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Rats

Organ Name	Tumor Name	0 mg Control N=65	50 mg Low N=65	100 mg Med N=65	200 mg High N=65	P_Value Dose Resp	P_Value L vs Com C	P_Value M vs Com C	P_Value H vs Com C
Adipose Tissue	B-Fibroma	0	1	0	0	0.4877	0.4940	.	.
Adrenal, Cortex	B-Adenoma	4	1	4	2	0.5843	0.8047	0.3698	0.5997
	M-Carcinoma	1	0	1	0	0.6069	0.4878	0.2529	0.4684
	M-Ganglioneuroma	0	1	0	0	0.4907	0.4878	.	.
Adrenal, Medull	B-Ganglioneuroma	0	1	1	0	0.4750	0.4878	0.5059	.
	B-Pheochromocytoma	3	2	2	2	0.5549	0.4766	0.5000	0.4393
	M-Neuroblastoma	0	0	1	0	0.4938	.	0.5059	.
Body, Whole/Cav	B-Hemangioma	0	0	0	2	0.0517	.	.	0.2162
	M-Hemangiosarcoma	1	0	0	2	0.1790	0.4878	0.5000	0.4520
	M-Histiocytic Sarcoma	1	2	0	1	0.5157	0.4815	0.5000	0.7205
	M-Lymphoma, Lymphocytic	0	0	1	0	0.4907	.	0.5000	.
Brain	M-Malignant Astrocytoma	0	0	1	0	0.4938	.	0.5059	.
Cavity, Abdomin	B-Hemangioma	0	0	0	1	0.2298	.	.	0.4684
Cavity, Thoraci	B-Lipoma	0	0	1	0	0.4938	.	0.5059	.
	M-Liposarcoma	0	0	1	0	0.4938	.	0.5059	.
Cervix	M-Carcinoma	0	0	1	0	0.4907	.	0.5000	.
Kidney	B-Lipoma	0	0	1	1	0.1724	.	0.5000	0.4684
	M-Carcinoma, Tubule Cell	0	1	0	0	0.4907	0.4878	.	.
	M-Malignant Renal Mesench	0	0	1	0	0.4938	.	0.5059	.
	M-Nephroblastoma	0	0	0	1	0.2298	.	.	0.4684
Liver	B-Adenoma, Hepatocellular	3	0	1	0	0.9360	0.8704	0.6921	0.8548
	M-Carcinoma, Hepatocellul	0	1	0	0	0.4907	0.4878	.	.
Mammary, Female	B-Fibroadenoma	29	25	24	22	0.7207	0.6503	0.7941	0.6656
	B-Fibroma	0	1	1	1	0.2683	0.4940	0.5000	0.4684
	B-Lipoma	0	0	0	1	0.2298	.	.	0.4684
	M-Carcinoma	17	12	19	18	0.2155	0.7530	0.4479	0.4017
Ovary	B-Lipoma	0	0	1	0	0.4938	.	0.5059	.
	M-Malignant Teratoma	0	0	1	0	0.4907	.	0.5000	.
	M-Mesothelioma	1	0	0	0	0.7391	0.4878	0.5000	0.4684
Pancreas	B-Adenoma, Islet Cell	3	1	0	1	0.8125	0.6830	0.8795	0.6428
	M-Carcinoma, Acinar Cell	0	0	1	0	0.4907	.	0.5000	.
	M-Carcinoma, Islet Cell	0	1	1	1	0.2683	0.4940	0.5000	0.4684
Parathyroid	B-Adenoma	0	0	0	1	0.2298	.	.	0.4684
Pituitary	B-Adenoma	50	46	53	50	0.1400	0.7692	0.5628	0.3141
	M-Carcinoma	1	0	0	0	0.7391	0.4878	0.5000	0.4684

Table 3B: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Rats

Organ Name	Tumor Name	0 mg Control N=65	50 mg Low N=65	100 mg Med N=65	200 mg High N=65	P_Value Dose Resp	P_Value L vs Com C	P_Value M vs Com C	P_Value H vs Com C
Skin/Subcutis	B-Keratoacanthoma	0	0	1	0	0.4938	.	0.5059	.
	B-Schwannoma	2	1	1	1	0.6031	0.4726	0.5000	0.4532
	M-Carcinoma, Baso Squamou	1	0	0	0	0.7391	0.4878	0.5000	0.4684
	M-Carcinoma, Mammary Glan	0	0	0	1	0.2298	.	.	0.4684
	M-Carcinoma, Squamous Cel	0	0	0	1	0.2298	.	.	0.4684
	M-Malignant Schwannoma	0	1	0	0	0.4877	0.4940	.	.
Thyroid	B-Adenoma, Follicular Cel	0	1	0	1	0.2873	0.4878	.	0.4684
	B-C-Cell Adenoma	8	10	6	4	0.8956	0.3740	0.6323	0.7590
	M-Carcinoma, C-cell	1	2	0	0	0.8312	0.4909	0.5000	0.4684
	M-Carcinoma, Follicular C	0	0	1	1	0.1724	.	0.5000	0.4684
Uterus	B-Polyp, Endometrial Stro	2	2	2	3	0.2921	0.6735	0.3169	0.4393
	M-Carcinoma	0	1	1	1	0.2683	0.4940	0.5000	0.4684
Vagina	B-Fibroma	1	0	0	0	0.7391	0.4878	0.5000	0.4684
	B-Polyp, Stromal	0	2	0	0	0.6052	0.2349	.	.

Table 4A: Intercurrent Mortality Rate in Male Mice

Week	0 mg|kg|day No. of Death	Cum. %	50 mg|kg|day No. of Death	Cum. %	150 mg|kg|day No. of Death	Cum. %	500 mg|kg|day No. of Death	Cum. %
0 - 52	4	6.15	5	7.69	9	13.85	13	20.00
53 - 78	15	29.23	15	30.77	9	27.69	17	46.15
79 - 91	7	40.00	15	53.85	8	40.00	9	60.00
92 - 104	12	58.46	7	64.62	14	61.54	7	70.77
Ter. Sac.	27	41.54	23	35.38	25	38.46	19	29.23
Total	N=65		N=65		N=65		N=65	

The terminal sacrifice week fo male mice in all treated groups was Week 105.

Mice Table 4B: Intercurrent Mortality Rate Female

Week	0 mg|kg|day No. of Death	Cum. %	50 mg|kg|day No. of Death	Cum. %	150 mg|kg|day No. of Death	Cum. %	500 mg|kg|day No. of Death	Cum. %
0 - 52	6	9.23	4	6.15	11	16.92	6	9.23
53 - 78	7	20.00	18	33.85	20	47.69	11	26.15
79 - 91	18	47.69	12	52.31	15	70.77	10	41.54
92 - 102	11	64.62	10	67.69	3	75.38	22	75.38
Ter. Sac.	23	35.38	21	32.31	16	24.61	16	24.62
Total	N=65		N=65		N=65		N=65	

The terminal sacrifice week for female mice in the medium dose group was Week 96, and that for female mice in other treatment groups was Week 103..

Table 5A: Intercurrent Mortality Comparison Male Mice

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.0464
Homogeneity	Log-Rank	0.1536

Table 5B: Intercurrent Mortality Comparison Female Mice

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.6141
Homogeneity	Log-Rank	0.0052

Table 6A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons
Male Mice

Organ Name	Tumor Name	0 mg Control N=65	50 mg Low N=65	150 mg Med N=65	500 mg High N=65	P_Value Dose Resp	P_Value L vs Control	P_Value M vs Control	P_Value H vs Control
Adrenal, Cortex	B-Adenoma, Cortical Cells	1	0	0	0	0.7256	0.4767	0.4944	0.4304
	B-Adenoma, Subcapsular Ce	0	4	2	1	0.5368	0.0500*	0.2471	0.4430
Body, Whole/Cav	B-Hemangioma	1	1	2	2	0.1987	0.7291	0.5000	0.4056
	M-Hemangiosarcoma	2	2	1	1	0.6281	0.6559	0.4915	0.4056
	M-Histiocytic Sarcoma	2	0	0	0	0.9259	0.7291	0.7472	0.6787
	M-Lymphosarcoma	3	8	5	3	0.6338	0.0983	0.3694	0.5356
Duodenum	B-Adenoma	0	0	1	0	0.2086	.	0.5000	.
Gl, Harderian	B-Adenoma	7	6	4	6	0.4027	0.4268	0.7392	0.5638
Kidney	B-Adenoma, Tubule Cell	0	0	1	0	0.2086	.	0.5000	.
	M-Carcinoma, Tubule Cell	0	0	1	0	0.2086	.	0.5000	.
Liver	B-Adenoma, Hepatocellular	5	4	6	3	0.6237	0.4401	0.5000	0.5118
	M-Carcinoma, Hepatocellul	9	4	3	1	0.9883	0.8575	0.9358	0.9792
Lung	B-Adenoma, Bronchiolar-Al	9	12	10	6	0.7801	0.2133	0.4573	0.5214
	M-Carcinoma, Bronchiolar-	10	4	1	3	0.8932	0.8827	0.9942	0.8887
Muscle, Other	M-Sarcoma	1	0	0	0	0.7256	0.4767	0.4944	0.4304
Pancreas	B-Adenoma, Islet Cell	1	0	1	0	0.5830	0.4824	0.7529	0.4359
Pituitary	B-Adenoma	0	0	1	0	0.2086	.	0.5000	.
Skin/Subcutis	M-Sarcoma	1	3	1	2	0.3643	0.2824	0.7472	0.4160
Testis	B-Interstitial Cell Tumor	4	1	0	1	0.7742	0.7818	0.9361	0.7122
Thymus	M-Malignant Thymoma	1	0	0	0	0.7256	0.4767	0.4944	0.4304
Thyroid	B-Adenoma, Follicular Cel	0	1	1	0	0.4463	0.4824	0.5000	.
	B-C-Cell Adenoma	0	0	0	1	0.2134	.	.	0.4430
	M-Carcinoma, C-cell	0	1	0	0	0.4785	0.4824	.	.
	M-Carcinoma, Follicular C	0	1	0	0	0.4785	0.4824	.	.
Zymbal Gland	M-Squamous Cell Carcinoma	0	0	0	1	0.2134	.	.	0.4430

Table 6B: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons
Female Mice

Organ Name	Tumor Name	0 mg Control N=65	50 mg Low N=65	150 mg Med N=65	500 mg High N=65	P_Value Dose Resp	P_Value L vs Control	P_Value M vs Control	P_Value H vs Control
Adrenal, Cortex	B-Adenoma, Cortical Cells	1	0	0	0	0.7205	0.4767	0.4079	0.4944
	M-Carcinoma, Cortical Cel	1	0	0	0	0.7205	0.4767	0.4079	0.4944
Adrenal, Medull	B-Pheochromocytoma	1	0	0	0	0.7205	0.4767	0.4079	0.4944
Body, Whole/Cav	B-Hemangioma	3	0	0	1	0.7031	0.8613	0.7982	0.6833
	M-Hemangiosarcoma	0	4	4	0	0.8518	0.0503	0.0287*	.
	M-Histiocytic Sarcoma	4	3	0	2	0.7515	0.4498	0.8750	0.6405
	M-Lymphosarcoma	15	16	18	11	0.8773	0.4722	0.1393	0.7341
Cervix	B-Granular Cell Tumor	0	0	0	1	0.2733	.	.	0.4944
	B-Leiomyoma	1	2	0	4	0.0735	0.4738	0.4079	0.1733
	B-Schwannoma	1	1	0	1	0.5551	0.7291	0.4079	0.7472
	M-Leiomyosarcoma	1	1	0	0	0.7753	0.7296	0.4026	0.4889
	M-Sarcoma, Endometrial St	0	0	0	1	0.2733	.	.	0.4944
Gl, Harderian	B-Adenoma	2	1	2	1	0.6059	0.4647	0.5410	0.4915
	M-Carcinoma	1	0	0	0	0.7205	0.4767	0.4079	0.4944
Liver	B-Adenoma, Hepatocellular	1	2	1	0	0.8135	0.4738	0.6526	0.4944
	M-Carcinoma, Hepatocellul	0	1	0	0	0.4630	0.4828	.	.
Lung	B-Adenoma, Bronchiolar-Al	10	7	2	11	0.2039	0.6301	0.9379	0.4546
	M-Carcinoma, Bronchiolar-	2	7	1	0	0.9835	0.0639	0.3651	0.7416
Mammary, Female	M-Adenoacanthoma	0	0	1	0	0.2716	.	0.4156	.
	M-Carcinoma	0	2	3	0	0.7627	0.2359	0.0678	.
Muscle, Bi Fem	M-Rhabdomyosarcoma	1	0	0	0	0.7205	0.4767	0.4079	0.4944
Ovary	B-Adenoma	0	1	1	1	0.2934	0.4767	0.4079	0.4944
	B-Cystadenoma	0	0	2	1	0.2115	.	0.1632	0.4944
	B-Granulosa/Theca Cell Tu	0	0	1	0	0.2716	.	0.4156	.
	B-Luteoma	3	0	0	0	0.9792	0.8613	0.7982	0.8750
Pancreas	B-Adenoma, Islet Cell	1	0	0	1	0.4731	0.4767	0.4079	0.7472
Pituitary	B-Adenoma	2	1	3	1	0.6390	0.4738	0.3406	0.4915
Skin/Subcutis	B-Keratoacanthoma	0	0	0	1	0.2733	.	.	0.4944
	B-Papilloma, Squamous Cel	0	1	0	0	0.4658	0.4767	.	.
	B-Trichoepithelioma	0	0	1	0	0.2716	.	0.4156	.
	M-Sarcoma	3	0	2	1	0.6453	0.8568	0.3311	0.6747
Stomach, Nongl	B-Papilloma, Squamous Cel	0	1	0	1	0.3194	0.4767	.	0.4944
Thyroid	B-Adenoma, Follicular Cel	0	0	0	1	0.2778	.	.	0.5000
Uterus	B-Polyp, Endometrial Stro	5	4	4	5	0.4226	0.4428	0.5341	0.5872
	B-Schwannoma	0	1	0	0	0.4658	0.4767	.	.
Uterus	M-Carcinoma	0	0	0	1	0.2733	.	.	0.4944
	M-Sarcoma, Endometrial St	0	0	0	1	0.2733	.	.	0.4944
Vagina	B-Fibroma	0	0	0	2	0.0734	.	.	0.2416

Figure 1A: Kaplan-Meier Survival Functions for Male Rats

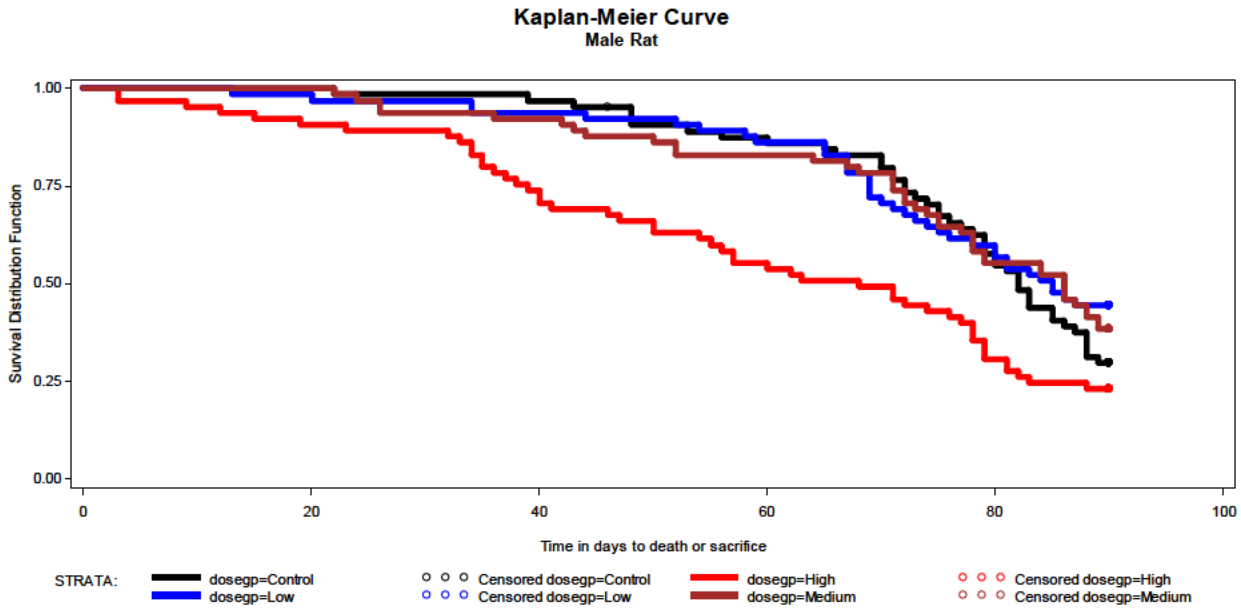


Figure 1B: Kaplan-Meier Survival Functions for Female Rats

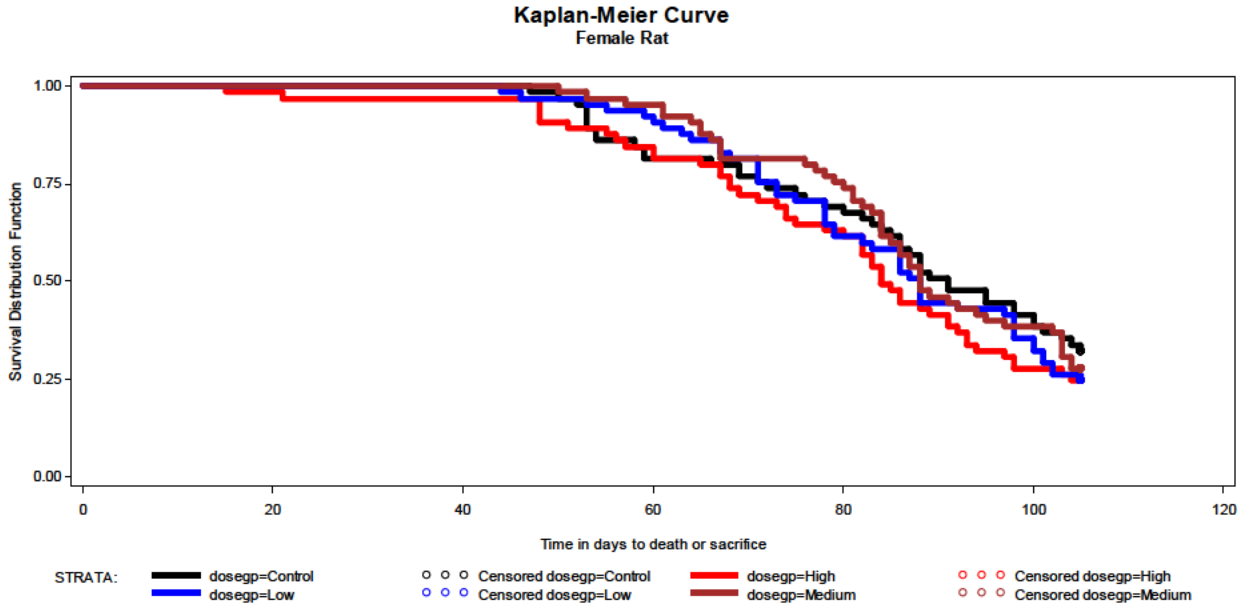


Figure 2A: Kaplan-Meier Survival Functions for Male Mice

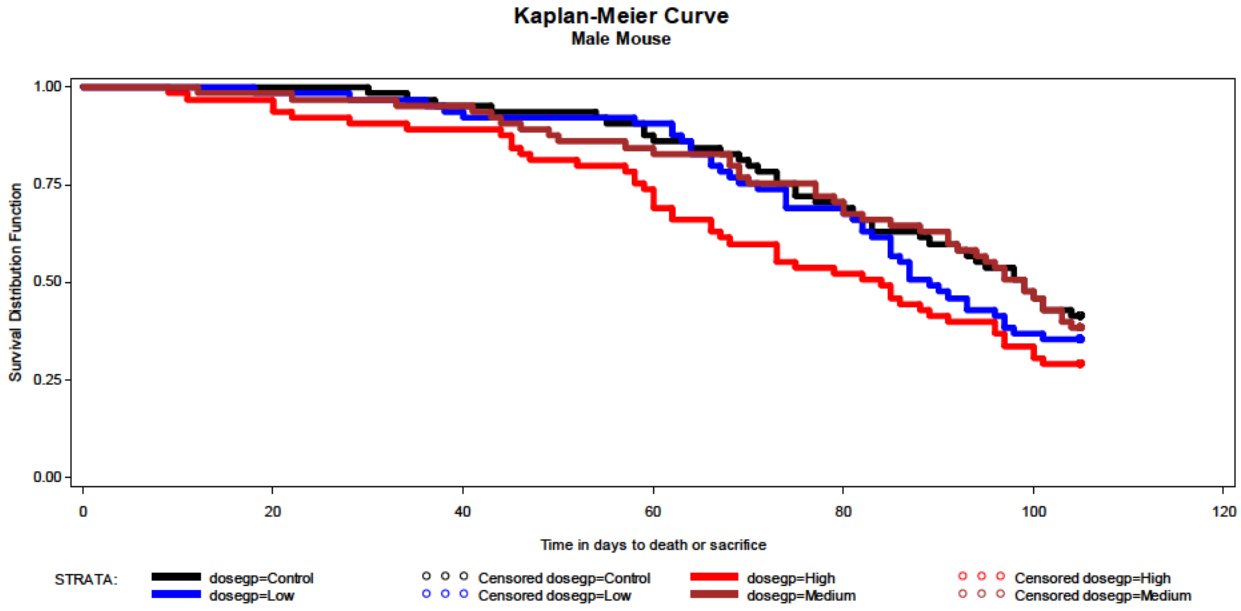
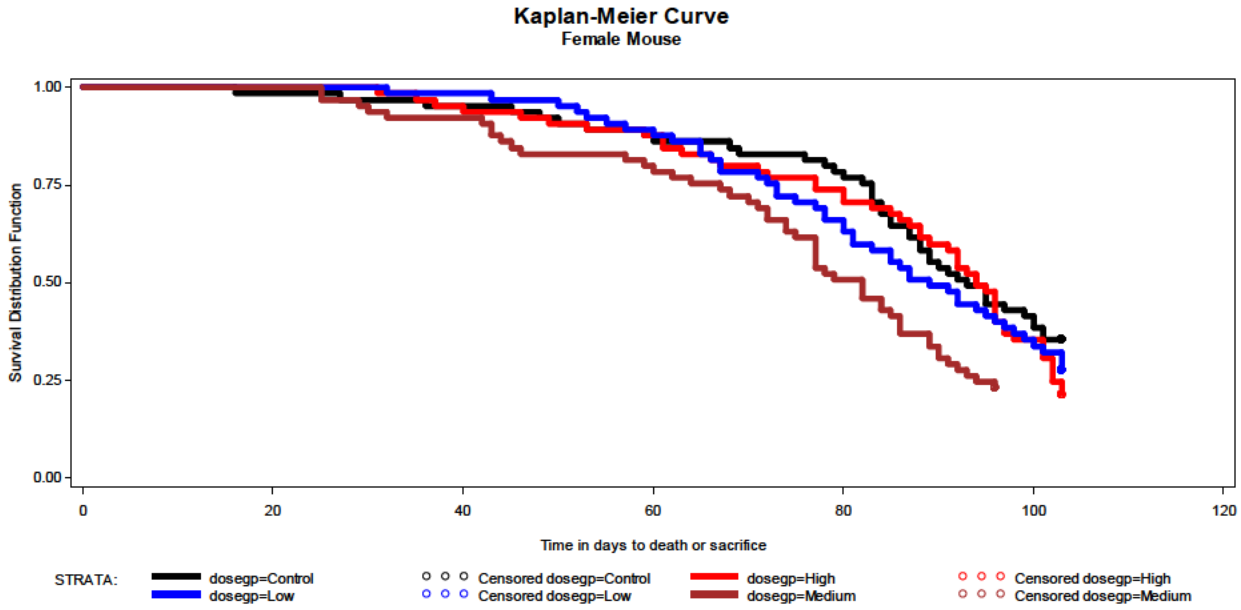


Figure 2B: Kaplan-Meier Survival Functions for Female Mice



7. References

1. Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva,* 311-426, 1980.
2. Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.
3. Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.
4. Tarone RE, "Test for trend in life table analysis", *Biometrika* 1975, 62: 679-82
5. Lin K.K. and Rahman M.A.," Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics,* 8(1), 1-15, 1998.
6. Haseman, J, "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339, 1983.
7. Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), May 2001.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

MOHAMMAD A RAHMAN
07/07/2014

KARL K LIN
07/08/2014
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 206316 **Applicant: Daiichi Sankyo** **Stamp Date: 1/8/2014**

Drug Name: Edoxaban **NDA/BLA Type: NDA, standard**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	X			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).		X		could not find define.pdf files. many datasets are close to 1 GB, cannot be open on laptop.

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? __Yes_____
If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.	X			
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.			X	don't know
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.		X		

Reviewing Statistician Date
Supervisor/Team Leader Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOHN P LAWRENCE
03/10/2014

HSIEN MING J HUNG
03/10/2014

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 206316 **Applicant:** Daiichi Sankyo **Stamp Date: January 8, 2014**

Drug Name: SAVAYSA (edoxaban tosylate) **NDA/BLA Type:** New original

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)			X	
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	X			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ____Yes____

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.	X			
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	X			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	X			

Comments:

File name: 5_Statistics Filing Checklist for a New NDA_BLA

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Yun Wang	13Feb2014
Reviewing Statistician	Date
Lei Nie	13Feb2014
Supervisor/Team Leader	Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YUN WANG
02/18/2014

LEI NIE
02/18/2014