CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

202155Orig1s000

MEDICAL REVIEW(S)

Subject:	Apixaban Clinical Review Addendum
NDA:	202155
Proposed Indication:	Reduction in the rate of stroke and systemic embolism in subjects with nonvalvular atrial fibrillation
Addendum Date:	December 17, 2012
Clinical Reviewers:	Martin Rose, M.D., J.D. (efficacy) and B. Nhi Beasley, Pharm.D. (safety)

Table of Contents

1 Introduction

This review addendum has two purposes:

- To address apixaban efficacy and safety issues raised in a special review by Dr. Thomas Marciniak, a medical team leader in DCRP who is not a member of the apixaban NDA review team but who has a strong interest in the safety issues described below.
- To discuss additional mortality analyses of ARISTOTLE, the study that is the primary support for the safety and efficacy of Apixaban for its proposed indication of reduction in the risk of stroke and systemic embolism in patients with non-valvular atrial fibrillation.

2 Discussion of Dr. Marciniak's Special Review

Dr. Marciniak on his own initiative filed a "special clinical review" of the apixaban NDA on December 11, 2012. The review focused on the issue of missing data (mostly missing follow-up information) in ARISTOTLE, primarily data for death. He also addressed data on bleeding and cancer in ARISTOTLE and APPRAISE-2. This addendum will focus on the issue of missing data for mortality. In connection with incomplete follow-up information, Dr. Marciniak recommended the following:

- The indication statement should include only stroke and systemic embolism.
- (b) (4)
- The Clinical Studies section of the label should include a discussion of the data quality problems in ARISTOTLE. It should summarize the dispensing errors and provide the missing follow-up statistics for both vital status and events. It should report that a change in one death eliminates the statistical significance of the death benefit and that, because of the missing data, we cannot have confidence in a death benefit.

In previously filed reviews, the clinical reviewers reached the conclusions concordant with Dr. Marciniak's first two recommendations: the indication should only include stroke and systemic embolism, and (b) (4) . Our views on those matters have not changed. However, we do not agree with his third recommendation concerning medication errors and the observed death benefit in ARISTOTLE.

We acknowledge that dispensing errors in ARISTOTLE were a major review issue. However, the Applicant's response to our CR letter, which included information derived from bottle labels, convinced us that the likelihood that the trial results were confounded by clinical events relating to medication errors was acceptably low (see our clinical review addendum of Dec. 10, 2012). Thus, the information that Dr. Marciniak would include in labeling regarding medication errors would be of negligible value in interpreting the findings of ARISTOTLE and would be more likely to confuse than to

edify practitioners. Consequently, we do not agree that information about the medication errors should be included in labeling.

Second, we do not agree that the labeling should include information regarding how many additional events in the apixaban arm or fewer events in the warfarin arm it would take to make the death finding not be statistically significant. A p-value of 0.0465 already implies that the results are close. One way to think about a statistically significant mortality finding is that there were so many fewer deaths in the apixaban arm than in the control arm (in the ITT analysis, 603 vs. 669, 66 fewer, **Table 1**) that it is very unlikely that the observed finding was due to chance. If this difference in deaths is reduced to 65, there are still many fewer deaths in the apixaban arm. In order to establish superiority, it is not appropriate to require both the "cushion" of fewer deaths needed to achieve statistical significance at the 0.05 level plus an additional cushion to take care of post hoc "what ifs." Moreover, the analysis that determines the number of events needed to overturn the mortality study finding (or stroke/se or major bleeding) is exploratory, highly conservative, non-random, and somewhat unrealistic.[1] It should not be used for labeling.

Table 1. ARISTOTLE Death and Primary Endpoint – ITT and On Treatment Analyses

	Apixaban		Warfarin		Δ	Apixaban vs. Warfarin		
Event[1]	n/N	%/yr	n/N	%/yr		Hazard Ratio	95% CI	p-value
Death ITT	603 / 9120	3.52	669 / 9081	3.94	66	0.89	(0.80, 1.00)	0.0465
Death Tx	265 / 9088	1.70	296 / 9052	1.94	31	0.87	(0.74, 1.03)	0.1130
Death TxLD+7	330 / 9088	2.10	372 / 9052	2.42	42	0.87	(0.75, 1.00)	0.0555
Death TxLD+30	429 / 9088	2.65	471 / 9052	2.97	42	0.89	(0.78, 1.01)	0.0763
Stroke SE ITT	212 / 9120	1.27	265 / 9081	1.60	53	0.79	(0.66, 0.95)	0.0114
Stroke SE Tx	176 / 9088	1.14	225 / 9052	1.49	49	0.77	(0.63, 0.93)	0.0080
Stroke SE TxLD+7	184 / 9088	1.18	236 / 9052	1.55	52	0.76	(0.63, 0.93)	0.0060
Stroke SE TxLD+30	218 / 9088	1.36	255 / 9052	1.62	37	0.84	(0.70, 1.00)	0.0526

Reviewer's analysis: erateHR create run eff tx txn7 txn30.sas, applicant's data: adefl, adbs2
Stroke SE is the primary endpoint. Δ=events in warfarin arm – events in apixaban arm
The period of analysis for the event was defined as:
ITT = randomization to January 30, 2011 (efficacy cut-off date)
Tx = first dose to last dose + 2 days (per protocol definition)
TxLD+7 = first dose to last dose + 7 days
TxLD+30 = first dose to last dose +30 days

[1] For example, to determine the number of apixaban-treated subjects needed to negate the statistically significant mortality finding, subjects are ordered by apixaban treatment, then censor date. Events are sequentially imputed to apixaban-treated subjects without events until the results are not statistically significant. Because the cox proportional hazards model is dependent on time, this analysis is highly conservative since the additional events are occurring early in the trial.

In addition, the missing data rate in ARISTOTLE cited by Dr. Marciniak (3.2% to 3.6% for vital status) is not especially large and more importantly, it is unclear whether the missingness was biased in favor of apixaban. The primary efficacy analysis in this study was the ITT analysis. In a large global study with many subjects who withdraw consent, it is possible to lose track of subjects that stop coming to a site before the cutoff date for the ITT analysis. Perhaps some were lost to follow-up because they had a stroke or died, which potentially biases the study results. However, there is no reason to believe that this was more likely in the apixaban arm than in the warfarin arm. We cannot directly address whether the missing data are biased in one direction. However, on-treatment analyses are less likely to have missing follow-up information during the period of analysis. In the on-treatment analyses shown in Table 1, events are counted if they occurred during the analysis period. If the ITT analyses were biased in favor of apixaban because of differential event rates in those whose data are missing, one would expect the on- treatment results for the primary endpoint and death to be less favorable for apixaban than the ITT results because of better follow-up while subjects are on treatment. If the assumed bias in ITT analysis were removed in this way, we would expect the hazard ratio for death or the primary endpoint to move in favor of warfarin compared to ITT, barring other effects. Instead, point estimates for the on-treatment analyses for the primary endpoint and death are both slightly lower – i.e., more favorable for apixaban – than the corresponding ITT results (**Table 1**). While this is not definitive proof of a lack of bias in the ITT analysis, it is reassuring and suggests that bias, if present, was not large. Given this reassurance from the on-treatment analyses and the lack of information to suggest bias in the ITT analysis, it would be confusing and potentially misleading to include data on follow-up statistics from ARISTOTLE in labeling.

3 Observed Persistence of the Effect of Apixaban on Death after end of Study Treatment

One other issue regarding the mortality finding should be mentioned. In the ITT analysis, there were 66 fewer all-cause deaths, but this was reduced to a difference of 31 deaths in the on-treatment analysis. However, for primary endpoint events, the analogous differences in event counts are 53 and 49, respectively. Thus, for primary endpoint events (which were mostly strokes), nearly all of the benefit of apixaban was established during the treatment period, as one might expect for an anticoagulant. However, for death, a substantial portion of the benefit of apixaban was established off treatment (**Table 1**). This might be interpreted to make the observed benefit of apixaban for all-cause mortality to be less credible.

However, there is an explanation for the observed persistence of the effect of apixaban on mortality after study drug is discontinued that does not undercut the strength of the overall finding in the ITT analysis. It relates to the large effect of apixaban on the rate of fatal stroke, the timing of death in fatal stroke cases, and the fact that physicians practice medicine conservatively, i.e., when a patient in a clinical study becomes

seriously ill, they are often taken off an investigational drug and treated with usual therapy for the patient's condition.

In ARISTOTLE, the difference between the treatment arms in fatal stroke strongly favored apixaban in the ITT analysis (38 vs. 65 deaths, HR=0.58, 95% CI: 0.39, 0.86). It is notable that while fatal stroke accounts for less than 10% of deaths in the ARISTOTLE ITT analysis, the difference in the number of fatal strokes between the treatment arms accounts for more than 40% of the overall difference in deaths. This relationship was also observed in RE-LY (see our review of Applicant's Complete Response dated 10 Dec. 2012). This difference suggests that the observed overall difference in favor of apixaban is not due to chance. It thus seems useful to examine the timing of stroke mortality in relation to the subjects last dose of study drug and last known stroke event.

Figure 1 is a display of the difference in days between the last dose of study drug and the last known adjudicated stroke in subjects with death adjudicated as a CV death due to stroke ("fatal stroke") in 102 of the 109 subjects with a fatal stroke in the ITT analysis.[2] Data for both treatment arms are combined in this figure and the 2 others that follow, but all of the trends in the data in the 3 figures discussed here were similar in the two treatment arms (data not shown).The majority of subjects with a fatal stroke (72 subjects, or 71%) had their last dose of study drug on the day of the stroke or one day earlier. This is consistent with the conservative practice of discontinuing an experimental drug when a patient becomes seriously ill, especially when other recommended therapies are available.

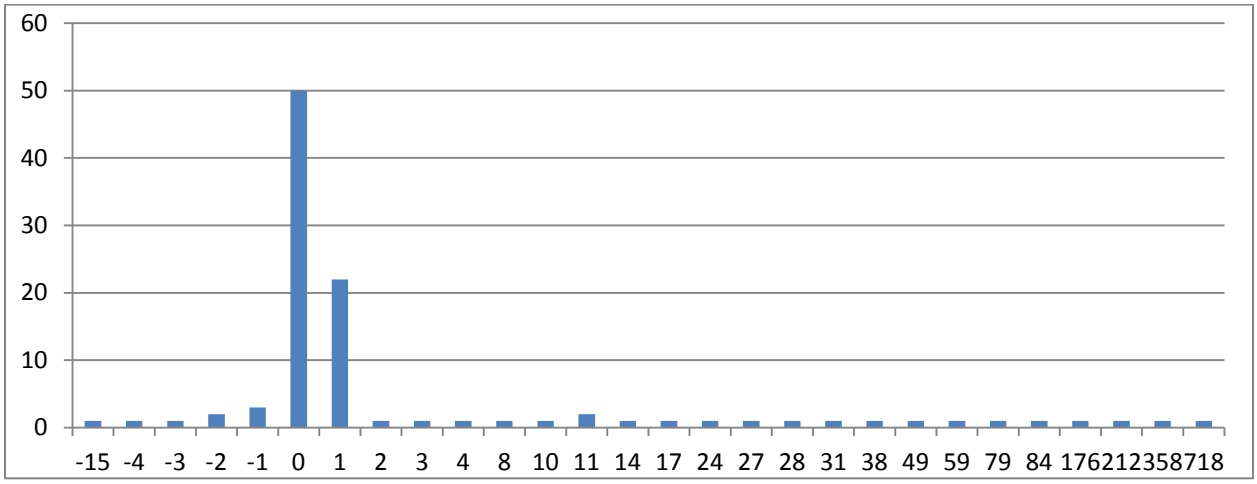
Figure 2 is a display of the difference in days between the last dose of study drug and death for the same 102 subjects, and **Figure 3** shows data for the days between the final stroke and death. For fatal stroke, which was the single largest contributor to the difference in deaths between the treatment arms, the data indicate only 30 of the subjects (29%) died "on treatment" (i.e., 0 to 2 days after their last dose of study drug, and 56 (55%) died within 7 days of their final last dose (**Figure 2**). Forty-eight subjects (47%) died within 2 days of their final stroke, 77 (75%) died within 7 days of their final stroke (**Figure 3**).

Thus, the data from the sample of subjects with fatal strokes indicate that a substantial part of the apparent late effect of apixaban on death has an explanation that does not undercut the observed beneficial effect of apixaban on mortality: In 29% of subjects with a known date for their final and fatal stroke, the stroke occurred after the on-treatment period. However, over half of deaths occurred more than 2 days after the final stroke and 1/4 of subjects in the sample died more than 7 days after their last stroke. Thus, it is not surprising that the effect of apixaban on mortality is not confined to the on-treatment period.

---

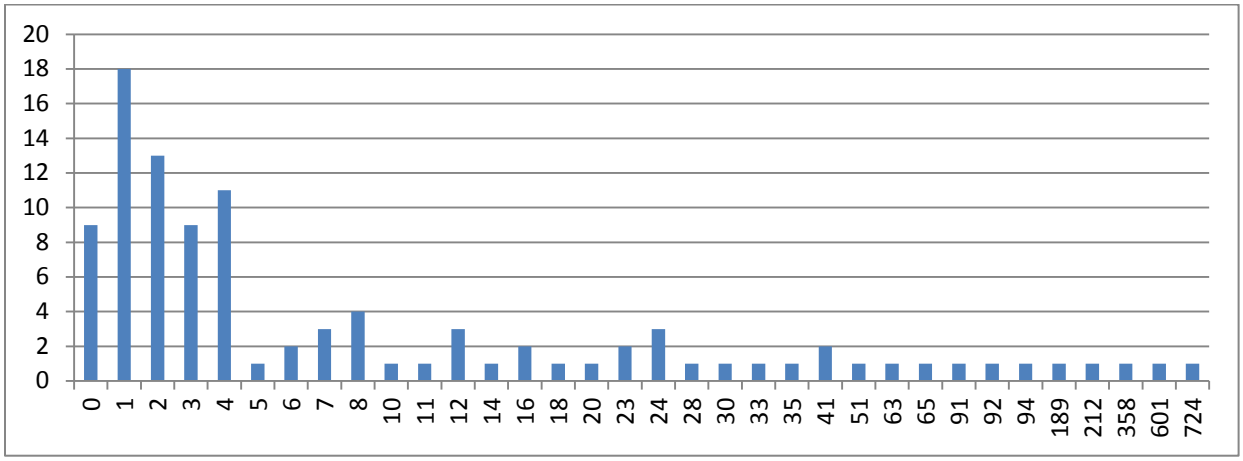
[2] The remaining 7 subjects with a fatal stroke (5 in the warfarin arm and 2 in the apixaban arm) did not have either a "stroke" in the study database nor a date for the stroke, but they had a date and adjudicated cause for mortality. While these patients were included in analyses of death, they were not included in analyses of stroke or the primary endpoint performed by the applicant or by FDA.

Figure 1. Distribution of Subjects with Fatal Stroke: Days from Last Dose of Study Drug to Final Stroke (N=102)



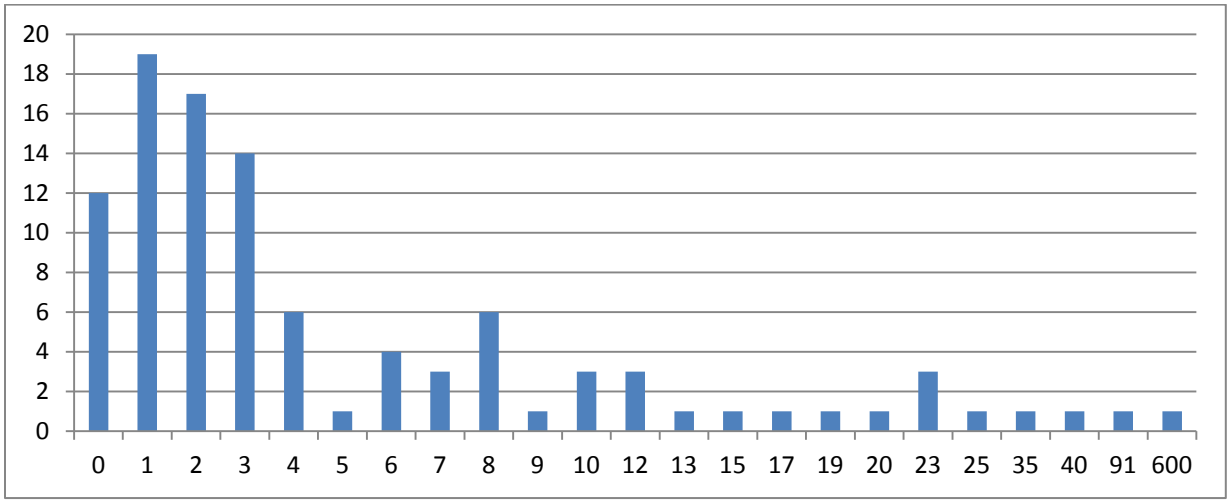
Horizontal axis is days between last dose of study drug and adjudicated final stroke event. Negative values mean stroke occurred before the last dose; positive values mean the last dose occurred before the stroke.
Vertical axis represents the number of subjects at each time point.

Figure 2. Distribution of Subjects with Fatal Stroke: Days from Last Dose of Study Drug to Death (N=102)



Horizontal axis is days between last dose of study drug and adjudicated CV death due to stroke.
Vertical axis represents the number of subjects at each time point.

Figure 3. Distribution of Subjects with Fatal Stroke: Days from Final Stroke to Death (N=102)



Horizontal axis is days between adjudicated final stroke and adjudicated death due to stroke. Vertical axis represents the number of subjects at each time point.

Notably, despite the narrowing of the difference in deaths between the treatment arms after treatment, the hazard ratio for death (apixaban vs. warfarin) is not worse (**Table 1**).

Also, while the following data from other trials do not directly address the issue of the prolongation of the apparent effect of apixaban, they are reassuring in that they indicated the pattern of timing with respect to treatment of deaths vs. primary endpoint events in ARISTOTLE is not unique.

In ARISTOTLE a large number of deaths occurred after the end of treatment, a pattern that was not observed for the primary endpoint events. The ratio of primary endpoint events in the on treatment period (i.e., first dose to last dose + 2 days) vs. those in the ITT period was 401:444, or 0.84. Because this ratio is reasonably close to unity, it should not be surprising that the differences in primary endpoint events between the treatment arms (warfarin minus apixaban) were similar in the two periods (53 and 49 in the ITT and on-treatment analyses, respectively). However, the ratio of deaths during the on treatment period vs. the ITT period was 561:1272, or 0.44.

It is not clear why the ratio of deaths in the on-treatment period vs. ITT period is higher for primary endpoint events than for deaths. Better ascertainment of death than stroke after treatment may account for some of the difference, but it seems likely that discontinuation of study drug because of a serious illness (such as a stroke) that eventually was fatal accounted for some of the difference (see **Figure 1** to **Figure 3** and associated text above).

A similar pattern regarding the proportion of primary endpoint events vs. deaths that occurred on-treatment was observed in ROCKET and RE-LY, which were warfarin-controlled trials of rivaroxaban and dabigatran, respectively, performed in patients with nonvalvular atrial fibrillation (**Table 2**). Primary endpoint events were defined similarly in all 3 trials. As in ARISTOTLE, in both ROCKET and RE-LY the ratio of total deaths

on treatment vs. deaths during the ITT period was substantially lower than the analogous ratio for primary endpoint events: 0.37 vs. 0.75 respectively for ROCKET and 0.55 vs. 0.74 respectively for RE-LY.

Table 2. Atrial Fibrillation Trials – Endpoint Events by Period

	ARISTOTLE Apixaban vs. W	**ROCKET Rivaroxaban vs. W**	**RE-LY Dabigatran 150 mg vs. W**
Deaths: n in ITT HR (95% CI)	1272 0.89 (0.80, 1.00)	1264 0.92 (0.82, 1.03)	923 0.88 (0.77, 1.00)
Deaths: n in Tx HR (95% CI)	561 0.87 (0.74, 1.03)	458 0.85 (0.70, 1.02)	506 NP
Ratio: n in Tx / n in ITT	**0.44**	**0.37**	**0.55**
Stroke/SE : n in ITT HR (95% CI)	477 0.79 (0.66, 0.95)	575 0.88 (0.74, 1.03)	336 0.65 (0.52, 0.81)
Stroke/SE: n in Tx HR (95% CI)	401 0.77 (0.63, 0.93)	432 0.79 (0.65, 0.95)	250 0.64 (0.50, 0.81)
Ratio: n in Tx / n in ITT	**0.84**	**0.75**	**0.74**

W = Warfarin
Stoke/SE = Primary endpoint: time to first stroke or systemic embolism in each study
Death = All cause death, analyzed as time to first event
ITT = Intent to treat period, counting events from randomization to the time that sites were notified that the event target had been reached and the study was to end.
Tx = On-treatment period, counting events from date of first dose to date of last dose + 2 days for ARISTOTLE and ROCKET and from first dose to last dose for RE-LY
HR is for the experimental drug vs. warfarin in each study; data for ROCKET from NDA review; data for RE-LY from review or from NDA for on treatment data only
NP= not provided.

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

/s/

BACH N BEASLEY
12/21/2012

MARTIN ROSE
12/21/2012

CLINICAL REVIEW (Updated)

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
DIVISION OF CARDIOVASCULAR AND RENAL PRODUCTS

Date: December 17, 2012

Reviewer: Thomas A. Marciniak, M.D.
Medical Team Leader

NDA: 202-155

Drug: apixaban (Eliquis®)

Indication: To reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation

Subjects: Completeness of follow-up and bleeding and cancer

Summary and Recommendations
Because of a special interest in and experience with two issues, completeness of follow-up and cancer, I reviewed the apixaban studies regarding these issues. I filed an initial review on December 11, 2012. I updated that review for a sponsor submission dated December 14, 2012, on study closeout procedures and for FDA discussions regarding the cancer issue and I corrected typos and minor inaccuracies. This updated review incorporates and completely replaces my initial review. My summary and recommendations remain unchanged:

I document below that completeness of follow-up and reporting of dates were poor in ARISTOTLE. Our confidence in the fragile alleged death benefit (with one additional death in the apixaban arm eliminating statistical significance) is destroyed by the missing vital status. Our confidence in the superiority of the stroke benefit to warfarin is also challenged by incomplete follow-up. Finally, the ARISTOTLE and APPRAISE-2 trials show an association between bleeding and solid cancers also seen in other anticoagulant and antiplatelet drug trials.

I recommend the following:

- The indication statement should include only stroke and systemic embolism.
- (b) (4).
- The Clinical Studies section of the label should include a discussion of the data quality problems in ARISTOTLE. It should summarize the dispensing errors and provide the

missing follow-up statistics for both vital status and events. It should report that a change in one death eliminates the statistical significance of the death benefit and that, because of the missing data, we can not have confidence in a death benefit.

- The data regarding bleeding and cancer should be presented and discussed at an advisory committee meeting. If the rigorous analysis of the ARB trials confirms a risk for ARBs and cancer, then the data regarding ARBs and cancer should also be presented and discussed at an advisory committee meeting. It may be advantageous to have both topics addressed at the same meeting.

- The data regarding bleeding and cancer should be included in the apixaban label and in the labels for all antiplatelet and anticoagulant drugs. If the advisory committee meeting discussing bleeding and cancer is scheduled promptly then the labeling regarding bleeding and cancer can be delayed until after the meeting.

Completeness of Follow-up and Fragility of Results in ARISTOTLE

Definition of Completeness of Follow-up

The clinical study report (CSR) for the ARISTOTLE trial of apixaban vs. warfarin in atrial fibrillation states that vital status could not be determined for 2.0% in the apixaban group and 2.2% in the warfarin group (380 patients in both groups total). The main study publication reported the same vital status statistics. (Granger, Alexander et al. 2011) However, the rates of discontinuation from the study were much higher, 25.3% in the apixaban group and 27.5%, with 10.1% of apixaban patients and 10.0% of warfarin patients discontinuing at their own request. While these reported statistics for completeness of follow-up are not good, my recent experience with other outcome trials suggests that the sponsor's reporting of completeness of follow-up is usually optimistic compared to analyses of the submitted datasets. Hence I analyzed the datasets for completeness of follow-up.

I assert that there is a straightforward definition of completeness of follow-up: Most outcome studies have a specified global study end date or censoring date for efficacy outcomes. A few have a pre-specified duration of follow-up from randomization such as two years. I assert that follow-up is complete if the patient has documented follow-up on or after the specified end date.

Per a statistical analysis plan appendix and the NEJM publication ARISTOTLE had a cutoff date for efficacy outcomes of January 30, 2011. However, I note that the December 14, 2012, submission refers to "this common efficacy cut-off date (31-Jan-2011)" and has the following detailed description:

> "By December, 2010 we had confirmed our view that we would reach 448 events in January, 2011. We therefore organized our CRO partners and site monitors to prepare them for the efforts that would be involved in the close out process, and asked them to communicate to sites our expectation that the efficacy cutoff date would be January 31, 2011. Furthermore, all visits for the cessation of study drug (and initiation of VKA or other antithrombotic agents) were to occur no sooner than 31-Jan-2011."

These statements make it unclear whether the pre-specified efficacy cutoff date was January 31, 2011, or January 30, 2011. Hence for ARISTOTLE follow-up for a patient is complete if the patient had documented follow-up on or after January 30, 2011--or January 31, 2011?

Determining completeness of follow-up has another complication: The type of follow-up typically varies in outcome studies. For patients who have a face-to-face study visit with the investigator on or after the study end date follow-up is complete for all study outcomes or events. However, for some patients final follow-up may consist of a phone call with the patient or a spouse or a primary physician, for others a report of a hospitalization, and for still others a newspaper obituary or a registry report of alive or dead. The level of detail available from these latter, non-face-to-face follow-ups varies and, while usually adequate for determining vital status, may not be adequate for ascertaining endpoints or adverse events. I recommend estimating two levels of completeness of follow-up: (1) vital status; and (2) events. For the former I accept any type of documented follow-up having a date unambiguously referencing the patient as alive. For the latter I accept reports of documented face-to-face visits, hospitalizations or other events, and phone calls with documented queries regarding events. Because of ambiguities in case report form (CRF) design and in reporting by sites determining completeness of event follow-up requires subjective judgments.

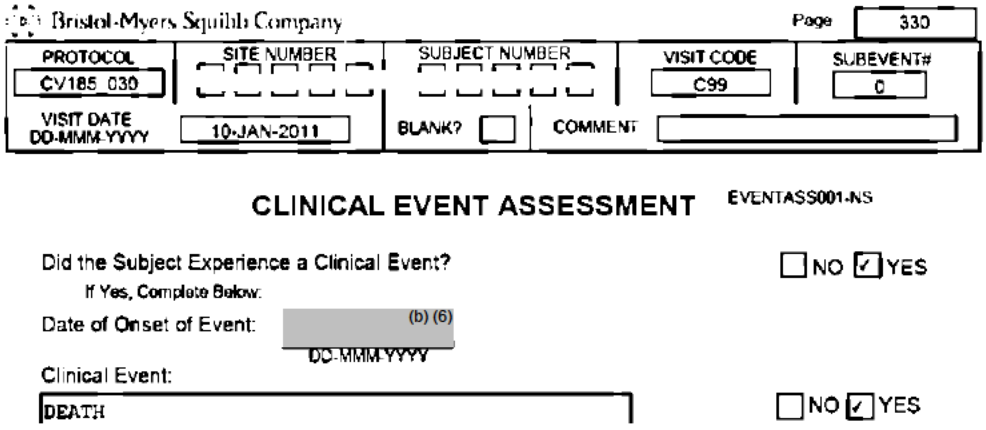
Completeness of Vital Status Follow-up

For one estimate of completeness of vital status follow-up in ARISTOTLE I used the sponsor's variable *e_cddn* described as "Censor Date for Death ITT Period" from the sponsor's ADEFS.XPT ("Efficacy Summary") dataset. Counting good vital status as either known dead at any time or a censor date of January 30, 2011, or later, 3.2% of patients (589) are missing vital status. For a second estimate I also analyzed all of the datasets for the maximum dates of events, procedures, vital sign recordings, and status reports to estimate a date of last follow-up for every patient. By this latter analysis of the raw data 3.6% of patients (659) are missing vital status. Either number, 589 or 659, greatly exceeds the sponsor's and investigators' reports of 380 missing vital status follow-up.

Date Recording Problems in ARISTOTLE

Both the sponsor's and my estimates of vital status follow-up are likely optimistic: Dates are not infrequently misrecorded or misinterpreted in ARISTOTLE. The death CRF excerpt in Figure 1 illustrates the possibility for misrecording or misinterpretation: The "visit date" for the death CRF precedes the death date. (Note that the reported death date was (b) (6) . This patient was one of the three warfarin patients reported to have died (b) (6) —compared to no apixaban patients. See the further discussion of problems with death reports below.)

Figure 1: Death in the Future



Bristol-Myers Squibb Company Page 330

PROTOCOL	SITE NUMBER	SUBJECT NUMBER	VISIT CODE	SUBEVENT#
CV185_030			C99	0

VISIT DATE DD-MMM-YYYY: 10-JAN-2011 BLANK? ☐ COMMENT

CLINICAL EVENT ASSESSMENT EVENTASS001-NS

Did the Subject Experience a Clinical Event? ☐ NO ☑ YES

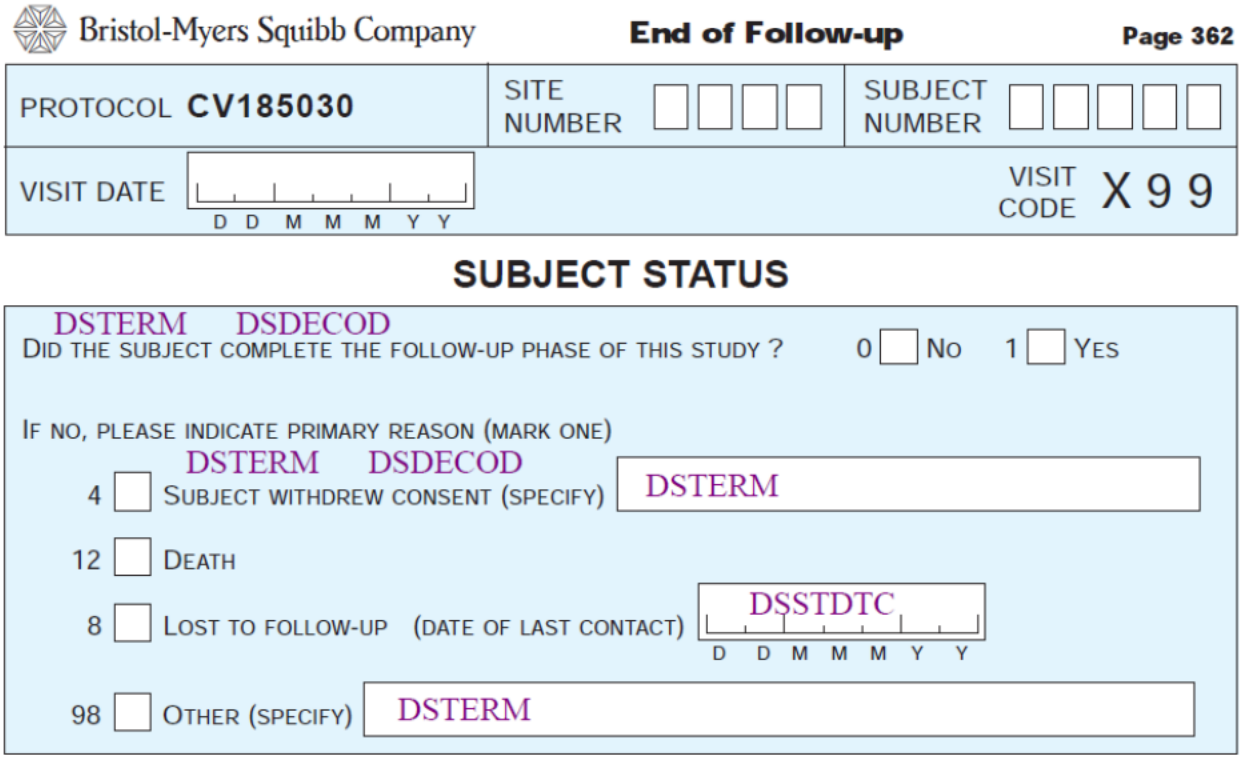
If Yes, Complete Below:

Date of Onset of Event: (b) (6) DD-MMM-YYYY

Clinical Event: DEATH ☐ NO ☑ YES

The "visit date" may have little connection to a date on which the patient was actually observed or contacted. In many cases there is no way of determining whether the "visit date" is an observation date or the date of recording or something else. However, both I and the sponsor based our dates of last vital status partially on "visit dates". In particular a critical form for follow-up, the End of Follow-up CRF shown in Figure 2, has this ambiguity regarding "visit date".

Figure 2: End of Follow-up CRF



Bristol-Myers Squibb Company **End of Follow-up** Page 362

PROTOCOL CV185030	SITE NUMBER ☐☐☐☐	SUBJECT NUMBER ☐☐☐☐☐
VISIT DATE D D M M M Y Y		VISIT CODE X 9 9

SUBJECT STATUS

DSTERM DSDECOD

DID THE SUBJECT COMPLETE THE FOLLOW-UP PHASE OF THIS STUDY ? 0 ☐ NO 1 ☐ YES

IF NO, PLEASE INDICATE PRIMARY REASON (MARK ONE)

DSTERM DSDECOD

4 ☐ SUBJECT WITHDREW CONSENT (SPECIFY) DSTERM

12 ☐ DEATH

8 ☐ LOST TO FOLLOW-UP (DATE OF LAST CONTACT) DSSTDTC D D M M M Y Y

98 ☐ OTHER (SPECIFY) DSTERM

What does the "visit date" on the End of Follow-up CRF represent? Only for lost to follow-up is the date of last contact to be recorded on this CRF. For death there is the death form with a field for date of death, but what about "withdrew consent" or "other" or even "complete"? It is easy to document that "visit date" for "withdrew consent" likely does not represent the date on which the patient visited the site or withdrew consent. For example, one patient discontinued treatment on 16jul00 with the last verifiable events on 18jun00. However, the disposition (DS) dataset has a "Start Date/Time of Disposition Event" for withdrew consent of 12apr11 and the sponsor counts the patient as completing follow-up, censoring on 30jan11. Withdrawing consent on 12apr11, long after the trial ended, is not rational and would not represent a withdrawal of consent during the ITT period, as the sponsor classifies this patient. Another patient is similar, with end of treatment and last events on 06may10 but withdrawal of consent allegedly on 24feb11 with sponsor's censoring on 30jan11. I count both of these patients (and other similar ones) as having incomplete vital status follow-up, partially explaining and justifying why my estimate of incompleteness of vital status follow-up is higher than the sponsor's.

There are other examples of anomalous dates, e.g., at least three patients have dates of last contact by "direct contact with subject" long after the patient was reported dead. In fact, about 65% of patients who died have a visit date or other date greater than the date of death. Furthermore, for patients who did not die during the study, we do not have an unequivocal last date against which to check the validity of reported dates. There is no good way to detect or resolve many of these date inconsistencies even with a painstaking manual review of the CRFs--and we do not have most of the CRFs. Any estimates of completeness of vital status follow-up are optimistic.

Fragility of the Death Benefit in ARISTOTLE

While the estimates of missing vital status follow-up are concerning (>3%) and likely higher, how do they relate quantitatively to the reported death benefit? The sponsor claims that there were statistically significantly fewer deaths in the apixaban arm than in the warfarin arm in ARISTOTLE based on the analysis in Table 1 excerpted from the clinical study report:

Table 1: Summary of Adjudicated Causes of Death during the Intended Treatment Period – Randomized Subjects (Excerpt from Sponsor's Table 7.2.1)

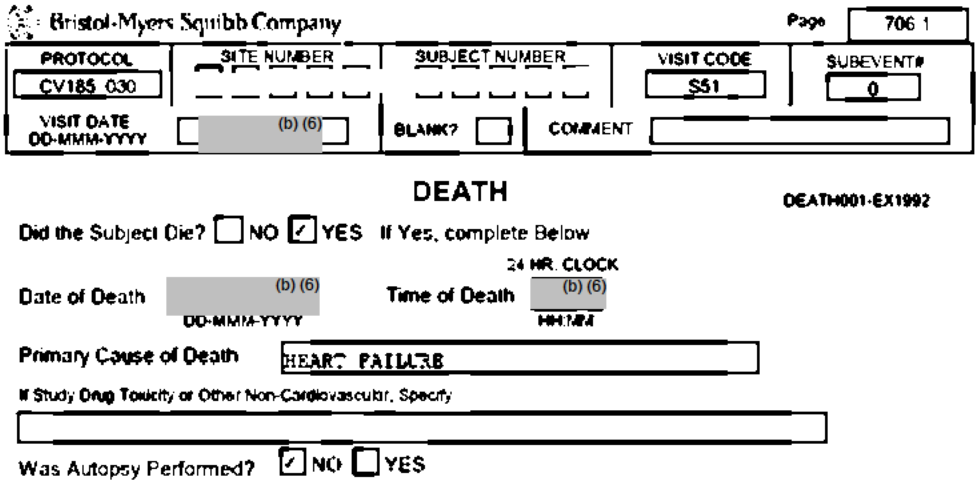
		Apixaban N = 9120	Warfarin N = 9081
ALL-CAUSE DEATH, n (%)		603 (6.61)	669 (7.37)
EVENT RATE (%/YR)		3.52	3.94
HAZARD RATIO (APIXABAN/WARFARIN)	0.89		
95% CI FOR HAZARD RATIO	(0.80, 1.00)		
P-VALUE	0.0465		

The p-value is close to 0.05 and the upper confidence limit for the hazard ratio is 1.00 so we know that this result is very fragile. How fragile? A change in only one death (one more with apixaban or one fewer with warfarin) could make this result nominally statistically insignificant. This critical number 1 is dwarfed by 317, the estimated—but still likely optimistic—number of patients with missing vital status in the apixaban arm. Even the total difference in deaths

between the two arms (66) is a small fraction of number of apixaban patients with missing vital status.

This numeric fragility may not be surprising but, in addition, the quality of the documentation of the deaths is fragile or suspect. I examined deaths around the time of the cutoff date (January 30, 2011) for efficacy outcomes. Three warfarin patients (and no apixaban patients) died (b) (6) One patient in Hungary discontinued treatment on 4jan11 for bleeding with the clinical event assessment shown in Figure 1 dated 10jan11 and reporting death on (b) (6). Another CRF shown in Figure 3 does appear to confirm the date of death as (b) (6) at virtually midnight.

Figure 3: Death near Midnight on Cutoff Date


Bristol-Myers Squibb Company

Page 706 1

PROTOCOL	SITE NUMBER	SUBJECT NUMBER	VISIT CODE	SUBEVENT#
CV185 030			S51	0

VISIT DATE DD-MMM-YYYY (b) (6) BLANK? COMMENT

DEATH

DEATH001-EX1992

Did the Subject Die? NO YES If Yes, complete Below

Date of Death (b) (6) DD-MMM-YYYY

Time of Death (b) (6) 24 HR. CLOCK HH:MM

Primary Cause of Death HEART FAILURE

If Study Drug Toxicity or Other Non-Cardiovascular, Specify

Was Autopsy Performed? NO YES

Figure 3 is another example of a "visit date" occurring after death. Usually the CRFs include for a death a handwritten narrative faxed by the investigator to the sponsor's monitors. Such a narrative was not provided for this death. Other confirmatory documentation, such as death certificates, were not submitted in general.

Conversely, an apixaban patient suffered a pulmonary embolism on (b) (6), was hospitalized on (b) (6) and died. The site reported the date of death as 31jan11 and hence the death is not counted as an apixaban death. Could this represent (b) (6) error?

Finally, a warfarin patient suffered a stroke on day 221, was diagnosed with lung cancer on day 229, and then was reported lost to follow-up on day 281. However, the investigator contacted the family on day 868 after the study ended and recorded that the family had reported the patient as having died from lung cancer but date unknown. For the death analysis the sponsor counts this patient as having died on the lost-to-follow-up date and I followed the sponsor's assignment for the primary death analyses in this and my initial review. However, because the date of death was completely unknown, I had classified this patient as having bad vital status follow-up in my initial review. In this review, to be consistent with the definition of good vital status as "known dead at any time during the study", I count the patient as having good vital status.

The problems with these death reports, like the problems with visit dates and the substantial missing vital status, destroy our confidence that apixaban decreases death rates compared to warfarin. Two sensitivity analyses regarding deaths from ARISTOTLE add to our lack of confidence in a death benefit for apixaban: If one censors at 31jan11 rather than (b) (6), the p value is 0.05. If one counts all reported deaths, not just those censored at (b) (6) or 31jan11, then the death difference is remote from nominal statistical significance, i.e., p = 0.07.

Warfarin and Deaths

While it appears unclear whether apixaban is superior to warfarin regarding death rates, a good question is whether warfarin is superior to placebo regarding death rates. We may be more impressed with marginal p values if the comparator is known to be active. While many clinicians may assume that warfarin reduces all-cause mortality because of its substantial effects upon stroke rates, a death benefit for warfarin compared to placebo is also unclear as I document below.

I extracted the death statistics from the six historical trials of warfarin vs. placebo in atrial fibrillation, i.e., the ones used for efficacy noninferiority analyses of new anticoagulants: AFASAK (Petersen, Boysen et al. 1989), BAATAF (BAATAF_Investigators 1990), CAFA (Connolly, Laupacis et al. 1991), EAFT (EAFT_Study_Group 1993), SPAF (SPAF_Investigators 1991), and SPINAF (Ezekowitz, Bridgers et al. 1992). AFASAK reported vascular deaths (on treatment?) by arm but not all-cause deaths by arm, reporting only that 71 patients died and that an ITT analysis of all deaths showed no difference in either vascular or total mortality. Using the *metan* package for Stata 12 I performed a random effects meta-analysis of the all-cause death risk ratios in these studies assigning neutral results for AFASAK. I show the meta-analysis results in Table 2.

Table 2: Meta-Analysis of Deaths in the Warfarin vs. Placebo Trials in Atrial Fibrillation

Study	RR	[95% Conf.	Interval]	% Weight
AFASAK	1.024	0.638	1.645	23.96
BAATAF	0.415	0.211	0.818	13.98
CAFA	1.277	0.515	3.164	8.52
SPAF	0.754	0.266	2.134	6.66
SPINAF	0.695	0.369	1.310	15.61
EAFT	0.953	0.649	1.398	31.28
D+L pooled RR	0.829	0.626	1.099	100.00

Heterogeneity chi-squared = 6.45 (d.f. = 5) p = 0.265
I-squared (variation in RR attributable to heterogeneity) = 22.4%
Estimate of between-study variance Tau-squared = 0.0276

Test of RR=1 : z= 1.30 p = 0.192

RR = risk ratio warfarin/placebo; D+L = DerSimonian-Laird method

Only one of the historical trials (BAATAF) shows a significant mortality benefit. While some may like to interpret the results as suggestive of a mortality benefit because of the favorable point estimate of the risk ratio, the p value of 0.192 does not support statistical significance or

firm conclusions of a mortality benefit for warfarin. The conservative interpretation of the mortality comparison in ARISTOTLE is that we are comparing apixaban to an inactive agent.

Completeness of Event Follow-up

The problem of completeness of follow-up is not limited to vital status. I estimated completeness of event follow-up based on events, procedures, vital sign recordings, and last direct contacts with the patient (i.e., the types of reports relevant to events, endpoint or adverse, or the absence thereof) but not counting the flawed status report visit dates. For these reports about 15% of patients, or over 2,700 patients, have incomplete follow-up. Compare 1,349, the number of apixaban patients with incomplete follow-up, to 53, the difference in primary endpoints. There is a vast amount of missing follow-up in which endpoints may be hidden or missed.

COMMENT: The alleged death benefit of apixaban compared to warfarin is fragile as reported by the sponsor, i.e., p = 0.046, a change in only one death rendering the difference statistically insignificant. Furthermore, the validity of this fragile benefit depends upon having 100% valid data. The substantial missing vital status follow-up, the problems with date recordings, and the lack of a significant death benefit for warfarin destroy confidence that apixaban reduces all-cause mortality. We might have some confidence that apixaban reduces stroke rates in atrial fibrillation: Stroke reduction with warfarin is substantial and allegedly apixaban improves upon the warfarin reductions. However, our confidence in the apixaban stroke benefit is also reduced by the substantial missing event follow-up as well as by the data quality issue identified by the primary reviewers, i.e., errors in the dispensing of the study drug or documentation thereof.

The sponsor's December 14, 2012, provides some details regarding the closeout procedures. While, as I note above, it confounded the question of the pre-specified efficacy cutoff date, the procedures it describes for closeout seem reasonable. However, there are two more limitations of that submission and the trial:

- *The submission includes, in a "SLIDE DECK – END OF STUDY PLAN" a "SURVIVAL FOLLOW-UP DATA / PRIMARY OUTCOME" CRF that would appear to have been able to alleviate some of the follow-up problems I described above because it has checkboxes for the source of the follow-up, e.g., "DIRECT CONTACT FOR SUBJECT IN PERSON", "MEDICAL RECORDS", etc. However, this CRF was not included in the BLANKCRF.PDF file in the original submission nor are variables from it defined in the DEFINE.PDF file. The latter omission appears to be a failure of the SDTM standard: There are source of follow-up variables from this CRF as data in the SUPPDS.XPT dataset that I used my analyses. There are obvious problems with these variables, e.g., these variables report three patients as having direct subject contact long after death (the anomalous dates I mentioned above.). Finally, sites used this CRF infrequently because only about 6% of patients who did not die during the study have a source of follow-up recorded.*

- *While the December 14, 2012, submission describes reasonable closeout procedures that theoretical should help to insure good follow-up, the follow-up statistics I have provided above document reality. The reality is that follow-up in ARISTOTLE was poor.*

I, like most FDA reviewers, would like to conclude that apixaban is effective in atrial fibrillation—we would like to have alternatives to warfarin. While there are problems with the ARISTOTLE data, the stroke results are reasonably in the right direction and our priors are that a drug with an anticoagulant pharmacodynamic action should be effective. I consider it to be very unfortunate that ARISTOTLE, like many other recent outcome trials, has substantial problems with data quality. Some of the responsibility for the data quality problems rests with us, the FDA: We have approved drugs ignoring similar data quality issues, granting superiority claims and not discussing in the labels the data quality issues. We must stop doing this.

If we approve apixaban I recommend the following for the labeling:

- *The indication statement should include only stroke and systemic embolism.*
- (b) (4).
- *The Clinical Studies section of the label should include a discussion of the data quality problems in ARISTOTLE. It should summarize the dispensing errors and provide the missing follow-up statistics for both vital status and events. It should report that a change in one death eliminates the statistical significance of the death benefit and that, because of the missing data, we can not have confidence in a death benefit.*

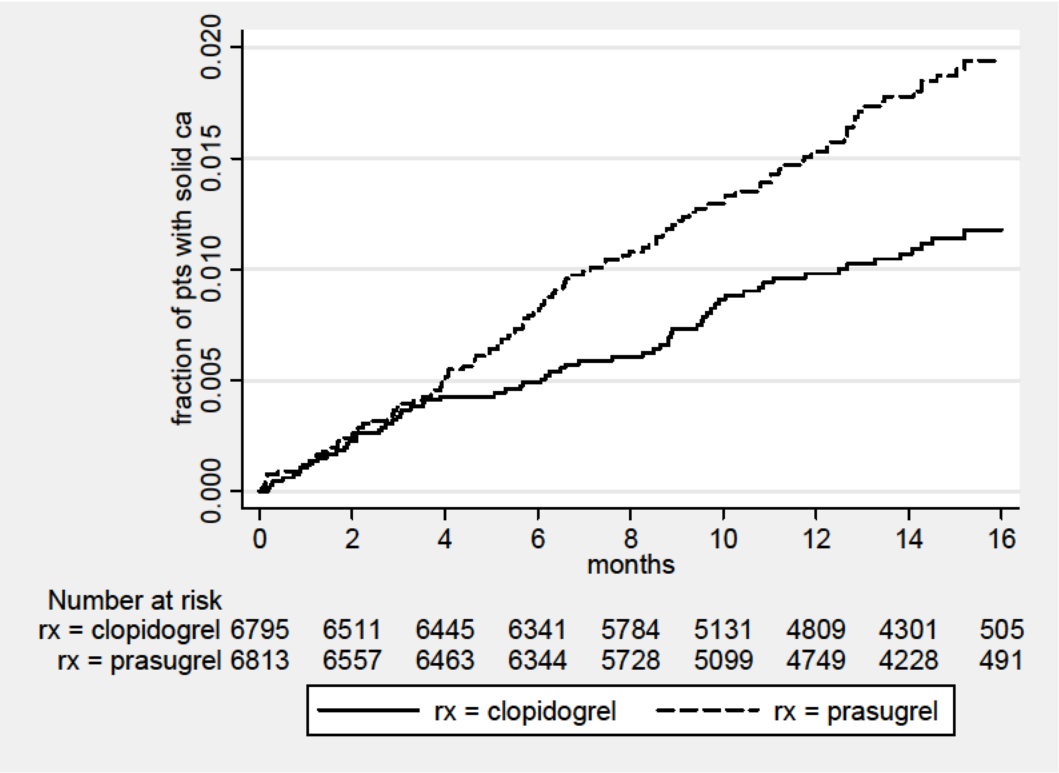
Bleeding and Cancer

Background

I raised the issue of whether a drug that affects bleeding might also affect cancer rates in my review of prasugrel, a platelet inhibitor. The details of the data and my discussion are available in the Medical Reviews, Parts 18 to 23, available at http://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/022307s000TOC.cfm. I summarize the findings with prasugrel below for ease of reference.

I analyzed solid cancer rates in the large TRITON outcomes trial of prasugrel vs. clopidogrel in acute coronary syndromes (Wiviott, Braunwald et al. 2007) because my interpretation of the prasugrel 24-month mouse carcinogenicity study was that prasugrel may be a tumor promoter for a wide variety of solid cancers. To my surprise the solid cancer event rates by arm in TRITON showed the strikingly different incidence curves shown in Figure 4: Times to First Solid Cancer* Events in the Prasugrel TRITON StudyFigure 4.

Figure 4: Times to First Solid Cancer* Events in the Prasugrel TRITON Study

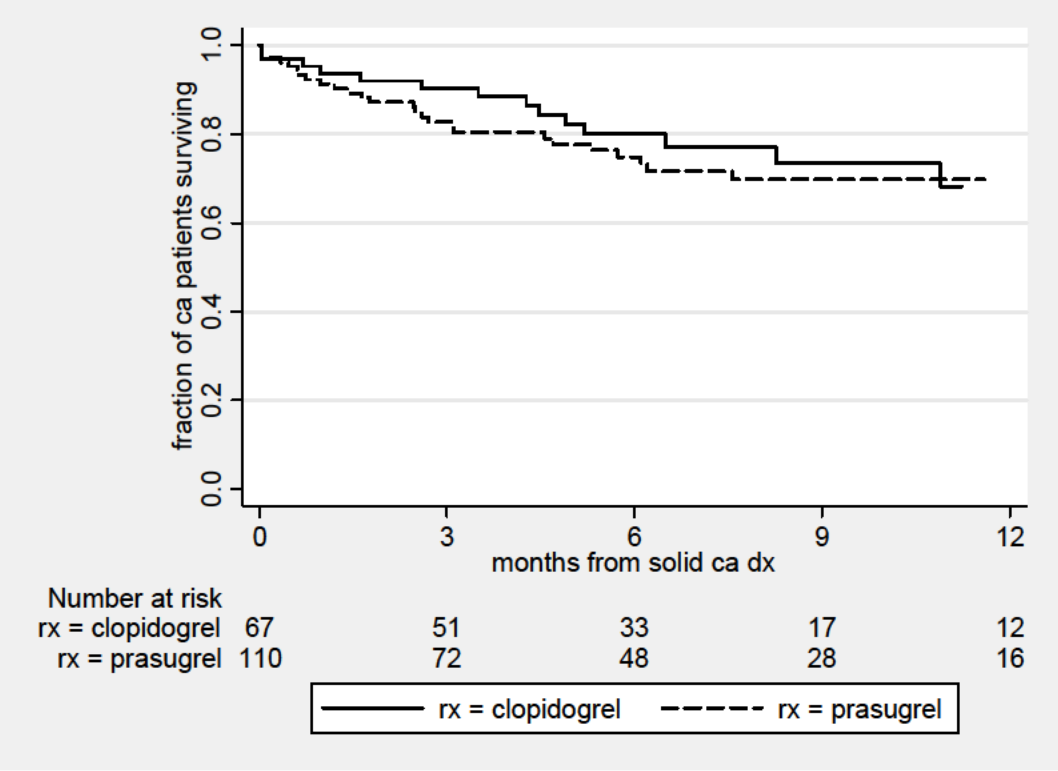


*excluding non-melanoma skin and brain; p = 0.0013 by log rank test

The solid cancer event rates begin to diverge at about 4 months and continued to diverge throughout 16-months of follow-up. The hazard ratio estimated by Cox regression is about 1.6 (95% CI 1.2-2.2). The absolute risk difference at 16 months is about 0.8%. While Figure 4 includes recurrent cancers as well as new cancers, the results limited to new solid cancers (excluding non-melanoma skin and brain) are similar although of lower statistical significance (p = 0.024).

Experiencing a solid cancer event was similarly deadly in both arms as shown in Figure 5.

Figure 5: Survival after First Solid Cancer* Events in the Prasugrel TRITON Study



*excluding non-melanoma skin and brain; p >0.5 by log rank test

Survival at 9 months is only about 70%. If anything survival was slightly worse for prasugrel patients with solid cancer events despite their greater numbers. There does not appear to be a lead-time bias or early detection bias that would lead to the appearance of lengthened survival after diagnosis.

Bleeding was more common in the prasugrel arm in TRITON. The prasugrel/clopidogrel hazard ratio for non-CABG-related TIMI major bleeding was 1.3 (95% CI 1.03-1.7) and for TIMI life-threatening bleeding was 1.5 (95% CI 1.08-2.1).

Variations on these findings were presented and discussed at a meeting of the Cardiovascular and Renal Drugs Advisory Committee on February 3, 2009. The variations presented by the sponsor were misleading and inaccurate: While I had prospectively excluded non-melanoma skin cancers based on the mouse carcinogenicity study results and because non-melanoma skin cancers are much less serious than other solid tumors and likely not to be reported completely, the sponsor included some skin cancers. The sponsor's presentations were misleading and inaccurate because they did not count skin cancers that they had miscoded to the MedDRA procedure system-organ class. The omitted skin cancers were predominantly in the prasugrel arm. I describe the details of the miscounts and the correct results in my prasugrel review referenced above.

The prasugrel TRITON cancer results alone do not help us understand whether solid cancer promotion is a peculiar effect of prasugrel, a class effect of $P2Y_{12}$ platelet inhibitors, a class

effect of all platelet inhibitors, or an effect of drugs that increase bleeding. While my preliminary analyses of older clopidogrel studies did not confirm a similar effect, I analyzed the FDA submissions for new, potent platelet inhibitors and for new anticoagulants. My preliminary analyses of the trials of new anticoagulants showed reasonably consistent results: Whatever arm had more bleeding had more solid cancer events and the solid cancers were deadly.

I had developed a rigorous methodology for ascertaining cancer events in CV outcome trials for a meta-analysis of angiotensin receptor blockers and cancer. FDA staff can access the review plan with the details of the methodology for this latter meta-analysis under Tracked Safety Issue 935 in a DARRTS communication filed August 31, 2012. Following as closely as possible the rigorous methodology I ascertained cancer events in the APPRAISE-2 and ARISTOTLE trials of apixaban. For the first version of this review I used analysis datasets for APPRAISE-2 provided by the sponsor in an early NDA submission. For this version I used the more detailed SDTM datasets provided in a later submission. The APPRAISE-2 cancer findings did not change for this second version, although from the SDTM datasets I did identify one additional apixaban patient with a gastrointestinal tumor likely malignant but not documented unequivocally as malignant.

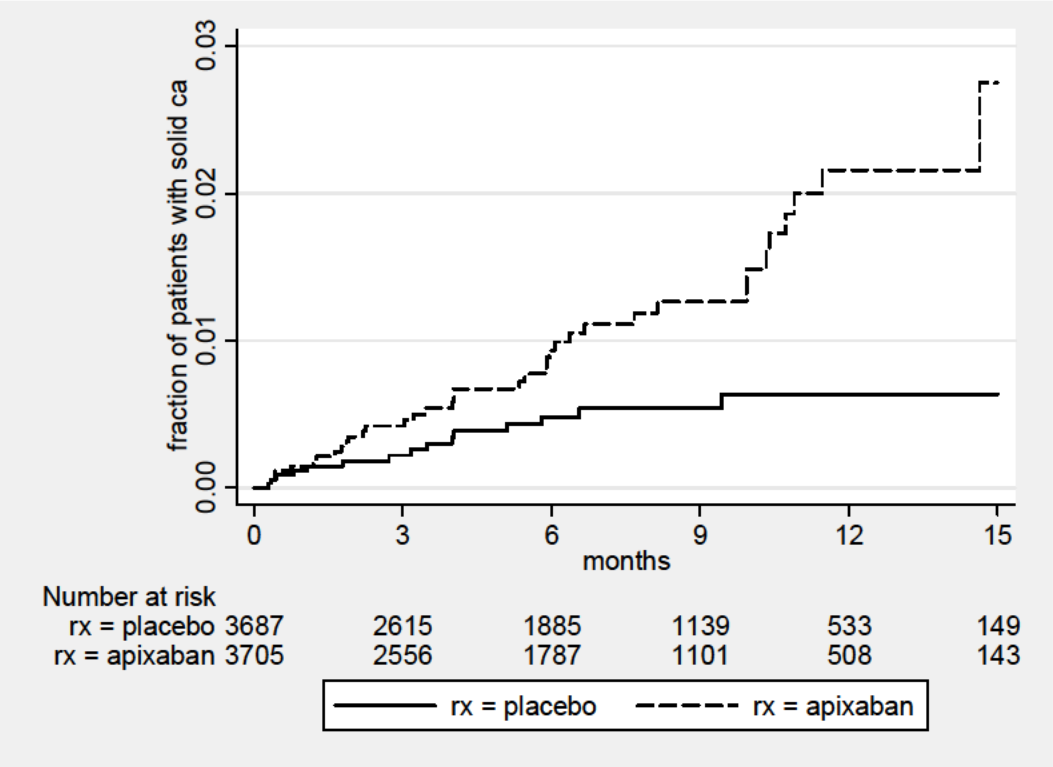
I have not yet performed a rigorous ascertainment of cancer events in a third large apixaban outcome trial, AVERROES. (Connolly, Eikelboom et al. 2011) AVERROES was a randomized, double-blind trial of apixaban vs. aspirin in patients with atrial fibrillation unsuitable for warfarin therapy. AVERROES was terminated early for a benefit favoring apixaban. Major bleeding rates were not significantly different between the apixaban and aspirin groups so the cancer findings may not be relevant to the question of bleeding and cancer. I will do a rigorous analysis of AVERROES in the future. I present the cancer findings for APPRAISE-2 and ARISTOTLE below.

Solid Cancers in APPRAISE-2

APPRAISE-2 was a double-blind, randomized, placebo-controlled trial of apixaban added to standard antiplatelet therapy in patients with a recent acute coronary syndrome. (Alexander, Lopes et al. 2011) The trial was terminated early because of a higher rate of major bleeding events in the apixaban arm without a counterbalancing reduction in ischemic events. The apixaban/placebo hazard ratio for TIMI major bleeding was 2.59 (95% CI 1.5-4.5).

Rates of solid cancer events were higher in the apixaban arm as shown in Figure 6**.**

Figure 6: Times to First Solid Cancer* Events in APPRAISE-2



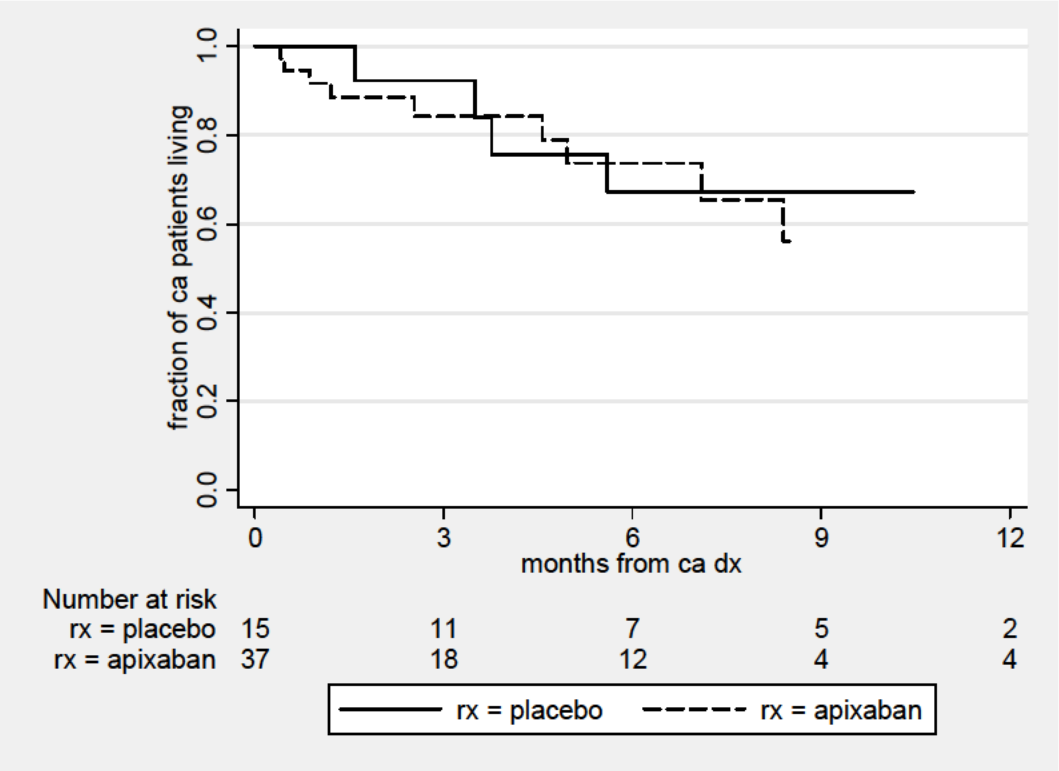
*excluding non-melanoma skin and brain; p = 0.0016 by log rank test

In APPRAISE-2 the solid cancer event incidence curves diverge early (about 1-2 months) and diverge more after 6 months. The hazard ratio for solid cancer events is 2.5 (95% CI 1.4-4.6). The sites contributing the most to the difference in solid cancer events were bladder, lung, and stomach.

After unblinding the cancer site assignments and examining the tumor descriptions, I noted that three of the apixaban solid tumors were of uncertain malignancy by the site-reported terms. Reassigning them to benign for a sensitivity analysis increases the p value to 0.0049 and reduces the hazard ratio to 2.3 (95% CI 1.2-4.3).

Survival was equally poor in the two arms after a solid cancer event as shown in Figure 7.

Figure 7: Survival after First Solid Cancer* Events in APPRAISE-2



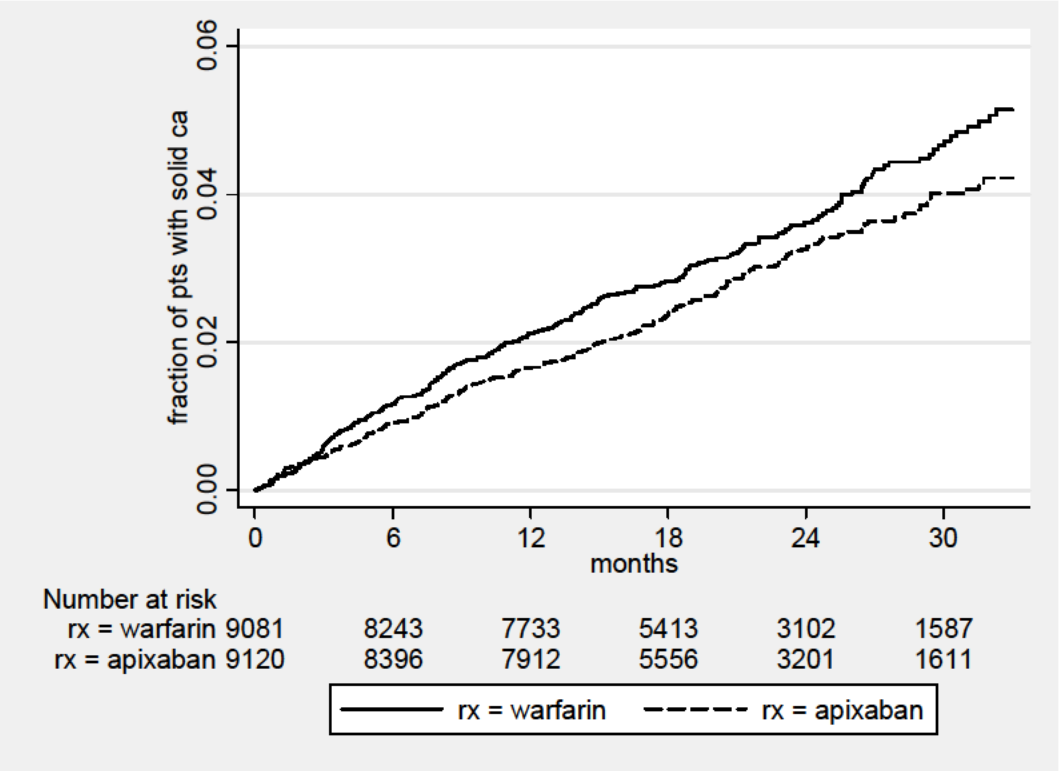
*excluding non-melanoma skin and brain; p > 0.6 by log rank test

As with prasugrel, the survival rates after a solid cancer event are similar with the new drug and control despite the greater number of patients with solid cancer events in the new drug arm.

Solid Cancers in ARISTOTLE

Per the FDA primary clinical review ISTH major bleeding was less in the apixaban arm compared to the warfarin arm, hazard ratio 0.69 (95% CI 0.6-0.8). Solid cancer events were also less frequent in the apixaban arm as shown in Figure 8.

Figure 8: Times to First Solid Cancer* Events in ARISTOTLE



*excluding non-melanoma skin and brain; p = 0.058 by log rank test

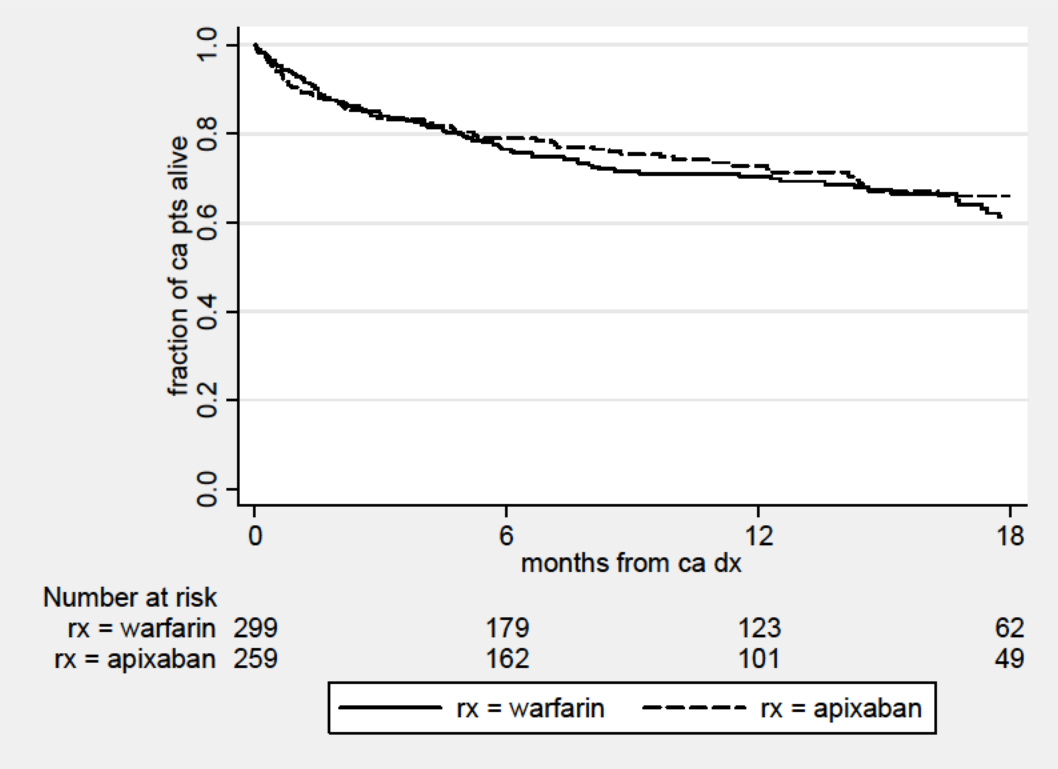
The curves diverge again at about 2-3 months but, given that the difference is not nominally statistically significant, we might ignore the divergence. However, because of our priors regarding the association of more bleeding with more solid cancers, we should have more confidence that the divergence is real. The apixaban/warfarin hazard ratio for solid cancer events is 0.85 (95% CI 0.7-1.0).

The cancer site with the largest difference in events between the two arms was bladder, a cancer site frequently manifested by bleeding. However, gastrointestinal tract cancer events were approximately evenly distributed between the two arms. The other sites substantially more frequent in the warfarin arm were unknown, lung, pancreas, and prostate.

The results above are for solid cancer events censored at January 30, 2011. If we count all solid cancer events reported, including those after the censoring date, the p value decreases to 0.027.

Survival after a first solid cancer event was equally poor in both arms as shown in Figure 9.

Figure 9: Survival after First Solid Cancer* Events in ARISTOTLE



*excluding non-melanoma skin and brain; p > 0.5 by log rank test

Survival in patients experiencing a first solid cancer event was about 72% at one year after the event and 63% at 18 months. We should not dismiss these solid cancer events lightly.

COMMENT: We can consider it reassuring, and consistent with the bleeding results, that solid cancer rates are similar or lower in ARISTOTLE with apixaban than with warfarin. However, the APPRAISE-2 results raise the concern that apixaban, probably like all other drugs that increase bleeding rates, has the potential for increasing solid cancer rates. While it would be reassuring to assume that the increased solid cancer rates represent a detection bias based on bleeding of a tumor leading to earlier diagnosis, the continued divergence of the incidence curves and the poor survival after the events argue against complacency. We should make clinicians aware that bleeding on an anticoagulant or antiplatelet drug may represent a deadly cancer and that it should not be dismissed as a simple drug side effect. We should include the bleeding and cancer findings in all anticoagulant and antiplatelet labels, beginning with the apixaban label.

For months I have distributed within the FDA my preliminary findings for a wide range of antiplatelet and anticoagulant drugs associating bleeding with solid cancers. I have been proposing that the data associating bleeding and solid cancers be presented and discussed at an advisory committee meeting. I have never received a response from my superiors regarding my proposal. The closest that I have received to an explanation for the inactivity is an email

responding to a related issue: The issue of whether angiotensin receptor blockers (ARBs) are associated with increased cancer rates, as raised by a 2010 published meta-analysis (Sipahi, Debanne et al. 2010), remains unanswered today. The June 2, 2011, FDA safety communication (available at http://www.fda.gov/Drugs/DrugSafety/ucm257516.htm) that cleared ARBs is based on an inadequately-specified and flawed FDA meta-analysis, including counting "malignant lung neoplasm" but not "lung carcinoma" as lung cancers. I documented the many problems with the FDA meta-analysis in reviews (available to FDA staff) filed in DARRTS on July 20, 2012, and August 31, 2012, under Tracked Safety Issue 935. I proposed analyzing the ARB trials rigorously but was discouraged by Dr. Ellis Unger, the Office Director, in the email reproduced in Attachment 1. The portion of that email most relevant to bleeding and cancer is the following quotation:

> "Third, if you were to go ahead with this and find a RR of, say 1.3, I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that. People would have various opinions on where the meaningful threshold is, but it might be worth asking for some input before you start."

I find it disturbing that Dr. Unger expressed a lack of concern about a drug increasing the risk of cancer by 30%. We routinely accept smaller efficacy benefits and our advisory committee has consistently refused to set a minimum magnitude for an acceptable benefit. We mandate that diabetic drugs not increase cardiovascular risk by 30%. That limit was partially based on practicality of trials—most clinicians would prefer a lower risk limit if the sizes of the resulting trials were feasible. I assert that any validated risk of cancer is concerning for a hypertensive drug used chronically for which there are many alternatives—or for anticoagulant and antiplatelet drugs used chronically. We should inform practitioners and patients about the risks of drugs so that they can make an informed decision and that they can institute follow-up measures to minimize the risks. I have proceeded in time available with the rigorous analysis of the ARB trials.

My review did invoke a response from Dr. Unger. I've included his response, and my response to it, as Attachment 2. I believe my response in Attachment 2 addresses well his reasons for not pursuing these critical cancer issues. My recommendations below remain unchanged.

I recommend the following:

- *The data regarding bleeding and cancer should be presented and discussed at an advisory committee meeting. If the rigorous analysis of the ARB trials confirms a risk for ARBs and cancer, then the data regarding ARBs and cancer should also be presented and discussed at an advisory committee meeting. It may be advantageous to have both topics addressed at the same meeting.*

- *The data regarding bleeding and cancer should be included in the apixaban label and in the labels for all antiplatelet and anticoagulant drugs. If the advisory committee meeting discussing bleeding and cancer is scheduled promptly then the labeling regarding bleeding and cancer can be delayed until after the meeting.*

References

Alexander, J. H., R. D. Lopes, et al. (2011). "Apixaban with antiplatelet therapy after acute coronary syndrome." N Engl J Med **365**(8): 699-708.

BAATAF_Investigators (1990). "The effect of low-dose warfarin on the risk of stroke in patients with nonrheumatic atrial fibrillation. The Boston Area Anticoagulation Trial for Atrial Fibrillation Investigators." N Engl J Med **323**(22): 1505-11.

Connolly, S. J., J. Eikelboom, et al. (2011). "Apixaban in Patients with Atrial Fibrillation." New England Journal of Medicine **364**(9): 806-817.

Connolly, S. J., A. Laupacis, et al. (1991). "Canadian Atrial Fibrillation Anticoagulation (CAFA) Study." J Am Coll Cardiol **18**(2): 349-55.

EAFT_Study_Group (1993). "Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke. EAFT (European Atrial Fibrillation Trial) Study Group." Lancet **342**(8882): 1255-62.

Ezekowitz, M. D., S. L. Bridgers, et al. (1992). "Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. Veterans Affairs Stroke Prevention in Nonrheumatic Atrial Fibrillation Investigators." N Engl J Med **327**(20): 1406-12.

Granger, C. B., J. H. Alexander, et al. (2011). "Apixaban versus warfarin in patients with atrial fibrillation." N Engl J Med **365**(11): 981-92.

Petersen, P., G. Boysen, et al. (1989). "Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study." Lancet **1**(8631): 175-9.

Sipahi, I., S. M. Debanne, et al. (2010). "Angiotensin-receptor blockade and risk of cancer: meta-analysis of randomised controlled trials." Lancet Oncol **11**(7): 627-36.

SPAF_Investigators (1991). "Stroke Prevention in Atrial Fibrillation Study. Final results." Circulation **84**(2): 527-39.

Wiviott, S. D., E. Braunwald, et al. (2007). "Prasugrel versus clopidogrel in patients with acute coronary syndromes." N Engl J Med **357**(20): 2001-15.

Attachment 1: Email explaining FDA reluctance to understand cancer risks

From: Unger, Ellis
Sent: Monday, August 20, 2012 11:04 PM
To: Marciniak, Thomas
Cc: Southworth, Mary Ross; Temple, Robert; Stockbridge, Norman L
Subject: RE: Emailing: ARB ca review plan v1p2.doc

Tom, et al,

I've gone through the protocol only fairly quickly, but I have a few comments.

First, this would represent a lot of man-hours, so I have to assume that there is a paucity of work in the Division at this point, or that you will be doing this mostly after hours.

Second, when we get into writing analytic plans, and specifically plans for adjudicating clinical endpoints, the plan/protocol might need to be reviewed at a high level – i.e., the OND IO or higher. There is a MAPP on this, I believe. You should consult that MAPP before you start any work to see if it applies here. If it applies, the protocol will need to go up to for review and comment before you begin.

Third, if you were to go ahead with this and find a RR of, say 1.3, I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that. People would have various opinions on where the meaningful threshold is, but it might be worth asking for some input before you start.

Finally, given you familiarity with some of the trial data, any decision YOU make regarding inclusion and exclusion of trials can be called into question after the fact. It doesn't matter that your criteria are reasonable and defensible, because you can know the effect that your criteria will have on the trials to be included/excluded before you begin.

Ellis

Attachment 2: Email discussion of FDA reluctance to understand cancer risks

From: Marciniak, Thomas
Sent: Sunday, December 16, 2012 3:08 PM
To: Unger, Ellis
Cc: Bai, Steven; Stockbridge, Norman L; Beasley, Bach Nhi t; Rose, Martin; Blaus, Alison; Temple, Robert; Southworth, Mary Ross; Grant, Stephen
Subject: RE: Finalized - NDA-202155 General Review (REV-CLINICAL-03)
Because you stated that "I think you'll agree" I have to respond. I do not agree:

I did not interpret your passage out of context but provided the complete context. I did extend your statement about a relative risk (RR) of 1.3 from one context, i.e., ARBs and cancer, to a second related one, i.e., bleeding and cancer. Because you have also applied your statement to the second context I don't believe I was wrong in doing so.

While now you state that you "certainly care about relative risks of 1.3" , now your itemized reasons for rejecting my proposals are "uncertainties". Why didn't you state that previously? Why aren't you supporting resolving those "uncertainties"?

However, your "uncertainties" are either not relevant or not different than all other drug issues regarding which the FDA has made regulatory decisions. Let's examine your uncertainties:

- "If the prospectively planned primary endpoint of a study has a statistically significant relative risk of 1.3 (i.e., benefit is 1/1.3), and that finding is substantiated with a second study, then we tend to believe it" and "where analyses are performed post hoc on an exploratory basis, it is not unreasonable to view a relative risk of 1.3 as a non-definitive." While the reality is that we act upon many post hoc-detected safety issues, both of these cancer risk issues were defined prospectively: I defined the bleeding and solid cancer risk prospectively based on my interpretation of the prasugrel mouse carcinogenicity studies. The association in bleeding and solid cancer risk is demonstrated with high statistical significance in the prasugrel TRITON trial and in not just a second but many other trials, i.e., the two large apixaban trials. I identified a risk of lung cancer with losartan in the LIFE trial and have proposed testing it in all the ARB trials. Both of these issues are better defined and pre-specified than most other safety issues upon which the FDA has acted, e.g., cardiac events with rofecoxib, adverse events in patients treated with vernakalant.

- ". . . neither cancer history nor cancer-related adverse events are obtained with great care . . .". How do you know that these are not "obtained with great care"? So all of the adverse event reporting in our clinical trials is worthless? If it is, then you and the rest of the FDA management should be taking urgent action to correct it. Cancer is a serious problem that patients are likely to report and investigators capture unless there is poor follow-up or an explicit study design

flaw that prevents capturing the identities of SAEs. We can eliminate trials with the latter problems. If the trials with reasonable SAE capture approaches and adequate follow-up consistently show a cancer risk, then we can have confidence that there is a real problem. Your alleged problems would represent noise that would likely obscure real risks and produce inconsistent results from trial to trial, not consistently show a risk. Finally, cancer is easier to document than the cardiac events that have proved useful in identifying CV risk—and efficacy!--of many drugs.

- ". . . where there is strong potential for ascertainment bias (bleeding can lead to tumor discovery) . . ." Note that this remark does not apply to ARBs and cancer, the original context of your 1.3 RR remark. I agree that there is a potential for ascertainment bias due to bleeding. However, that potential does not appear to explain the cancer risks observed because the cancer incidence curves continue to diverge for the durations of all studies and the cancer survival in each study is the same in all arms despite the increased cancer rate in the arm with more bleeding. If ascertainment bias played a major role we would expect the cancer incidence curves to converge and survival to be prolonged in the arm with more bleeding, if only by a lead time bias if not also because of earlier detection with higher rates of cure. The question of how much ascertainment bias plays a role is precisely the reason why we need to analyze the vast amounts of data we have more thoroughly and present and discuss the results openly at an advisory committee meeting.

- ". . . where there is no reasonably plausible underlying mechanism of action . . ." The FDA has never required, and should not require, a plausible mechanism of action for a safety issue. However, there are plausible mechanisms of action for both of these cancer risk issues. I would be happy to discuss them.

Finally, you have identified potentially legitimate but unsupported "uncertainties" in my proposals. However, there is one issue about which there is no uncertainty: Not counting lung carcinomas as lung cancers in the FDA meta-analysis of ARBs and cancer is wrong. The FDA safety communication of June 2, 2011, that cleared ARBs is based on that poorly executed and seriously flawed meta-analysis. That safety communication is long overdue for correction and the relationship of bleeding to cancer is long overdue for characterization and public dissemination.

Tom

From: Unger, Ellis
Sent: Thursday, December 13, 2012 1:47 PM
To: Marciniak, Thomas
Cc: Bai, Steven; Stockbridge, Norman L; Beasley, Bach Nhi t; Rose, Martin; Blaus, Alison; Temple, Robert; Southworth, Mary Ross; Grant, Stephen
Subject: RE: Finalized - NDA-202155 General Review (REV-CLINICAL-03)

Tom,

Thanks for your review. You make some good points.

I would note, however, that the passage taken out of context from my August 20, 2012 email to you was, in fact, interpreted out of context:

> "Third, if you were to go ahead with this and find a RR of, say 1.3, I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that. People would have various opinions on where the meaningful threshold is, but it might be worth asking for some input before you start."

I/we certainly care about relative risks of 1.3, if the absolute risk is appreciable. For cancer, the absolute risk is appreciable; therefore, a relative risk of 1.3 would constitute an important public health issue. I want to be clear about that.

But all relative risks of 1.3 are not created equal, even when p-values and confidence intervals are identical. If the prospectively planned primary endpoint of a study has a statistically significant relative risk of 1.3 (i.e., benefit is 1/1.3), and that finding is substantiated with a second study, then we tend to believe it – especially if there is mechanistic/non-clinical support for the drug effect.

But when there are uncertainties – and we have discussed these at length over the years – a relative risk of 1.3 can be less compelling, and not necessarily a public health issue. I'm not minimizing your careful findings, but I think you'll agree that findings such as these, in acutely ill patients where neither cancer history nor cancer-related adverse events are obtained with great care, where there is strong potential for ascertainment bias (bleeding can lead to tumor discovery), where there is no reasonably plausible underlying mechanism of action, and where analyses are performed post hoc on an exploratory basis, it is not unreasonable to view a relative risk of 1.3 as a non-definitive, hypothesis-generating signal. And that is why "…I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that."

Ellis

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

/s/

THOMAS A MARCINIAK
12/17/2012

CLINICAL REVIEW

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
DIVISION OF CARDIOVASCULAR AND RENAL PRODUCTS

Date: December 11, 2012

Reviewer: Thomas A. Marciniak, M.D.
Medical Team Leader

NDA: 202-155

Drug: apixaban (Eliquis®)

Indication: To reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation

Subjects: Completeness of follow-up and bleeding and cancer

Summary and Recommendation
Because of a special interest in and experience with two issues, completeness of follow-up and cancer, I reviewed the apixaban studies regarding these issues. I document below that completeness of follow-up and reporting of dates were poor in ARISTOTLE. Our confidence in the fragile alleged death benefit (with one additional death in the apixaban arm eliminating statistical significance) is destroyed by the missing vital status. Our confidence in the superiority of the stroke benefit to warfarin is also challenged by incomplete follow-up. Finally, the ARISTOTLE and APPRAISE-2 trials show an association between bleeding and solid cancers also seen in other anticoagulant and antiplatelet drug trials.

I recommend the following:

- The indication statement should include only stroke and systemic embolism.
- (b) (4).
- The Clinical Studies section of the label should include a discussion of the data quality problems in ARISTOTLE. It should summarize the dispensing errors and provide the missing follow-up statistics for both vital status and events. It should report that a change in one death eliminates the statistical significance of the death benefit and that, because of the missing data, we can not have confidence in a death benefit.

- The data regarding bleeding and cancer should be presented and discussed at an advisory committee meeting. If the rigorous analysis of the ARB trials confirms a risk for ARBs and cancer, then the data regarding ARBs and cancer should also be presented and discussed at an advisory committee meeting. It may be advantageous to have both topics addressed at the same meeting.

- The data regarding bleeding and cancer should be included in the apixaban label and in the labels for all antiplatelet and anticoagulant drugs. If the advisory committee meeting discussing bleeding and cancer is scheduled promptly then the labeling regarding bleeding and cancer can be delayed until after the meeting.

Completeness of Follow-up and Fragility of Results in ARISTOTLE

Definition of Completeness of Follow-up

The clinical study report (CSR) for the ARISTOTLE trial of apixaban vs. warfarin in atrial fibrillation states that vital status could not be determined for 2.0% in the apixaban group and 2.2% in the warfarin group (380 patients in both groups total). The main study publication reported the same vital status statistics. (Granger, Alexander et al. 2011) However, the rates of discontinuation from the study were much higher, 25.3% in the apixaban group and 27.5%, with 10.1% of apixaban patients and 10.0% of warfarin patients discontinuing at their own request. While these reported statistics for completeness of follow-up are not good, my recent experience with other outcome trials suggests that the sponsor's reporting of completeness of follow-up is usually optimistic compared to analyses of the submitted datasets. Hence I analyzed the datasets for completeness of follow-up.

I assert that there is a straightforward definition of completeness of follow-up: Most outcome studies have a specified global study end date or censoring date for efficacy outcomes. A few have a pre-specified duration of follow-up from randomization such as two years. I assert that follow-up is complete if the patient has documented follow-up on or after the specified end date. ARISTOTLE had a cutoff date for efficacy outcomes of January 30, 2011. (Granger, Alexander et al. 2011) Hence for ARISTOTLE follow-up for a patient is complete if the patient had documented follow-up on or after January 30, 2011.

Determining completeness of follow-up has one complication: The type of follow-up typically varies in outcome studies. For patients who have a face-to-face study visit with the investigator on or after the study end date follow-up is complete for all study outcomes or events. However, for some patients final follow-up may consist of a phone call with the patient or a spouse or a primary physician, for others a report of a hospitalization, and for still others a newspaper obituary or a registry report of alive or dead. The level of detail available from these latter, non-face-to-face follow-ups varies and, while usually adequate for determining vital status, may not be adequate for ascertaining endpoints or adverse events. I recommend estimating two levels of completeness of follow-up: (1) vital status; and (2) events. For the former I accept any type of documented follow-up having a date unambiguously referencing the patient as alive. For the latter I accept reports of documented face-to-face visits, hospitalizations or other events, and phone calls with documented queries regarding events. Because of ambiguities in case report

form (CRF) design and in reporting by sites determining completeness of event follow-up requires subjective judgments.

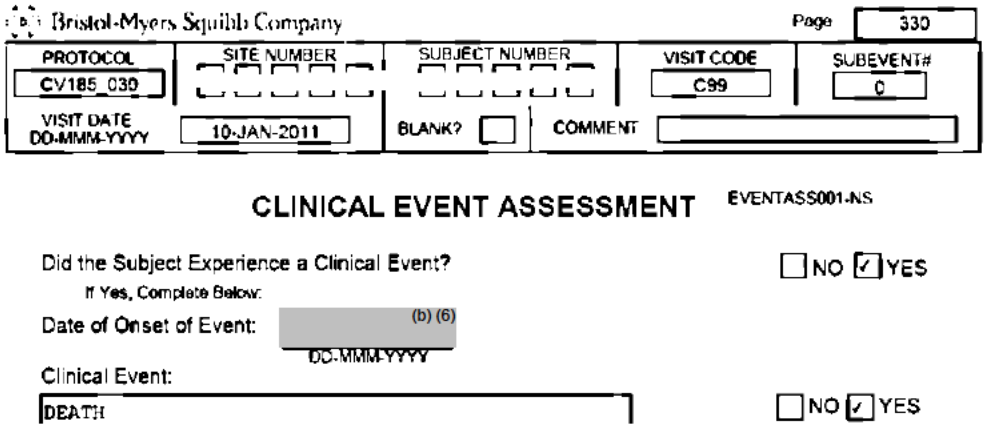
Completeness of Vital Status Follow-up

For one estimate of completeness of vital status follow-up in ARISTOTLE I used the sponsor's variable *e_cddn* described as "Censor Date for Death ITT Period" from the sponsor's ADEFS.XPT ("Efficacy Summary") dataset. Counting good vital status as either known dead at any time or a censor date of January 30, 2011, or later, 3.2% of patients (590) are missing vital status. For a second estimate I also analyzed all of the datasets for the maximum dates of events, procedures, vital sign recordings, and status reports to estimate a date of last follow-up for every patient. By this latter analysis of the raw data 3.6% of patients (660) are missing vital status. Either number, 590 or 660, greatly exceeds the sponsor's and investigators' reports of 380 missing vital status follow-up.

Date Recording Problems in ARISTOTLE

Both the sponsor's and my estimates of vital status follow-up are likely optimistic: Dates are not infrequently misrecorded or misinterpreted in ARISTOTLE. The death CRF excerpt in Figure 1 illustrates the possibility for misrecording or misinterpretation: The "visit date" for the death CRF precedes the death date. (Note that the reported death date was (b) (6) (b) (6) —compared to no apixaban patients. See the further discussion of problems with death reports below.)

Figure 1: Death in the Future



Bristol-Myers Squibb Company Page 330

PROTOCOL	SITE NUMBER	SUBJECT NUMBER	VISIT CODE	SUBEVENT#
CV185_030			C99	0

VISIT DATE DD-MMM-YYYY: 10-JAN-2011 BLANK? ☐ COMMENT

CLINICAL EVENT ASSESSMENT EVENTASS001-NS

Did the Subject Experience a Clinical Event? ☐NO ☑YES

If Yes, Complete Below:

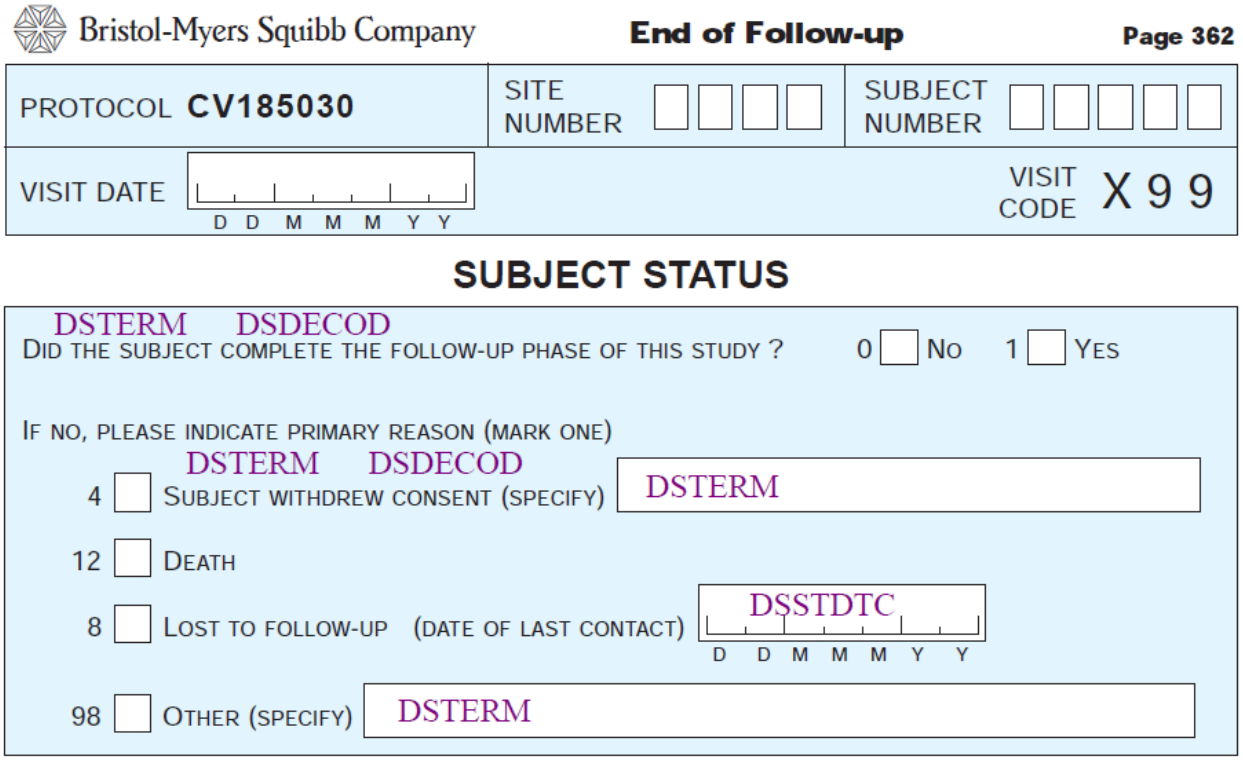
Date of Onset of Event: (b) (6) DD-MMM-YYYY

Clinical Event:

DEATH ☐NO ☑YES

The "visit date" may have little connection to a date on which the patient was actually observed or contacted. In many cases there is no way of determining whether the "visit date" is an observation date or the date of recording or something else. However, both I and the sponsor based our dates of last vital status partially on "visit dates". In particular a critical form for follow-up, the End of Follow-up CRF shown in Figure 2, has this ambiguity regarding "visit date".

Figure 2: End of Follow-up CRF



Bristol-Myers Squibb Company

End of Follow-up Page 362

PROTOCOL **CV185030** | SITE NUMBER | SUBJECT NUMBER

VISIT DATE (D D M M M Y Y) | VISIT CODE X 9 9

SUBJECT STATUS

DSTERM DSDECOD
DID THE SUBJECT COMPLETE THE FOLLOW-UP PHASE OF THIS STUDY ? 0 ☐ NO 1 ☐ YES

IF NO, PLEASE INDICATE PRIMARY REASON (MARK ONE)

DSTERM DSDECOD
4 ☐ SUBJECT WITHDREW CONSENT (SPECIFY) DSTERM

12 ☐ DEATH

8 ☐ LOST TO FOLLOW-UP (DATE OF LAST CONTACT) DSSTDTC (D D M M M Y Y)

98 ☐ OTHER (SPECIFY) DSTERM

What does the "visit date" on the End of Follow-up CRF represent? Only for lost to follow-up is the date of last contact to be recorded on this CRF. For death there is the death form with a field for date of death, but what about "withdrew consent" or "other" or even completed? It is easy to document that "visit date" for "withdrew consent" likely does not represent the date on which the patient visited the site or withdrew consent. For example, one patient discontinued treatment on 16jul00 with the last verifiable events on 18jun00. However, the disposition (DS) dataset has a "Start Date/Time of Disposition Event" for withdrew consent of 12apr11 and the sponsor counts the patient as completing follow-up, censoring on 30jan11. Withdrawing consent on 12apr11, long after the trial ended, is not rational and would not represent a withdrawal of consent during the ITT period, as the sponsor classifies this patient. Another patient is similar, with end of treatment and last events on 06may10 but withdrawal of consent allegedly on 24feb11 with sponsor's censoring on 30jan11. I count both of these patients (and other similar ones) as having incomplete vital status follow-up, partially explaining and justifying why my estimate of incompleteness of vital status follow-up is higher than the sponsor's.

There are other examples of anomalous dates, e.g., three patients have last contact dates by "direct contact with subject in person" long after the patient was reported dead. In fact, about 65% of patients who died have a visit date or other date greater than the date of death. Furthermore, for patients who did not die during the study, we do not have an unequivocal last date against which to check the validity of reported dates. There is no good way to detect or resolve many of these date inconsistencies even with a painstaking manual review of the CRFs--and we do not have most of the CRFs. Any estimates of completeness of vital status follow-up are optimistic.

Fragility of the Death Benefit in ARISTOTLE

While the estimates of missing vital status follow-up are concerning (>3%) and likely higher, how do they relate quantitatively to the reported death benefit? The sponsor claims that there were statistically significantly fewer deaths in the apixaban arm than in the warfarin arm in ARISTOTLE based on the analysis in Table 1 excerpted from the clinical study report:

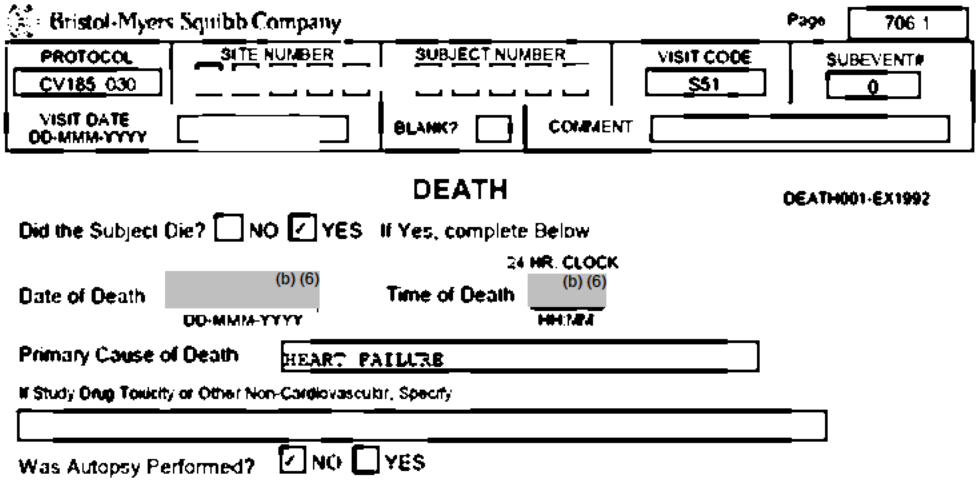
Table 1: Summary of Adjudicated Causes of Death during the Intended Treatment Period – Randomized Subjects (Excerpt from Sponsor's Table 7.2.1)

		Apixaban N = 9120	Warfarin N = 9081
ALL-CAUSE DEATH, n (%)		603 (6.61)	669 (7.37)
EVENT RATE (%/YR)		3.52	3.94
HAZARD RATIO (APIXABAN/WARFARIN)	0.89		
95% CI FOR HAZARD RATIO	(0.80, 1.00)		
P-VALUE	0.0465		

The p-value is close to 0.05 and the upper confidence limit for the hazard ratio is 1.00 so we know that this result is very fragile. How fragile? A change in only one death (one more with apixaban or one fewer with warfarin) could make this result nominally statistically insignificant. This critical number 1 is dwarfed by 317, the estimated—but still likely optimistic—number of patients with missing vital status in the apixaban arm. Even the total difference in deaths between the two arms (66) is a small fraction of number of apixaban patients with missing vital status.

This numeric fragility may not be surprising but, in addition, the quality of the documentation of the deaths is fragile or suspect. I examined deaths around the time of the cutoff date (January 30, 2011) for efficacy outcomes. Three warfarin patients (and no apixaban patients) died (b) (6) One patient in Hungary discontinued treatment on 4jan11 for bleeding with the clinical event assessment shown in Figure 1 dated 10jan11 and reporting death on (b) (6). Another CRF shown in Figure 3 does appear to confirm the date of death as (b) (6) at virtually midnight.

Figure 3: Death near Midnight on Cutoff Date



Bristol-Myers Squibb Company — Page 706 1

PROTOCOL	SITE NUMBER	SUBJECT NUMBER	VISIT CODE	SUBEVENT#
CV185_030			S51	0

VISIT DATE DD-MMM-YYYY | BLANK? | COMMENT

DEATH

DEATH001-EX1992

Did the Subject Die? ☐ NO ☑ YES If Yes, complete Below

Date of Death (b) (6) (DD-MMM-YYYY) Time of Death (b) (6) (24 HR. CLOCK HH:MM)

Primary Cause of Death: HEART FAILURE

If Study Drug Toxicity or Other Non-Cardiovascular, Specify:

Was Autopsy Performed? ☑ NO ☐ YES

Figure 3 is another example of a "visit date" occurring after death. Usually the CRFs include for a death a handwritten narrative faxed by the investigator to the sponsor's monitors. Such a narrative was not provided for this death. Other confirmatory documentation, such as death certificates, were not submitted in general.

Conversely, an apixaban patient suffered a pulmonary embolism on (b) (6), was hospitalized on (b) (6)), and died. The site reported the date of death as 31jan11 and hence the death is not counted as an apixaban death. Could this represent (b) (6) error?

The problems with these death reports, like the problems with visit dates and the substantial missing vital status, destroy our confidence that apixaban decreases death rates compared to warfarin. One final statistic regarding deaths from ARISTOTLE adds to our lack of confidence in a death benefit for apixaban: If one counts all reported deaths, not just those censored at (b) (6), then the death difference is insignificant, i.e., p = 0.077.

Warfarin and Deaths

While it appears unclear whether apixaban is superior to warfarin regarding death rates, a good question is whether warfarin is superior to placebo regarding death rates. We may be more impressed with marginal p values if the comparator is known to be active. While many clinicians may assume that warfarin reduces all-cause mortality because of its substantial effects upon stroke rates, a death benefit for warfarin compared to placebo is also unclear as I document below.

I extracted the death statistics from the six historical trials of warfarin vs. placebo in atrial fibrillation, i.e., the ones used for efficacy noninferiority analyses of new anticoagulants: AFASAK (Petersen, Boysen et al. 1989), BAATAF (BAATAF_Investigators 1990), CAFA (Connolly, Laupacis et al. 1991), EAFT (EAFT_Study_Group 1993), SPAF (SPAF_Investigators 1991), and SPINAF (Ezekowitz, Bridgers et al. 1992). AFASAK reported vascular deaths (on treatment?) by arm but not all-cause deaths by arm, reporting only that 71 patients died and that an ITT analysis of all deaths showed no difference in either vascular or

total mortality. Using the *metan* package for Stata 12 I performed a random effects meta-analysis of the all-cause death risk ratios in these studies assigning neutral results for AFASAK. I show the meta-analysis results in Table 2.

Table 2: Meta-Analysis of Deaths in the Warfarin vs. Placebo Trials in Atrial Fibrillation

Study	RR	[95% Conf.	Interval]	% Weight
AFASAK	1.024	0.638	1.645	23.96
BAATAF	0.415	0.211	0.818	13.98
CAFA	1.277	0.515	3.164	8.52
SPAF	0.754	0.266	2.134	6.66
SPINAF	0.695	0.369	1.310	15.61
EAFT	0.953	0.649	1.398	31.28
D+L pooled RR	0.829	0.626	1.099	100.00

Heterogeneity chi-squared = 6.45 (d.f. = 5) p = 0.265
I-squared (variation in RR attributable to heterogeneity) = 22.4%
Estimate of between-study variance Tau-squared = 0.0276

Test of RR=1 : z= 1.30 p = 0.192

RR = risk ratio warfarin/placebo; D+L = DerSimonian-Laird method

Only one of the historical trials (BAATAF) shows a significant mortality benefit. While some may like to interpret the results as suggestive of a mortality benefit because of the favorable point estimate of the risk ratio, the p value of 0.192 does not support statistical significance or firm conclusions of a mortality benefit for warfarin. The conservative interpretation of the mortality comparison in ARISTOTLE is that we are comparing apixaban to an inactive agent.

Completeness of Event Follow-up

The problem of completeness of follow-up is not limited to vital status. I estimated completeness of event follow-up based on events, procedures, vital sign recordings, and last direct contacts with the patient (i.e., the types of reports relevant to events, endpoint or adverse, or the absence thereof) but not counting the flawed status report visit dates. For these reports about 15% of patients, or over 2,700 patients, have incomplete follow-up. Compare 1,351, the number of apixaban patients with incomplete follow-up, to 53, the difference in primary endpoints. There is a vast amount of missing follow-up in which endpoints may be hidden or missed.

COMMENT: The alleged death benefit of apixaban compared to warfarin is fragile as reported by the sponsor, i.e., p = 0.046, a change in only one death rendering the difference statistically insignificant. Furthermore, the validity of this fragile benefit depends upon having 100% valid data. The substantial missing vital status follow-up, the problems with date recordings, and the lack of a significant death benefit for warfarin destroy confidence that apixaban reduces all-cause mortality. We might have some confidence that apixaban reduces stroke rates in atrial fibrillation: Stroke reduction with warfarin is substantial and allegedly apixaban improves upon the warfarin reductions. However, our confidence in the apixaban stroke benefit is also reduced

by the substantial missing event follow-up as well as by the data quality issue identified by the primary reviewers, i.e., errors in the dispensing of the study drug or documentation thereof.

I, like most FDA reviewers, would like to conclude that apixaban is effective in atrial fibrillation—we would like to have alternatives to warfarin. While there are problems with the ARISTOTLE data, the stroke results are reasonably in the right direction and our priors are that a drug with an anticoagulant pharmacodynamic action should be effective. I consider it to be very unfortunate that ARISTOTLE, like many other recent outcome trials, has substantial problems with data quality. Some of the responsibility for the data quality problems rests with us, the FDA: We have approved drugs ignoring similar data quality issues, granting superiority claims and not discussing in the labels the data quality issues. We must stop doing this.

If we approve apixaban I recommend the following for the labeling:

- *The indication statement should include only stroke and systemic embolism.*
- (b) (4).
- *The Clinical Studies section of the label should include a discussion of the data quality problems in ARISTOTLE. It should summarize the dispensing errors and provide the missing follow-up statistics for both vital status and events. It should report that a change in one death eliminates the statistical significance of the death benefit and that, because of the missing data, we can not have confidence in a death benefit.*

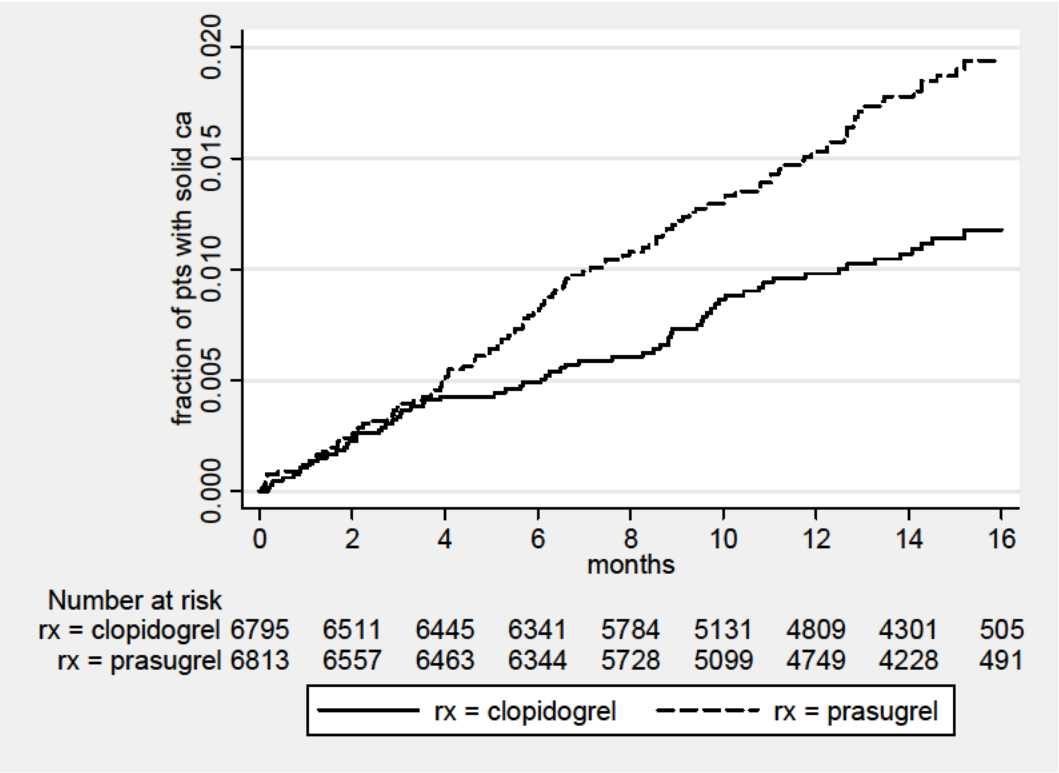
Bleeding and Cancer

Background

I raised the issue of whether a drug that affects bleeding might also affect cancer rates in my review of prasugrel, a platelet inhibitor. The details of the data and my discussion are available in the Medical Reviews, Parts 18 to 23, available at http://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/022307s000TOC.cfm. I summarize the findings with prasugrel below for ease of reference.

I analyzed solid cancer rates in the large TRITON outcomes trial of prasugrel vs. clopidogrel in acute coronary syndromes (Wiviott, Braunwald et al. 2007) because my interpretation of the prasugrel 24-month mouse carcinogenicity study was that prasugrel may be a tumor promoter for a wide variety of solid cancers. To my surprise the solid cancer event rates by arm in TRITON showed the strikingly different incidence curves shown in Figure 4: Times to First Solid Cancer* Events in the Prasugrel TRITON StudyFigure 4.

Figure 4: Times to First Solid Cancer* Events in the Prasugrel TRITON Study

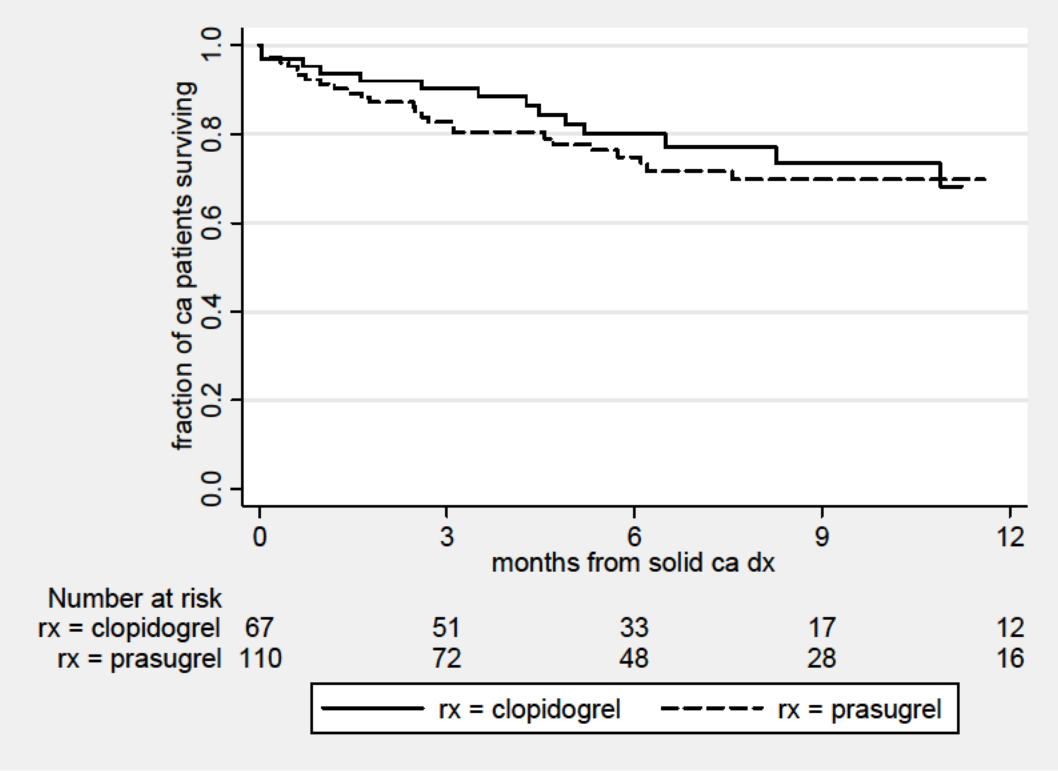


*excluding non-melanoma skin and brain; p = 0.0013 by log rank test

The solid cancer event rates begin to diverge at about 4 months and continued to diverge throughout 16-months of follow-up. The hazard ratio estimated by Cox regression is about 1.6 (95% CI 1.2-2.2). The absolute risk difference at 16 months is about 0.8%. While Figure 4 includes recurrent cancers as well as new cancers, the results limited to new solid cancers (excluding non-melanoma skin and brain) are similar although of lower statistical significance (p = 0.024).

Experiencing a solid cancer event was similarly deadly in both arms as shown in Figure 5.

Figure 5: Survival after First Solid Cancer* Events in the Prasugrel TRITON Study



*excluding non-melanoma skin and brain; p >0.5 by log rank test

Survival at 9 months is only about 70%. If anything survival was slightly worse for prasugrel patients with solid cancer events despite their greater numbers. There does not appear to be a lead-time bias or early detection bias that would lead to the appearance of lengthened survival after diagnosis.

Bleeding was more common in the prasugrel arm in TRITON. The prasugrel/clopidogrel hazard ratio for non-CABG-related TIMI major bleeding was 1.3 (95% CI 1.03-1.7) and for TIMI life-threatening bleeding was 1.5 (95% CI 1.08-2.1).

Variations on these findings were presented and discussed at a meeting of the Cardiovascular and Renal Drugs Advisory Committee on February 3, 2009. The variations presented by the sponsor were misleading and inaccurate: While I had prospectively excluded non-melanoma skin cancers based on the mouse carcinogenicity study results and because non-melanoma skin cancers are much less serious than other solid tumors and likely not to be reported completely, the sponsor included some skin cancers. The sponsor's presentations were misleading and inaccurate because they did not count skin cancers that they had miscoded to the MedDRA procedure system-organ class. The omitted skin cancers were predominantly in the prasugrel arm. I describe the details of the miscounts and the correct results in my prasugrel review referenced above.

The prasugrel TRITON cancer results alone do not help us understand whether solid cancer promotion is a peculiar effect of prasugrel, a class effect of $P2Y_{12}$ platelet inhibitors, a class

effect of all platelet inhibitors, or an effect of drugs that increase bleeding. While my preliminary analyses of older clopidogrel studies did not confirm a similar effect, I analyzed the FDA submissions for new, potent platelet inhibitors and for new anticoagulants. My preliminary analyses of the trials of new anticoagulants showed reasonably consistent results: Whatever arm had more bleeding had more solid cancer events and the solid cancers were deadly.

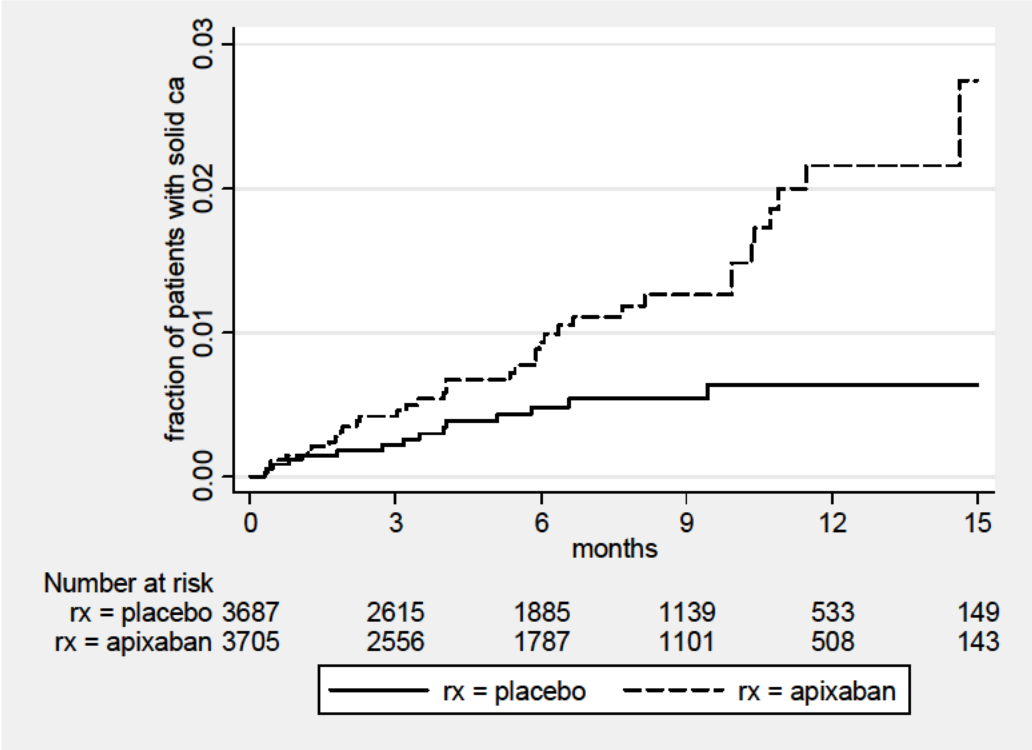
I had developed a rigorous methodology for ascertaining cancer events in CV outcome trials for a meta-analysis of angiotensin receptor blockers and cancer. FDA staff can access the review plan with the details of the methodology for this latter meta-analysis under Tracked Safety Issue 935 in a DARRTS communication filed August 31, 2012. Following as closely as possible the rigorous methodology I ascertained cancer events in the APPRAISE-2 and ARISTOTLE trials of apixaban. (For APPRAISE-2 the sponsor submitted selected analysis datasets rather than all raw datasets. Because APPRAISE-2 was terminated early with good follow-up I believe the cancer findings for it are valid.) I present the cancer findings for both trials below.

Solid Cancers in APPRAISE-2

APPRAISE-2 was a double-blind, randomized, placebo-controlled trial of apixaban added to standard antiplatelet therapy in patients with a recent acute coronary syndrome. (Alexander, Lopes et al. 2011) The trial was terminated early because of a higher rate of major bleeding events in the apixaban arm without a counterbalancing reduction in ischemic events. The apixaban/placebo hazard ratio for TIMI major bleeding was 2.59 (95% CI 1.5-4.5).

Rates of solid cancer events were higher in the apixaban arm as shown in Figure 6.

Figure 6: Times to First Solid Cancer* Events in APPRAISE-2



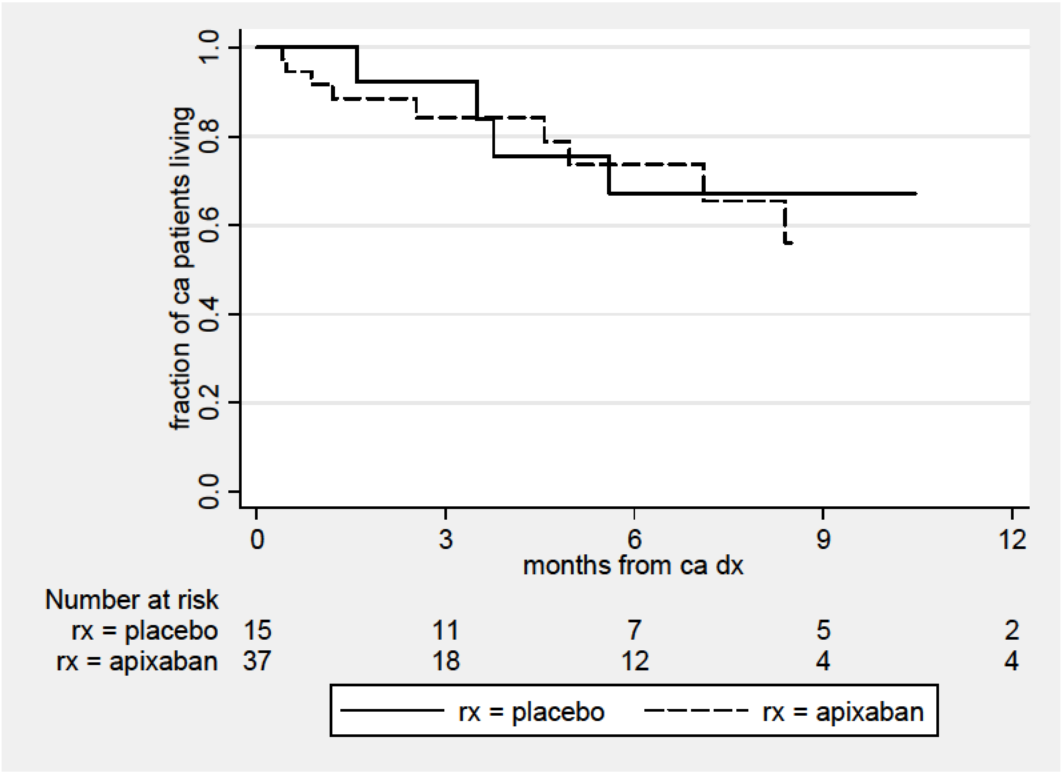
*excluding non-melanoma skin and brain; p = 0.0016 by log rank test

In APPRAISE-2 the solid cancer event incidence curves diverge early (about 1-2 months) and diverge more after 6 months. The hazard ratio for solid cancer events is 2.5 (95% CI 1.4-4.6). The sites contributing the most to the difference in solid cancer events were bladder, lung, and stomach.

After unblinding the cancer site assignments and examining the tumor descriptions, I noted that three of the apixaban solid tumors were of uncertain malignancy by the site-reported terms. Reassigning them to benign for a sensitivity analysis increases the p value to 0.0049 and reduces the hazard ratio to 2.3 (95% CI 1.2-4.3).

Survival was equally poor in the two arms after a solid cancer event as shown in Figure 7.

Figure 7: Survival after First Solid Cancer* Events in APPRAISE-2



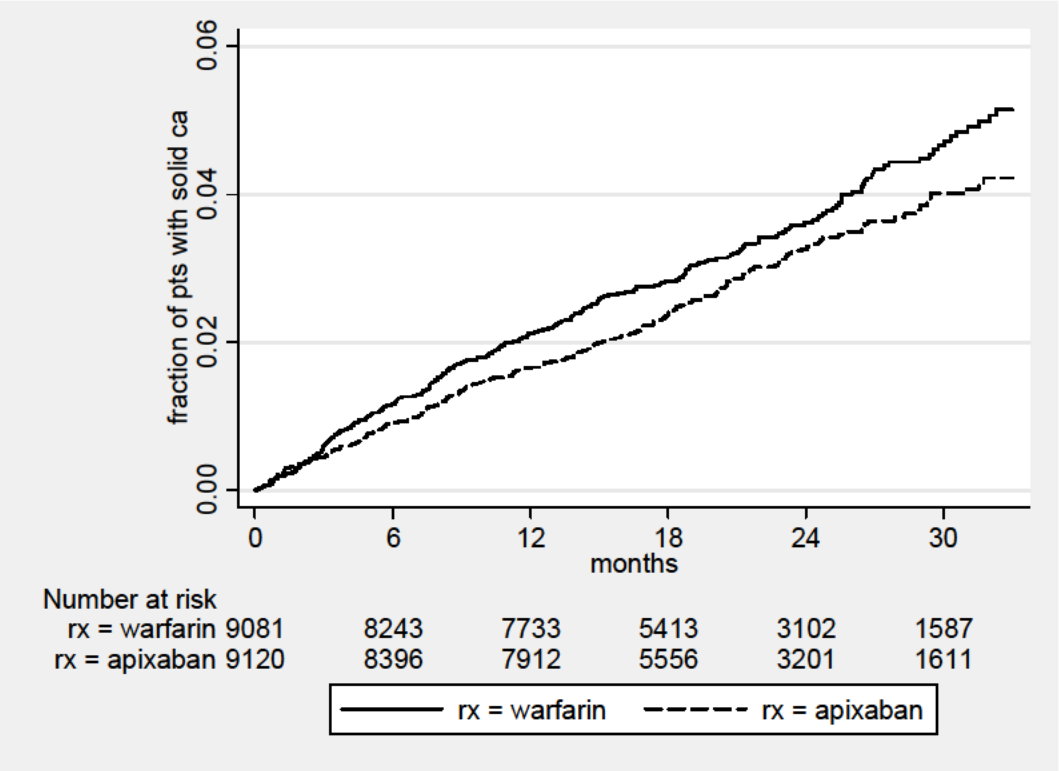
*excluding non-melanoma skin and brain; p > 0.6 by log rank test

As with prasugrel, the survival rates after a solid cancer event are similar with the new drug and control despite the greater number of patients with solid cancer events in the new drug arm.

Solid Cancers in ARISTOTLE

Per the FDA primary clinical review ISTH major bleeding was less in the apixaban arm compared to the warfarin arm, hazard ratio 0.69 (95% CI 0.6-0.8). Solid cancer events were also less frequent in the apixaban arm as shown in Figure 8.

Figure 8: Times to First Solid Cancer* Events in ARISTOTLE



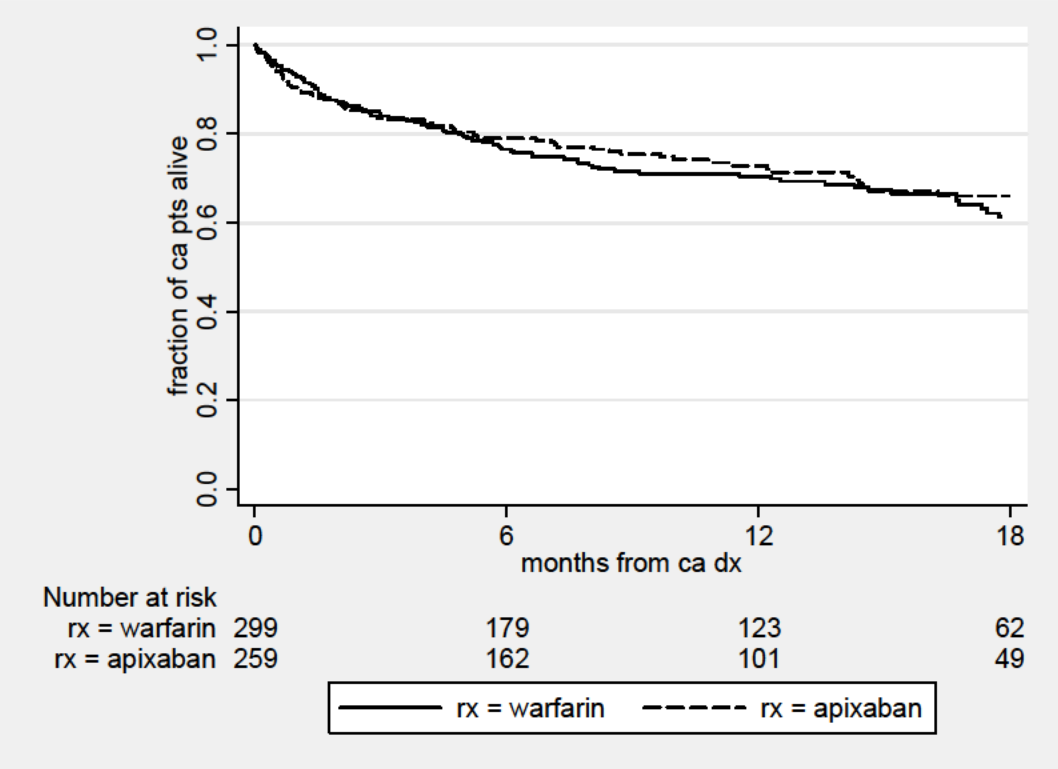
*excluding non-melanoma skin and brain; p = 0.058 by log rank test

The curves diverge again at about 2-3 months but, given that the difference is not nominally statistically significant, we might ignore the divergence. However, because of our priors regarding the association of more bleeding with more solid cancers, we should have more confidence that the divergence is real. The apixaban/warfarin hazard ratio for solid cancer events is 0.85 (95% CI 0.7-1.0).

The cancer site with the largest difference in events between the two arms was bladder, a cancer site frequently manifested by bleeding. However, gastrointestinal tract cancer events were approximately evenly distributed between the two arms and colorectal cancer first events were the same in both arms. The other sites substantially more frequent in the warfarin arm were unknown, lung, pancreas, and prostate.

Survival after a first solid cancer event was equally poor in both arms as shown in Figure 9.

Figure 9: Survival after First Solid Cancer* Events in ARISTOTLE



*excluding non-melanoma skin and brain; p > 0.5 by log rank test

Survival in patients experiencing a first solid cancer event was about 72% at one year after the event and 63% at 18 months. We should not dismiss these solid cancer events lightly.

COMMENT: We can consider it reassuring, and consistent with the bleeding results, that solid cancer rates are similar or lower in ARISTOTLE with apixaban than with warfarin. However, the APPRAISE-2 results raise the concern that apixaban, probably like all other drugs that increase bleeding rates, has the potential for increasing solid cancer rates. While it would be reassuring to assume that the increased solid cancer rates represent a detection bias based on bleeding of a tumor leading to earlier diagnosis, the continued divergence of the incidence curves and the poor survival after the events argue against complacency. We should make clinicians aware that bleeding on an anticoagulant or antiplatelet drug may represent a deadly cancer and that it should not be dismissed as a simple drug side effect. We should include the bleeding and cancer findings in all anticoagulant and antiplatelet labels, beginning with the apixaban label.

For months I have distributed within the FDA my preliminary findings for a wide range of antiplatelet and anticoagulant drugs associating bleeding with solid cancers. I have been proposing that the data associating bleeding and solid cancers be presented and discussed at an advisory committee meeting. I have never received a response from my superiors regarding my proposal. The closest that I have received to an explanation for the inactivity is an email

responding to a related issue: The issue of whether angiotensin receptor blockers (ARBs) are associated with increased cancer rates, as raised by a 2010 published meta-analysis (Sipahi, Debanne et al. 2010), remains unanswered today. The June 2, 2011, FDA safety communication (available at http://www.fda.gov/Drugs/DrugSafety/ucm257516.htm) that cleared ARBs is based on an inadequately-specified and flawed FDA meta-analysis, including counting "malignant lung neoplasm" but not "lung carcinoma" as lung cancers. I documented the many problems with the FDA meta-analysis in reviews (available to FDA staff) filed in DARRTS on July 20, 2012, and August 31, 2012, under Tracked Safety Issue 935. I proposed analyzing the ARB trials rigorously but was discouraged by Dr. Ellis Unger, the Office Director, in the email reproduced in the Attachment. The portion of that email most relevant to bleeding and cancer is the following quotation:

> "Third, if you were to go ahead with this and find a RR of, say 1.3, I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that. People would have various opinions on where the meaningful threshold is, but it might be worth asking for some input before you start."

I find it disturbing that Dr. Unger expressed a lack of concern about a drug increasing the risk of cancer by 30%. We routinely accept smaller efficacy benefits and our advisory committee has consistently refused to set a minimum magnitude for an acceptable benefit. We mandate that diabetic drugs not increase cardiovascular risk by 30%. That limit was partially based on practicality of trials—most clinicians would prefer a lower risk limit if the sizes of the resulting trials were feasible. I assert that any validated risk of cancer is concerning for a hypertensive drug used chronically for which there are many alternatives—or for anticoagulant and antiplatelet drugs used chronically. We should inform practitioners and patients about the risks of drugs so that they can make an informed decision and that they can institute follow-up measures to minimize the risks. I have proceeded in time available with the rigorous analysis of the ARB trials.

I recommend the following:

- *The data regarding bleeding and cancer should be presented and discussed at an advisory committee meeting. If the rigorous analysis of the ARB trials confirms a risk for ARBs and cancer, then the data regarding ARBs and cancer should also be presented and discussed at an advisory committee meeting. It may be advantageous to have both topics addressed at the same meeting.*

- *The data regarding bleeding and cancer should be included in the apixaban label and in the labels for all antiplatelet and anticoagulant drugs. If the advisory committee meeting discussing bleeding and cancer is scheduled promptly then the labeling regarding bleeding and cancer can be delayed until after the meeting.*

References

Alexander, J. H., R. D. Lopes, et al. (2011). "Apixaban with antiplatelet therapy after acute coronary syndrome." N Engl J Med **365**(8): 699-708.

BAATAF_Investigators (1990). "The effect of low-dose warfarin on the risk of stroke in patients with nonrheumatic atrial fibrillation. The Boston Area Anticoagulation Trial for Atrial Fibrillation Investigators." N Engl J Med **323**(22): 1505-11.

Connolly, S. J., A. Laupacis, et al. (1991). "Canadian Atrial Fibrillation Anticoagulation (CAFA) Study." J Am Coll Cardiol **18**(2): 349-55.

EAFT_Study_Group (1993). "Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke. EAFT (European Atrial Fibrillation Trial) Study Group." Lancet **342**(8882): 1255-62.

Ezekowitz, M. D., S. L. Bridgers, et al. (1992). "Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. Veterans Affairs Stroke Prevention in Nonrheumatic Atrial Fibrillation Investigators." N Engl J Med **327**(20): 1406-12.

Granger, C. B., J. H. Alexander, et al. (2011). "Apixaban versus warfarin in patients with atrial fibrillation." N Engl J Med **365**(11): 981-92.

Petersen, P., G. Boysen, et al. (1989). "Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study." Lancet **1**(8631): 175-9.

Sipahi, I., S. M. Debanne, et al. (2010). "Angiotensin-receptor blockade and risk of cancer: meta-analysis of randomised controlled trials." Lancet Oncol **11**(7): 627-36.

SPAF_Investigators (1991). "Stroke Prevention in Atrial Fibrillation Study. Final results." Circulation **84**(2): 527-39.

Wiviott, S. D., E. Braunwald, et al. (2007). "Prasugrel versus clopidogrel in patients with acute coronary syndromes." N Engl J Med **357**(20): 2001-15.

Attachment: Email explaining FDA reluctance to understand cancer risks

From: Unger, Ellis
Sent: Monday, August 20, 2012 11:04 PM
To: Marciniak, Thomas
Cc: Southworth, Mary Ross; Temple, Robert; Stockbridge, Norman L
Subject: RE: Emailing: ARB ca review plan v1p2.doc

Tom, et al,

I've gone through the protocol only fairly quickly, but I have a few comments.

First, this would represent a lot of man-hours, so I have to assume that there is a paucity of work in the Division at this point, or that you will be doing this mostly after hours.

Second, when we get into writing analytic plans, and specifically plans for adjudicating clinical endpoints, the plan/protocol might need to be reviewed at a high level – i.e., the OND IO or higher. There is a MAPP on this, I believe. You should consult that MAPP before you start any work to see if it applies here. If it applies, the protocol will need to go up to for review and comment before you begin.

Third, if you were to go ahead with this and find a RR of, say 1.3, I doubt there would be much enthusiasm for basing a regulatory decision (labeling or otherwise) on that. People would have various opinions on where the meaningful threshold is, but it might be worth asking for some input before you start.

Finally, given you familiarity with some of the trial data, any decision YOU make regarding inclusion and exclusion of trials can be called into question after the fact. It doesn't matter that your criteria are reasonable and defensible, because you can know the effect that your criteria will have on the trials to be included/excluded before you begin.

Ellis

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

/s/

THOMAS A MARCINIAK
12/11/2012

Subject:	Review of Apixaban Complete Response
NDA:	202155
Proposed Indication:	Reduction in the rate of stroke and systemic embolism in subjects with nonvalvular atrial fibrillation
Resubmission Date:	September 17, 2012
Clinical Reviewers:	Martin Rose, M.D., J.D. (efficacy) and B. Nhi Beasley, Pharm.D. (safety)

Table of Contents

1 Executive Summary

The Applicant has responded to our Complete Response (CR) letter of 6/22/2012 and addressed each of the 6 items in that letter. Most importantly, the Applicant has provided the requested information relating to medication errors from an additional 20% of bottle labels from ARISTOTLE study sites randomly selected by FDA. When combined with the 12% of bottle labels from randomly selected sites previously collected to satisfy the EMA, these bottle labels comprise the Applicant's "Combined Random Sample" of bottle labels. Due to conservative sampling practices, the Combined Random Sample includes 155,038 labels, representing about 35.5% of the bottles dispensed at all sites in the study.

The Applicant compared the serial number on each legible label to the IVRS record of bottles that were to be dispensed to the subject and generated rates of "Type 1" errors involving dispensing a bottle with a serial number not assigned to the subject (regardless of the bottle contents) and also a subset of Type 1 errors termed "Type 2" errors, in which the unassigned bottle contained the wrong medication or the wrong dose of active apixaban. All dispensations associated with illegible labels or missing labels were imputed to be Type 1 errors and either 50% or 100% of such dispensations were imputed as Type 2 errors. Type 2 errors had potentially serious implications for subjects, because nearly all Type 2 errors resulted in a subject receiving either two placebos or two active drugs for the period in which the subject took study drug from the bottle dispensed in error, thereby increasing the subject's risk of stroke or bleeding, respectively.

Labels representing over 99% of dispensations at the sampled sites in the Combined Random Sample were collected, and over 99% of collected labels were legible. The bottle-level error rate among bottles with legible labels was about 0.18% and 0.25% in the apixaban and warfarin arms, respectively. When imputed errors for illegible or missing labels are included, the bottle level error rate was 0.9% and 1.1% in the apixaban and warfarin arms, respectively. With the absolute worst case assumption that each missing or illegible label corresponded to a Type 2 error, the subject level error rate (i.e., the percentage of subjects who received at least one Type 2 error bottle) was 12.2% and 12.4% in the apixaban and warfarin arms, respectively.

Using worst case models of the effects of errors on outcomes (i.e., assuming that all errors either increased the rate of the primary efficacy endpoint, time to stroke or systemic embolism rate in the warfarin arm, or alternatively, increased the rate of primary safety endpoint, time to ISTH major bleeding, in the warfarin arm), we compared the worst case per subject error rates of about 12% to the error rates that would be required to negate superiority for apixaban to warfarin for the primary efficacy and safety endpoints. Dr. S. Bai (Office of Biostatistics) calculated that 13 fewer strokes or systemic emboli in the warfarin arm would negate superiority for the efficacy endpoint, and 86 fewer major bleeds in the warfarin arm would negate superiority for ISTH major bleeding. If all errors occurred in the warfarin arm, and all modeled error-induced events were eliminated from the analysis, it would take per-subject Type 2 error rates of over 20% (i.e., considerably greater than our worst case error estimate) to negate either superiority for either the primary efficacy or primary safety endpoint,

assuming that the rate of the relevant endpoint during the period that the subject took study drug from the erroneously dispensed bottle was 5 times the observed rate in the study for the warfarin arm. Thus, superiority of apixaban over warfarin for the primary efficacy and safety endpoints persisted despite highly unlikely, worst case models of the effects on event rates of highly unlikely, worst case error rates.

We also requested information regarding manual changes to an IVRS database held by a contractor that involved changing treatment assignment in the database after medication errors were discovered. We were concerned that these changes might have affected the randomization database. The sponsor submitted convincing information that the randomization data were not affected.

The CR letter also included requests for information relating to:

- Unblinding of the scratch-off panel of the bottle labels
- The intensity of monitoring of key eCRF fields, including drug accountability information
- Database inconsistencies with eCRFs observed in connection with our review of the medication error information
- Inconsistencies in the adverse event database, consisting of multiple entries for what appeared to be the same adverse event

The responses provided by the Applicant were sufficient to allay our concerns about the implications of these issues for the approvability of the application.

Recommendation: The application should be approved. We recommend that labeling describe that apixaban was superior to warfarin for the primary efficacy and safety endpoints as well as mortality. A revised draft of labeling is the eRoom.

In addition to our review of the Applicant's complete response, this review includes discussions of several safety issues not raised in the original clinical review of May 22, 2012. None of these issues affects our recommendation regarding approval. There is also an expanded discussion of the mortality data in the application, which in the opinion of this reviewer (MR) supports inclusion of data and language in Sec. 14 of labeling indicating that apixaban was significantly superior to warfarin for all-cause mortality in ARISTOTLE.

2 Background

The Applicant submitted this NDA for the reduction in the rate of stroke and systemic embolism in subjects with nonvalvular atrial fibrillation on 9/28/2011. The proposed indication (b) (4). After a 3 month extension of the review clock

because of a major submission late in the review cycle, we sent a Complete Response (CR) letter on 6/22/2012, with the following comments and requests:

"We have completed our review of this application, as amended, and have determined that we cannot approve the application in its present form. We have described our reasons for this action below and provided recommendations to address these issues.

1. As you are aware, some subjects in ARISTOTLE were given the wrong study drug (e.g. active instead of placebo and vice-versa). Knowledge of the study drugs actually dispensed to subjects is crucial to understanding the outcomes of the study. While we recognize your efforts to better define the rate and nature of dosing errors, which suggest a lower rate than initially reported, we believe the information you have submitted for us to review has not adequately characterized the frequency of errors in dispensing study drugs. Before we can approve your NDA you will need to submit reliable information regarding the following:

    - The frequency of a subject in ARISTOTLE receiving the wrong study medication, specifically:

        a. dispensing active warfarin instead of placebo warfarin to a subject randomized to apixaban;

        b. dispensing active apixaban instead of placebo apixaban to a subject randomized to warfarin;

        c. dispensing placebo apixaban instead of apixaban to a subject randomized to apixaban;

        d. dispensing placebo warfarin instead of warfarin to a subject randomized to warfarin, and

        e. dispensing the wrong dose of apixaban to a subject randomized to apixaban

    - The frequency of dispensing to a subject a bottle with a serial number other than the one assigned by the interactive voice response system (IVRS).
    - The frequency with which the serial number on the tear-off label from a study drug bottle did not match the IVRS assigned serial number.
    - The frequency with which the IVRS assigned serial number did not match each of the following and any one of the following: the eCRF entry of the serial number of dispensed study drug bottles, the eCRF entry of the serial number of returned study drug bottles, and the eCRF entry of the serial number of study drug bottles brought in to a visit but not returned.

    You should use whatever sources of information you have available to respond to our requests. We believe that the best source of information for responding to the first two bulleted items requested above is the tear off labels on which are printed the serial numbers of the study drug bottles. You may choose to collect all of these labels from investigative sites. However you recently submitted to us a summary report (but not the full report) you prepared for the European Medicines Agency (EMA), detailing your assessment of the errors in dispensing

study drugs in a random 12% sample of subjects. The full report of this assessment might constitute an adequate response to our requests above, or may serve as a template for designing a response to our requests. A brief review of the summary report raises some questions that will need to be addressed if the full report is used to respond to our requests. For example, your analysis included only legible labels whereas it seems to us that a bottle with a difficult to read label is more likely to be dispensed in error. You assert in the report you prepared for EMA that the frequency of a wrong study drug being dispensed to a subject was less than 0.1%. If the frequency is much greater than that, we may have additional requests for information. One of those will be identification of all primary endpoint events, deaths, and ISTH major bleeds that occur at times a subject may have been on two active study drugs or no active study drug. Another will be determining the frequency with which site monitoring identified and did not identify a subject whose eCRF indicates they may have been dispensed a bottle with an incorrect serial number.

2. Investigators in the ARISTOTLE study were supposed to report all instances in which they scratched off a coating on the tear-off labels to unblind a subject. We do not believe you have checked tear-off labels to verify that all instances of unblinding by investigators were reported in the 12% sample. If you have not done so, submit a plan to us to determine the frequency of unblinding by the investigator not reported to you. If the frequency is more than minimal, we may have additional requests.

3. In submissions to your NDA, you describe the intensity with which certain eCRF fields were monitored and checked during the conduct of ARISTOTLE. For example, you have told us that the entry on the eCRF listing the serial number of the bottle dispensed was subjected to more intense monitoring and edit checks than the entry listing the serial number of the bottle returned. To help us understand these different procedures, you need to provide us with all plans used to monitor and verify the accuracy of the entries on the eCRF. We note that the final monitoring plan you submitted in your NDA was dated after data-lock, and so may not have been a working plan actually used during the conduct of ARISTOTLE. You should provide the initial monitoring plan with all subsequent changes to that plan, both formal and informal. Indicate when and how changes in monitoring were implemented. Include all communications to sites, and identify the organizations responsible for all monitoring in the original monitoring plan and any changes made.

4. You recently informed us that manual changes were made to a dataset containing the serial numbers of bottles assigned by the IVRS to subjects in ARISTOTLE. You stated that these changes were made in response to information provided by investigators directly to the IVRS vendor, (b) (4) to ensure that a bottle dispensed in error would be removed from inventory so it could not be assigned.

   Manual changes to the IVRS concern us because of the possibility of alterations to the randomization dataset. You therefore should provide the following:

   - All agreements between BMS and (b) (4) concerning the role of (b) (4) in the conduct of ARISTOTLE.

- All SOPs from (b) (4) related to the conditions under which manual changes to data from the IVRS could be made and the documentation required to do so.
- An IVRS dataset that flags all subjects whose original IVRS-assigned bottle serial number was later changed, the serial number originally assigned and the altered serial number, and the reason for the change. Original IVRS datasets with codes to create the dataset (kitassgn) that was provided to the Agency may be helpful.
- Most importantly, an audit trail of the changes indicating who, when and why manual changes were made to the data set containing IVRS assigned bottle serial numbers. If the changes were made in response to information provided by an investigative site, please include the communication from the site documenting the information provided.
- A statement signed by responsible individuals that no changes were made to the randomization dataset.

5. We are concerned that the trial datasets submitted in your NDA do not accurately reflect the information in the eCRFs. In our brief review of your medication error dataset (smed.xpt) (used for most of your medication error analyses), we identified an observation with a valid date in the eCRF that was misrepresented by a period in the dataset, indicating that a valid date was missing. We also found medication data (indicating that drug was dispensed and taken until the end of treatment) for which there were no corresponding eCRFs in one subject. We can provide more details concerning these mismatches on request.

   We are concerned about these errors because they were found after a cursory examination of these datasets, leading us to believe that there may be important errors in the datasets used for critical analyses. You should explain how these and similar errors, if any, occurred. If you believe that the datasets for important analyses are accurate, please provide the basis for your belief.

6. Some subjects have a unique adverse event listed multiple times as both non-serious and serious. This appears to be because the site personnel completed a non-serious AE CRF and a serious adverse event (SAE) CRF for the same event. You should prepare an adverse event analysis dataset (adae.xpt) in which all adverse events are listed a single time with the correct designation as serious or non-serious.

We expect that you will anticipate and answer any reasonable questions we are likely to have after reviewing the information you submit in your complete response."

We met with the Applicant on July 16, 2012 (minutes dated August 2) to discuss these requests and their plans for a response. We also sent a follow-up general advice letter on August 9. The sponsor's response to the CR was substantially complete on September 17, and on September 26 we sent a letter classifying the September 17 submission as a Class 2 Resubmission, with a goal date of March 17, 2013. This is primarily a review of the resubmission.

For additional information about the rationale for the CR letter and the above requests for information, see the initial Clinical Review of this NDA dated 5/22/2012, the Cross Disciplinary Team Leader review and the Division Director's memo, both dated 6/22/2012.

3 Review of Submitted Responses

Request 1

Request 1 related to the frequency of dispensing errors. These errors fall into 2 general types:

> The subject received a bottle with a serial number not identical to one allocated to the subject by the IVRS system (called "Type 1"). The bottle may or may not have contained the wrong medication for the subject.
>
> The subject received a bottle of the wrong medication (or the wrong dose of apixaban) (called "Type 2"). The bottle would necessarily bear a serial number different from the one allocated to the subject by the IVRS. Type 2 errors are a subset of Type 1 errors.
>
> Type 1 errors may or may not be medically important, but in all cases reflect dispensing errors at the site. Type 2 errors would be potentially medically important to the subject who received the wrong bottle. This was a double dummy study in which each subject received an active drug (either apixaban or warfarin) and a placebo for the other treatment. Warfarin and apixaban bottles and labels appeared quite different. Most Type 2 errors involved substitution of a placebo for an assigned active drug (e.g., warfarin placebo instead of warfarin active) or vice versa. A subject who received a placebo version of the assigned active drug would be taking 2 placebos for a period of several weeks to 3 months if the error was not caught, and would be at higher risk for stroke or systemic embolism during this period. A subject who received an active drug instead of the assigned placebo would be taking 2 actives and be at higher risk of bleeding. A subject in the active apixaban arm who received the wrong active dose could be at greater risk of stroke or bleeding, depending on whether the subject was supposed to receive the 5 mg dose (but received 2.5 mg tablets) or the 2.5 mg dose (but received 5 mg tablets), respectively.

Sources of Data

The discussion below focuses primarily on error information derived from the partial set of bottle label tear-off panels (which include the bottle serial number and a bar code) collected by the Sponsor from randomly selected study sites. We will refer to these as "labels." The Sponsor collected all available labels from randomly selected sites in two waves. The "First Random Sample" included about 12% of all dispensed labels, and was collected late in the initial review cycle. It was done in response to a request by the EMA. After that, the "Confirmatory Random Sample" was collected in response to the CR letter and was intended to include another 20% of dispensed labels. The combination of these two samples, termed the "Combined Random Sample," is the primary focus of this analysis. Both samples consisted of labels from sites randomly selected to provide the target number of labels. All labels at selected sites were to be submitted. In the case of First Random Sample, BMS ran the site selection program. In the case of the Confirmatory Random Sample, FDA ran the site selection program. Due to conservative sampling practices intended to assure that at least the targeted number of legible labels were collected, the Combined Random Sample included 155,038 labels, representing about 35.5% of the 436,182 bottles dispensed at all sites in the study, somewhat more than the combined target of 32% of labels.

In addition to these labels, the Sponsor had another 8% of labels, termed the "Convenience Sample" that were provided to the Sponsor in error, with more than 80% of these coming from sites in Russia. Presumably, the CRO that serviced sites in Russia provided erroneous information to the sites about the need to submit these labels to the sponsor. These labels were available prior to the CR letter, and analyses of error data from the Convenience Sample are included in our initial medical review of 5/22/2012. These labels are not included in most of the analyses discussed below because of the non-random nature of the sample.

Information on the processes used to collect these samples and enter data is found in **Table 1**, copied from the sponsor's submission. The "investigations" of bottle label mismatches and missing labels mentioned in the table are discussed further **below**.

Table 1. Processes Used to Collect Bottle Labels and Record Data

Sample	Collection	Entry	Illegible	Mismatch	Missing/ irretrievable
Convenience	Original Labels available at BMS	Excel – Single Entry of Original labels only	Yes – barcode reader	Yes – Investigation	No follow up
First Random	Original Labels – Random Sampling of Sites by BMS – all labels from randomized subjects	Excel – Double Entry of Original labels only	Yes – barcode reader and investigations	Yes - Investigation	Yes – Investigation
Confirmatory Random	Original or Copy of label pages - Random Sampling of Sites provided by FDA – all labels at site	Oracle Clinical – copies accepted if no original is available	Yes – barcode reader and investigations	Yes - Investigation	Yes – Investigation

Although we requested analyses of error rates based on several sources of information regarding what was dispensed to the subject, the reviewers believe that the most accurate source of information regarding dispensing is the tear-off portion of the bottle label that was removed from the label at the time of dispensing and retained at the site on either a hard copy of CRF page 800 or in the subject file. The sponsor regarded these labels as the source data for information entered by the site into the eCRF dispensing field and the drug dispensing logs that were kept at the sites.

Table 2 is a display of the makeup of the First, Confirmatory and Combined Random Samples, with overall study information for sites, bottles dispensed, and all treated subjects for comparison. Sites that were in the First Random Sample could not be in the Confirmatory Random Sample. Sites that contributed to the Convenience Sample could not be in either random sample. The Convenience Sample had 35,859 legible labels from 1508 subjects at 97 sites in 15 countries, but over 80% of labels were from sites in Russia.

Table 2. Random Samples of Labels: Composition

	First Random Sample	Confirmatory Random Sample	Combined Random Sample	Entire Study
Contributing sites	123	235	358	1033 [1]
Contributing countries[2]	29	39	40	40
Bottles Dispensed	56,343	99,774	156,117	436,182
Legible labels collected	55,937	98,949	154,886	(a)
Subjects with ≥ 1 bottle dispensed	2327	4197	6520	18,140
Subjects with ≥ 1 legible label	2324	4190	6511	(a)

1 Sites that treated at least 1 subject.
2 reviewer analysis: resub\country labels, applicant datasets: fdasites.xpt, rsample.xpt submission 73
(a) All known collected and analyzed labels were either in the Combined Random Sample or the Convenience Sample (see **above**)

Data for disposition of the labels in the Combined Random Sample are shown in **Table 3.** The collected labels accounted for more than 99% of the bottles dispensed at the sampled sites, and 99.9% of the collected labels had a discernible bottle number, determined either visually, or if not visually legible, by reading the bar code on the label with an electronic device.[1]

The geographic origin of the bottles in the Combined Random Sample was reasonably consistent with the origin of subjects in the study overall. Also about 69% of the labels were from bottles of warfarin or matching placebo and 31% were from bottles of apixaban or matching placebo, consistent with the overall dispensing data for the trial (**Table 4**).

[1] The sponsor provided information on the barcode readers used to read visually illegible labels (all were Wasp Model 3509 Barcode Readers manufactured by Wasp Barcode Technologies). For validation, each of the 9 barcode readers purchased by BMS was used to read 100 visually legible ARISTOTLE bottle labels (different labels were used to validate each reader). Three readers were tested at each BMS site that received bottle labels (2 in the US and one in Belgium). The visually legible bottle serial number on each label was entered into a database and compared to the number obtained using the barcode reader, which was entered into the database electronically. Each reader was 100% accurate in matching the visually read serial numbers. The NDA submission suggests that the individually validated barcode readers were then used to read labels for the analyses described in this review.

Table 3. Combined Random Sample: Label Disposition

	APIXABAN ARM	WARFARIN ARM
BOTTLES DISPENSED	80229	75888
LABELS COLLECTED, n (%) [1]	79717 (99.36)	75321 (99.25)
LEGIBLE LABELS, n (%) [2]	79640 (99.90)	75246 (99.90)
VISUALLY LEGIBLE, n (%) [2]	79434 (99.64)	75038 (99.62)
BARCODE READABLE, n (%) [2]	206 (0.26)	208 (0.28)
ILLEGIBLE LABELS, n (%) [2]	77 (0.10)	75 (0.10)
ABSENT LABLES, n (%) [1]	435 (0.54)	492 (0.65)
MISSING LABELS, n (%) [1, 3]	512 (0.64)	567 (0.75)

1 Denominator for percentages is the number of bottles dispensed at the sites included in the Combined Random Sample.
2 Denominator for percentages is the number of labels collected.
3 “Missing labels” are those that are absent (based on the number of bottles dispensed minus number of labels collected) plus those that are illegible

Table 4. Combined Random Sample - Additional Information on Labels and Dispensed Bottles

	Legible Labels (L=154,886)	Missing Labels (L=1231)	Bottles Dispensed (D=156,117)	Overall Study Data (D=436,182, S=18,140)
GEOGRAPHIC REGION: ASIA/PACIFIC, n (%)	20028 (12.93)	297 (24.13)	20325 (13.02)	(16) [1]
EUROPE, n (%)	58188 (37.57)	488 (39.64)	58676 (37.58)	(40) [1]
NORTH AMERICA, n (%)	49164 (31.74)	334 (27.13)	49498 (31.71)	(25) [1]
LATIN AMERICA, n (%)	27506 (17.76)	112 (9.10)	27618 (17.69)	(19) [1]
CONTAINER TYPE: APIXABAN/PLACEBO, n (%)	47585 (30.72)	357 (29.00)	47942 (30.71)	132680 (30.4)[2]
WARFARIN/PLACEBO, n (%)	107301 (69.28)	874 (71.00)	108175 (69.15)	303502 (69.6)[2]

1 Percentage of subjects in region. Geographic data on dispensing were not provided.
2 Percentage of bottles dispensed
D=No. of bottles dispensed
L=No. of labels
S=No. of treated subjects

The applicant did not provide additional analyses on the demographics and disease-related parameters in subjects included in the Combined Random Sample. However, such analyses are of modest interest only, because the primary focus of these analyses is dispensation.

Note that site 386 in Belgium was one of the sites selected to be in the Confirmatory Random Sample. This site reportedly sent its bottle labels to BMS but they were lost in shipment. No copies were made prior to shipment. BMS considered these labels to be "non-informative missing data," and they were excluded from all analyses.

Reviewer Comment: The Applicant's treatment of the data from site 386 is acceptable.

Analytic Methods

In general, bottle based error rates were calculated as E/D, where E represents the number of bottles erroneously dispensed (including actual errors and, in most analyses, also imputed errors) and D represents the number of dispensed bottles at the relevant sites included in the sample.

The method for determining whether an error occurred differed across analyses and will be described in connection with the analyses presented below.

The method for determining the count of dispensed bottles was consistent across most analyses. Unless otherwise indicated, a bottle was considered dispensed if one or more of the following events occurred:

a label existed or a site indicated that a container was dispensed

a container was assigned by IVRS

a container was entered in the eCRF as dispensed, even if it was not assigned by the IVRS.

Reviewer Comment: This method for determining dispensing events would tend to produce a larger denominator for error rate analyses than one based on only one source of data, such as the IVRS database. While this might slightly reduce the observed actual error rates, it would probably notably increase the rate of imputed errors, because in FDA's preferred analysis the number of imputed errors is equal to the number of dispensations minus the number of obtained legible labels. We think this is a reasonable worst-case approach since labels may not be missing at random, and because we think a missing or illegible label may be more likely to be associated with an error than one that is available and legible. This approach also gives the sponsor a strong incentive to minimize the number of uncollected labels or illegible labels.

In general, subject-based error rates were calculated as P_e/N, where P_e represents the number of subjects with one or more dispensing errors (again, usually including imputed errors) and N is the number of subjects who received one or more bottles of study drug at the sites included in the sample.

Results of Dispensing Error Analyses Requested by FDA

The analyses of dispensing errors based on bottle labels that were requested by FDA in general had the following characteristics:

Analyses based on bottle labels attempted to account for missing and illegible bottle labels, assuming that such labels should not be considered to be "missing (or illegible) at random". In any case where a bottle was dispensed that could not be matched to a legible bottle label, it was assumed that a bottle with the wrong serial number was dispensed. Note that the vast majority of known dispensing errors involved the same type of bottle (i.e., if the IVRS-allocated bottle was active warfarin, another bottle of "warfarin" would in most cases be dispensed in the event of an error), with a 50% chance of the subject getting either placebo or active study drug. Accordingly, each imputed error involving dispensing of the wrong bottle was considered to result in 0.5 imputed dispensing errors involving the wrong medication.

A dispensing error based on a non-matching, missing, or illegible bottle label was counted as an error, even if other sources of information, such as the case report form or a medication log, indicated that the correct bottle might have been dispensed. Our rationale for this choice was that the bottle label represented the source document for what was dispensed. The CRF and the medication logs were filled in by site staff, perhaps the same site staff that dispensed a wrong bottle. It is possible that the site filled out these forms based on what was supposed to have been dispensed, not what was actually dispensed. Note the site received the allocated bottle number from the IVRS through a computer generated voice on the telephone and then within minutes to hours later, via a fax or email. If the fax or email was used as a source document, the site-generated record would match the IVRS regardless of what bottle was actually dispensed.

The sponsor provided analyses consistent with FDA's preference. The sponsor provided additional analyses in which other sources of information were used to negate errors based on the bottle labels.

In addition to the analyses of dispensing errors based on bottle labels, FDA requested analyses based on the eCRF records of bottles dispensed, returned, or "validated" (i.e., brought to the site for counting of tablets, and then taken home again by the subject) as the source of dispensing information (see **above**). The sponsor also used other sources of information to negate dispensing errors found in these analyses.

Data for bottle-level type errors based on bottle labels (i.e., the subject received a bottle not assigned by the IVRS, regardless of what drug was contained in the bottle) is displayed in **Table 5**. About 0.2% of dispensations in each arm were associated with legible labels indicating that a subject received a bottle with the wrong serial number. Another 0.73% to 0.85% of dispensations in the apixaban and warfarin arms, respectively, were associated with missing labels (by the Applicant's definition, the term "missing labels" included illegible labels). In the worst case, if all missing labels are

imputed to be associated with bottle-level errors), the error rate is 0.9% in the apixaban arm and 1.1% in the warfarin arm. In the best case, if none of the missing labels were associated with errors, the error rates would be 0.18% and 0.25% in the apixaban and warfarin arms, respectively.

Table 5. Combined Random Sample: Type 1 Bottle Level Errors

	APIXABAN ARM D=80,229	WARFARIN ARM D=75,888
Legible labels, n (%)	79640 (99.27)	75246 (99.15)
Bottles not matching IVRS assignment	142 (0.18)	188 (0.25)
Missing labels, n (%)	589 (0.73)	642 (0.85)
Composite of estimated errors, n(%) [1]	731 (0.91)	830 (1.09)

D is the number of bottles dispensed; all percentages use D as the denominator
1. Composite n= bottles not matching IVRS assignment + missing labels (worst case)

Table 6 provides bottle-level data on the number of Type 2 errors (dispensing of the wrong medication to a subject). Over 99.0% of dispensations in each arm were associated with labels that decoded to the correct treatment. About 0.10% vs. 0.11% of labels for apixaban and warfarin subjects arms, respectively, decoded to the wrong treatment. Errors in the dispensing of warfarin (which in the table represents the incorrect active treatment for apixaban arm subjects and the incorrect placebo treatment for warfarin arm subjects) accounted for about 2/3 of the errors, consistent with the fact that more than twice as many bottles containing warfarin active or placebo were dispensed than bottles of apixaban active or placebo. Errors consisting of dispensing the wrong dose of active apixaban (5 mg tablets for 2.5 mg tablets, or vice-versa) were exceedingly rare. About 0.73% and 0.85% of labels were missing for the subjects in the apixaban and warfarin arms, respectively. If each of these missing labels was associated with a bottle not matching the IVRS allocated bottle, then about 50% of these labels would be associated with the wrong medicine. This follows because the typical Type 2 error was substitution of active for placebo of the same type or vice-versa. Thus, 50% of missing labels were imputed to be associated with dispensing of the wrong medication. If the observed errors are combined with the imputed errors, the composite bottle level Type 2 error rate was 0.47% vs. 0.54% in the apixaban and warfarin arms, respectively.

Table 6. Combined Random Sample: Type 2 Bottle Level Errors

	APIXABAN ARM D=80,229	WARFARIN ARM D=75,888
Legible labels, n (%)	79640 (99.27)	75246 (99.15)
Total containers decoding to correct treatment, n (%)	79560 (99.17)	75159 (99.04)
Total containers decoding to incorrect treatment, n (%)	80 (0.10)	87 (0.11)
Incorrect active containers, n (%)	50 (0.062)	28 (0.037)
Incorrect placebo containers, n (%)	25 (0.031)	59 (0.078)
Wrong apixaban strength, n (%)	5 (0.006)	NA
Missing labels, n (%)	589 (0.73)	642 (0.85)
Composite of type 2 errors, n (%) [2]	374.5 (0.47)	408 (0.54)

D is the number of bottles dispensed; all percentages use D as the denominator
n = bottles not matching IVRS assignment + 0.5 x missing labels

*Reviewer Comment: The rate of observed bottle level Type 2 errors in the Combined Random Sample (third row of data in **Table 6**) is reasonably close to the previously reviewed analogous error rates in the Convenience Sample, which were 0.10% and 0.15% in the apixaban and warfarin arms, respectively. Note that the Applicant's analysis of the Convenience Sample errors did not count substitution of one active apixaban dose for another as an error, but there were very few such errors in the Combined Random Sample, so the error rates seem similar in the Convenience Sample and Combined Random Sample.*

Table 7 provides data on subject level errors, including both Type 1 errors (subject received at least one bottle not assigned by the IVRS, regardless of the type of medication) and Type 2 errors (subject received at least one bottle containing the wrong study medication).

Type 1 errors – subject level:

Over 99.8% subjects in each arm had at least one legible label. About 3.4% and 4.2% of subjects in the apixaban and warfarin arms, respectively, had at least one label with a serial number not matching an IVRS assigned serial number for that subject. About 10.8% of subjects in each arm had at least one missing label. No analysis on the overlap between subjects with non-matching labels and missing labels were provided. However, the information for Type 2 errors suggests that at least 17 subjects in the apixaban arm and 11 subjects in the warfarin arm had both observed medication errors and missing labels. If all missing labels are imputed to represent non-matching labels and only the 28 subjects described in the previous sentence had both a missing label and a non-matching label, then one can derive an estimate of subject-level Type 1

errors. This estimate yields error rates of 13.7% and 14.6% in the apixaban and warfarin arms, respectively.

Type 2 errors – subject level:

When missing labels are ignored, over 97.8% of subjects in each arm had at least one legible label and received no labels corresponding to bottles containing the wrong medication. However, 10.79% and 10.75% of subjects in the apixaban and warfarin arms, respectively, had at least one missing label. In this case, we know the number of subjects who had at least one missing label and who received no known bottle with the medication (**Table 7**). If we assume that 50% of missing labels are associated with the wrong medication, and that no person had more than one missing label, then the rate of Type 2 errors was 7.03% and 7.24% in the apixaban and warfarin arms, respectively. FDA agreed to accept the 50% rate based on verbal information from the sponsor that very few subjects had more than one missing label. However, it is now clear that a substantial number of subjects had more than one missing label. For example, there were 589 missing labels in the apixaban arm subjects, and 353 subjects had at least one missing label. Among those subjects, the mean number of missing labels was 589/353, or 1.67, but we did not calculate the distribution. In the warfarin arm, the analogous mean value is 1.83. For any individual, as the number of missing labels increases, the expected probability of receiving a bottle containing the wrong medication would increase.[2] Because we were not provided with the distribution of the missing labels, in addition to creating a composite of observed medication errors + imputed errors using a 50% rate for imputation, we also calculated the composite assuming that 100% of missing labels were associated with dispensing the wrong medicine. This absolute worst case scenario yielded Type 2 error rates of 12.16% and 12.44% in the apixaban and warfarin arms, respectively.

[2] If the probability of receiving the wrong medication associated with each missing label is p, then the probability of receiving at least one bottle of wrong medication for a subject with n missing labels is $1-(1-p)^n$ For example, if the probability of receiving the wrong medication associated with each missing label is 0.5, then a subject with either 1, 2, or 3 missing labels would have a 50%, 75%, or 87.5% likelihood, respectively, of receiving at least one bottle of the wrong medication.

Table 7. Combined Random Sample - Subject Level Summary of Medication Errors and Missing Labels

	Apixaban Arm S=3273	Warfarin Arm S=3247
Type 1 Errors		
Subjects with at least one legible label, n (%)	3269 (99.88)	324 (99.85)
Subjects with at least one legible label not matching IVRS assignment, n (%)	111 (3.39)	136 (4.19)
Subjects with at least one missing label, n (%)	353 (10.79)	349 (10.75)
Estimate of subjects with at least one missing label and no legible labels not matching IVRS assignment	336 (10.27)	338 (10.41)
Estimate of composite type 1 error rate, n (%) [1]	447 (13.66)	474 (14.60)
Type 2 Errors		
Subjects receiving all containers of correct type, n (%)	3207 (97.98)	3176 (97.81)
Subjects receiving at least one container of incorrect type, n (%)	62 (1.89)	66 (2.03)
Active container of incorrect type, n (%)	45 (1.37)	28 (0.86)
Placebo container of incorrect type, n (%)	24 (0.73)	51 (1.58)
Wrong apixaban strength, n (%)	3 (0.09)	-NA-
Subjects with at least one missing label, n (%)	353 (10.79)	349 (10.75)
Subjects with at least one missing label and without study medication error in legible labels, n (%)	336 (10.27)	338 (10.41)
Composite of observed type 2 errors and missing data, n (%) [2]	230 (7.03)	235 (7.24)
Composite of observed type 2 errors and missing data, n (%) (worst case) [3]	398 (12.16)	404 (12.44)

1 n = subjects with at least one legible label not matching IVRS + estimate of subjects with at least one missing label but without observed mismatch with IVRS assignments
2 n = subjects receiving a container of incorrect type + 0.5 x subjects with at least one missing label and without a study medication error in legible labels.
3 n = subjects receiving a container of incorrect type + 1.0 x subjects with at least one missing label and without a study medication error in legible labels
S = number of subjects with at least one bottle dispensed, used as the denominator for all percentages
NA – not applicable

FDA also requested a bottle-level analysis of mismatches between bottle serial numbers entered into 3 eCRF fields vs. the IVRS allocated bottles. The 3 eCRF fields represented:

- dispensed bottles,
- returned bottles, and
- "verified" bottles (i.e., bottles containing study medication brought in for tablet counts and then taken home again by the subject).

Data entry into these fields was performed manually at the site. Sites were not supplied with bar code readers.

Table 8 provides bottle-level data on Type 1 errors found when comparing bottle serial numbers in these 3 fields to the IVRS allocation for each treated subject.

Table 8. All Treated Subjects – Bottle-Level Data for Type 1 Errors based on eCRF Fields

Discrepancies between IVRS allocation and eCRF data in -	Apixaban Arm n / N (%)	Warfarin Arm n / N (%)
Dispensed field	109 / 218863 (0.05)	135 / 206747 (0.07)
Returned field	1601 / 203586 (0.79)	1530 / 191940 (0.80)
Verified field [1]	57 / 59108 (0.10)	74 / 57991 (0.13)
Any category above	1667 / 224271 (0.74)	1607 / 211911 (0.76)

1 A bottle might be brought in to be verified more than once.
Denominator for each percentage is the number of containers collected from eCRF for each category and each treatment group.

Per bottle Type 1 error rates varied widely depending on which of the 3 fields was compared to the IVRS file. The lowest error rates – 0.05% to 0.07% – were observed with the dispensed field. Error rates in the returned were more than 10 x error rates in the dispensed field. Error rates in the verified field were about 2 x the rates in the dispensed field. Note that these data were reviewed in our initial NDA review. It is also notable that all error rates in **Table 8** are lower than the composite bottle-level Type 1 error rates based on bottle labels in **Table 5**.

The sponsor noted that about ¾ of the errors represented in **Table 8** involved bottles apparently returned to the site that were never shipped to the site in question (based on the sponsor's shipping records). They argue that it was physically impossible for such a bottle to be dispensed from or returned to the site. They performed an analysis where such bottles were eliminated from the numerator and denominator (**Table 9**). Error rates for all 3 eCRF fields were reduced from those displayed in **Table 8**, with the largest reduction observed in the returned field.

Table 9. All Treated Subjects – Bottle-Level Data for Type 1 Errors based on eCRF Fields

(Counting Only Containers Shipped to the Relevant Site)

Discrepancies between IVRS allocation and eCRF data in -	Apixaban Arm n / N (%)	Warfarin Arm n / N (%)
Dispensed field	97 / 218851 (0.04)	119 / 206731 (0.06)
Returned field	359 / 202344 (0.18)	369 / 190779 (0.19)
Verified field [1]	23 / 59074 (0.04)	36 / 57953 (0.06)
Any category above	383 / 222987 (0.17)	398 / 210702 (0.19)

1 A bottle might be brought in to be verified more than once.
Denominator for each percentage is the number of containers collected from eCRF for each category and each treatment group.

Reviewer Comment: The reviewers agree that a bottle not shipped to a site was probably never dispensed to or returned by a subject at that site.[3] However, it is not known what bottle was actually dispensed or returned in cases where the eCRF indicates that bottle shipped to another site was given to or returned by a subject. It seems possible that some other bottle at the site was mistakenly dispensed in cases where a bottle not shipped to the site appears in an eCRF field.

Additional Dispensing Error Analyses Submitted by the Applicant

We agreed that the Applicant could submit additional analyses beyond those that we requested, including those based on bottle labels that included only legible bottle labels and those that allowed elimination of errors in cases where site-created records (i.e., the eCRF or the dispensing logs) indicated that the correct bottle was dispensed, even if the bottle label indicated otherwise or was missing or illegible. They also corrected what they termed to be "label placement errors," which occurred when a bottle was correctly dispensed but the corresponding label was placed in the record of another subject at the site. The evidence supporting such corrections consisted of:

eCRF dispensing data

Protocol Deviation Reports by site

Site Monitoring Visit Reports

Site Correspondence regarding label pages

Master Drug Logs

[3] When the sponsor first submitted this analysis, we did an analysis of the approximately 35,000 labels in the Convenience Sample database provided by the Sponsor. Not one label was found at a site where the corresponding bottle was not shipped according to the shipping records. This is one reason why we believe that the bottle labels are the best source of evidence regarding what was actually dispensed.

Individual Subject Drug Logs

The following decision rule was used:

> If the container number is uniquely identified in a Protocol Deviation Report, Site Monitoring Visit Report or other Site Correspondence, the subject identified as receiving the container in these records was assigned the container.
>
> Otherwise, if the drug logs, the dispensed field of the eCRF data, and IVRS data are all consistent with this container being dispensed to another subject at the same site, and there is no evidence in the drug logs, dispensed field of the eCRF, and IVRS that the container was dispensed to the subject identified on the label page, the container was assigned to this other subject.
>
> If a container cannot be uniquely re-assigned to a subject according to this process, it remained as entered in the label database associated with the subject as reported by the site.

In addition, the sponsor used the following techniques to eliminate alleged false positive bottle error determinations:

> The sponsor used the eCRF entries and site drug logs to help "deduce" the serial number on illegible labels.
>
> Sites were queried about all missing labels in attempt to determine if a bottle was actually dispensed in connection with a dispensing event that could not be matched to a label. The sites confirmed that 161/589 missing labels in the apixaban group and 128/462 missing labels in the warfarin group were associated with supposed dispensing events that did not occur.
>
> The sponsor also made more favorable assumptions than FDA about the consequences of a missing label. FDA imputed all missing labels to be associated with Type 1 errors and initially 50% to be associated with dispensing of the wrong medication on a per bottle basis. The sponsor made a series of assumptions with Type 2 error rates ranging from 1 to 10 x the observed error rate for legible labels.

These various techniques led to error rates lower than those reported in the tables in this review. However, it is not necessary to provide additional details regarding the sponsor's analyses, because even if one accepts FDA's worst case assumptions, the error rates are very unlikely to have affected the treatment effects on the primary efficacy endpoint and the primary safety endpoint in ARISTOTLE, as discussed in the next section of this review.

Potential Impact of Type 2 Errors on Key Study Outcomes

As noted in Sec. 3.1.1.2 of the initial Medical review, Dr. Steven Bai, the statistical reviewer for apixaban, calculated that it would take 13 fewer primary endpoint events (strokes or systemic emboli) in warfarin arm subjects to negate the finding that apixaban was superior to warfarin for the primary endpoint. **Table 10**, copied from page 52 of the initial NDA review, reports the results of a simple model intended to explore the effects

of the additional primary endpoint events (strokes or systemic emboli) resulting from a worst case scenario of warfarin arm subjects receiving placebo for warfarin instead of warfarin. Such subjects were assumed to receive two placebos for 38 days, the mean duration of treatment with one bottle of warfarin. This was assumed to increase the event rate to either 3X or 5X the rate actually observed in the warfarin arm of ARISTOTLE (1.6 events/100 subject-years) for 38 days. With the expected 3X event rate (consistent with a 66% reduction in event rate with therapy compared to placebo, comparable to the results of the placebo-controlled trials of warfarin), it would require a per-subject error rate of >40% in the warfarin arm, with no analogous errors (i.e., a subject receiving 2 placebos) in the apixaban arm, to negate the finding of superiority of apixaban for the primary endpoint. If a 5X event rate were to result from taking 2 placebos, it would take a >20% per-subject error rate in the warfarin arm, with no analogous errors in the apixaban arm to negate superiority for the primary endpoint.

Table 10 Extreme Worst-Case Modeling of Effects of Medication Errors on Primary Endpoint Events in ARISTOTLE

Warfarin arm subjects with errors, % [1] N=9052 [2]	Rate of primary endpoint during error period, multiple of observed rate [3]	Error-induced additional events, per 100 pt-yr [4]	Additional events
10%	3x	3.2	3.0
40%	3x	3.2	12.1
100%	3x	3.2	30.2
20%	5x	6.4	12.1
100%	5x	6.4	60.3

1. All errors were assumed to be substitution of placebo warfarin for active warfarin in warfarin arm subjects (see text).
2. All treated subjects in warfarin arm
3. The primary endpoint occurred at a rate of 1.6 events per 100 subject-years in the warfarin arm in the ITT analysis.
4. Effect of error was assumed to last for 38 days before returning to the baseline rate in each case.

However, the highest per-subject error rates calculated in connection with this review, even using worst case (and very unlikely) assumptions regarding missing data, are displayed in **Table 7**. The composite rate of observed and imputed Type 2 subject-level errors was 12.16% and 12.44% in the apixaban and warfarin arms, respectively. This is below the threshold for losing statistical significance for the primary endpoint finding even if one assumes a primary efficacy event rate in subjects with errors that is 5x the observed event rate in the warfarin arm. Also note that the composite error rate in **Table 7** is quite similar in the 2 treatment arms, rather than the nearly impossible (and exceedingly unfavorable for the warfarin arm) asymmetrical pattern that the model assumes. Thus, the observed superiority of apixaban for the primary endpoint persists despite worst-case assumptions about the rate of errors and their distribution.

Table 11 Extreme Worst-Case Modeling of Effects of Medication Errors on ISTH Major Bleeding in ARISTOTLE

Warfarin arm subjects with errors, % [1] N=9052 [2]	Rate of major bleeding during error period, multiple of observed rate [3]	Error-induced additional events, per 100 pt-yr [4]	Additional events
10%	3x	6.2	15.4
40%	3x	6.2	61.5
100%	3x	6.2	154.7
20%	5x	12.4	61.5
100%	5x	12.4	307.3

1. All errors were assumed to be substitution of active apixaban for placebo apixaban in warfarin arm subjects (see text).
2. All treated subjects in warfarin arm
3. The rate of major bleeding was 3.1 events per 100 subject-years in the warfarin arm in the safety analysis.
4. Effect of error was assumed to last for 100 days before return to the baseline rate in each case.

Analogous data for the primary safety endpoint, ISTH major bleeding, are shown in **Table 11**. Dr. Bai calculated that a minimum of 86 fewer bleeds in the warfarin arm would negate superiority of apixaban. If we assume that all errors were substitution of active apixaban for placebo apixaban in warfarin arm subjects, resulting in a subject taking two active drugs for 100 days, then a per subject error rate appreciably greater than 40% would be required for 86 additional bleeds in warfarin arm subjects if the bleeding rate during the error period was 3 x the observed overall major bleeding rate of 3.1%/year. If we assume a bleeding rate 5 x the observed overall rate during the error period, then an error rate appreciably greater than 20% would be required to negate superiority of apixaban. However, the worst case error rate was less than 13%, and the actual error probably substantially less than that. Note that errors in warfarin/placebo dosing were considerably more common than errors in apixaban/placebo, and this model assumes that all errors involved apixaban dosing in the warfarin arm, which clearly was not the case. Thus, the observed superiority of apixaban for ISTH major bleeding also persists despite worst-case modeling assumptions.

The finding of superiority for mortality in ARISTOTLE is more fragile than either the primary efficacy or safety endpoint findings. A swing of one death in the wrong direction would result in a p-value greater than 0.05 for the difference in mortality rates. However, given that extensive analyses show roughly similar error rates in the two

treatment arms, and the observed rates of stroke and death due to stroke[4] are substantially reduced in the apixaban arm, it seems speculative and illogical to infer an effect on death that disfavors apixaban in an analysis adjusted for medication errors. For a more thorough discussion of this issue, see the discussion of mortality in Sec. **4**.

Request 2

This request concerned additional cases of unblinding discovered in connection with the bottle label analyses conducted by the sponsor. As noted in the initial review, the bottle label tear off panel had a scratch-off coating that covered information on the identity of the study drug within the bottle. Investigators were supposed to report all cases in which they scratched off this coating.

The sponsor examined all collected labels in the Convenience, First Random and Confirmatory Random Samples. There were no unreported unblindings in the Convenience and First Random Samples. In the Confirmatory Random Sample, there were 4 labels from 3 subjects (each at a different site) discovered to have been scratched off, breaking the blind for the relevant subject, that were not known to BMS prior to examining the labels. This amounts to 3 out of 6511 subjects in the Combined Random Sample, or 0.046%, with previously undocumented unblinding. At this rate, one would expect a total of about 8.4 such unblindings in subjects in the entire study population of 19,140 treated subjects, which is not of concern. For more information about the subjects with previously unreported unblinding in the Confirmatory Random Sample, see **Table 12**.

Table 12. Listing of Subjects in the Confirmatory Random Sample with Previously Unreported Unblinding

Country	Site ID	Subject ID	Container Number(s)
France	0838	05697	564165
Germany	0956	02308	150230 176911
United Kingdom	1679	11498	433357

Reviewer comment: One might be concerned about missing labels with regard to unreported unblinding, because it is possible that an unreported, scratched off label would be more likely to be not submitted to BMS by a site involved in one of the two random samples than some other label. Overall, in the Combined Random Sample, there were 927 labels that were missing in the sense of being absent (435 and 492 in the apixaban and warfarin arms, respectively, see ***Table 3***). However, only 2 of these absent labels were at sites where there were unreported unblindings in the Combined Random Sample (one at each of 2 sites, 0838 and 0956). This is somewhat, but not

[4] There were 38 vs. 65 deaths due to stroke (HR=0.58, 95% CI, 0.39, 0.86), HR This compares favorably to a HR of 0.79 both for the primary endpoint and for stroke of all kinds **(Table 20).**

completely reassuring. However, the hypothesis of informative missingness suggested in the first sentence of this comment is speculative.

Request 3

This request related to the monitoring plans for ARISTOTLE.

As noted in our CR letter, the site final monitoring plan (SMP) was not signed until after data lock (see **above**). However the sponsor had multiple draft (unsigned) and "final" (signed) versions of the monitoring plan in place during the study, as follows:

> A global SMP, Ver. 3.0, dated Dec 11, 2006, was in place for the ARISTOTLE study at the time of the first subject visit on Dec 19, 2006. Key details of this plan are discussed below. Earlier versions of the SMP were not implemented.
>
> 2 other draft versions were implemented on Jan 29, 2007 and Jul 11, 2007.
>
> 6 final, signed versions were implemented on dates ranging from Dec 15, 2007 to Jul 7, 2011. The last "final" version, 6.0, was signed after data lock, which occurred on Jun 10, 2011.

The remainder of this portion of the review will focus on monitoring of drug accountability (DA), because medication errors were the primary issue that led to the CR.

Initially, all subjects were subject to 100% monitoring of drug accountability and other parts of the case record. However, ARISTOTLE had a reduced source data verification (RSDV) plan, which was first described in Ver. 4.0 of the SMP (Jan 29, 2007). Under this plan, once a site met specified data quality criteria,[5] 1 out of 2 subjects at the site would have 100% source data verification (SDV). In addition, after 5000 subject had been randomized, at sites where RSDV was allowed, only 1 in 5 subjects had 100% SDV. Other subjects would have RSVD, consisting of verification of:

1. Informed Consent Forms (ICFs).
2. Serious Adverse Events (SAEs).
3. Screening and randomization visits
4. If the site participated in the Pharmacogenetic Study: SDV of ICFs for Pharmacogenetic (PGx) Samples, the CRA was to ensure that ICF was signed prior to collection of the PGx Sample.

Note that DA information was not one of the items that were monitored in the subjects (ultimately 80% of subjects at qualifying sites) who had RSDV. However, other information provided to us suggests that DA may have been monitored in RSDV subjects. This information includes CRA training slides for a training session held on October 5, 2006 (about 2 months prior to enrollment of the first study subject). The

[5] The quality criteria included quality checks of "critical data modules," which included the drug dispensing CRF pages for both apixaban/placebo and warfarin/placebo. There were both initial quality criteria and ongoing quality criteria to maintain RSVD eligibility.

deck includes a slide regarding interim monitoring visits (IMV, i.e., visits between initiation and closeout) that reads:

On the basis of this information, the sponsor states that the container ID in the CRF was compared to the bottle label at each visit for each subject.

However, there was a major change in the monitoring plan that was rolled out in 2 stages in 2008 and 2009. This involved "reduced drug accountability" (RDA). RDA was first implemented in North America on Mar 18, 2008 following training of study monitors. It was first mentioned in the SMP in Final version 2.0, dated June 1, 2009, when it was extended globally. RDA was implemented at sites that qualified for RSVD. At these sites, 1 out of 4 subjects had "Full DA" review; others, i.e., 75% of subjects, had "No DA" review. (Note: quotes are from a RDA training slide deck for CROs dated "January 2008"). The description of RDA in final ver. 2.0 indicates states,

> "4. The current RDA Program requires that Drug Accountability is conducted on 1 in 4 subjects at a site. The remaining 3 of 4 subjects at that site do not require any Drug Accountability review. (Don't count any tablets).
>
> "5. Drug Accountability will be conducted on every other 100% SDV subject; remaining subjects will be eliminated from drug accountability review." [6]

Neither the slides nor the short description of RDA in the SMP specifies whether bottle labels would routinely be compared to the eCRF in the 75% of subjects with RDA. However, the language in the slides and the SMP suggested that this comparison might not have been made in subjects with RDA.

While a clarification of whether eCRF vs. bottle label monitoring was performed in RDA subjects would be some interest, we believe that further questioning of the Applicant is not warranted at this time. Regardless of whether such monitoring occurred in ARISTOTLE, we are confident that the rate of medication errors was too low to negate the findings of superiority for the primary efficacy endpoint or the primary safety

[6] A later document indicates that 1/5 of subjects at eligible sites, not ¼, were subject to complete DA.

endpoint in ARISTOTLE. Thus, the outcome of an inquiry regarding DA monitoring would be very unlikely to affect our view of the study results.

Request 4

This request dealt with manual changes of the databases following medication errors, and possible corruption of the randomization code resulting from such changes.

(b) (4) Processes

(b) (4) (a contractor for the Applicant) was responsible for the IVRS system and study drug inventory management. As requested, the sponsor provided us with SOPs relating to (b) (4) processes in place during ARISTOTLE for maintaining databases related to randomization and drug inventory.

Briefly, manual changes were allowed but an audit trial (either electronic or hard copy) was to be generated. In the case of changes affecting either randomization or drug dispensing, notification and approval of the sponsor (without breaking the blind) was to occur. The ICTI working instructions on data modifications state:

"5.3.1. As a general rule, ICTI should not make changes to randomization data as the ICTI database should always reflect what actually occurred." (Note: ICTI is now known as (b) (4)

Both (b) (4) and the sponsor maintain that the randomization database for ARISTOTLE was never modified due to medication errors, although randomizations were cancelled if subjects were randomized in error. Our analysis supports this assertion (see **below**).

The working instructions on data modification state the following regarding medication errors:

> 5.3.8 If a kit has been dispensed incorrectly to a subject in a double-blind study, generally the kit that was assigned by the IVRS but not given to the subject should not be put back into circulation, as this may cause partial unblinding/biasing.
>
> 5.3.8.1 The client must be notified of the incorrectly dispensed kit and the resolution discussed and documented.

In addition, the IVRS SAS Data Set Specifications for (b) (4) (dated Oct 24, 2011, after data base lock) indicate that in the case of changes to the inventory database for related to incorrectly dispensed medication, errors, the randomization file was not to be changed.

The sponsor provided an audit trail summary relating to each of the 107 manual changes of the database. None of the audit trail entries suggest that the randomization was altered.

FDA's Comparisons of Databases Relevant to Randomization

To independently confirm that the randomization database was not altered, we made the following comparisons of databases relevant to randomization in ARISTOTLE, focusing on subject identifying numbers and the corresponding treatment code:

- We compared the BMS-generated randomization file that was sent to (b) (4) before any subjects were randomized (rmrsch.xpt) to the randomization file executed by (b) (4) ((b) (4).xpt). The files were concordant with respect to entries for each randomized subject.

- We compared the (b) (4) randomization file to the randomization file used by BMS to unblind the study and create treatment-based data tables in the study report for ARISTOTLE, as evidenced by the ADDM.xpt analysis file included in Module 5 of the NDA. The files were concordant with respect to each randomized subject.

- We obtained from BMS the code for "RAINMAN" (the SAS-based program used by BMS to generate the randomization scheme for ARISTOTLE and other studies), along with instructions for its use and the seed used to generate the randomization for ARISTOTLE. Dr. Steve Bai (OB) then ran RAINMAN using the provided seed and generated a new version of the ARISTOTLE randomization file. He compared this new file with the file provided by BMS to (b) (4) described in the first bullet (rmrsch.xpt). The two files were concordant.

These comparisons and the demonstrated concordance of the various databases support the statements by BMS and (b) (4) that the study randomization was not affected by the inventory-related changes to records made by (b) (4)

Request 5

This request pertains to discrepancies between the medication error dataset and the eCRFs found after a cursory review of the data that the applicant submitted in response to our questions about the treatment imbalance in medication errors. We were concerned that important datasets used for the primary and secondary analyses may also be inaccurate. The FDA-identified errors and the Applicant's investigation of these errors are summarized in **Table 13**.

Table 13. Specific Examples of Discrepancies Between the Data and eCRF and the Applicant's Resolution

Applicant's table 5.1, submission 66, 8/22/2012

The Applicant explained that the analysis datasets and programmed statistical outputs (primarily developed by PPD) were double programmed by independent statisticians or programmers. Discrepancies between programmed outputs were resolved through review of the issues. Duke Clinical Research Institute (DCRI) independently generated analysis datasets and programming to support the ARISTOTLE publication in the New England Journal of Medicine.[1]

> *Reviewer's comment: The Applicant investigated all submitted CRFs (5,452) and found that the data in all truncated CRFs (44) were accurate. This investigation and DCRI's independent duplication of key analyses provides reassurance of the integrity of the data used for important analyses.*

Request 6

This request was for a dataset that indicates the correct designation of the adverse event as either serious or non-serious (NSAE). Unique adverse events that appeared multiple times in the AE dataset (as both an SAE and NSAE) were collapsed into one record if the records contained the same MedDRA Preferred Term (PT) and the same onset date. In addition, if information regarding the study drug differed between the 2+ records that were collapsed, then the most conservative value was kept (e.g., The most conservative value for "Relationship to study drug" is "most related". The most

conservative value for “Action taken regarding study drug” is “drug discontinued, followed by drug interrupted, followed by dose change”).

The revised dataset, ADAE2.xpt, appeared to correctly designate unique AEs as either serious or NSAE, however the dataset still contained multiple NSAE records for a unique event. (e.g., two NSAE records for one subject are identical except that the drug was discontinued in one record and no action was taken with the drug for the other record). The effect of this would have an impact if the NSAE records were mapped to different MedDRA terms.

Reanalysis of the SAE data with the revised dataset did not alter the SAE results, as suspected. As mentioned in Section 7.3.2.2 (Other Nonfatal SAEs) of our review, there were more SAEs for syncope and related terms in subjects on apixaban compared to warfarin. The reviewer cannot offer an explanation for this finding, but the treatment difference and concern of falls in this population warrants inclusion of syncope in labeling.

Table 14. ARISTOTLE – SAE of syncope and related terms

Preferred Term (PT)	Apixaban N=9088	(%)	Warfarin N=9052	(%)
Subjects	130	(1.4)	95	(1.0)
Syncope	79	(0.9)	49	(0.5)
Vertigo	21	(0.2)	13	(0.1)
Dizziness	18	(0.2)	13	(0.1)
Presyncope	15	(0.2)	12	(0.1)

Reviewer’s analysis: analysis\ae\common.sas
Subjects include those with SAE from Day 1 to Day 30 post dose with the PTs presyncope, vertigo, dizziness and PTs mapped to the SOC of “Disturbances in consciousness”

The analysis of adverse events leading to treatment discontinuation with the revised dataset was similar to that reported in our original review. The reasons for treatment discontinuation were similar between treatment groups. The most common High Level Group Term (HLGT) for nonfatal AE leading to treatment discontinuation was central nervous system vascular disorders followed by gastrointestinal hemorrhages, and then epidermal and dermal conditions.

Significant nonfatal AEs that led to dose interventions are shown for HLGT. Based on the ARISTOTLE protocol, it is expected that warfarin treated subjects would have more dose changes.

Table 15. ARISTOTLE - Significant AE (AE leading to a dose intervention)

High Level Group Terms	Apixaban N=9088 (%)		Warfarin N=9052 (%)	
Dose Interrupted	1505	16.6	1847	(20.4)
Infections - Pathogen Unspecified	159	(1.7)	180	(2.0)
Dental And Gingival Conditions	139	(1.5)	148	(1.6)
Urinary Tract Signs And Symptoms	115	(1.3)	155	(1.7)
Injuries	102	(1.1)	152	(1.7)
Upper Respiratory Tract Disorders (Excluding Infections)	100	(1.1)	154	(1.7)
Dose Decrease	161	(1.8)	429	(4.7)
Dose Increase	11	(0.1)	20	(0.2)

Reviewer's analysis: analysis\ae\aedc.sas
Subjects counted only once in each HLGT.

Non-serious AEs that were severe or greater in intensity occurred in ~ 8% of subjects and were generally similar in both treatment arms. **Table 16.** ARISTOTLE - Non-serious AE with intensity ≥ severe highlights the SOCs that occurred in ≥ 1% of apixaban treated subjects. Selected HLGT within each SOC are shown if they occurred in at least 0.5% of subjects on apixaban.

Table 16. ARISTOTLE - Non-serious AE with intensity ≥ severe

System Organ Class **High Level Group Terms**	Apixaban N=9088 (%)		Warfarin N=9052 (%)	
Subjects	702	(7.7)	768	(8.5)
Musculoskeletal And Connective Tissue Disorders	139	(1.5)	137	(1.5)
Joint Disorders	77	(0.8)	52	(0.6)
Musculoskeletal And Connective Tissue Disorders	47	(0.5)	65	(0.7)
Gastrointestinal Disorders	95	(1.0)	91	(1.0)
Infections And Infestations	88	(1.0)	95	(1.0)
Infections - Pathogen Unspecified	62	(0.7)	70	(0.8)

Reviewer's analysis: analysis\ae\severe.sas
Subjects counted only once in each MedDRA level analysis.

Common AE with the revised dataset produced results identical to that of the Applicant's, and the results were discussed in Section 7.4.1 of the original review.

These data provide no evidence of a signal of a difference in non-bleeding AEs between warfarin and apixaban.

4 Additional Review Issues

AVERROES and Major Bleeding

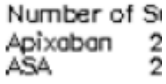
AVERROES was stopped early because of a significant benefit in stroke and SE on apixaban compared to aspirin. There was an acceptable safety profile. Bleeding outcomes are shown in **Table 17**. The annual rate of major bleeding was higher on apixaban than on aspirin, but there no significant relative difference.

Table 17. AVERROES - Bleeding Endpoints

	Apixaban N=2798		Aspirin N=2780		Apixaban vs. Aspirin	
Event	(n)	%/yr	(n)	%/yr	HR	95% CI
Major Bleed	45	1.41	29	0.92	1.54	(0.96, 2.45)
Fatal Bleed	5	0.16	5	0.16	0.99	(0.29, 3.41)
Major Bleed or CRNM	140	4.46	101	3.24	1.38	(1.07, 1.78)
CRNM	98	3.12	74	2.37	1.31	(0.97, 1.78)
Minor Bleed	200	6.55	153	5.02	1.30	(1.05, 1.61)
Any Bleed	325	10.9	250	8.32	1.30	(1.10, 1.53)

Reviewer's analysis: analysis\averroes\erateHR runs bleed.sas, first event
CRNM=clinically relevant non-major

Figure 1. AVERROES - Time to First ISTH Major Bleed (on treatment period)



Applicant's Figure 8.2 from AVERROES Clinical Study Report

The number of intracranial hemorrhages was the same (11) between treatments. Similar to ARISTOTLE, there were more major intraocular bleeds on apixaban compared to the comparator (6 vs. 0).

Concomitant Use of Drugs that Cause Bleeding

Almost 40% of subjects were taking aspirin during the trial. The applicant determined bleed event rates in subjects that took aspirin and those who did not. The analysis confirms that bleeding is worse in subjects on aspirin compared to no aspirin; bleeding is worse on warfarin compared to apixaban.

Table 18. ARISTOTLE - Bleeding Rates with Concomitant Aspirin Use

EVENT RA
95% CI
RATE DIFF
95% CI

PT-YRS
EVENT
95% CI
RATE D
95% CI

who are
The denc
each tre

Source:

In ARISTOTLE, 27% of subjects on apixaban also took NSAIDS (mostly acetaminophen) for some duration during the trial. We have asked the applicant to conduct an analysis similar to their aspirin concomitant medication analysis.

Outcome Events in Subjects with Emergent Procedures

The lack of a reversal agent for the new OAC appears to cause consternation with providers and has been cited as a reason for not prescribing one of the new OAC. Although warfarin has reversal agents, it is unclear if their use has had a significant impact on important outcomes. We requested information on outcome events in subjects with emergent and elective procedures in ARISTOTLE and AVERROES. If useful information can be gleaned from that information, an addendum will be written.

Mortality in ARISTOTLE

As noted above, the information on medication errors does not negate the observed findings of superiority to warfarin for the primary efficacy endpoint, time to the composite of stroke and systemic embolism, and the primary safety endpoint, time to ISTH major bleeding.

The mortality finding (superiority for all-cause death) is not nearly as robust as the findings for the primary endpoint and major bleeding. **Table 19** is adapted from the original clinical review of the NDA. Note that the HR for all-cause death favors apixaban with a point estimate of 0.89, but the upper limit of the 95% CI for all-cause death is 1.00 and the p is 0.465. Dr. Bai calculated that 1 less death in the warfarin arm would negate statistical significance for superiority of apixaban. While one could imagine that medication errors might have contributed to one or more deaths in the warfarin arm, such errors might also have contributed to deaths in the apixaban arm.

It is noteworthy that in addition to being superior to warfarin for the primary endpoint of time to stroke or systemic embolism, apixaban was superior to warfarin for time to stroke of any kind (199 vs. 250 events, HR =0.79, 95% CI 0.66, 0.95) and fatal stroke (38 vs. 65, HR=0.58. 95% CI 0.39, 0.86), and nearly so for non-fatal stroke (161 vs. 185, HR=0.84, 95% CI 0.68, 1.04). The low HR for fatal stroke seem consistent with the advantage of apixaban over warfarin in terms of hemorrhagic strokes (40 vs. 78), which tend to be more severe and more often fatal than non-hemorrhagic strokes. It is quite plausible that an agent that is superior to another in reducing the rate of stroke, and in particular, fatal stroke, would have an advantage in terms of overall mortality, as long as it was at least neutral in terms of other important causes of death, which apixaban seems to be.

Thus, data from ARISTOTLE for fatal outcomes other than all-cause death and for serious non-fatal outcomes lend credibility to the observed finding of superiority of apixaban over warfarin for all-cause death. In addition, the strong trend for a mortality advantage for apixaban over aspirin in AVERROES, which was terminated after an interim analysis showed superiority for the primary endpoint of time to stroke or SE, also lends support to the mortality finding in ARISTOTLE. [7] Note that while aspirin is not approved to reduce either the rate of stroke or mortality in subjects with nonvalvular atrial fibrillation, a meta-analysis of study data suggests that it does reduce the rate of stroke in patients with atrial fibrillation.[2] It thus seems likely that the rate of death in aspirin-treated subjects with atrial fibrillation is not increased over no treatment, and that superiority of apixaban over aspirin for all-cause death would support an effect of apixaban on mortality.

[7] Rates of all cause death in the apixaban and aspirin arms in AVERROES were, respectively, 3.51 vs. 4.42 %/year (HR, 0.79, 95% CI, 0.62, 1.02, p=0.068).

Table 19 ARISTOTLE –- Adjudicated Deaths in the ITT Population

Endpoint	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR(95% CI)	2-sided p (superiority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
All-cause death	603 (6.61)	3.52	669 (7.37)	3.94	0.89 (0.80, 1.00)	0.0465
CV death (Caused by ↓)	308 (3.38)	1.80	344 (3.79)	2.02	0.89 (0.76, 1.04)	-
Stroke	38 (0.42)	-	65 (0.72)	-	0.58 (0.66, 0.95)	-
Systemic embolism	1 (0.01)	-	2 (0.02)	-	-	-
MI	21 (0.23)	-	17 (0.19)	-	-	-
Sudden death	126 (1.38)	-	129 (1.42)	-	-	-
Heart failure	76 (0.83)	-	92 (1.01)	-	-	-
Other CV cause	23 (0.25)	-	22 (0.24)	-	-	-
Unobserved death	23 (0.25)	-	17 (0.19)	-	-	-
Non-CV death (Caused by ↓)	196 (2.15)	1.14	208 (2.29)	1.22	0.93 (0.77, 1.13)	-
Bleeding	15 (0.16)	-	17 (0.19)	-	-	-
Malignancy	60 (0.66)	-	66 (0.73)	-	-	-
Infection	67 (0.73)	-	52 (0.57)	-	-	-
Trauma	7 (0.08)	-	13 (0.14)	-	-	-
Respiratory failure	19 (0.21)	-	35 (0.39)	-	-	-
Other non-CV cause	28 (0.31)	-	25 (0.28)	-	-	-
Unknown cause of death	99 (1.09)	0.58	117 (1.29)	0.69	0.84 (0.64, 1.09)	-

It is notable that the mortality data from the RE-LY study of dabigatran vs. warfarin in subjects with atrial fibrillation show a similar pattern to those of ARISTOTLE. While not in the same pharmacologic class, both drugs inhibit factors in the final common coagulation pathway. The first step in the pathway is the activation of Factor X to Factor Xa. Factor Xa is inhibited by apixaban and the approved agent rivaroxaban. Factor Xa triggers the activation of Factor II (prothrombin) to Factor IIa (thrombin); thrombin is inhibited by dabigatran. Thrombin triggers the conversion of fibrinogen to fibrin, the third step in the pathway.

In RE-LY, the data for the comparison of dabigatran 150 mg bid vs. warfarin yielded a HR for mortality of 0.88 (95% CI, 0.77, 1.00). A hazard ratio for fatal stroke is not available, but the risk ratio for dabigatran 150 mg vs. warfarin was 0.52. Thus the apixaban mortality data are quite similar to those of dabigatran, a closely related product for the same indication (**Table 20**).

In sum, apixaban and dabigatran are closely related drugs that inhibit adjacent steps near the end of the coagulation cascade. Both drugs are superior to warfarin in terms of reducing the rates of stroke and systemic embolism, and show similar reductions in the rate of all-cause death and fatal stroke compared to warfarin.

It is also notable that in ARISTOTLE, while fatal strokes make up only 8% of total deaths (103/1272) the difference between the treatment arms in fatal strokes, 27, is 41% of the difference between the two arms in total deaths (66). In RE-LY, fatal strokes made up 7% of total deaths, while the difference between the treatment arms in fatal strokes, 21, was 43% of the difference between the arms in deaths (49). Thus, in each study, fatal strokes made up less than 10% of overall deaths, but the difference between the treatment arms in fatal strokes made up over 40% of the overall difference between the arms in deaths. Thus, the difference in mortality between the arms was due in substantial part to a reduction in the rate of fatal stroke, which is a not surprising finding for drugs that reduce the rate of stroke. Note that both apixaban and dabigatran were associated with large reductions in the rate of hemorrhagic stroke, which tends to be more serious and more often fatal than ischemic stroke. This too would also be consistent with an effect on overall mortality (**Table 20**).[8]

Thus, the similarities in the mortality data in ARISTOTLE and RE-LY, and the close proximity of Factor Xa and IIa in the coagulation cascade, suggest that the two studies could be considered to support each other's findings.[9] There is precedent for this practice. Losartan and irbesartan were both studied (in separate placebo-controlled trials) for use in prevention of progression of diabetic nephropathy based on a composite that included changes in serum creatinine. Each product was directly supported by one trial, but two trials would ordinarily be required to support an approval based on a surrogate. However, the studies and the drugs were similar enough so that the studies could be considered to support each other, and both drugs were approved for the diabetic nephropathy indication.[10] In that case, the drugs were in the same class (angiotensin receptor blockers) and the sponsors gave each other a right of reference to the relevant data. These conditions are not applicable here. However, the mortality data in ARISTOTLE and RE-LY are very similar, and the drugs are closely related, as both are inhibitors of clotting factors in the final common pathway of the coagulation

[8] Dabigatran, unlike apixaban, also significantly reduced the rate of ischemic stroke compared to warfarin. To date, dabigatran is the only novel anticoagulant to have shown this benefit in subjects with AF.
[9] ROCKET-AF, a blinded trial comparing rivaroxaban to warfarin in subjects with AF, had a directionally similar pattern of efficacy findings, but rivaroxaban was not superior to warfarin in the ITT population for the primary endpoint or for mortality.
[10] See NDA 20-757, Review by E. Fromm, file date June 7, 2002

cascade. The data from both studies are available to the public in published reports,[3,4] and for RE-LY, in the medical review of the dabigatran NDA (on FDA's website).[11]

Table 20. Selected Parameters Related to Mortality in ARISTOTLE and RE-LY

	ARISTOTLE (Apixaban vs. warfarin) HR (95% CI) N=18,201	**RE-LY** (Dabigatran 150 mg bid vs. warfarin) HR (95% CI) or RR [1] N=12,098
All-cause death	603 vs. 669 deaths 0.89 (0.80, 1.00) p=0.0465	438 vs. 487 deaths 0.88 (0.77, 1.00) p=0.052
Fatal stroke	38 vs. 65 deaths 0.58 (0.39, 0.86)	23 vs. 44 deaths 0.52 [1]
Composite of stroke and systemic embolism	212 vs. 265 subjects 0.79 (0.66, 0.95)	134 vs.202 subjects 0.65 (0.52, 0.81)
Stroke (all types)	199 vs. 250 subjects 0.79 (0.66, 0.95)	122 vs. 186 subjects 0.64 (0.51, 0.81)
Hemorrhagic stroke	40 vs. 78 subjects 0.51 (0.35, 0.75)	12 vs. 45 subjects 0.26 (0.14, 0.49)
Ischemic stroke	162 vs. 175 subjects [2] 0.92 (0.74, 1.13)	103 vs.134 subjects 0.75 (0.58, 0.97)

1 RR was calculated when HR was not provided
2 Includes strokes of undetermined type in ARISTOTLE

Reviewer Comment: Accordingly, this reviewer concludes that the mortality results of ARISTOTLE should be displayed in Section 14 of labeling along with a clear indication that there was superiority for apixaban over warfarin for all-cause death as well as for the primary efficacy endpoint. (b) (4)

[11] Available at http://www.accessdata.fda.gov/drugsatfda_docs/nda/2010/022512Orig1s000MedR.pdf, accessed Nov. 2, 2012.

Reference List

(1) Granger CB, Alexander JH, McMurray JJ et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med 2011;365:981-992.

(2) Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007;146:857-867.

(3) Granger CB, Alexander JH, McMurray JJ et al. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981-992.

(4) Connolly SJ, Ezekowitz MD, Yusuf S et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009;361:1139-1151.

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

/s/

MARTIN ROSE
12/10/2012

BACH N BEASLEY
12/10/2012

Deputy Office Director Decisional Memo

Date	(electronic stamp)
From	Robert Temple, MD
Subject	Deputy Office Director Decisional Memo
NDA/BLA # **Supplement #**	202155
Applicant Name	Bristol-Myers Squibb
Date of Submission	September 28, 2011
PDUFA Goal Date	June 8, 2012
Proprietary Name / **Established (USAN) Name**	Eliquis (apixaban) tablets
Dosage Forms / Strength	Tablet 5 mg and 2.5 mg
Proposed Indication(s)	1. Atrial Fibrillation
Action:	

Material Reviewed/Consulted OND Action Package, including:	**Names of discipline reviewers**
Medical Officer Review	Nhi Beasley, PharmD Martin Rose, M.D., JD
Statistical Review	Steve Bai
Pharmacology Toxicology Review	Pat Harlow, Ph.D.
CMC Review/OBP Review	Charles Jewell, William Adams, Young Wang
Microbiology Review	
Clinical Pharmacology Review	Jim Ping Lai, Tzu McDonald
OPDP	Emily Baker – Full Product Labeling Zarna Patel – Patient Labeling
OSI	Sharon Gershon
CDTL Review	Stephen M. Grant, M.D.
OSE/DEpi	
OSE/DMEPA	Morgan Walker, Ray Ford
OSE/DRISK	Danielle Smith
Other – Div Dir Review Dep Dir for Safety Review	

OND=Office of New Drugs
OPDP=Office of Prescription Drug Products
DSI=Division of Scientific Investigations
CDTL=Cross-Discipline Team Leader
OSE= Office of Surveillance and Epidemiology
DEPi= Division of Epidemiology
DMEPA=Division of Medication Error Prevention and Analysis
DRISK=Division of Risk Management

I. Introduction

NDA 202155 is for apixaban (Eliquis), an anticoagulant (Factor Xs inhibitor) intended to reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular Atrial Fibrillation. (b) (4)

Dr. Rose has discussed the study supporting effectiveness, ARISTOTLE, in detail, in his May 22, 2012 (completion date) review. It was a non-inferiority trial of apixaban 5 mg bid (2.5 mg if age > 50, weight < 60 kg, or SCr > 1.5, vs warfarin titrated to an INR range of 2-3, conducted at 1053 sites in 40 countries in patients with non-valvular AF or Aflutter at enrollment or at least twice in the 12 months prior to enrollment. Patients were to have at least one risk factor (age > 75); prior stroke, TIA, or SE; CHF or LVEF < 40%, diabetes, or treated hypertension). It was a randomized double-blind event driven trial (target 446 adjudicated events) and patients were to be treated until the study ended. The NI margin (M_2) to be used for the primary endpoint (stroke plus non-CNS systemic embolism, adjudicated) was 1.38, accepted by FDA, representing ruling out a loss of > 50% of the effect of warfarin.

Two other anticoagulants have recently been approved for this use, dabigatran 150 mg and rivaroxaban. In both cases the NI-designed trials clearly showed effectiveness (dabigatran was superior to warfarin). One issue in such trials is how well the control (warfarin) was used (i.e., time in range for INR), an issue with rivaroxaban. For both drugs the clearest benefit was reduced hemorrhagic stroke, but dabigatran also showed a clear reduction in thromboembolic stroke.

The results in ARISTOTLE, as reported/presented, are favorable for time to first event.

ARISTOTLE Reported

		Apixaban n = 9120	Warfarin n = 9081			
	n (%)	events/ pt-yr	n (%)	events/pt-yr	HR	P
Primary endpoint	212 (2.3)	1.27	285 (2.9)	1.60	0.79	0.0114
Isch stroke	159 (1.7)		173 (1.9)			
Hemorrhagic	38 (0.4)		76 (0.8)			
Syst Embol	15 (0.2)		16 (0.2)			

As with the other anticoagulant drugs, decreased hemorrhagic stroke is a prominent effect.

Death was a secondary endpoint and a marginally significant reduction was reported (HR 0.89, p = 0.0465); CV death was similar and most of the benefit related to stroke. Apixaban also appeared to have lower rates of severe bleeding.

II. CR Determination

The above results are not discussed further because it would be premature given our major concern with the application. As described in detail in the clinical reviews and in Dr. Grant's CDTL review, the principal concern in the apixaban review has been what seemed initially to be a very high rate of medication errors (wrong active drug, placebo instead of active, active instead of placebo), a rate initially thought to be as great as 10%. It now is contended by the sponsor that these were mainly reading errors,

so that the bottle number placed in the CRF was often wrong, not reflecting what was actually given to the patient. BMS has reviewed stickers from the bottles, which were placed on the CRF (early part of study) or kept in the file (later) and believes the true error rate is far lower, perhaps 0.1-0.2%. The documentation of their analysis of a random 12% of patients (in response to an EMA request) has not been formally or fully submitted to us. The informal submission remains under review. Whether this 12% (plus another 8% analyzed previously using a convenience sample of available stickers) will settle the issue remains to be seen.

It is acknowledged that the apixaban results seem strong and we have not found significant lesions in the analysis. Nonetheless, I have concluded that approval would not be appropriate without strong assurance that we know what drugs the patients were given. We hope the CR will elicit a prompt and complete response so that evaluation of a promising drug can be completed without undue delay.

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

/s/

ROBERT TEMPLE
09/12/2012

Deputy Office Director Decisional Memo

Date	(electronic stamp)
From	Robert Temple, MD
Subject	Deputy Office Director Decisional Memo
NDA/BLA # **Supplement #**	202155
Applicant Name	Bristol-Myers Squibb
Date of Submission	September 28, 2011
PDUFA Goal Date	June 8, 2012
Proprietary Name / **Established (USAN) Name**	Eliquis (apixaban) tablets
Dosage Forms / Strength	Tablet 5 mg and 2.5 mg
Proposed Indication(s)	1. Atrial Fibrillation
Action:	

Material Reviewed/Consulted OND Action Package, including:	**Names of discipline reviewers**
Medical Officer Review	Nhi Beasley, PharmD Martin Rose, M.D., JD
Statistical Review	Steve Bai
Pharmacology Toxicology Review	Pat Harlow, Ph.D.
CMC Review/OBP Review	Charles Jewell, William Adams, Young Wang
Microbiology Review	
Clinical Pharmacology Review	Jim Ping Lai, Tzu McDonald
OPDP	Emily Baker – Full Product Labeling Zarna Patel – Patient Labeling
OSI	Sharon Gershon
CDTL Review	Stephen M. Grant, M.D.
OSE/DEpi	
OSE/DMEPA	Morgan Walker, Ray Ford
OSE/DRISK	Danielle Smith
Other – Div Dir Review Dep Dir for Safety Review	

OND=Office of New Drugs
OPDP=Office of Prescription Drug Products
DSI=Division of Scientific Investigations
CDTL=Cross-Discipline Team Leader
OSE= Office of Surveillance and Epidemiology
DEPi= Division of Epidemiology
DMEPA=Division of Medication Error Prevention and Analysis
DRISK=Division of Risk Management

I. Introduction

NDA 202155 is for apixaban (Eliquis), an anticoagulant (Factor Xs inhibitor) intended to reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular Atrial Fibrillation. (b) (4)

Dr. Rose has discussed the study supporting effectiveness, ARISTOTLE, in detail, in his May 22, 2012 (completion date) review. It was a non-inferiority trial of apixaban 5 mg bid (2.5 mg if age > 50, weight < 60 kg, or SCr > 1.5, vs warfarin titrated to an INR range of 2-3, conducted at 1053 sites in 40 countries in patients with non-valvular AF or Aflutter at enrollment or at least twice in the 12 months prior to enrollment. Patients were to have at least one risk factor (age > 75); prior stroke, TIA, or SE; CHF or LVEF < 40%, diabetes, or treated hypertension). It was a randomized double-blind event driven trial (target 446 adjudicated events) and patients were to be treated until the study ended. The NI margin (M_2) to be used for the primary endpoint (stroke plus non-CNS systemic embolism, adjudicated) was 1.38, accepted by FDA, representing ruling out a loss of > 50% of the effect of warfarin.

Two other anticoagulants have recently been approved for this use, dabigatran 150 mg and rivaroxaban. In both cases the NI-designed trials clearly showed effectiveness (dabigatran was superior to warfarin). One issue in such trials is how well the control (warfarin) was used (i.e., time in range for INR), an issue with rivaroxaban. For both drugs the clearest benefit was reduced hemorrhagic stroke, but dabigatran also showed a clear reduction in thromboembolic stroke.

The results in ARISTOTLE, as reported/presented, are favorable for time to first event.

ARISTOTLE Reported

		Apixaban n = 9120	Warfarin n = 9081			
	n (%)	events/ pt-yr	n (%)	events/pt-yr	HR	P
Primary endpoint	212 (2.3)	1.27	285 (2.9)	1.60	0.79	0.0114
Isch stroke	159 (1.7)		173 (1.9)			
Hemorrhagic	38 (0.4)		76 (0.8)			
Syst Embol	15 (0.2)		16 (0.2)			

As with the other anticoagulant drugs, decreased hemorrhagic stroke is a prominent effect.

Death was a secondary endpoint and a marginally significant reduction was reported (HR 0.89, p = 0.0465); CV death was similar and most of the benefit related to stroke. Apixaban also appeared to have lower rates of severe bleeding.

II. CR Determination

The above results are not discussed further because it would be premature given our major concern with the application. As described in detail in the clinical reviews and in Dr. Grant's CDTL review, the principal concern in the apixaban review has been what seemed initially to be a very high rate of medication errors (wrong active drug, placebo instead of active, active instead of placebo), a rate initially thought to be as great as 10%. It now is contended by the sponsor that these were mainly reading errors,

so that the bottle number placed in the CRF was often wrong, not reflecting what was actually given to the patient. BMS has reviewed stickers from the bottles, which were placed on the CRF (early part of study) or kept in the file (later) and believes the true error rate is far lower, perhaps 0.1-0.2%. The documentation of their analysis of a random 12% of patients (in response to an EMA request) has not been formally or fully submitted to us. The informal submission remains under review. Whether this 12% (plus another 8% analyzed previously using a convenience sample of available stickers) will settle the issue remains to be seen.

It is acknowledged that the apixaban results seem strong and we have not found significant lesions in the analysis. Nonetheless, I have concluded that approval would not be appropriate without strong assurance that we know what drugs the patients were given. We hope the CR will elicit a prompt and complete response so that evaluation of a promising drug can be completed without undue delay.

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

/s/

ROBERT TEMPLE
06/22/2012

Addendum 2 to Clinical Review for NDA 202155

Drug: Apixaban (Eliquis)
Sponsor: Bristol Myers Squibb
Indication: Prevention of stroke and systemic embolism in atrial fibrillation

Division: Division of Cardiovascular and Renal Products
Reviewers: Nhi Beasley and Martin Rose

Subject: Applicant's programmed edit checks, label analysis, impact of medication errors on endpoint events, data integrity
Date: June 22, 2012

The purpose of this addendum is to address information relating to medication errors submitted by the sponsor late in the review cycle that was not addressed in our final review: the applicant's programmed edit checks, the applicant's method of producing the container label dataset, the potential impact on endpoint events. The addendum also adds to the discussion of data integrity, an issue that was described in our review.

Programmed edit checks

While we refer to the applicant's less than diligent monitoring at the sites, we did not fully explain the rationale for our comment. This addendum provides our thoughts on the applicant's programmed validation checks (submission 48) that ran in the background of the eCRF; these edit checks ran after values were entered into the eCRF and saved in the database. Discrepant fields were highlighted in red for Site personnel to review.

There were ten programmed edit checks used for study medication.[1] Four (#2, 3, 4, and 5) of these checked that the container number dispensed or verified/returned were within the allowable range of container numbers (100001-875128). The problem with these checks is that the container number entered in these fields was not checked against the number assigned by the IVRS. Of the edit check discrepancy messages found in the CRFs, the most common one was for edit check #7 which compared the apixaban/placebo container number verified at a visit with the apixaban/placebo container number dispensed at the previous visit. The problem with check #7 is that the apixaban/placebo bottle was not always dispensed at the previous visit, yet this edit check message appeared. So while this discrepancy message appears many times for a subject, it was often meaningless. Check #9 compares the container number dispensed at a visit to all the container numbers dispensed at any subsequent visit to catch duplication. While this check was at the subject level only, had it also been done for the entire system it would have caught transcription errors. The reviewer did not find

[1] In the descriptions of the edit checks, "apixaban/placebo" and "warfarin/placebo" refer to blinded study medication bottles containing apixaban or apixaban placebo and warfarin or warfarin placebo, respectively.

many examples of discrepancy text for edit checks #1, 6, 8, 10 [2]; however three of these four only applied to apixaban/placebo. Thus, edit checks for apixaban/placebo were more rigorous than for warfarin/placebo. This may explain why the rate of discrepancy between the verified/return fields and IVRS database was considerably higher for warfarin/placebo than for apixaban/placebo. However, it does not prove that the correct bottle was actually dispensed in the case of such discrepancies.

Container label dataset

The applicant defined illegible label in an email communication to Dr. Grant on June 19, 2012. We disagree with their definition and believe that an illegible label should be a label that cannot be read (i.e., the container number is torn off the label, or the label is sufficiently smudged and cannot be read). Moreover, during FDA inspection of the labels at BMS, Ms. Blaus and Dr. Rose were able to read labels that the applicant deemed illegible. Labels that are difficult to read are more prone to errors; the applicant should consider all illegible labels a medication error. We note each bottle label tear-off panel also has a bar code; we were informed by the sponsor that this is a unique bar code. It might be possible use these bar codes to ascertain the bottle number when the numbers on the label are not legible.

Observed effects of medication errors on elinical events

The sponsor did several analyses of clinical events that occurred during periods that a patient took medication from an erroneously administered medication bottle of the wrong type; i.e., one that contained the wrong study medication (not including dispensing of the wrong dose of apixaban). In an analysis provided in the Sponsor's "White Paper" on medication errors that included events occurring while taking tablets from an error bottle and for 90 additional days afterwards, there were 4 primary efficacy or safety outcomes:

- One patient in the apixaban arm had a non-fatal ischemic stroke during a 9 day period of dual placebo therapy.
- One patient in the apixaban arm experienced a non-fatal major bleed during a 90day period of dual active therapy,
- One patient in the apixaban arm had a non-fatal major bleed within 90 days of a period of dual placebo therapy, and
- One patient in the warfarin arm experienced a nonfatal major bleed within 90 days of a 31 day period of dual active therapy.

  Reviewer Comment: the Sponsor did not specify how errors were detected and or how many errors occurred. We were orally informed that these errors were found by comparing the eCRF dispensing field with the IVRS data.

[2] These edit checks are summarized in the Appendix.

The third of these events, a major bleed during dual placebo therapy, is an unexpected finding. The other 3 events are consistent with expectations: patients taking two placebos would have in increased risk of ischemic stroke, while those taking two actives would have an increased risk of bleeding. In any event this small number of events is not informative. However, as explained in the main review, we are concerned that some errors in dispensing may not have been recognized because the site entered the IVRS assigned bottle number into the eCRF dispensing field instead of the bottle that was really dispensed to the patient. We think the bottle labels are a more accurate record of dispensing.

The sponsor did an analogous analysis in patient with errors identified by using the dispensing, verified, and returned fields. This produced a much larger number of errors. Tables provided below show rates of stroke/systemic embolism and death during the time the patient took tablets from the error bottle and for 90 days afterwards:

Table 1 Patients Who Received Bottles of the Wrong Medication Type: Rates of Primary Endpoints

	APIXABAN ARM	WARFARIN ARM
STROKE/SYSTEMIC EMBOLISM, n/N (%)	9/ 774 (1.16)	9/ 711 (
EVENT RATE (%/YR)	6.47	6.94
ALL-CAUSE DEATH, n/N (%)	10/ 783 (1.28)	11/ 717 (1.53)
EVENT RATE (%/YR)	7.08	8.31
ISTH MAJOR BLEEDING, n/N (%)	10/ 767 (1.30)	5/ 708 (0.71)
EVENT RATE (%/YR)	7.18	3.71

Table 2 Patients Who Received a Placebo Bottle of Wrong Type: Rates of Primary Endpoints

	APIXABAN ARM	WARFARIN ARM
STROKE/SYSTEMIC EMBOLISM, n/N (%)	3/ 133 (2.26)	5/ 623 (
EVENT RATE (%/YR)	6.31	5.504
ALL-CAUSE DEATH, n/N (%)	5/ 134 (3.73)	4/ 627 (4.37)
EVENT RATE (%/YR)	10.33	4.37
ISTH MAJOR BLEEDING, n/N (%)	4/ 132 (3.03)	3/ 619 (0.48)
EVENT RATE (%/YR)	8.34	3.20

Table 3 Patients Who Received an Active Bottle of the Wrong Type: Rates of Primary Endpoints

	APIXABAN ARM	WARFARIN ARM
STROKE/SYSTEMIC EMBOLISM, n/N (%)	6/ 655 (0. 92)	4/ 107 (
EVENT RATE (%/YR)	5.84	8.72
ALL-CAUSE DEATH, n/N (%)	5/ 663 (0.75)	7/ 109 (6.42)
EVENT RATE (%/YR)	10.33	14.59
ISTH MAJOR BLEEDING, n/N (%)	7/ 649 (1.08)	2/ 108 (1.85)
EVENT RATE (%/YR)	6.85	4.14.

Reviewer Comment: Tables with differing lengths of follow-up after the patient stopped taking medication from the error bottle were provided to us (0, 30, 60 and 90 days), but we are presenting the 90 day data here; it generally has the highest event rates and permits capture of events during the period that warfarin control is reestablished after an interruption. Table 1 includes patients who received any study medication bottle of the wrong type, i.e., one that contained a drug other the one they were supposed to receive, except that substitution of one dose of active apixaban for another was not counted as an error. Patients represented in Table 2 all received a placebo bottle in place of an active bottle. Thus, unless they experienced two errors at the same time, they took two placebos for some period of time instead of one active and one placebo, and would have been at higher risk of ischemic stroke. The patients represented in Table3 all received an active in place of placebo. Unless they experienced two errors at the same time, they would have taken two actives for some period of time, and would have been at increased risk of bleeding.

However, the expected pattern of events was not observed here, although there may have been too few events to make these tables interpretable. One would expect higher rates of stroke/SE in patients who received two placebo bottles (Table 2) than in those who received two actives (Table 3). However, the opposite was observed. One would expect higher rates of bleeding in Table 3 than in Table 2, but the results are fairly similar in the two tables. Note that most errors represented in these tables were errors in returned bottles. The sponsor argues that these were mostly transcription errors; to some unknown extent that might be true. We continue to believe that the best source of error information is the bottle labels.

Data integrity

This addendum cites another mismatch issue identified with the datasets and CRF. The data for one subject, ID 185030-0252-20280, shows that the subject continued treatment until the end of the study (Month 13), while the eCRF does not have data (including medication and disposition data) past Month 4. This error was identified rather easily in a dataset that contains over 1,000,000 lines of observations. One possibility is that the applicant left important pages off of the eCRF, but another possibility is that the data do not match the eCRF, an issue that was cited in the main review for an identified valid date.

Appendix – Edit Checks

Check#1 compared container number dispensed at Visit 1 with container number dispensed at subsequent visits to check for duplication of dispensed container.

Check #6 compared the apixaban/placebo container number verified to that verified at the previous visit.

Check#8 compares the apixaban/placebo container number dispensed on one date to all the apixaban/placebo container numbers verified at any subsequent visit. This checks that drug was dispensed and later verified.

Check #10 compares the apixaban/placebo container number verified at one visit to all dispensed container numbers at previous visits for a match.

Summary of Sponsor's Table of Definitions of Edit Checks:

#	Rule Message	Description
1	Container %Kit_ID% was dispensed at Visit C01 and Visit %VISIT_CODE%. Please reconcile responses or explain.	This validation as designed to look at the Container number dispensed at the C01 Visit for Apix/Placebo and Warfarin/Placebo and check if it is a duplicate of a container number dispensed at a subsequent Visit It does this by taking the KITID variable and comparing it to all KITID2 variables at subsequent visits to check for duplication of dispensed containers
2	Container Number %KIT_ID2% for %SMED_NAME% dispensed on %SMED_DISP_D% does not match the container # assigned by IVRS. Please verify the container #, enter correct Container Number and send this query back to DM review.	This validation is designed to check that the Container Number dispensed for Apix/Placebo and Warfarin/Placebo is within the allowable range of numbers available by IVRS (100001-875128) It does this by taking the KITID2 variable and comparing it to the range of IVRS numbers above to check for incorrectly recorded numbers

3	Container Number %KIT_ID1% for %SMED_NAME% verified* on %SMED_DISP_D% does not match the container # assigned by IVRS. Please verify the container #, enter correct Container Number and send this query back to DM review.	This validation is designed to check that the Container Number Verified for Apix/Placebo is within the allowable range of numbers available by IVRS (100001-875128) It does this by taking the KITID1 variable and comparing it to the range of numbers above to check for incorrectly recorded numbers
4	Container Number %KIT_ID3% for %SMED_NAME% RETURNED on :%SMED_RETURN_D% does not match the container # assigned by IVRS. Please verify the container #, enter correct Container Number and send this query back to DM review.	This validation is designed to check that the Container Number Returned for Warfarin/Placebo is within the allowable range of numbers available by IVRS (100001-875128) It does this by taking the KITID3 variable and comparing it to the range of numbers above to check for incorrectly recorded numbers
5	Container Number %KIT_ID% for %SMED_NAME% dispensed on %SMED_DISP_D% does not match the container # assigned by IVRS. Please verify the container #, enter correct Container Number and send this query back to DM review	This validation is designed to check the Container Number Dispensed for Apix/Placebo or Warfarin/Placebo at C01 Visit is within the allowable numbers available by IVRS (100001-875128) It does this by taking the KITID variable and comparing it to the range of numbers above to check for incorrectly recorded numbers
6	'Was study medication verified?' is answered YES, however the container number %KIT_ID1% verified on %SMED_D%does not match the container number %KIT_ID1% verified on %SMED_D% at previous visit. Please provide the correct container number.	This validation is designed to compare the container number verified for Apix/Placebo to the container number verified at the previous Visit as Apix/Placebo was only redispensed every 3 months. It does this by taking the KITID1 variable and comparing it to the KITID1 variable from the previous Visit
7	Was study medication verified? is answered YES, but container number %KIT_ID1% verified on %SMED_D% does not match the container number %KIT_ID2% dispensed on %SMED_DISP_D% at previous visit. Please provide the correct container number.	This validation is designed to compare the Container number verified for Apix/Placebo at this Visit with the container number for Apix/Placebo dispensed at the previous Visit. It does this by looking at the KITID1 variable and going to the previous visit and comparing it to the KITID2 variable

8	Container number %KIT_ID2% was dispensed on %SMED_DISP_D% however it does not match ANY container number verified %KIT_ID1% at ANY subsequent visits. Please record this container number as verified on the appropriate eCRF in TAO.	This validation is designed to take the Container number dispensed for Apix/Placebo on this date and check against all Apix/Placebo container numbers verified at any subsequent visit for a match It does this by taking the KITID2 variable and checking all KITID1 variables at subsequent visits for a match.
9	Container %Kit_ID2% was dispensed at Visit %VISIT_CODE% and Visit %VISIT_CODE%. Please reconcile responses or explain.	This validation is designed to check the Apix/Placebo or Warfarin /Placebo Container numbers dispensed at this Visit and compare to all container numbers dispensed to check for duplicates to ensure it wasn't dispensed again It does this by taking the KITID2 variable by VISIT and comparing it to the KITID2 variables at all other visits
10	Container number %KIT_ID1% was verified on %SMED_D% however it does not match ANY container number dispensed at previous visits. Please record this Container number as dispensed on the appropriate eCRF in TAO.	This validation is designed to check that the Container number Verified for Apix/Placebo at this Visit and compare it to all dispensed container numbers for Apix/Placebo at subsequent visits for a match It does this by taking the KITID1 variable on that date and comparing it to any KITID2 at previous visits for a match

---

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

/s/

MARTIN ROSE
06/22/2012

BACH N BEASLEY
06/22/2012

Cross-Discipline Team Leader Review/ Divisional Memo

Date	21 June 2012
From	Stephen M Grant Deputy Director Division of Cardiovascular and Renal Products
Subject	Cross-Discipline Team Leader Review/ Divisional Memo
NDA/BLA #	202155
Applicant Name	Bristol-Myers Squibb
Date of Submission	28 September 2011
PDUFA Goal Date	28 June 2012
Proprietary Name / Established (USAN) Name	ELIQUIS/ apixaban
Dosage Forms / Strength	Tablet: 5 mg and 2.5 mg
Proposed Indication	To reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation
Recommended	Complete Response

Material Reviewed/Consulted OND Action Package, including:	**Names of discipline primary reviewers**
CMC	Charles Jewell, William M. Adams, Yong Wang
Biopharmaceutics	Sandra Suarez Sharp
Pharmacology Toxicology	Patricia P. Harlow
Clinical Pharmacology	Ju-Ping Lai, Tzu-Yun McDowell
Clinical	Martin Rose, Nhi Beasley
Statistical	Steve Bai
OSI	Sharon Gershon
OPDP	Emily Baker, Zarna Patel
OSE/DMEPA	Morgan Walker, Ray Ford
OSE/DRISK	Danielle Smith
Project Management	Alison Blaus

OND=Office of New Drugs
OSI=Office of Scientific Investigations
OPDP=Office of Prescription Drug Promotion
OSE= Office of Surveillance and Epidemiology
DMEPA=Division of Medication Error Prevention and Analysis
DRISK=Division of Risk Management

1. Introduction

Bristol-Myers Squibb (BMS) has submitted NDA 202155 seeking authorization to market apixaban (proposed trade name ELIQUIS) for the following indication:

> "ELIQUIS is a factor Xa inhibitor indicated to reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation. (b) (4) "

To support the efficacy and safety of apixaban for this indication, the application relies on the results of two large international, randomized, active comparator-controlled trials: the principal trial **ARISTOTLE** ("A Phase 3, Active (Warfarin)-Controlled, Randomized, Double-blind, Parallel Arm Study to Evaluate Efficacy and Safety of Apixaban in Preventing Stroke and Systemic Embolism in Subjects with Non-valvular Atrial Fibrillation") and a supportive trial **AVERROES** ("Apixaban Versus Acetylsalicylic Acid to Prevent Stroke in Atrial Fibrillation Patients Who Have Failed or Are Unsuitable for Vitamin K Antagonist Treatment").

- In **ARISTOTLE** apixaban was compared to investigator adjusted vitamin K antagonist (VKA, mostly warfarin) administered double-dummy for reducing the incidence of a composite of stroke and non-CNS systemic embolism in patients with atrial fibrillation (AF) with at least one additional risk factor for stroke and without mitral stenosis. The applicant asserts that data from this trial demonstrate that apixaban was superior at a two-sided p-value = 0.011 with a hazard ratio (HR) of 0.79, primarily due to a reduction in the incidence of hemorrhagic stroke. In addition the applicant asserts that apixaban reduced the risk of major bleeding classified using the International Society on Thrombosis and Hemostasis (ISTH) criteria at a two-sided p-value < 0.0001 with a HR of 0.69 and also reduced the risk of all-cause mortality at a two-sided p-value = 0.047 with a HR of 0.89.
- In **AVERROES** apixaban was compared to aspirin for its effect on the incidence of the same composite of events in patients with AF with at least one additional risk factor for stroke and without mitral stenosis who were unwilling or unable to take a VKA. This trial was terminated early by its DMC for overwhelming efficacy. The applicant asserts that data from this trial demonstrates that apixaban was superior at a two-sided p-value <0.00001. The Agency concluded that AVERROES could not provide the principal support for an NDA because aspirin was an inadequate comparator (although it does support the effectiveness of apixaban for this condition). The availability of dabigatran (which was approved years after AVERROES was initiated) provided a drug demonstrated superior to warfarin for patients unable to take warfarin.

Clearly, a drug that is more effective and safer than warfarin for this indication and that also significantly reduced all-cause mortality in this population would be an attractive proposition, attractive enough that the application was assigned a priority review. And the presentation of the ARISTOTLE results at the European Society of Cardiology Congress in Paris on August 2011 engendered a great deal of enthusiasm among the clinical cardiology community, if the press is to be believed.

Although noting that the statistical persuasiveness of the mortality finding is not robust (p-value barely less than 0.05), the clinical reviewers agree that the data submitted support a finding that relative to warfarin, apixaban has significantly better efficacy with significantly less bleeding. Nonetheless, they recommend not approving the application at this time

because of their concern about errors in dispensing of study drug; at a frequency that is not yet clear, some subjects were given the wrong study drug (e.g., active instead of placebo and vice-versa). Sensitivity analyses that remove all subjects who may have been dispensed incorrect study drug do not undermine the conclusion that the benefit of apixaban in ARISTOTLE is greater than the risk and that apixaban was noninferior to warfarin in reducing the risk of stroke with significantly less bleeding. Nonetheless, the reviewers believe that knowing as well as possible the study drug each subject received throughout the course of the trial is important for approval.

The reviewers also found that monitoring of the trial at least with respect to dispensing of study drug was clearly inadequate. Although site monitoring did detect some instances of subjects receiving bottles of study drugs different from those assigned, errors in dispensing study drug continued during the conduct of the trial with no action taken to lessen their frequency. And finding error in something as important as proper dispensing of study drug raises a question about the quality of other aspects of trial monitoring.

There has been extensive internal discussion among Agency personnel about these issues and the signatory authority in the Office of Drug Evaluation 1 has indicated he is inclined to issue a complete response for this application. The focus of this memo therefore will be on the Division's understanding of the errors in dispensing study drug and the rationale for not approving this application at this time. A detailed analysis of the results and the significance of ARISTOTLE and AVERROES will not be included in this memo because the applicant will be asked to supply more reliable data about which study drug subjects actually received as well as about the monitoring of ARISOTLE. The assessment of trial outcomes will be addressed in a subsequent memo when the applicant re-submits the application.

2. Background

Stroke in Atrial Fibrillation

A stroke is a sudden neurological deficit of vascular origin lasting more than 24 hours or associated with infarction on brain imaging. Stroke remains a major source of morbidity and mortality in the USA. Each year about 795,000 Americans have a stroke, of who about 135,000 die, making stroke the third leading cause of death in the USA (*AHA Heart Disease and Stroke Statistics-2012 Update*). About 15% of all strokes are attributable to AF. The prevalence of AF is strongly associated with age; the prevalence in the general population is about 1% but in those over 80 the prevalence approaches 10%. In those over 80 years AF is the single leading cause of major stroke.

Atrial fibrillation results in uncoordinated atrial contraction with resulting stasis of blood, especially in the left atrial appendage. Stasis predisposes to thrombus formation; embolization of thrombus from the left atrial appendage into the cerebral circulation results in stroke. Strokes in patients with AF are often large and disabling.

The baseline risk of stroke in patients with AF can vary widely depending on the presence or absence of various concomitant conditions. In the absence of any of these conditions ('lone atrial fibrillation'), the risk is about 1% per year. Hence anticoagulants are generally prescribed only to patients with AF whose baseline risk for stroke is high enough to offset the risk of bleeding (major bleeding rate of about 2% per year). The risk for stroke in patients with

AF is commonly estimated using the $CHADS_2$ (**C**HF, **H**ypertension, **A**ge, **D**iabetes, **S**troke or TIA) stroke risk classification. In this classification scheme, the presence of CHF, hypertension, age ≥ 75, and diabetes are assigned a score of one and a history of stroke or TIA are assigned a score of two. The score is the sum of all component scores. The current *2011 ACCF/AHA/HRS Focused Updates Incorporated into ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation* recommend that all patients with a $CHADS_2$ score ≥ 2 should be treated with an anticoagulant.

Warfarin, a vitamin K antagonist, has been marketed for reduction of the risk of stroke in patients with AF for many years. The results of placebo-controlled trials of administering warfarin to patients with atrial fibrillation indicate that warfarin markedly reduces the relative risk of stroke, by about 2/3. Although very effective in reducing the risk of stroke, warfarin is difficult and burdensome to administer well. Its pharmacological action can be affected by a number of drug-drug, drug-disease, and food-drug interactions. As a result, its pharmacodynamic effect (measured as international normalized ratio or INR) is typically monitored at least monthly so that the dose can be adjusted if needed. Because of the difficulty in dosing warfarin and because it is an anticoagulant, major bleeding is common (2% per year) in patients with AF taking warfarin. Warfarin-related toxicity causes more than 43,000 emergency room visits each year making it the second most common drug (after insulin) implicated in emergency room visits in the United States.

In the last two years, two single anticoagulation factor inhibitors have been approved for marketing in the USA for reduction of the risk of stroke in patients with AF. Dabigatran, a direct thrombin inhibitor, was approved in October 2010 "to reduce the risk of stroke and systemic embolism in patients with non-valvular atrial fibrillation." The current PI indicates that a dose of 150 mg bid is superior to warfarin with a two-sided p-value = 0.0001 and a HR of 0.65 with significant reduction in the risk of ischemic as well as hemorrhagic stroke and with a similar rate of major bleeding. Rivaroxaban, a factor Xa inhibitor (like apixaban), was approved in November 2011 "to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation." The current PI indicates that a dose of 20 mg qPM is noninferior but not superior to warfarin with a similar rate of major bleeding.

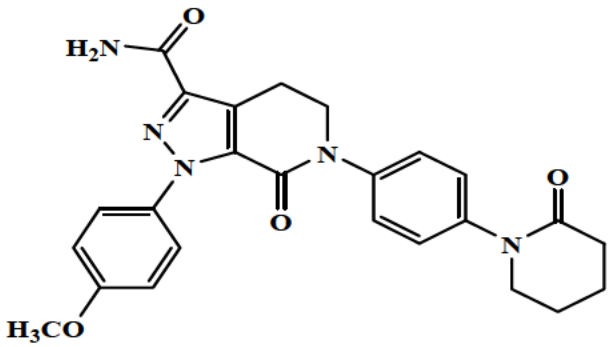
Although not included in the indication section of the PI for aspirin (21 CFR 343.80), the *2011 ACCF/AHA/HRS Focused Updates Incorporated into ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation* states "aspirin, 81 - 325 mg, is recommended as an alternative to vitamin K antagonists…in those with contraindications to oral anticoagulation." Meta-analyses have indicated that aspirin reduces the relative risk of stroke in patients with atrial fibrillation by about 20-25% per year; i.e., it is much less effective than warfarin for this indication.

3. CMC

The CMC reviewers concluded that the applicant's proposed manufacturing (and associated analytic methods) of the drug product and drug substance are acceptable. Manufacturing site inspections were acceptable. Stability testing supports an initial expiry of 36 months. There are no remaining outstanding CMC issues.

A. General product quality considerations

ELIQUIS (apixaban), the drug substance, is chemically described as 1-(4-methoxyphenyl)-7-oxo-6-[4-(2-oxopiperidin-1-yl)phenyl]-4,5,6,7-tetrahydro-1H-pyrazolo[3,4-c]pyridine-3-carboxamide, with a molecular weight of 459.5. Apixaban has the following structural formula:



Apixaban is a white to pale-yellow powder. At physiological pH (1.2-6.8), apixaban does not ionize; its aqueous solubility across the physiological pH range is ~0.04 mg/mL. The drug has no chiral centers.

Apixaban is synthesized as (b) (4) . However, they are present in the final drug substances at levels well below those that are acceptable. All identified impurities are similarly well controlled. The drug substance shows little if any degradation under long term and accelerated storage conditions, and is not sensitive to light.

The proposed to-be-marketed drug product is an immediate release, film coated tablet for oral administration in strengths of 2.5 mg and 5 mg. The excipients are not novel. The applicant proposal to supply the drug product packaged in HDPE and in blister packs is acceptable. Based on the available stability data an initial expiration date of 36 months is proposed for commercial apixaban tablets.

B. Facilities review/inspection

Facilities inspections have been completed and are acceptable.

4. Nonclinical Pharmacology/Toxicology

The pharmacology/toxicology review concludes that there are no outstanding pharm/tox issues that preclude approval. The applicant and the Division have reached agreement on labeling related to nonclinical studies in sections 8 (Use in Specific Populations) and 13 (Nonclinical Toxicology) of the PI.

A. General nonclinical pharmacology/toxicology considerations

In purified in vitro systems, apixaban inhibits FXa with high affinity and selectivity compared with related proteases involved in coagulation, fibrinolysis, and digestion. Pharmacology studies demonstrated that apixaban inhibits both venous and arterial thrombosis in a dose-dependent manner.

The doses administered in most nonclinical studies submitted to support the NDA were limited by bleeding due to the anticoagulant activity of apixaban. Nonetheless, standard repeat dose toxicology studies and safety pharmacology studies indicated that the safety margins for all toxicities except bleeding were acceptable. Humans have a significantly higher concentration of one metabolite (O-desmethyl apixaban sulfate) than in the usual nonclical species. Nonetheless, the total exposure in the dogs was high enough to conclude that this human metabolite is toxicologically qualified. Apixaban appears to be retained within the eye for a prolonged period after discontinuation of dosing but is not phototoxic *in vitro*.

There were two notable findings in the nonclinical studies:

1) Up to 12% of the maternal dose of apixaban is excreted in milk, posing a risk to nursing babies of unintended anticoagulation with the attendant risk of bleeding. This risk will be disclosed in the PI.

2) Juvenile studies indicate that apixaban may cause degeneration of the developing testicular seminiferous tubules. This finding appears reversible. This finding (b) (4) will need to be further explored prior to the conduct of any pediatric studies.

Apixaban did not demonstrate genotoxic potential in the Ames Assay, in vivo chromosome aberration assay (Chinese hamster ovary cells), or rat micronucleus assay.

B. Carcinogenicity

Acceptable carcinogenicity studies were performed in mice and rats. The Executive CAC reviewed these studies and concluded they did not demonstrate evidence of carcinogenicity.

C. Reproductive toxicology

Reproductive toxicology was evaluated in a series of studies in rats, mice, and rabbits. Apixaban did not affect mating or fertility and did not induce fetal toxicity or malformation. Although no maternal deaths occurred during parturition in the pre/postnatal development study, the incidence of bleeding was higher in apixaban treated rats.

5. Clinical Pharmacology

The clinical pharmacology review concludes that there are no outstanding clinical pharmacology issues that preclude approval. No agreement has been reached with applicant about final labeling.

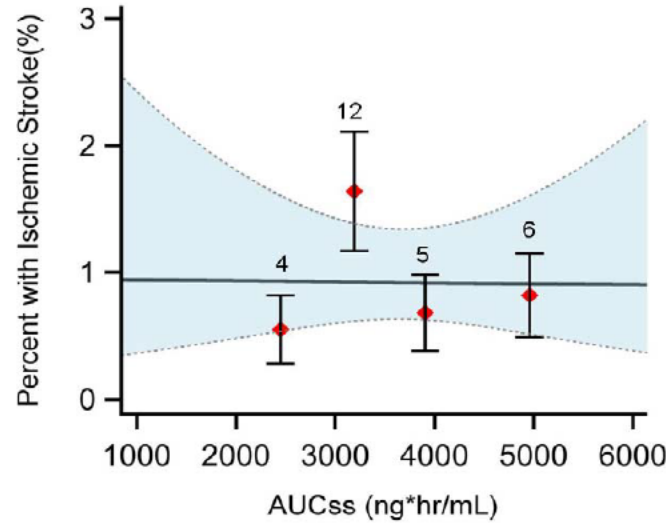
A. General clinical pharmacology

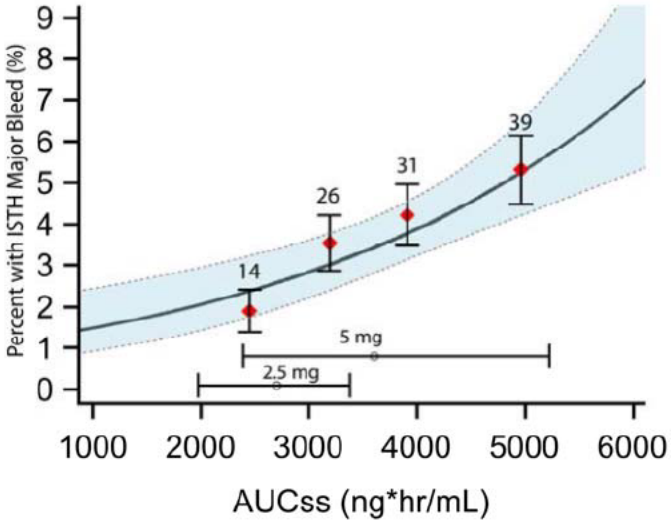
Apixaban is an orally active, selective inhibitor of the coagulation factor Xa (FXa) that directly and reversibly binds to the active site of FXa, and exerts anticoagulant and antithrombotic effects by diminishing the conversion of prothrombin to thrombin. Apixaban is a potent inhibitor of human FXa with a high degree of selectivity over other coagulation proteases and structurally-related enzymes involved in digestion and fibrinolysis.

Apixaban is absorbed with a T_{max} of three to four hours. Bioavailability is about 50%. Exposure is dose proportional for doses 10 mg or less (in the confirmatory trial ARISTOTLE most subjects were administered 5 mg bid) and is not significantly impacted by food. Apixaban is not extensively metabolized and no metabolites have anticoagulant activity. The half-life of apixaban is about twelve hours with an accumulation factor of < 2 after twice daily dosing. About 25% is renally excreted.

B. Dose considerations

The dose was selected based on studies in patients with DVT and the PK/PD properties of apixaban. PK data collected in the confirmatory trial ARISTOTLE suggest that increasing exposure did not decrease the risk of ischemic stroke but did increase the risk of bleeding, as shown in the following figures taken from the pharmacometric review:





The confidence intervals for the exposure ischemic stroke relationship are wide (the numbers above the confidence intervals for each quartile indicate the number of events observed) and so it is far from certain that there is no relationship between exposure and outcome. Nonetheless if there is any effect of these exposures on outcome, it is not large. So it appears likely that the dose administered in ARISTOTLE resulted in exposures above those needed to significantly

reduce the risk of ischemic stroke. It may be that a lower dose of apixaban would result in a similar reduction of the risk of CNS and non-CNS embolism with less risk of bleeding. Of course it should be noted that relative to warfarin, apixaban appears to have reduced the risk of bleeding.

In ARISTOTLE subjects with two or more of three factors (age ≥ 80 years, weight ≤ 60 kg, and serum creatinine ≥ 1.5 mg/dl) were administered a dose of 2.5 mg bid. This dose adjustment was made to minimize bleeding risk and was not based on PK considerations. Average exposure in the subjects administered 2.5 mg bid was about 25% less than those administered 5 mg bid.

C. Drug-drug interactions

About 20% of apixaban dose is metabolized by CYP3A4. Exposure about doubles when co-administered with a strong CYP3A4 inhibitor so the dose administered clinically should be halved if administration with a strong inhibitor such as ketoconazole is required. Conversely, exposure is halved when co-administered with a strong CYP3A4 inducer so they should be avoided.

D. Intrinsic factors potentially affecting elimination: age, gender, hepatic insufficiency and renal impairment.

Age, body weight, and gender as single factors have minimal impact on exposure. Similarly, mild and moderate renal failure also have minimal impact on exposure. Severe renal impairment increases exposure by about a third. Hepatic impairment has minimal impact on exposure.

E. Thorough QT study

A 'Thorough QT study' was performed and the results were evaluated by the FDA Interdisciplinary Review Team for QT studies. The study was judged to demonstrate that apixaban does not significantly prolong the QT interval.

6. Clinical Microbiology

N/A

7. Clinical/Statistical- Efficacy

A. Design and General Conduct of ARISTOTLE

ARISTOTLE was randomized, double-blind, double-dummy active-controlled non-inferiority study comparing administration of apixaban to warfarin in patients with non-valvular atrial fibrillation and at least one additional of the $CHADS_2$ risk factors for stroke. The primary endpoint was time to the first occurrence of stroke or non-CNS systemic embolism. The trial was initiated in December 2006 and the last subject visit occurred in May 2011. The final clinical study report indicates that 18,201 subjects recruited at over 1000 sites in 40 countries were randomized 1:1 to either apixaban (generally at a dose of 5 mg po bid) or investigator titrated warfarin. 477 endpoint events were analyzed in the primary endpoint analysis.

B. Errors in Dispensing Study Drugs

Page 88 of the clinical study report of ARISTOTLE contains the following statement:

> "The difference in the proportion of subjects with relevant or significant deviations is driven by error in treatment assignment where 7.3% of subjects in the apixaban group and 1.2% of subjects in the warfarin group received, at some point during the study, a container of the wrong type."

So the clinical study report states that apixaban subjects were six times more likely to be dispensed an incorrect study drug than were warfarin subjects and that a rather large number of all subjects were dispensed a study drug to which they were not randomized at some point during their participation in ARISTOTLE. While the discrepancy and rate of error in dispensing study medication apparently did not occasion any serious inquiry by the applicant prior to submission of the NDA, it was noticed by the clinical and statistical reviewers of this NDA. Inquiries from the Agency resulted in the applicant providing information that could be characterized as evolving over time but also could be characterized as contradictory.

These dispensing errors were the result of the complicated procedures employed by the applicant for dispensing study drug, which provided many avenues for human error (both in dispensing and recording what was dispensed). As should have been anticipated, everywhere human error could have occurred it did occur (cf. Murphy's Law). The following are the steps that had to be completed related to dispensing of study drug:

1. The investigator or his/her designee had to enter the correct identifying information for the subject in the interactive voice response system (IVRS).
2. They had to remember or record the IVRS assigned serial number of the bottle to be dispensed.
3. They had to either choose the correct bottle from wherever study drug was being stored (which varied from site to site) or depend on someone else to have the correct information to do so.
4. They had to tear off a label from the bottle and place it in the correct subject's paper CRF 800 (at least during the first half of the study. For reasons apparently not contemporaneously documented in any trial related document, the bottle stickers in the second half of the trial were to be retained but not necessarily on a paper CRF and were no longer intended to be collected by the applicant at the end of the study).
5. They had to give the bottle to the correct subject.
6. They had to enter the correct serial number on the eCRF (which might happen at a time remote from dispensing; initially the system used for recording information on the eCRF frequently was not available).
7. They subsequently had to enter the correct bottle serial number on the eCRF each time the subject came in for a visit.
8. They had to enter the correct bottle serial number on the eCRF when (and if) the subject returned the bottle.

In addition, there were two logs at each investigative site supplied by the sponsor/applicant: 1) an inventory log in which the serial number of each bottle shipped to the site and each bottle dispensed was recorded and 2) a log for each subject in which the serial number of each bottle dispensed to that particular subject was recorded.

Each subject was dispensed bottles of apixaban or apixaban placebo and separate bottles of warfarin or warfarin placebo. Each of the two study drugs were dispensed independently and therefore when a dispensing error resulted in an incorrect study drug being dispensed, usually only one of the two study drugs was dispensed incorrectly resulting in a subject being administered two active study drugs or two placebos until the next time study drug was dispensed. The dispensing of two placebo study drugs results in a subject not being treated with any anticoagulant putting them at unnecessary risk for stroke. The dispensing of two active study drugs exposes a subject to unnecessary risk of bleeding.

The applicant had four sources of data available for determining which study drugs were dispensed to each subject:

1. The serial numbers entered in the eCRFs.
2. The serial numbers recorded by the IVRS.
3. The labels torn from the bottles and retained after being dispensed.
4. Two drug accountability logs.

The applicant apparently used only the first source to come to the conclusion stated on page 88 of clinical study report of ARISTOTLE. And that conclusion reflected only subjects who appeared to receive active study drug instead of placebo (e.g., if assigned to apixaban but received active warfarin instead of the placebo for warfarin). In response to FDA requests the applicant examined the second source and finally, from the third source, an 8% sample of tear off labels (mostly from Russian sites) BMS had in-house (despite the decision in 2009 not to collect the paper CRFs to which the bottle stickers had been affixed). During this process the applicant disclosed to FDA that manual changes had been made to the IVRS system to change the serial number of the bottle assigned. According to the applicant these changes were made in response to information provided by investigators directly to the IVRS vendor, (b) (4) to ensure that a bottle dispensed in error would be removed from inventory so it could not be assigned. It should be noted here that manual changes to an IVRS should be made only under very limited circumstances and must always have a full audit trail indicating who, why, and when the changes were made.

In response to a request from EMA, the applicant eventually collected an additional 12% of tear off labels from sites selected at random in order to clarify the extent of errors in dispensing study drug. The summary report of the findings (but not the full report) was only recently submitted to FDA. This report appears to represent the current thinking of the applicant on this issue. The applicant asserts in this report that the actual number of bottles dispensed that contained an incorrect medication (not counting incorrect doses of apixaban) was likely 0.03% and the number of subjects who received an incorrect study drug at some time during the study was 0.65 % (and because there was 50% chance that a bottle dispensed with an incorrect serial number contained the incorrect study medication, it follows that 1.3% of subjects received a bottle with an incorrect serial number at some point during the study.) The summary report raises a number of issues, among them:

- A random sample of 12% may plausibly characterize the bottle error rate but it is not clear how the sample size was determined.
- The appropriateness of including the convenience sample 8% already collected is not satisfactorily explained.

- The analyses the applicant performed utilized legible labels; it is unclear how many illegible labels there were and how these missing data may have affected the conclusions. Intuitively, it seems that a bottle with a difficult to read label is more likely to be dispensed in error.
- The sponsor asserts that most of the apparent errors were actually due to incorrect recording in the eCRF of the serial number of the bottle dispensed. The errors were found mostly in the fields of the eCRF in which the serial numbers of the bottles returned were recorded. The applicant states this field was not subjected to as rigorous an edit check as the field in the eCRF in which serial numbers of dispensed bottles were recorded. The applicant does not provide an explanation of why transcription errors were so frequent at the time when the investigative site actually had the bottle and so could consult it to determine the serial number.

It should be noted again that the submission of even the summary report came toward the end of the the 3 month extension of the initial PDUFA goal date, far too late to be fully reviewed before the final PDUFA goal date.

The information provided to the Agency is not all the information available for determining the actual number of dispensing errors and the number of subjects who actually received an incorrect bottle of study drug (i.e., a bottle with the serial number other than the one assigned via the IVRS). As noted above, other sources for determining when and to whom study bottles were dispensed remains unexamined, most importantly the 80% of bottle tear-off labels that remain at investigative sites. Depending on assumptions made, more than 10% of subjects may have received at least one bottle of study drug different from the one assigned by the IVRS. And the number of bottles dispensed that contained incorrect study medication (regardless of whether the serial number on the bottle was the one actually assigned) may be much less than 0.1%.

With the currently available information the following can be confidently concluded:

- The percentage of subjects who either were dispensed a bottle with an incorrect serial number or who had at least one error on a CRF related to study drug dispensed is in excess of 10%.
- The number of bottles dispensed that contained incorrect study drug (defined as receiving a study drug of the wrong type or the wrong dose of study drug) is much less than 1.0%.
- Using any reasonable assumptions, the number of bottles dispensed that contained incorrect study drug is not large enough to overturn the finding that apixaban was noninferior to warfarin in ARISTOTLE.
- No conclusive evidence of harm to a subject as a result of an error in dispensing study drug has been found.

Regarding the last point, however, the failure to identify the problems with dispensing study drugs during the trial could have increased subjects' risks of stroke and bleeding. FDA could refuse to review a study submitted in support of an NDA if the safety of the subjects in the study was not adequately protected. So we need a clearer picture of exactly how dosing was monitored.

The reviewers identify an additional issue related to the errors in dispensing study drugs: why they were not detected during the course of the trial and corrected. The monitoring performed

by the applicant did identify many subjects in whom entries in the CRFs and IVRS were discrepant and so may have been dispensed a wrong bottle of study drug (the trial was blinded so study monitors should not have known if a study bottle dispensed with an incorrect serial number contained the assigned study drug). And there are instances of subjects being notified they had received incorrect bottles and returning them to the site. Nonetheless the applicant does not seem to have been aware during the conduct of ARISTOTLE that many subjects were dispensed an incorrect bottle of study drug. And the applicant has not provided any evidence that they took action to remedy the complicated procedures utilized to dispense study drug, which were the probable source of the errors.

8. Safety

No detailed discussion of the safety of administering apixaban to patients with AF at the doses used in ARISOTLE is made in this memo for the same reason efficacy is not assessed in the section above, *viz.* that the Agency is not yet persuaded that the data submitted in the NDA is the final data upon which a decision is to be made.

9. Advisory Committee Meeting

This application was not the subject for an advisory committee (AC). Two other oral anti-coagulants (dabigatran and rivaroxaban) have been approved for marketing for stroke prevention in patients with AF and so the issues presented by this NDA did not seem particularly novel or controversial at the time of filing. Trial conduct issues were identified late in the review cycle (the application was a priority review). The possibility of an advisory committee was then re-considered but it was concluded that the issues raised could not be usefully addressed by an AC with the information currently available.

10. Pediatrics

The Pediatric Review Committee met on 7 December 2011 and waived pediatric assessment requirements. Atrial fibrillation is rare in the pediatric population and so a clinical trial for this indication in children is not feasible.

11. Other Relevant Regulatory Issues

OSI Inspections/ Fraud in China

On 30 December BMS notified the Agency of GCP misconduct by one of its employees at site 1200 in Shanghai China. The original PDUFA goal date of 28 March 2012 was extended by three months to allow for adequate time for OSI to complete its assessment of this event.

Although BMS contracted with a Contract Research Organization, PPD, to provide site monitoring for ARISTOTLE, PPD did not have a presence in the People's Republic of China when the trial was initiated in PRC; BMS initially used its own employees for monitoring. One BMS employee along with at least one other individual altered subject records after being notified the site would be inspected by OSI. OSI inspected eight clinical sites worldwide after

becoming aware of this action. Additionally, after errors in dispensing study drug became an issue, BMS and PPD, a CRO involved in conducting and monitoring ARISTOTLE, were inspected specifically to review the issue of trial oversight and monitoring. OSI concludes that the study appears to have been conducted and monitored adequately. They did recommend that data from sites in China be excluded because the employee who committed the GCP violation in China was involved in the conduct of the trial at all Chinese sites.

Two observations in the OSI Clinical Inspection Summary dated 11 May 2012 are of note:

- Two instances were found of investigators unblinding subjects for reasons other than medical emergency. Investigators could unblind subjects by scratching off the stickers in their possession, i.e., without notifying the applicant. Because the applicant did not collect most of the stickers torn off the study drug bottles, it is unclear if all, or even most, instances of unblinding are known.
- One investigator in India knowingly maintained subjects' INRs in a range below that specified in the protocol. Apparently this practice was not identified and corrected by monitoring.

12. Labeling

A. General labeling comments

General agreement has been reached with the applicant on the portions of the label with information related to the chemistry, non clinical toxicology and clinical pharmacology of apixaban. The Agency has not discussed other portions of the label with the applicant.

If the NDA were to be approved at this time, then the Division's recommendation would be to indicate in the label that administration of apixaban reduces the risk of stroke at an acceptable rate of bleeding without including any discussion of specific trial results and p-values. The uncertainty that remains about which study drug some subjects were actually taking precludes reaching a conclusion with sufficient precision to state in the label.

B. Medication guide

If approved, a medication guide is recommended to inform patients

- Of the risk of bleeding and
- Not to discontinue apixaban without discussing the advisability of doing so with their health care providers because discontinuation will result in increased risk for stroke.

These recommendations are consistent with those contained in the medications guide for dabigatran and for rivaroxaban.

C. Proprietary name

Reviewers in the Division of Medication Error Prevention and Analysis evaluated the proprietary name proposed by the applicant, ELIQUIS. They found it does not pose a risk for confusion and is not unacceptably promotional.

13. Recommended Regulatory Action

The drug dispensing errors do not appear at this time frequent enough to undermine a finding that apixaban is safe and effective for reducing the risk of stroke and non-CNS systemic emboli in patients with nonvalvular atrial fibrillation. Nonetheless, it is a matter of concern that the actual study drugs given to subjects remain uncertain because the best source for this information, the study drug bottle tear off labels, have not been fully examined. If the Agency had been aware of the statement on page 88 of the clinical study report for ARISTOTLE prior to filing the application, we would have refused to file it; MAPP 6010.5 indicates that "substantive deficiencies … that appear to have been inadequately addressed in the application…" are a basis for refusal to file. And among the reasons to refuse to approve an NDA listed in 21CFR 314.125 is the lack of "explanation of the omission of other information…pertinent to the application."

The decision to be made then is a difficult and subjective one, balancing the utility of making apixaban available now and the undesirability of approving an NDA when there remains uncertainty as to what study drug subjects were given, an uncertainty that has not yet been adequately addressed. And there is hardly anything more fundamental to understanding the outcomes of a trial than knowledge of the study drugs actually administered. Clearly it would be unreasonable to require the applicant to try to determine with 100% accuracy what every subject in ARISTOTLE was taking at all times. But that is not the same as requesting the applicant to use all available sources for determining as well as possible which study drugs subjects were dispensed. It appears the applicant has collected a 12% random sample of the study drug bottle tear off labels for an EMA-requested analysis and that may be sufficient to place an acceptable upper bound on the rate of dispensing incorrect study drug.

Beyond these questions, the Division believes that there are two reasons the NDA should not be approved at this time:

1. Without more clarity about when subjects received either no active study drug or two active study drugs, it is not possible to query whether subjects suffered harm as a result of the applicant's inaction while ARISOTLE was being conducted.
2. The applicant has not yet provided the audit trail for the manual changes to the IVRS to demonstrate that all changes to the IVRS were appropriate.

Finally, the applicant does not appear to have provided the monitoring plan actually used during the conduct of ARISTOTLE. There was a preliminary monitoring plan at initiation of the trial. And the applicant submitted a final monitoring plan, but it is dated after data lock so it is not clear if it was the plan used while the trial was being conducted. The Agency should require submission of the monitoring plan actually used during the trial, i.e. documentation of all changes made to the initial monitoring plan during the trial for the following reasons:

- The sponsor's most favorable interpretation of the data indicates that more than 1% of subjects in the trial received a study drug bottle with a serial number other than the one assigned by the IVRS and site monitoring identified a number of these subjects. The failure to correct this problem could have led to errors in interpretation of the trial and could have exposed subjects to unnecessary harm. The Agency should review the plan to determine if it is likely that monitoring was inadequate in other important areas.

- The applicant has made many assertions about the frequency and kinds of edit checks that were made during the conduct of the trial. For example, the document responding to EMA's concerns about drug dispensing errors, the applicant states "programmed listings were also run to identify potential data errors, primarily in the dispensed fields of the eCRF…" Without the monitoring plan, the applicant's assertions can not be verified.

Acknowledgements

The Division's conclusions rest entirely upon the excellent review performed by this review team.

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

/s/

STEPHEN M GRANT
06/22/2012

Addendum to Clinical Review of NDA 202155

Drug: Apixaban (ELIQUIS)
Sponsor: BMS
Indication: Prevention of stroke and systemic embolism in patients with nonvalvular atrial fibrillation

Division: Division of Cardiovascular and Renal Products
Reviewers: Martin Rose

Subject: Comparative Disposition Data
Date: June 11, 2012

Reviewer's Conclusions

The rate of discontinuation of follow-up for the primary efficacy endpoint in ARISTOTLE was similar in the apixaban 5 mg bid and warfarin treatment arms, 4.8%, at the time the primary endpoint was assessed. This endpoint, the time to the composite of stroke and systemic embolism, was assessed at the end of the Intended Treatment Period (ITP), which was the expected date of attainment of the study's event target. The rate of early discontinuation of follow-up was slightly higher than the analogous rates in the RE-LY trial (4.1% and 4.2% for the dabigatran 150 mg bid and warfarin arms, respectively), but it was considerably lower than the analogous rates in the ROCKET trial (7.4% and 7.0% for rivaroxaban 20 mg od and warfarin, respectively). Both dabigatran and rivaroxaban were approved for the same indication sought for apixaban on the basis of the studies cited above. This reviewer believes that the rate of early discontinuation of follow-up in ARISTOTLE is not in itself problematic in interpreting the results of the trial.

Background

ARISTOTLE is the primary support for the approval of apixaban for its proposed indication, the reduction in risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation. It was a large (18,201 subjects), double-blind (double dummy), randomized, global, event driven trial comparing apixaban 5 mg bid (with a lower dose for patients with at least 2 of 3 pre-specified risk factors for bleeding) to warfarin titrated to a target INR of 2.0 to 3.0. The primary efficacy endpoint was non-inferiority of apixaban to warfarin for time to the composite of stroke or systemic embolism. This endpoint was met. In an additional analysis that was allowed under the study's pre-specified hierarchical analysis plan, apixaban was also superior to warfarin.

A question has arisen regarding how the data for early discontinuation of follow-up in ARISTOTLE compare to analogous data from studies in NDAs submitted to FDA since 2009 for other drugs for the same indication as apixaban. These studies include the warfarin-controlled trials of dabigatran (RE-LY) and rivaroxaban (ROCKET). Like

ARISTOTLE, these were large, global trials in patients with non-valvular atrial fibrillation and had the same endpoint as apixaban.

Assessment of early discontinuation of follow-up may be complicated by the fact that some patients discontinue follow-up early (for example, for withdrawal of consent), but nonetheless may have known vital status at the end of a study, typically obtained from a database of deaths or a report from a family member. It is not rare for sponsors to consider these patients as not having discontinued follow-up early. However, while their vital status is known, the investigator may not know whether the patient has had an event such a stroke (which was part of the primary endpoint in all these studies). Thus, the better practice is usually to request discontinuation data that identifies these patients as having early discontinuation of follow-up.

Such data was requested from the Applicant. Table 1 below is copied from the clinical review of the apixaban NDA, dated May 22, 2012. It contains discontinuation of follow-up data from ARISTOTLE that treats patients as lost to follow-up if they discontinue follow-up alive for the primary endpoint and other endpoints, without regard to knowledge of vital status at some time subsequent to the time that follow-up for other endpoints was discontinued. It indicates that 4.8% in each arm of ARISTOTLE had early discontinuation of follow-up.

Table 1 ARISTOTLE - Early Discontinuation of Follow-up
Before end of ITP, ITT Pop.

	APIXABAN N=9120 n (%)	**WARFARIN N=9081 n (%)**
COMPLETED ITP	**8105 (88.9)**	**8000 (88.1)**
DID NOT COMPLETE ITP	**1015 (11.1)**	**1081 (11.9)**
DEATH	575 (6.3)	643 (7.1)
DISCONTINUED ALIVE	**440 (4.8)**	**458 (4.8)**
WITHDREW CONSENT	260 (2.9)	259 (2.9)
LOST TO FOLLOW-UP	180 (2.0)	179 (2.0)

Similar data were requested and obtained for early discontinuation of follow-up in RE-LY and ROCKET. Summary data on early discontinuation of follow-up from ARISTOTLE and the two other studies are displayed in Table 2. RE-LY was a three-arm trial, but data for only the approved dabigatran 150 mg bid dose and the warfarin comparator are shown; data for the dabigatran 110 mg bid dose (not approved in the US) are not shown.

Table 2 Summary of Early Discontinuation of Follow-up Across SPAF Trials Submitted 2009-2011

Patients Not Followed to End of the "Intended Treatment Period" or Analogous Study Milestone

Trial (experimental drug)	Population	Experimental Drug n/N (%)	Warfarin n/N (%)
ARISTOTLE (apixaban 5 mg bid)	ITT	440 / 9120 (4.8)	458 / 9081 (4.8)
ROCKET (rivaroxaban 20 mg od)	Safety	525 / 7111 (7.38)	501 / 7125 (7.03)
RE-LY (dabigatran 150 mg bid)	ITT	251 / 6976 (4.1)	250 / 6022 (4.2)

SPAF: Stroke prevention in patients with atrial fibrillation
n=Number of subjects who had not died before the discontinuation of follow-up and who withdrew consent for follow-up, were lost to follow-up, or had some other reason for early discontinuation of follow-up prior to the end of the ITP or an analogous study milestone.

In RE-LY, the early discontinuation rate was 4.1% and 4.2% in the dabigatran 150 mg bid and warfarin arms, respectively, slightly lower than in ARISTOTLE. However, in ROCKET, the early discontinuation rate was 7.4% and 7.0%, respectively, in the rivaroxaban 20 mg od and warfarin arms, substantially higher than in ARISTOTLE.

Dabigatran and rivaroxaban were approved on the basis of the RE-LY and ROCKET trials, respectively, and the rate of early discontinuation in neither trial was considered to be an impediment to approval. The rate of early discontinuation in ARISTOTLE is intermediate to the rates of early discontinuation in RE-LY and ROCKET, but it is substantially closer to the low rate of RE-LY than to ROCKET. This reviewer does not consider the early discontinuation rate in ARISTOTLE to be an impediment to approval.

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

/s/

MARTIN ROSE
06/11/2012

CLINICAL REVIEW

Application Type	NDA Type 1
Application Number	202155
Priority or Standard	Priority
Submit Date	28 September 2011
Received Date	28 September 2011
Original PDUFA Goal Date	28 March 2012
Revised PDUFA Goal Date (after major amendment)	28 June 2012
Division / Office	DCaRP/ODE1/OND
Reviewer Names	Nhi Beasley (safety), Martin Rose (efficacy)
Review Completion Date	May 22, 2012
Established Name	Apixaban
Trade Name	Eliquis®
Therapeutic Class	Anticoagulant (Factor Xa inhibitor)
Applicant	BMS
Formulation	Oral tablets – 2.5 & 5 mg
Dosing Regimen	2.5 or 5 mg twice daily
Proposed Indication	"To reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation. (b) (4) "
Intended Population	Adults

Template Version: March 6, 2009

Note to Readers

In this review, a high level summary of the efficacy and safety data is found in Section **1.2**. Individual summaries of the efficacy and safety data are found at the beginning of Section **6** and Section **7**, respectively. Internal hyperlinks to other parts of the review are in **bolded blue font.** Entries in the Table of Contents (below), Table of Tables (p. **5**) and Table of Figures (p. **8**) are also hyperlinked to their targets.

Table of Contents

Table of Tables

Table of Figures

1 Recommendations/Risk Benefit Assessment

1.1 Recommendation on Regulatory Action

Based on our review of the clinical data, we recommend a complete response (CR). Reasons for this recommendation include:

1. This application rests entirely on the ARISTOTLE study as support for the efficacy of apixaban for its target indication, the prevention of stroke and systemic embolism in patients with non-valvular atrial fibrillation (AFib). In addition, ARISTOTLE supplies most of the NDA's safety data. We do not have sufficient confidence in the ARISTOTLE study data to approve the application.

   ARISTOTLE was a double dummy study in which subjects randomized to apixaban received active apixaban and placebo warfarin, and subjects randomized to warfarin received active warfarin and placebo apixaban. We discovered a substantial issue involving medication errors in ARISTOTLE, whereby subjects received either two active drugs or two placebos for varying amounts of time. The former would primarily increase the risk of bleeding, while the latter would primarily increase the risk of ischemic stroke and other thrombotic events. Some of the data indicate that the medication error rate, affecting up to about 8% of subjects in each arm, were driven by errors in the dispensing of warfarin/placebo (refers to active warfarin or placebo warfarin) bottles. If this is the case, one would expect that some warfarin arm subjects might receive placebo warfarin and placebo apixaban, and that some apixaban arm subjects might receive active apixaban and active warfarin. The potential consequence of these errors is a higher ischemic stroke rate in the warfarin arm, and a higher bleeding and or hemorrhagic stroke rate in the apixaban arm.

   We are still not sure of the number and type of errors that occurred, despite considerable effort on the parts of the Applicant and the review team. Similarly, we are not sure of how these errors affected study outcomes.

   The Applicant has told us that they were unaware of the scope of the medication errors during the trial or even when they submitted the NDA. This was due in part to deficiencies in centralized monitoring while data were accruing and less than diligent monitoring at the sites. Notably, there is no evidence that the sponsor initiated effective procedural changes to ameliorate the rate of medication errors, such as increasing the intensity of monitoring or the intensity of its centralized data checking procedures.

   Both the Applicant and the review team have modeled the data to account in various ways for assumed medication errors. Even at assumed error rates that represent presumptive worst case scenarios, the observed findings of superiority

of apixaban to warfarin for the primary efficacy endpoint (time to stroke or systemic embolism) and for the primary safety endpoint of ISTH major bleeding[a] probably would not be overturned, however, the Applicant's laxity in conducting ARISTOTLE makes us uneasy about other aspects of the study that remain to be elucidated. There may be substantial problems about which the review team and the Applicant are not aware.

We do not expect trials to be conducted perfectly. However, the trial infrastructure should include processes and staffing to detect and understand problems that threaten the integrity of the trial. Trial management should then use this information to implement effective corrective measures to ensure trial integrity. It is notable that another Sponsor of a large global trial of a cardiovascular drug recently discovered a medication error problem during their trial and responded by substantially increasing the intensity of study monitoring and by changing drug dispensing procedures. This was not done during the ARISTOTLE trial.

Apixaban is proposed to treat a life-threatening condition for which warfarin and two newer oral agents are already approved. The Applicant should convince us of the rigor of its study processes and the integrity of the ARISTOTLE study data before we approve apixaban as the fourth orally available drug for this indication.

2. If we were inclined to approve apixaban at this time, it would be difficult to describe the results of ARISTOTLE in labeling due to the unknown rate of medication errors and their effects on the outcomes of the study. Notably, the Applicant has the ability to substantially ameliorate the current uncertainty about the rate of medication errors by collecting the bottle labels still outstanding. These bottle labels, which bear a unique code that identifies the medication, were to be torn from each bottle prior to dispensing and then affixed to a page of the paper case report form (CRF 800) for each subject. The Applicant initially planned to collect all CRF 800s at the end of the study. However, over half way through the trial the Applicant decided to not collect these labels at the end of the study, although investigators were instructed to continue to collect the labels and maintain them at the study sites.[b] The Applicant informed us that they possess only 8% of the labels; the rest are presumably still at the study sites. In our view, these labels are the best available evidence of the medication dispensed to a subject.

   The clinical reviewers believe that in the case of controlled trials that are the

a ISTH =International Society on Thrombosis & Hemostasis
b The reason given to us (during interviews with BMS executives) for this change was that the project team determined that collecting the bottle labels was unnecessary because the IVRS system kept track of assigned bottles. The change in procedure was made in July 2009, and implemented through a series of emails to operations staff. The monitoring plan itself was not modified until July 2011, after the end of the data collection window for the ITT analysis.

basis of decisions regarding safety and efficacy, an Applicant should ordinarily be required to submit the best available information regarding the study medication administered to each subject. This is especially important when, as here, there is only one trial for the key comparison of safety and efficacy, or when the sponsor [redacted] (b) (4). Because we believe the bottle labels generally represent the best source of information regarding what drug was dispensed to any given subject, we believe that prior to approval, the Applicant should attempt to collect all of the outstanding bottle labels and thoroughly and accurately describe the medication error rate and its impact on important efficacy and safety parameters in ARISTOTLE.[c] We believe that collecting a fraction of the labels will not suffice since knowing the treatments subjects received is fundamental to the interpretation of a trial, and that here the medication errors could confound both the evaluation of efficacy and safety.[d]

Our concerns about the conduct of this study are sufficiently great such that we believe that (1) there is a lack of substantial evidence from adequate and well-controlled clinical investigations that apixaban has the effects its labeling purports it to have (21 CFR Sec. 314.125(b)(5)) and (2) there is insufficient information to determine whether the drug is safe for use as proposed in labeling (Sec. 314.125(b)(4)). In addition, we believe the proposed labeling is misleading in that it fails to describe and account for the problems in study conduct. (Sec. 314.125(b)(6)) Given the uncertainties of the effect of medication errors and trial conduct, we are unable to recommend useful, informative labeling. Thus, we are recommending a Complete Response.

1.1.1 Comments for Complete Response Letter

There are a number of actions that the Applicant might consider to restore the integrity in ARISTOTLE. The reviewers recommend the following for consideration in the CR letter:

"The trial database for ARISTOTLE suggests that there was an unacceptable rate of medication errors in the trial, affecting about 8% of patients overall. In response to our initial inquiries about medication errors, you provided us with new information that suggests that some of the data points that you initially indicated as representing medication errors may be spurious. However, the true rate of medication errors is

c We are not asking for perfect execution. The Applicant should make a diligent attempt to collect all the bottle labels. We expect that not all will be available for collection, and that some will be illegible. We also recognize that an individual bottle label in a patient's file may not accurately represent the medication actually received. However, in the aggregate, we think the bottle labels represent the best evidence of what was dispensed.

d In a non-inferiority trial, poor study execution would ordinarily tend to reduce differences between treatments, making it easier to win. While the nominal results of this trial support superiority of apixaban to warfarin, the trial was planned as a non-inferiority trial at a time when it was not known whether any new anticoagulant would be superior to warfarin in efficacy. Thus, poor execution might be expected to favor the Applicant's desired outcome.

not known. This is problematic because medication errors would be expected to affect the safety and efficacy findings of the trial.

We are also concerned that the findings in the study database did not lead you to investigate the apparent errors or induce you to take corrective actions to reduce the error rate during the trial. It appears to us that such an investigation did not occur until we started questioning you during the NDA review about medication errors. Likewise, you did not undertake any systemic corrective actions during the trial to reduce the error rate, such as increasing the intensity of monitoring or improving your centralized data checking procedures, which were substantially more rigorous for finding dispensing errors for apixaban than for warfarin.

We consider your lack of responsive action during the trial to represent a failure of trial monitoring and other quality assurance functions that are required to be in place pursuant to the International Conference on Harmonization (ICH) Guidance - E6 Good Clinical Practice. This failure of quality assurance leads us to be concerned that additional aspects of the trial may not have been executed well. There may be substantial problems in the conduct of the trial about which you and FDA are not aware. Our uncertainty about the conduct of the trial engenders concern about the validity of the results that you and your investigators have reported.

Before this application can be approved, you will need to provide us with reliable information regarding the rate of medication errors in ARISTOTLE. You will also need to convince us that the quality assurance measures that were in place in ARISTOTLE were adequate to detect other problems in trial conduct that might have occurred. Specifically, you will need to:

1. Collect all of the ARISTOTLE study medication bottle label panels (including all panels affixed to CRF 800 or in other locations). Using the actual label (not a scan or fax), double data enter and create an accurate dataset of all ARISTOTLE subjects' bottle numbers and the study medication contained in the dispensed bottles. Thoroughly and concisely report the medication errors in ARISTOTLE and its impact on efficacy and safety. Analyses of safety should include ISTH major bleeding, GUSTO severe bleeding, TIMI major bleeding, and adverse events. Analyses of efficacy should include the time to the primary endpoint (along with information on components of the primary endpoint) and time to all-cause death. Medication errors should be defined in two ways: as (1) any bottle number associated with a subject that was not randomly assigned to that subject by the IVRS (i.e., manually altered IVRS data would likely count as an error), and 2) any bottle number described in (1) that is also of a different treatment type than that in the bottle assigned to the subject. For definition (2) receiving the wrong apixaban dose should count as an error. You should use the following sources of information:
   - IVRS,
   - eCRF (both dispensed and returned/verified bottle fields),

- drug shipment,
- drug log, and
- any other information pertinent to understanding this issue.

We expect you to use the multiple sources of data on medication errors to create multiple versions of the requested information regarding errors and their impact on clinical outcomes.

We note that you may have to convert the handwritten drug log information into a dataset.

2. Some of your earlier responses to our questions about medication errors required subsequent correction multiple times, so you should use appropriate quality control measures and take the time that you need to provide a thorough, intelligent, and accurate resubmission.

3. Provide a detailed and thoughtful assessment of the adequacy of the monitoring, Quality Assurance, and Quality Control in ARISTOTLE to assure integrity of the study data and protection of human subjects.

4. Explain why monitoring failed to identify the medication errors that were discovered after we questioned the imbalance in medication errors.

5. Explain why no actions were taken to remedy the problem when the errors were discovered during the trial.

We recently learned that the (b) (4) IVRS data were manually altered to account for errors in dispensing. You disclosed this information to us after we repeatedly asked about obvious illogical entries in the IVRS dataset. It is our understanding that you did not know about the alterations in the IVRS dataset until we started probing. The IVRS alterations raise concerns about the integrity of the randomization dataset and your late understanding of the alterations increases our concerns about the quality assurance processes in ARISTOTLE.

6. Explain whether the manual manipulations of the IVRS dataset impacted the analysis of ARISTOTLE. Assure us that the randomization in ARISTOTLE was preserved.

7. Provide a thorough and complete explanation for all of the manual manipulations of the IVRS data. Include an IVRS dataset that flags all subjects with manual manipulations and the reason for the manual manipulation.

Reports from OSI indicate that unblinding may have occurred for reasons other than a medical emergency. From the materials you provided about your monitoring plan (which included a revision of the plan that was dated July 2011, which was after the

date of database lock, June 10, 2011) we were unable to determine whether the bottle label panels for all study subjects were to be checked for unblinding.

8. During your creation of the bottle number dataset, you should include a flag for subjects that were unblinded (indicated by the label being scratched off).

9. Using all available sources of information (e.g., bottle labels, monitoring notes) to identify subjects unblinded early, diligently search available records of unblinded patients for
   a) Unreported SAEs and
   b) Potential study endpoints that were not sent for adjudication

10. Provide a table with information on unblinding by treatment arm.

11. Perform sensitivity analyses to determine the potential impact of unblinding on the key time to event study analyses (i.e., primary endpoint, all-cause death, and ISTH major bleeding). These analyses should be of two general types: those that include any additional events found as result of the search described in paragraph (9), and those that exclude subjects whose treatment assignment was unblinded.

We are concerned that the trial datasets do not match the information in the CRF. In our review of four observations of data out of over one million observations in your medication error dataset (smed.xpt), used for most of your medication error analyses, we found an observation with a valid date in the CRF that was misrepresented by a period in the dataset. The period indicates that a valid date was missing. This error in the dataset has the possible effect of reducing the reported medication error rate since the observation was subsequently excluded from the count of medication errors. We can provide more details concerning this mismatch on request.

12. You should explain how this error and similar errors, if any, occurred and how you fixed them in the resubmission.

The identified date problem, found with little effort, is worrisome since your important analyses, such as the primary endpoint analysis, are time to event analyses. You should assure us that the datasets for these analyses are accurate and describe why you believe that they are. If your data cleaning processes were different for these datasets than they were for the medication error datasets, then you should apply similar processes to clean your medication error datasets.

13. Provide information to alleviate our concern that there may be other unidentified aspects of the datasets that do not match the CRF.

Some subjects have the same unique event listed multiple times as both non-serious and serious. This appears to be because the site personnel completed a non-serious

AE CRF and a serious adverse event (SAE) CRF for the same unique event. Monitoring did not appear to catch this.

14. You should clean the adverse events analysis dataset (adae.xpt).

In both ARISTOTLE and AVERROES there was an increased rate of stroke in subjects who completed the study in the apixaban arm. This raises the question of whether patients who discontinue apixaban after long-term treatment may be at increased risk for thrombotic events over their baseline risk.

15. You should provide a plan for evaluating the function of the coagulation system in patients who discontinue long-term treatment with apixaban."

1.2 Risk Benefit Assessment

Apixaban (ELIQUIS®) is an orally available, reversible, direct inhibitor of Factor Xa. The Applicant, Bristol Myers Squibb, has submitted NDA 202155 for apixaban on behalf of itself and Pfizer, its development partner, for the following proposed indication:

> " … *to reduce the risk of stroke, systemic embolism,* (b) (4) *in patients with nonvalvular atrial fibrillation.* (b) (4) "

Study conduct issues possibly affecting outcomes
In our view, resolution of the medication errors and study conduct issues in ARISTOTLE is essential prior to approval since they may have affected the efficacy and safety outcomes of ARISTOTLE. There were study design features that increased the likelihood of medication errors, and this was compounded by inadequately rigorous quality assurance measures. Moreover, during the trial the Applicant failed to take corrective measures after medication errors were identified. One could argue that identified errors were too few to justify a change in study design. However, the treatments subjects receive are fundamental to a study, especially in a non-inferiority trial and in a trial where substitution of placebo for active drug could increase the stroke rate or substitution of active drug for placebo could increase the bleeding rate.

The medication error issue is quite complex and ever evolving. For several months we were in a protracted loop of constant communication back and forth with the Applicant whereby we had more questions because the Applicant failed to anticipate additional issues that were apparent from their answers. There were several occasions when the Applicant informed us that their earlier responses were invalid or incorrect. We are still unsure of the scope of the medication errors and its effect on study outcome. However, we believe that the best evidence of the subjects' medication is the bottle labels (92% of which are still at the sites).

We recently learned that the (b) (4) IVRS data were manually altered to account for errors in dispensing. This raises concerns about the integrity of the randomization dataset. The Applicant only disclosed this information to us after we repeatedly asked about obvious illogical entries in the IVRS dataset. (There were two lines of observation that contained a treatment without an IVRS assigned bottle number.) It is our understanding that the Applicant did not know about the alterations in the IVRS dataset until we started probing. This increased our concerns about the quality assurance processes in ARISTOTLE.

While we do not expect trials to be conducted perfectly, we do expect Applicants to have a quality system in place to find errors and then to take corrective action. Our investigation into the medication error issue suggests that the quality assurance process did not work. We have a deeper concern that the quality assurance process in ARISTOTLE did not find other issues of which we and the Applicant are unaware.

For further discussion of these issues and their potential impact on the interpretation of the study results, see **Section 3.1.1.**

Efficacy Overview

The primary support for the proposed indication is the results of the warfarin-controlled ARISTOTLE trial. In addition, the Applicant conducted the aspirin-controlled AVERROES trial. Because warfarin has been convincingly demonstrated to be superior to antiplatelet agents (including aspirin alone or in combination with clopidogrel) in preventing thrombotic events in patients with nonvalvular AFib, and because aspirin is not indicated to prevent stroke in this population, the AVERROES trial should be considered merely supportive.

ARISTOTLE was a large (>18,000 subjects), randomized, double blind (double dummy), event-driven, warfarin-controlled, non-inferiority trial in adults with nonvalvular atrial fibrillation or atrial flutter (AFl) with at least one additional risk factor for thrombotic events. The dose of apixaban was 5 mg by mouth (po) twice daily (2.5 mg twice daily for those with pre-specified risk factors for bleeding); warfarin was to be titrated to a target INR range of 2.0 to 3.0.

The primary endpoint was time to a composite of stroke and systemic embolism (SE). The Applicant's designated primary endpoint analysis was for non-inferiority (with a margin of 1.38 for the hazard ratio (HR)) in the ITT (all randomized patient) population. This analysis included events occurring during the "intended treatment period" (ITP), which extended from each patient's date of randomization to January 30, 2011, the estimated date of attainment of the study's target of 448 primary endpoint events. This analysis will be referred to as the "ITT/ITP" analysis. As the initial analysis in a 4-step hierarchical analysis plan, it was used to evaluate (1) non-inferiority to and then (2) superiority to warfarin if non-inferiority was achieved. Steps 3 and 4 in the hierarchy are discussed below.

The ITT/ITP analysis of the primary endpoint yielded a hazard ratio (apixaban vs. warfarin) of 0.79, with a 95% CI of 0.66 to 0.95, p (superiority) = 0.0114, thus supporting both non-inferiority and superiority. However, if the standard for a superiority claim based on a single study is a $p \leq 0.01$, one could arguably deny a superiority claim. Additional analyses of the primary efficacy endpoint included ones of "evaluable" patients (i.e., the per-protocol population), counting events occurring during the on-treatment period (first dose of study drug to last dose + 2 days), and two other analyses based on the on-treatment period but with event windows extending to 7 or 30 days after the last dose of study drug. All these analyses of the primary endpoint results supported superiority of apixaban to warfarin, with p-values ranging from 0.011 in the worst case (the ITT/ITP analysis) to < 0.001 in the best case (for both the on-treatment and last dose + 7 days analyses in the per-protocol population). Additional analyses of the primary endpoint in all treated patients using various event windows were all supportive of non-inferiority to warfarin, but not all supported superiority.

These positive results in the ITT/ITP analysis were preserved across major subgroups of patients in the ITT/ITP analysis, including each gender, the elderly, subjects previously treated with a VKA, subjects with a prior history of stroke, TIA or systemic embolism, subjects in each of the 4 specified geographic regions, those who qualified for the lower dose, and those enrolled at US sites.

The primary endpoint findings were also supported by numerical imbalances for most important secondary efficacy endpoints that favored apixaban over warfarin in the ITT/ITP analysis. These endpoints included the rates of strokes (all types combined), hemorrhagic strokes, fatal strokes, systemic emboli, vascular deaths, and non-vascular deaths. The results for myocardial infarction also favored apixaban. The results for death also support the primary endpoint findings, and are discussed under a separate heading **below**.

There was a small imbalance of pure ischemic strokes in favor of warfarin in the ITT population during the ITP (140 vs. 136 events occurring during the ITP). However, when ischemic strokes with hemorrhagic conversion are also included, the results slightly favored apixaban (152 vs. 156), and the additional inclusion of strokes of uncertain type also favored apixaban (166 vs. 177 for the combined categories). There were few systemic emboli and the difference in rate between the treatment arms was small. Thus the primary endpoint results favoring apixaban were driven mostly by an excess of hemorrhagic strokes in the warfarin arm (40 vs. 78). While all strokes were counted as efficacy events, hemorrhagic stroke is a risk of anticoagulation, not something that is prevented by anticoagulation. Thus, if apixaban is superior to warfarin in terms of stroke, it is superior because it causes less hemorrhagic stroke than warfarin. There is a modest (non-significant) difference between apixaban and warfarin in terms of reducing the rate of ischemic stroke, the primary reason for giving anticoagulants to patients with atrial fibrillation.

In the pre-specified, 4-step hierarchical analysis plan, step 3 was a superiority analysis for time to ISTH major bleeding, the primary safety endpoint. This analysis was robustly successful and is discussed further in the safety summary. Step 4 in the hierarchy was an analysis of superiority of apixaban for time to all-cause death, conducted in a manner similar to that of the primary endpoint ITT/ITP analysis. This was just barely successful: the HR was 0.89, with a 95% CI of 0.80 to 1.00 (p=0.0465). Hazard ratios for CV death and non-CV death differed little from each other and from the all-cause death HR.

The following issues are relevant to the interpretation of the efficacy results of the trial:

Study medication errors and other trial execution issues in ARISTOTLE:

The clinical reviewers are concerned that study medication errors and deficiencies in monitoring and the data quality assurance process may have affected outcomes in ARISTOTLE. This complex issue is described in depth in Sections **3.1.1** and **3.2**.

(b) (4)

(b) (4), there is explicit language in Sec. 14 of the Applicant's proposed PI regarding superiority to warfarin.

The primary support (b) (4) is the just-barely significant reduction in time to all-cause death in the ITT/ITP analysis (HR=0.89, 95% CI, 0.80 to 1.00, p=0.0465). However, one additional death in the apixaban arm or one fewer death in the warfarin arm would negate the statistical significance of this finding. An analysis with a variable cutoff date that includes a 30 day follow-up period for treated patients after the last dose of study drug (rather than the ITT/ITP analysis, in which all completing patients (about 75% of the total) were still on treatment at the analysis cutoff date) was slightly less favorable for apixaban, with p = 0.08 and an upper limit of the HR of 1.01 (**Table 46**).

In addition, 590 patients (3.2% of those randomized) discontinued follow-up alive during the study and had no information on vital status at the cut-off date for the ITT/ITP analysis of death; they were censored on the date of their last contact prior to the cut-off date. As discussed in Sec. **3**, there were systemic blinding issues in this study that might have led to unblinding of individuals or even most patients at a site. If such unblinding occurred, ascertainment bias might have affected the vital status tracking of dropouts, thus potentially biasing the mortality results.

Also, the results of analyses based on site-specific INR control suggest that unlike for the primary endpoint results, better INR control was associated with less favorable results for apixaban for all-cause death. The Applicant's analyses show at sites above the median TTR, the HR for all cause death was 0.93, and sites in the top quartile of

TTR, it was 1.2 (see **6.1.10.1.3**). Thus, apixaban may not be superior to warfarin when time in therapeutic range is high.

In addition, the Sponsor notified us of fraud at site 1200 in China. The BMS manager responsible for this site (and all sites in China) altered patient records after she learned that DSI was inspecting site 1200. When the data from site 1200 are excluded from the analysis of death, which we think is an appropriate sensitivity analysis, statistical significance for death is lost. The HR for death becomes 0.90 (95% CI, 0.80-1.00, p=0.565). One other site in China was also inspected, but there was no evidence of fraud there. If all study sites in China are removed from the ITT/ITP analyses of the primary endpoint and all-cause mortality, the results of each analysis becomes more favorable for apixaban than the results of the analyses in which no sites are removed. For additional information on site 1200, see Sec. **3.1.2.**

Finally, we are still unsure of the number of medication (i.e. drug dispensing) errors that occurred in ARISTOTLE. Medication errors, coupled with other deficiencies in Good Clinical Practice, might have affected outcomes. One way to deal with medication errors in comparative analyses of safety and efficacy is to censor a subject at the time of the subject's first dispensing error. However, at this time, we are unsure who and when to censor due to medication errors. It is possible that censoring could change the hazard ratio in any given analysis, including the non-robust all-cause death analysis showing superiority of apixaban over warfarin. Also, is seems likely that the confidence intervals of the hazard ratios in key analyses will be widened due to loss of events and time of exposure. Thus, we are unsure of the final results of the trial, making it difficult to write descriptive labeling, even if we believe that apixaban is safe and effective for its target indication. For further information about this issue, see Sec. **3.1**.

(b) (4)

Efficacy events occurring after discontinuation of study drug in completers:

Approximately 3/4 of subjects in ARISTOTLE in each arm continued taking study drug until the end of this event-driven study. In these subjects, blinded study medication was stopped, and the investigator was to transition subjects to alternative anticoagulant therapy, usually a vitamin K antagonist such as warfarin. Investigators were urged to implement a short period of dual therapy with study drug and open-label warfarin (the latter was to continue long-term) for subjects in the apixaban arm to continue effective anticoagulation during the lag period of INR control at the start of warfarin therapy. In ARISTOTLE, this was done using blinded apixaban study drug, which was to be administered for four additional (twice-daily) doses. Blinding of study drug was maintained, so that warfarin arm subjects received apixaban placebo for their transition medicine. About 60% of completing subjects in each arm received this transition

regimen; about 84% received a VKA during the 30 days after the end of the double blind treatment period.

However, as in the ROCKET study of rivaroxaban, in the 30 days after the last dose of blinded study drug, there were significantly more primary endpoint events (mostly ischemic strokes) in the apixaban arm than in the warfarin arm, with a HR of about 4 (21 vs 5 events). Events were distributed throughout the 30 day period and not concentrated at the beginning of the transition period. In the apixaban arm, there were 3 hemorrhagic strokes, all in subjects who received open-label VKA treatment, and all occurring in the second half of the 30 day post-dose period. Although speculative, these 3 events could have been related to warfarin use. INR information during this period was not routinely collected, and was absent for the vast majority of subjects. However, less than 50% of subjects were receiving a VKA at the start of the study, suggesting that some investigators did not customarily manage warfarin in a manner consistent with practice guidelines used in the U.S. and the European Union (EU). Accordingly, it may be that INR control was suboptimal following the end of the study in many subjects.

There was an analogous finding (i.e., a significantly increased rate of stroke/SE in the 30 days after the last dose of study drug in the apixaban arm compared to the warfarin arm) in completing patients in AVERROES, the aspirin controlled study of apixaban in patients with AFib who had failed or were deemed unsuitable for VKA therapy. In that study, no completing patient was known to have received VKA treatment, so this explanation for the difference in stroke rate is not applicable.

The Applicant did not collect information on the concentration and function of clotting system constituents after cessation of long term treatment with apixaban, and nothing is known about the pharmacodynamics of warfarin that is initiated in this setting. Given the findings in ARISTOTLE and AVERROES, it would be desirable for the Applicant to collect information on this issue.

For more information regarding the rate of events after discontinuation of study drug in ARISTOTLE see Sec. **6.1.10.2.2**, and for information on the Applicant's proposed instructions for the transition from apixaban to warfarin see Sec. **5.3.1.5.4**.

Dosing regimen:

The Applicant evaluated one dosing regimen in its pivotal trial: 5 mg of apixaban bid for most patients, 2.5 mg bid for patients with at least 2 of 3 specified risk factors for bleeding. The Applicant established that this regimen overall was superior to warfarin for the primary efficacy and safety endpoint. Only 5% of subjects qualified for the 2.5 mg bid dose. Safety and efficacy results in this subgroup of subjects were slightly more favorable for apixaban than in the much larger subgroups that qualified for the 5 mg bid dose, so the dose reduction criteria seem reasonable. However, criteria for dose reduction other than those selected by the Applicant may also be reasonable. The

same regimen was used in AVERROES, an aspirin controlled trial of apixaban in patients with AFib who were not candidates for warfarin therapy.

This dosing regimen was based primarily on the results of a parallel-arm dose finding study in post-operative orthopedic surgery patients at risk for VTE. At the EOP2 meeting for the AFib program, we reviewed the relevant data and agreed to the use of this regimen in both ARISTOTLE and AVERROES prospectively. (Sec. **6.1.8**).

Adequacy of anticoagulation in the warfarin treatment arm and constancy assumption issues:

Although the primary efficacy analysis in this study involves non-inferiority to warfarin, satisfaction of the constancy assumption would not be a critical issue if we accept the nominal superiority of apixaban over warfarin in the prespecified primary endpoint analysis. Nonetheless, the constancy assumption was satisfied for ARISTOTLE with respect to the historical trials that established the effectiveness of warfarin for preventing stroke in patients with AFib. For additional information on the constancy assumptions and other aspects of warfarin use in ARISTOTLE, see Sec. **6.1.10.1**).

Quality of INR control in the warfarin arm in ARISTOTLE was in between that obtained in the ROCKET study of rivaroxaban (another study satisfying the constancy assumption with respect to use of warfarin) and the better results obtained in RE-LY, but was closer to RE-LY than to ROCKET in this regard. Thus, anticoagulation in ARISTOTLE was good enough so that the overall findings of non-inferiority or superiority to warfarin should not be rejected due to inadequate dosing in the control arm. However, for subjects at sites with TTR above the median, and in the best quartile of TTR, results for all-cause death suggest lack of superiority over warfarin, which might be a labeling issue (see text **above** and Sec. **6.1.10.1.3**).

Safety Overview
Most of the safety data for apixaban comes from ARISTOTLE (trial discussed in **Sec 5.3.1**). This was the largest randomized trial for the reduction in risk of stroke in AFib to date. As mentioned earlier, over 18,000 subjects with at least one risk factor for stroke were randomized to warfarin or apixaban in this non-inferiority trial; of the subjects randomized to apixaban, >95% of subjects received 5 mg po BID. Apixaban 2.5 mg po BID was for subjects with two out of three risk factors for bleeding at baseline. The mean duration of exposure, ~ 1.7 years, was adequate and similar to that of other large antithrombotic trials. The information from this trial alone is adequate to characterize the safety of apixaban. However, for specific events of interest (drug induced liver injury (DILI), rare serious neurologic adverse events, and concomitant antiplatelet therapy) the reviewer also analyzed the data in other trials.

The primary safety adjudicated endpoint was ISTH major bleed (i.e., bleeding leading to a 2 unit transfusion of packed red blood cells (PRBC) qualifies); a bleed that can be insignificant and readily reversible. However, the ISTH definition has historically been

used in patients receiving long-term anticoagulation and was the primary safety endpoint in the last two NDAs for this indication. Apixaban was consistently superior to warfarin for ISTH major bleed as well as for other serious bleeding (see **Figure 13**). While we do not like to compare across studies, the reviewer did examine across studies to look for consistency (albeit there were differences between RE-LY, ROCKET-AF and ARISTOTLE). Although the other antithrombotics were not superior to warfarin on ISTH major bleed, they each had a significant reduction in ICH, similar to apixaban.

The bleeding superiority findings in ARISTOTLE are robust (see **Table 81** Sensitivity Analyses of Superiority Findings for Bleeding Endpoints). To overturn the results, an additional 87 ISTH major bleeds in the apixaban arm are needed, or 93 fewer ISTH major bleeds in the warfarin arm are needed.

The site of major bleeds was mostly in the gastrointestinal (GI) tract (not a critical organ), followed by intracranial, then by intra-ocular (**Table 86**). Like dabigatran and rivaroxaban, the site of most major bleeds was the GI tract, however, in contrast to both antithrombotics, the annual rate of major GI bleed was lower on apixaban compared to warfarin (by 0.1%) and relative to warfarin, there was no difference in major GI bleed. With the exception of intracranial and intra-ocular bleeds, the site of major bleeds was similar between subjects on either treatment. There were numerically more intra-ocular major bleeds on apixaban (n=33) compared to warfarin (n=22). There are pre-clinical data to support that there may be a pharmacologic basis for this; apixaban radioactivity was still present in the eyes of rats at 168 hours post dose while it was last measureable in plasma at 24 hours. If apixaban is approved, the reviewer recommends that information on major bleeding at this critical site be collected and reported to the Agency in their Safety updates. Attempts should be made to collect adequate information on the subject's medical history to assess if a particular patient population may be at greater risk for major intra-ocular bleeding.

The reviewer conducted many analyses of all ARISTOTLE bleeding definitions and their relationship with TTR (time in range, time above range, time below range), a measure of warfarin control and safety. The overall mean TTR in ARISTOTLE was 62% (median 66%).[e] The many analyses did not show a consistent relationship between bleeding and TTR, suggesting that using TTR might not be a sensitive marker of individual warfarin control and ultimately bleeding events.

Analysis of countries by TTR quartiles shows that more than half of all US sites were above the median TTR of 66%; the same was true for Canada (the third highest enrolling country) (see **Figure 22**). Russia was the second highest enrolling country, but more than half of the subjects were in the lowest quartile of TTR, suggesting that warfarin control was poor in most subjects treated in Russia. The Ukraine, China, and India were the 9th, 10th and 11th top enrolling countries, but more than 75% of subjects in those countries were below the trial mean TTR of 62%.

e This is slightly worse than RE-LY, whose mean TTR was 64.4%.

Subgroup analyses of ISTH major bleed were generally consistent with the overall findings. A rather novel finding was that females on apixaban had less major bleeding than males and the relative benefit over warfarin was greater in females than in males. This finding is somewhat contrary to the results from a dedicated pharmacokinetic (PK) study that showed females have 15-18% higher concentrations than males, so one would not expect less bleeding in females based on PK. This reviewer is cautious about making conclusions based on un-prespecified subgroup analyses since findings can be spurious, and so would not recommend that this finding be highlighted in labeling. Subgroup findings are hypothesis generating. As age increased, so did the risk of major bleeds, but relative to warfarin there was less major bleeding on apixaban in subjects ≥ 75 years old. While there were 5 subgroups[f] on apixaban that appeared to lose their bleeding advantage over warfarin, for many of these subgroups the effect size was still close to the overall trial effect size **[ISTH major bleed apixaban vs. warfarin HR (95%CI): 0.69 (0.60, 0.80)]**, so the reason may be due in part to smaller number of subjects. A group worth noting where this might not be the case was in diabetics [HR (95%CI): 0.96 (0.74, 1.25)]. Many of the intra-ocular bleeds were in diabetics and if apixaban is approved, the Applicant should collect and analyze demographic and concomitant disease information in patients that have intra-ocular bleeding. Warfarin treated subjects that were VKA experienced did not have lower bleeding rates compared to VKA naïve subjects. Compared to the overall trial, the rates of major bleeding in the US were higher (+0.7%) and the relative benefit of apixaban over warfarin was slightly less, HR 0.75 (95%CI, 0.56, 1.00), nominal p-value=0.0497.

Apixaban 2.5 mg BID, given to subjects having at least 2 of 3 risk factors for bleeding (age ≥ 80 years, weight ≤ 60 kg or serum creatinine ≥1.5 mg/dL) at baseline, was safe and effective. This dose adjustment was not based on PK (exposure in these subjects treated with 2.5 mg was 25% lower than (not equivalent) subjects treated with 5 mg). Most subjects qualified by age and weight. Although <5% of subjects received the lower dose, apixaban 2.5 mg was superior to warfarin on major bleed and stroke/SE (see **Table 94**). Major bleeding was higher in subjects with at least 2 out of 3 risk factors (despite the lower apixaban concentrations) compared to those without 2 out of 3 risk factors, suggesting that bleeding in this subgroup was likely more due to the population rather than apixaban. The dose adjustment translated into a greater relative benefit over warfarin [(HR (95%CI) for apixaban 2.5 mg/warfarin, 0.50 (0.29, 0.86) versus apixaban 5 mg/warfarin, 0.71(0.61, 0.82)]. While the rate of stroke/se was also higher in this population, apixaban 2.5 mg had greater relative benefit over warfarin [(HR (95%CI) for apixaban 2.5 mg/warfarin, 0.40 (0.18, 0.88) versus apixaban 5 mg/warfarin, 0.80 (0.65, 0.98)]. If apixaban is approved, the lower dose should be approved and prescribed as used in ARISTOTLE.

f The subgroups were subjects < 65 years old, subjects with CHADS2 score ≥4, diabetics, subjects with previous stroke, and African Americans.

Of the other serious adverse events (SAEs), there was numerically more syncope on apixaban (n=77) than on warfarin (n=47). The reviewer cannot explain this, but it should be noted in labeling.

Generally, there were no differences in reasons for treatment discontinuation (by System Organ Class) between treatment arms. However in the trial as a whole, more subjects discontinued for an AE in the warfarin arm (8.4% vs. 7.6%, respectively). This was driven in part by a nearly three-fold higher number of warfarin-treated subjects with discontinuations for injury, poisoning, and procedural complications (63 (0.7%) vs. 22 (0.2%) subjects, respectively). The most common reason for treatment discontinuation was in the SOC of nervous system disorders (1.5% of subjects on apixaban vs. 1.7% on warfarin) which consisted mostly of stroke/TIA events, followed by gastrointestinal disorders. There were numerically more major bleeds after apixaban discontinuation (n=44) compared to after warfarin discontinuation (n=29). The excess bleeds did not appear to be from the inability to properly initiate warfarin treatment. Unfortunately data collection after drug discontinuation was too sparse to definitively determine the cause of the excess bleeds.

Apixaban does not appear to cause drug induced liver injury (DILI). There was one fatal case of hepatic failure that occurred on apixaban that independent, blinded, hepatologists judged as possibly related to apixaban or another drug (tianeptine). Otherwise, there were no probable cases and the number of potential Hy's Law cases was balanced between arms.

Apixaban does not appear to cause serious neurologic adverse events. In a Phase 2 (P2) dose ranging trial there was 1 case of Guillain-Barre Syndrome (GBS) and 1 case of amyotrophic lateral sclerosis (ALS) in subjects taking apixaban 5 mg QD and 10 mg QD, respectively. Following the two reports, the Applicant enhanced surveillance for neurological events, obtained neurological consultations for specific events, and instituted external, blinded, independent neurologist assessments of specific SAEs. There were a total of 6 GBS cases, 3 ALS cases and 7 cases of acute polyneuropathy identified in the NDA and Safety Update Report (SUR). All but one case were blindly reviewed; of these, all consensus assessments were "unlikely to be drug-related". One subject treated with apixaban did not have a consensus assessment as of late April 2012, but the three individual neurologists' assessments were 2 unlikely, 1 possible. Serious neurologic AEs occurred infrequently and were balanced between treatment arms in ARISTOTLE. Based on the totality of the data, the reviewer believes that apixaban does not cause serious neurologic AEs such as GBS, ALS or acute polyneuropathy.

The most common adverse event was bleeding. Minor bleeding and clinically relevant non-major bleeding was lower in the apixaban arm than in the warfarin arm (see **Table 101**). The reviewer was unable to complete the analysis of common adverse events because the sponsor's AE dataset contained errors that were likely created by an investigator filling out two or more CRFs for one unique event, a SAE CRF and a NSAE

(non-serious adverse event) CRF. Monitoring did not appear to catch this. There was no systematic pattern for how this error happened, nor was there a way to easily fix the dataset since the AE term was sometimes mapped to different higher MedDRA terms. The Applicant's analysis of common AEs indicates that the frequency of AEs were similar between apixaban and warfarin.

There were no significant laboratory findings. Thrombocytopenia was similar between treatment arms. There were no significant effects on vital signs or ECGs. The Thorough QT study was negative.

Renal elimination does not play a large role in the excretion of apixaban since its elimination is multimodal. However, it is known that subjects with renal impairment are inherently at risk for more adverse events, including bleeds and strokes. Event rates of both major bleeding and stroke/se increase in both treatment arms as level of renal impairment worsens (**Figure 26** and **Figure 27**). Relative to warfarin, apixaban has less major bleeding in subjects with mild-severe renal impairment. For stroke/se, there was no suggestion of worse outcome on apixaban relative to warfarin.

Apixaban is a substrate for CYP3A4 and the drug efflux transporter proteins, p-glycoprotein (P-gp) and breast cancer resistance protein (BCRP). The Clinical Pharmacology reviewers recommend, and the reviewer agrees, to avoid concomitant use with strong CYP3A4/P-gp inducers, and reduce the apixaban dose by half when coadministered with a strong CYP3A44/P-gp inhibitor (**Figure 6**).

Data from APPRAISE-2, an ACS trial where randomization was stratified by single or dual antiplatelet therapy shows that bleeding rates are 2-8 times greater on apixaban compared to placebo. The bleeding risk was ~2 times greater in subjects on dual antiplatelet treatment and apixaban compared to single antiplatelet treatment and apixaban. Although the population in APPRAISE-2 is different from those in ARISTOTLE, to put the bleeding in perspective, the rate of ISTH bleeding in APPRAISE-2 on apixaban plus single antiplatelet therapy is almost 3 times greater, and with dual antiplatelet therapy is 6 times greater, than in ARISTOTLE where apixaban was used alone. The rate of bleeding (TIMI major and ISTH major) on apixaban with single antiplatelet therapy in APPRAISE-2 was similar to the rate of bleeding on warfarin in ARISTOTLE. So in patients that need both an anticoagulant and an antiplatelet it may be reasonable to use apixaban with a single antiplatelet agent. Because of the risk of bleeding, the reviewer discourages the use of apixaban with dual antiplatelet therapy.

1.3 Recommendations for Postmarket Risk Evaluation and Mitigation Strategies

If the application is approved, a medication guide should be required. In addition to information about bleeding risk and other information for patients, it should warn against

the risk of stroke if patients with AFib stop treatment without consulting their physician, similar to the medication guide for rivaroxaban.

1.4 Recommendations for Postmarket Requirements and Commitments

If the application is approved the Applicant should be asked to characterize the effects of cessation of long-term treatment with apixaban on the components and function of the coagulation system. If the application is approved, this should be a PMR. It is possible that patients in the AVERROES long-term open label extension or some other ongoing study of long-term use of apixaban could provide blood samples for testing. Appropriate testing is to be determined.

In addition, the Applicant should continue to collect and aggregate the information on risk of intra-ocular bleeding.

2 Introduction and Regulatory Background

2.1 Product Information

Apixaban (Eliquis®) is an orally available direct inhibitor of activated Factor X (Factor Xa or FXa). Its proposed indication is the prevention of stroke and systemic embolism in patients with nonvalvular atrial fibrillation. (b) (4) The product is being developed through a joint collaboration between BMS and Pfizer.

The chemical structure of apixaban and its key attributes are provided in **Figure 1**, and additional product information is found in **Table 1**.

Figure 1 Chemical Structure of Apixaban

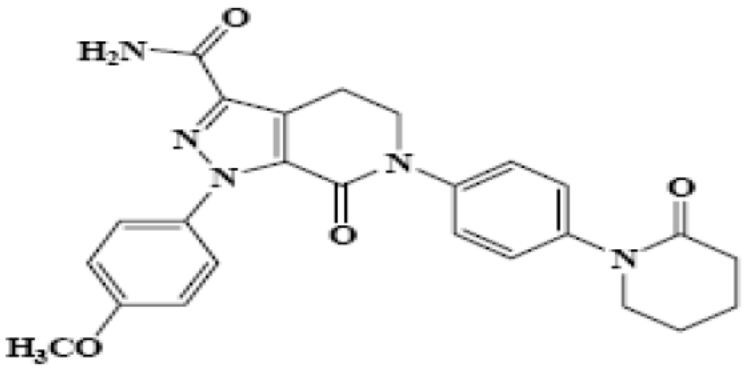


Table 1 Apixaban Product Information

Attribute	Description
Chemical Name	1-(4-Methoxyphenyl)-7-oxo-6-[4-(2-oxopiperidin-1-yl)phenyl]-4,5,6,7-tetrahydro-1Hpyrazolo[3,4-c]pyridine-3-carboxamide
Appearance	White to pale yellow, non hygroscopic crystalline powder
Molecular Formula	$C_{25}H_{25}N_5O_4$
Molecular Weight	459.50 daltons
Stereochemistry	Apixaban has no chiral centers, and thus has no stereoisomers.
Dosing Regimen	2.5 or 5 mg po bid
Proposed Age Group	Adults (a complete Pediatric Waiver has been requested)
Dosage Forms	2.5 and 5 mg film-coated oral tablets

2.2 Currently Available Treatments for Proposed Indication

2.2.1 Overview of Atrial Fibrillation and Stroke

Atrial fibrillation is the most common cardiac arrhythmia. It is estimated that 2.5 million Americans have AFib.[2] The rate of hospitalization for AFib has increased in recent years, possibly due to the aging of the population and an increased prevalence of chronic heart disease. AFib prevalence rises with age, and reaches about 8% after the age of 80, with a somewhat higher rate in men than women. The median age of AFib patients is about 75 years. [3]

The rate of ischemic stroke in AFib patients is ~5% year, 2 to 7 times the rate in persons without AFib.[3] Thirty-day stroke mortality in AFib patients has been estimated at 24% [2]. Non-cerebral embolic events also occur at an increased rate.

There is a body of literature on the risk factors for stroke in patients with AFib. Probably the most widely recognized risk factors are the 5 that are components of the $CHADS_2$ risk score: **C**ongestive heart failure, **H**ypertension, **A**ge > 75 years, **D**iabetes mellitus, and prior history of **S**troke or TIA. The last factor is worth 2 points in the score, and the other 4 are worth one point; the $CHADS_2$ score thus ranges from 0 to 6. More recently identified risk factors include female gender, age > 65 years, and history of vascular disease other than stroke. [4]

The most common source of emboli in AFib patients is believed to be the left atrial appendage. [3]

2.2.2 Currently Available Treatments

Warfarin (a pre-1962 product) is approved for and has long been used for the prevention of thromboembolic complications of non-valvular atrial fibrillation, the target indication of apixaban. There are 6 placebo-controlled studies of warfarin in patients with AFib. Taken together, they provide compelling evidence that warfarin is effective in reducing the risk of strokes in this setting, with FDA's meta-analysis showing a 64% reduction in the stroke rate.[g] A published meta-analysis suggests a more modest reduction in the rate of all-cause death using these same studies,[1] but this finding is not recognized in the warfarin label.

Aspirin provides only modest protection against stroke and is recommended as an alternative to warfarin in patients with low stroke risk or in patients in whom warfarin is contraindicated in the AHA/ACC/ESC consensus guidelines on the management of atrial fibrillation.[3] However, warfarin has been shown to be substantially superior to aspirin, especially in patients at higher risk of stroke, [1, 3] and this use of aspirin is not included in its label.

In October 2010, the Agency approved dabigatran, a Factor IIa inhibitor, based on RE-LY, a global, non-inferiority trial that compared two blinded doses of dabigatran to open label warfarin. Dabigatran 150 mg was superior to warfarin in reducing the risk of stroke/SE (primary endpoint) and no different from warfarin on ISTH major bleeding (**see Table 2**). Dabigatran 150 mg had a significant benefit for both hemorrhagic and ischemic strokes. Dabigatran 110 mg was non-inferior to warfarin on stroke/SE and superior to warfarin on ISTH major bleeding. Dabigatran 110 mg was significantly better than warfarin on hemorrhagic strokes, but not on ischemic strokes. Moreover, the absolute risk of ischemic stroke was higher on dabigatran 110 mg than it was on dabigatran 150 mg or warfarin. Relative to each other, dabigatran 150 mg caused more bleeding than dabigatran 110 mg, HR 1.16 (95%CI, 1.00, 1.34), but the higher dose was significantly better at reducing the risk of stroke/SE compared to the lower dose HR 0.72 (95%CI 0.58, 0.90).

Given that the threshold for an ISTH major bleed, the primary safety endpoint in RE-LY, was low (only a 2 g/dL reduction in Hg qualified), the prevention of a stroke was weighted heavier than the causation of a major bleed. Thus, the data favored approval of dabigatran 150 mg only. That said, the Agency searched for a population where the benefit outweighed the risk on dabigatran 110 mg and was unable to find one. The Agency approved dabigatran 150 mg po BID for patients with CrCl > 30 mL/min and 75 mg po BID for patients with CrCl 15-30 mL/min. The Agency did not approve the 110 mg dose.

g CDER/DBER Guidance for Industry - Non-Inferiority Clinical Trials (Draft) March 2010, accessed on 02/17/2012 at
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM202140.pdf (

Table 2 RE-LY - Overall results
ITT analysis, Followed to attainment of event target

Event	Dabigatran 150 mg N= 6076		Warfarin N=6022		Dabigatran 150 mg vs. Warfarin		
	(n)	%/yr	(n)	%/yr	HR	95% CI	p-value superiority
Stroke/SE*	134	1.1	202	1.7	0.65	(0.52, 0.81)	0.0001
Ischemic stroke	103	0.9	134	1.1	0.75	(0.58, 0.97)	0.030
Hemorrhagic stroke	12	0.1	45	0.4	0.26	(0.14, 0.49)	< 0.0001
ISTH major bleed*	399	3.3	421	3.6	0.93	(0.81, 1.07)	0.31[a]
Intracranial hemorrhage	38	0.3	90	0.8	0.41	(0.28, 0.60)	<0.0001[a]
Event	Dabigatran 110 mg N= 6015		Warfarin N=6022		Dabigatran 110 mg vs. Warfarin		
	(n)	%/yr	(n)	%/yr	HR	95% CI	p-value superiority
Stroke/SE*	171	1.4	202	1.7	0.90	(0.74, 1.10)	0.29
Ischemic stroke	152	1.3	134	1.1	1.13	(0.89, 1.42)	0.31
Hemorrhagic stroke	14	0.1	45	0.4	0.31	(0.17, 0.56)	0.0001
ISTH major bleed*	342	2.9	421	3.6	0.80	(0.70, 0.93)	0.0026[a]
Intracranial hemorrhage	27	0.2	90	0.8	0.30	(0.19, 0.46)	<0.0001[a]

a = p-value, *Primary endpoints, FDA non-inferiority margin=1.38
Yearly event rate (%/yr)= # subjects with event/subject-years*100
Subject years=sum(date of study termination-date of randomization+1) of all randomized subjects/365.25
Source: Dabigatran NDA 022512 clinical review by Drs. Beasley and Thompson.

In November 2011, the Agency approved rivaroxaban (Xarelto®) to reduce the risk of stroke and SE in adults with AFib, on the basis of the warfarin-controlled ROCKET AF trial. The recommended dose to be taken with the evening meal is 20 mg once daily for those with CrCL > 50 mL/min and 15 mg once daily for those with CrCl 15-50 mL/min. In the ITT population analysis to attainment of the event target, the hazard ratio of rivaroxaban vs warfarin was 0.88 (95% CI, 0.74, 1.03). Thus, rivaroxaban was non-inferior, but not superior to warfarin. The indication for rivaroxaban states that, "There are limited data on the relative effectiveness of XARELTO and warfarin in reducing the risk of stroke and systemic embolism when warfarin therapy is well-controlled…" Warfarin control, based on the time in therapeutic range (TTR), in ROCKET was substantially worse than in RE-LY (overall TTR of 56% vs. 64%) and in other recent warfarin-controlled AFib trials. There was substantial reduction in the rate of hemorrhagic stroke with rivaroxaban compared to warfarin, but the rates of ischemic stroke were similar.

In addition, rivaroxaban labeling includes a boxed warning pertaining to use in patients with AFib: "Discontinuing XARELTO places patients at an increased risk of thrombotic events. An increased rate of stroke was observed following XARELTO discontinuation in clinical trials in atrial fibrillation patients. If anticoagulation with XARELTO must be discontinued for a reason other than pathological bleeding, consider administering another anticoagulant…."." Another boxed warning, similar to one in the labeling for

dabigatran, describes risk of local bleeding in patients receiving neuraxial anesthesia or undergoing spinal puncture.

2.3 Availability of Proposed Active Ingredient in the United States

Apixaban is not approved for use in the US and (b) (4)

2.4 Important Issues with Consideration to Related Drugs

The most important safety risk of anticoagulant drugs is pathological bleeding. Anticoagulant agents affecting the intrinsic and/or extrinsic coagulation cascade may have their bleeding risks potentiated by anti-platelet co-therapies. For a discussion of this topic with apixaban, see Sec. **7**

Warfarin is associated with a myriad of drug-drug, and drug-food (and alcohol) interactions. Many mechanisms exist for these interactions, including metabolic inhibition of multiple cytochrome P450 isozymes (but primarily CYP2C9), drug reduction of the gut flora thereby reducing Vitamin K, and drug displacement of warfarin (a highly protein bound drug). Subjects may also metabolize warfarin at different speeds. Additionally, warfarin activity is determined partially by genetic factors; polymorphisms in the genes VKORC1 and CYP2C9 explain in part the difficulty in dosing warfarin to a therapeutic INR. The PK or PD interactions with warfarin may result in over- or under-anticoagulation and associated problems of bleeding or thrombosis, respectively.[5] While many of these mechanisms are specific to warfarin, those mechanisms that affect the clotting cascade could potentially affect apixaban.

Maintenance of target levels of anticoagulation in patients taking warfarin is highly variable across regions, individual study sites or practices, and patients. In the global RE-LY trial of dabigatran vs. warfarin, which supported approval of dabigatran for the apixaban proposed indication, an analysis of quartiles of site-specific levels of time in therapeutic range (TTR) of INR showed an inverse relationship between quartiles of TTR (with the 4th quartile having the highest TTR) and the rate of efficacy and safety events in the warfarin study arm. An inverse relationship between INR control and warfarin arm thrombotic event rate was also observed in the ROCKET AF trial of rivaroxaban, and additionally has been reported in the literature.[7]

Ximelagatran, an oral thrombin inhibitor, was associated with serious drug induced liver injury (DILI), and a possible increased risk of serious coronary events. It was not approved in the United States. Bleeding and DILI are discussed extensively in the review of safety in Section **7.**

2.5 Summary of Presubmission Regulatory Activity Related to Submission

The AFib IND for apixaban is IND 68,598. Key regulatory documents and decisions are described below. Issues raised that were resolved prior to the NDA filing or which refer to scenarios that are irrelevant to this review are omitted.

1. Pre-IND Meeting, 9/22/2004: The purpose of this meeting was to discuss study design and the general plan of an AFib program for apixaban, which was (b) (4) Key FDA recommendations included:

- Phase 2 DVT studies could be used for dose selection for AFib studies.
- It would be preferable to study more than one dose in Phase 3.

2. Additional pre-IND meeting, 9/21/2005: The Applicant had detailed questions on study design. Key FDA recommendations included:

- A program consisting of a double-blind superiority study of apixaban vs. aspirin (in subjects in whom VKA are contraindicated or who are unwilling to take them or whose INR cannot be controlled) and an open-label NI study of apixaban vs. warfarin could support approval for an AFib indication if both trials met their primary endpoints. Dr. Temple suggested that a trial in "lower-risk patients normally treated with aspirin" might also be done. Also, if the warfarin-controlled study failed to meet its endpoint, but apixaban was superior to aspirin in the other study, the latter finding could support approval.
- A single study, a double-blind warfarin controlled NI trial, could also support approval if FDA agreed to the NI and the analysis plan.
- VKAs other than warfarin could be used in a global trial, but the majority of control arm subjects should be on warfarin.
- In the study of apixaban vs. aspirin, the reason for not taking a VKA should be recorded for each patient.
- The proposed dose of aspirin (300-325 mg) was acceptable (note: this is not the dose of aspirin that was used in AVERROES, which was 81- 324 mg daily).

3. EOP2 meeting, 10/2/2006: This meeting dealt with additional details of the Phase 3 AFib program:

- The agency agreed with the proposed apixaban dose for Phase 3 of 5 mg BID, based on the results of the VTE prevention study CV185010.
- The NI margin in a warfarin controlled trial should be 1.38, not (b) (4) as proposed by the Applicant.
- A claim of "similarly effective" as warfarin would require a substantially smaller margin than 1.38.

- A proposed secondary endpoint consisting of the composite of stroke/SE/major bleeding was not appropriate if stroke and major bleeding had the same weight.
- Results in a warfarin-naïve subgroup might be clinically relevant.
- FDA agreed to the Applicant's proposal not to file expedited safety reports for major bleeding events unless they are fatal.
- Pediatric studies may be waived.
- The Applicant's planned aspirin controlled study was not necessary for approval.
- (b) (4) .

4. SPA response letter ("no agreement"), 12/11/2006 (for ARISTOTLE, protocol CV185030, the warfarin-controlled study in patients with AFib):

- The Division agreed that the study would provide the information necessary to make a regulatory decision regarding approval of apixaban for stroke and other thromboembolic complications of AFib.
- VKA naïve patients constitute a clinically meaningful subpopulation
- The division did not respond to two questions about two composite endpoints including efficacy events and bleeding because no details were provided about how events would be weighted.
- The Division recommended using a boundary of 0.001 or O'Brien-Fleming boundary for the interim analysis.
- The Division indicated that retaining "on the order of 90%" of the effect of warfarin in the NI analysis would be required to get a labeling claim of "as effective as warfarin".
- The Division did not object to the use of ISTH bleeding criteria in the study, but urged the Applicant to present results using other bleeding criteria. Any statements about superiority of major or other bleeding would be based on the strength and consistency of the findings.
- The primary endpoint should be limited to events collected on treatment or within a short period of time, e.g., 15-30 days after discontinuation.

5. Fast Track designation granted, 7/11/2007.

6. Advice Letter, 7/1/2008

- Dr. Stockbridge noted that it remains controversial as to whether a statistical plan conserves alpha error if one conducts a superiority test on the primary endpoint and then test an unrelated secondary endpoint.
- The Division expressed concern about the use of an endpoint that combined decreased thrombotic events with serious bleeding without weighting.

7. AVERROES Top Line Data Meeting, 1/24/2011

AVERROES is the aspirin controlled trial of apixaban. The Division had previously told the Applicant that this trial was not needed for approval, but it was conducted. The trial's Data Monitoring Committee (DMC) stopped AVERROES early for efficacy reasons. Note that this meeting occurred after the approval of dabigatran. Dabigatran was more effective than warfarin in a trial that dosed warfarin fairly well. FDA gave the following advice:

- Dr. Temple stated, "The Office's position, therefore, that an NDA for apixaban for treatment of atrial fibrillation, for patients unsuitable for warfarin because of a difficulty in reaching a suitable INR or high risk of bleeding, should not be submitted until the final results of ARISTOTLE are available."
- In summary, the Agency indicated that it believes that the primary registration study for apixaban should be ARISTOTLE, and that AVERROES is a supportive study. Like ARISTOTLE, the AVERROES study would be described in the label, but as part of the Clinical Studies section, section 14, and would be unlikely to result in an additional claim.

8. ARISTOTLE Top Line Data Meeting, 7/18/2011

FDA gave the following advice to the Applicant:

- It is unlikely that MI will be included in a table of efficacy outcomes because it was not part of the event hierarchy.
- Priority review is likely given the mortality data.

2.6 Other Relevant Background Information

2.6.1 Foreign Approvals

The European Commission approved apixaban May 2011 for use in 27 countries for the prevention of venous thromboembolic events (VTE) in adult patients after elective hip or knee replacement surgery (marketed June 2011). The recommended dose is 2.5 mg twice daily, starting 12-24 hours after surgery. The recommended duration of therapy is 32-38 days for hip replacement and 10-14 days for knee replacement. It may be taken with or without food. The Summary of Product Characteristics states that routine monitoring of exposure is not required, but the Rotachrom® anti-FXa assay may be useful in exceptional situations where knowledge of apixaban exposure may help to inform clinical decisions, e.g., overdose and emergency surgery."

Contraindications include:

- Hypersensitivity to the active substance or to any of the excipients.

- Clinically significant active bleeding.
- Hepatic disease associated with coagulopathy and clinically relevant bleeding risk

Warnings/precautions include:

- Risk of hemorrhage,
- Renal disease:
  - Lack of data in patients with CrCl < 15 mL/min – use not recommended
  - CrCl 15-29 mL/min: limited clinical data: plasma concentrations are increased, therefore use with caution alone or with aspirin
  - Mild-moderate impairment – no dose adjustment
- Use with caution in elderly patients also taking aspirin
- Hepatic impairment:
  - Not recommended in those with severe hepatic impairment
  - Use with caution in patients with mild or moderate hepatic impairment.
- Drug-drug interactions
  - Not recommended with strong inhibitors of both CYP3A4 and p-glycoprotein (P-gp)
  - 50% decrease of exposure with strong inducers of both CYP3A4 and P-gp
  - Care should be taken when used with NSAIDs, aspirin,
  - Not recommended with other antiplatelet agents or anticoagulants.
- Risk of hematoma with spinal/epidural anesthesia or puncture.

The SPC notes that "Clotting tests (e.g., PT, INR, and aPTT) are affected as expected by the mechanism of action of apixaban. Changes observed in these clotting tests at the expected therapeutic dose are small and subject to a high degree of variability."

The overdosage section recommends discontinuation of treatment, use of surgical hemostasis, and transfusion of fresh frozen plasma. Activated charcoal (up to 3 hours after administration) and recombinant FVIIa are suggested on the basis of pre-clinical data.

3 Ethics and Good Clinical Practices

3.1 Submission Quality and Integrity

Two main issues in the conduct of ARISTOTLE arose during our review that necessitated substantial time to dissect and comprehend: fraud in China at site 1200 and medication errors globally. For these reasons the Agency issued a major amendment letter on February 29, 2012, thereby extending the user fee goal date to June 28, 2012. Both issues are discussed extensively in the sections that follow.

While the reviewers believe that the fraud in China was isolated to site 1200, we cannot be certain. Nevertheless, the unsettling information on the medication errors was the driving force in our decision to issue a complete response.

3.1.1 Medication Errors in ARISTOTLE

3.1.1.1 Background and treatment dispensing

Medication errors are defined as dispensing the wrong type of study medication to a subject. In this study, of subjects randomized to apixaban, 5% (considered at high risk for bleeding), were to receive active apixaban 2.5 mg and placebo warfarin tablets; the rest were to receive active apixaban 5 mg and placebo warfarin. Subjects randomized to warfarin were to receive active warfarin 2 mg tablets (to be titrated to the target INR) and placebo apixaban.

Thus, medication errors could conceivably result in a patient having concomitantly:

- Two different active products (warfarin and apixaban)
- Two placebos
- The wrong apixaban tablet (2.5 mg instead of 5, or vice versa) and a placebo
- Rarely, two bottles of active warfarin or two bottles of active apixaban
- Rarely, the wrong active medication and a placebo

In the sponsor's analysis two scenarios were not counted as errors: 1) having the wrong apixaban dose, and 2) having a bottle with the wrong container number if the treatment was correct. This is because the type of medication was correct. Apixaban bottles (and their matching placebos) will be referred to as "apixaban/placebo". Warfarin (and their matching placebos) will be referred to as "warfarin/placebo". Apixaban/placebo bottles were noticeably larger than the warfarin/placebo bottles (insert FIG). Each apixaban/placebo strength was packaged in similar bottles, but the strength was clearly labeled on the bottle. Of course, placebo and active drug had similar labeling and bottles.

Figure 2 ARISTOTLE - Apixaban/placebo and warfarin/placebo bottles

From left to right, the bottles contain apixaban 2.5 mg, placebo apixaban 5 mg, apixaban 5 mg, placebo warfarin, and warfarin. The picture shows the label (and scratch off) that was to be affixed to CRF 800.

Perhaps adding to the complexity of this trial was that the days supply for apixaban/placebo and warfarin/placebo differed. Apixaban/placebo contained 200 tablets, a 100 day supply, whereas warfarin/placebo bottles contained 100 tablets of the 2 mg strength. Since warfarin was titrated to an INR between 2 and 3, the days supply of a bottle was variable. If the average dose was 6 mg/day (3 tablets), a bottle would last about 33 days, and the subject might need 3-4 bottles of warfarin/placebo prior to needing another bottle of apixaban/placebo. Thus, the consequence of receiving an erroneous apixaban bottle might be more lasting than the consequences of receiving an erroneous warfarin bottle, but one might expect more errors with warfarin/placebo bottles because more were dispensed.

Each time a bottle was dispensed to a patient, the following sequence of events was supposed to occur:

1. The site entered specified information into the telephonic IVRS system, including the patient number and the last 3 warfarin (active or placebo) doses in mg.
2. The IVRS system ((b) (4) replied in a computer-generated voice with the unique container number(s) for the bottle(s) to be dispensed to the subject. Each bottle of study drug had a unique container number.
3. The IVRS system also faxed or emailed the container number(s) to the site within 24 hours. The IVRS system retained an electronic record of the event.

4. Site staff procured the appropriate bottles from the **site**'s supply. Note that an open stock system was used; bottles were **not** stocked in "kits" for identified subjects. Instead, any bottle at a site conceivably could be dispensed to any subject.
5. Site staff removed the tear-off portion of the bottle label, which had the unique container number, and affixed it to CRF page 800 or another piece of paper kept at the site. From the start of the study in December 2006 until July 2009, it was intended that these sheets would be collected by site monitors and sent to BMS at the end of the study; copies were to be retained at the site. During this period, the labels were affixed to CRF page 800. In July of 2009 BMS determined that the labels were to be affixed to a piece of paper kept at the site (not a CRF page), and these source documents would not be sent to BMS, but were to be retained at the site.
6. The patient was given the bottle(s).
7. Site staff manually entered into the eCRF system the unique container number of each bottle dispensed.
8. When the patient returned with a bottle, the investigator manually entered the unique container number into the eCRF system on the appropriate page.

Some of the opportunities for errors to occur in this process are highlighted:

- The wrong subject ID could be entered into the IVRS system.
- The wrong bottle could be pulled from the shelf and given to a subject. Note that each bottle had a unique bar code, but it was not used.
- Site staff could have entered into the eCRF a number not matching the container number on the subject's bottle. This could occur in a "dispensed" field as well as in "returned" and "verified" fields (when the subject returned to clinic). Site staff could have used the IVRS fax to enter the container number and not the actual bottle (which could have been mistakenly dispensed). If this method of entering the container number was used and the subject had the wrong container number, then the eCRF would match the data in the IVRS and indicate that the subject had the correct bottle, but the subject would actually have the wrong bottle.
- The tear off-label or paper on which it was affixed might be lost or displaced, or the label could be affixed to the wrong subject record.

3.1.1.2 Medication Error Data

On January 23, 2012 the reviewers asked the Applicant to explain the 6:1 imbalance in medication errors found in the ARISTOTLE study report because it seemed too large to be random. There were 664 (7.3%) and 109 (1.2%) of subjects in the apixaban and warfarin arms, respectively with medication errors. The Applicant responded by email, stating that they

"….don't have an explanation for this difference… but given the study results and the sensitivity analyses performed our assessment was that this did not impact the results of the study….We would like to note that we used the Evaluable Population to perform sensitivity analysis… and the number of randomized subjects excluded from this population was similar for both treatment group [sic] (602 subjects excluded from the apixaban group and 606 excluded from the warfarin group). The Evaluable population excluded from the randomized population those subjects who… received a container of the incorrect treatment type BEFORE having a primary efficacy event (i.e., ...if the subject received an incorrect container type prior to having an event, then subject would be excluded from the Evaluable population)"

Reviewer comment: This reply started a cycle of questions between the Agency and the Applicant whereby the Applicant's attempts to address our questions only left more questions. This dialogue continued until late April when we continued to have more questions and discovered that the dataset the Applicant was using for the majority of their analyses did not match the information in the CRF. Although we now know considerably more about the medication error issue than we did in January, it remains a moving target and we still have questions.

The Applicant indicated that the reported numbers were for subjects that received the wrong active medication; since more warfarin/placebo bottles were dispensed, then there were more chances for error with the warfarin/placebo bottles (as the study report indicates). Thus, there would be less active treatment errors in subjects randomized to warfarin because these subjects had fewer occasions to receive apixaban/placebo treatment (since this bottle should only need replacing after 100 days).

We asked the Applicant to redo its analysis of errors, and count either erroneous dispensing of placebo or active drug as an error. We also asked for data on a per bottle basis, rather than a per patient basis, and data on what was actually supplied in the case of each error.

1. The Applicant provided the information we requested. Counting erroneous dispensing and returns of active or placebo, there were 8.6% and 7.9% of patients in the apixaban and warfarin arms, respectively, that had medication errors. Additional information on errors is provided in **Table 3**.

Table 3 Summary of Containers Dispensed or Returned of the Incorrect Type, by Treatment Group - Treated Subjects

APIXABAN ARM (active apixaban, placebo warfarin)	**Subjects Treated (S=9088) ***	**Bottles dispensed (B=224,271 **)**
Active-Warfarin Dispensed in Error, n (%)	664 (7.3)	723 (0.32)
Placebo-Apixaban Dispensed in Error, n (%)	134 (1.5)	136 (0.06)
Total Errors, n (%)	**784 (8.6) a**	**859 (0.38)**
WARFARIN ARM (active warfarin, placebo apixaban)	**Subjects Treated (S=9052)**	**Bottles dispensed (B=211,911)**
Placebo-Warfarin Dispensed in Error, n (%)	629 (6.9)	684 (0.32)
Active-Apixaban Dispensed in Error, n (%)	109 (1.2)	111 (0.05)
Total Errors, n (%)	**719 (7.9) a**	**795 (0.38)**

* The denominator to calculate each percentage is the number of subjects treated (S) in each randomized treatment group
** The denominator to calculate each percentage in the number of bottles dispensed (B) in each randomized treatment group
Errors represent mismatches between the IVRS record of bottles (i.e., unique bottle numbers) to be dispensed and either (1) the eCRF record of bottles dispensed or (2) the eCRF record of bottles returned.

Note that the per-bottle error rate was similar in the two arms, at 0.38%. Notably, there were about 6 times as many errors involving erroneous dispensing of warfarin active/placebo as apixaban active/placebo. However, the total of warfarin active/placebo bottles dispensed was only 2.3 times the number of apixaban bottles dispensed, meaning that the error rate for warfarin active/placebo bottles was more than twice as high as for apixaban active/placebo bottles. This discrepancy was not initially explained.

The medication errors represented in **Table 3** were discovered by comparing the IVRS derived database of medications to be dispensed with the eCRF-derived database of bottles dispensed or bottles returned (including bottles “verified”, i.e., brought back the site for counting of tablets but then taken home again by the patient). This comparison would not initially catch an error resulting from dispensing an erroneous bottle if the site entered the IVRS generated-bottle number into the eCRF instead of the number of the bottle actually dispensed in error, although the error might be picked up when the bottle was returned. .

Additional data were provided with respect to what was dispensed in cases of known dispensing errors. The most common substitution for active apixaban was placebo apixaban, and vice versa. The most common substitution for active warfarin was placebo warfarin, and vice versa. However, substitutions of active for active, one type of placebo for another, the opposite active for placebo, and one apixaban strength for another occurred. Thus, some patients were dispensed two actives of different types, some two placebos, some two actives of the same type, and some the wrong active or the wrong dose (of apixaban) with placebo.

The Applicant performed analyses of key endpoint (stroke/SE, all-cause death, major bleeding) in which patients were censored at the time of their first medication error. The result of these analyses were each slightly more favorable for apixaban than the ITT analyses of the same endpoints.

Because the data provided did not eliminate the possibility that a site simply entered into the CRF the number of the bottle they were supposed to dispense rather than the one they actually did dispense, we asked the Applicant to review the CRF 800s that they have in house (both original and scanned/faxed copies from the site) and compare the container numbers to the IVRS and the eCRF data. We also asked for:

- analyses of TTR as a metric of the effect of medication errors on control of INR
- a sensitivity analysis of the data, taking into account the error rate of medication errors already observed and how many more errors would have to be observed, , to overturn the non-inferiority and the superiority findings.
- a description of monitoring activities with respect to medication errors and corrective actions taken when errors were found, both at the site level and systemically.
- data for the event rate data for ISTH major bleed, stroke/se, and all cause death by the treatments the patient was actually receiving at the time of the event, 30 days, 60 days and 90 days after the incorrect treatment was received, along with data sets and the SAS codes used for the analyses.

2. The Applicant responded in part to the above request on 2/21/2012.

    a. Comparison of IVRS data to bottle labels: The Applicant attempted to enter data on dispensed bottles using scanned copies of the bottle labels in house, but there were substantial legibility issues with the scanned copies. We also had difficulties reading the bottle number from scanned copies of the labels. Accordingly, the Applicant used only legible original bottle labels to create a database of bottles dispensed, and matched it to the IVRS database. The Applicant had a total of 35,859 legible original

bottle labels in- house, about 8.2% of the total of 436,182 bottles dispensed during the study. (There were 187 additional labels that were not legible). Data for mismatches between the bottle labels and the IVRS record of bottles to be dispensed are displayed in **Table 4.** A mismatch occurred when a bottle label number associated with a patient was not one of the bottles numbers assigned by the IVRS system to the same patient. The table also includes for purposes of comparison the data on mismatches between the clinical database and the IVRS system from **Table 3**. Note that substitution of one active dose of apixaban for another (i.e., a 2.5 mg active tablet substituted for a 5 mg active tablet, or vice versa) were not counted as errors by the Applicant.

Table 4 Summary of Containers Dispensed of the Incorrect Type, by Treatment Group – Treated Subjects

APIXABAN ARM (active apixaban, placebo warfarin)	**Clinical Database (B=224,271) ***	**Original Container Labels (L= 18,001) ****
Active-Warfarin Dispensed in Error, n (%)	723 (0.32)	14 (0.08)
Placebo-Apixaban Dispensed in Error, n (%)	136 (0.06)	4 (0.02)
Total Errors, n (%)	**859 (0.38)**	**18 (0.10)**
WARFARIN ARM (active warfarin, placebo apixaban)	**Clinical Database (B=211,911) ***	**Original Container Labels (L= 17,858) ****
Placebo-Warfarin Dispensed in Error, n (%)	684 (0.32)	17 (0.10)
Active-Apixaban Dispensed in Error, n (%)	111 (0.05)	9 (0.05)
Total Errors, n (%)	**795 (0.38)**	**26 (0.15)**

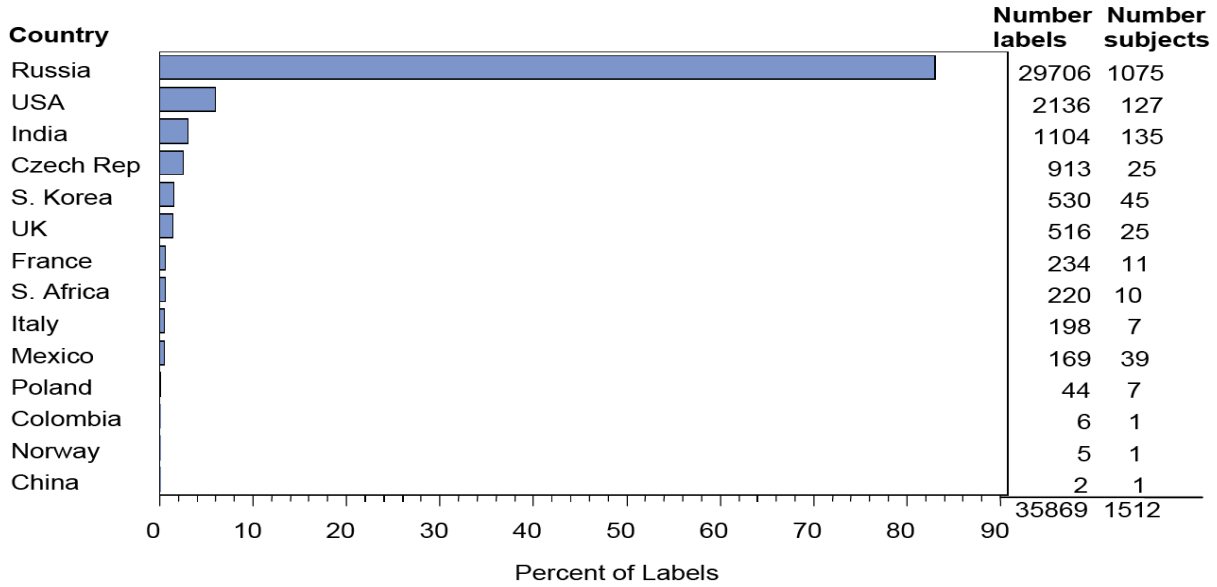
* The denominator to calculate each percentage in the number of bottles dispensed (B) in each randomized treatment group
** **The denominator to calculate each percentage in the number of original labels (L) available in each randomized treatment group. Note: The total in the 2 arms, 35, 859, is the same as the number of "legible" bottle labels in house provided on p. 4 of the submission with the data in the table above. Errors represent mismatches between the IVRS record of bottles (i.e., unique bottle numbers) to be dispensed and either (1) the eCRF record of bottles dispensed or (2) the eCRF record of bottles returned.

The data using the bottle labels to identify dispensed drug suggest a lower per-bottle error rate than was calculated from the comparison using the clinical database as the source of what was actually dispensed. The error rate using the bottle labels was 0.1% and 0.15% in the apixaban and warfarin arms, respectively.

Reviewer Comment: The initial monitoring plan provides for collection of the CRF pages with the bottle labels by BMS. We understand that this was to be done at the end of the study. However, after an amendment to the monitoring plan in July 2009, at which point more than 12,000 study patients had been enrolled and treated, the monitoring plan was amended so that the original CRF pages with the bottle labels would be kept at the site.

Nonetheless, the Applicant argues that the 8% of original labels collected are an unbiased (but certainly not random) sample of the total. The Applicant provided data indicating that in patients with collected bottle labels, information about error rates based on the clinical database was similar to the data for the study population as a whole, suggesting that the bottle label sample is representative. However, as indicated in **Figure 3***, 29,706 (83%) of the 35,859 collected bottle labels, from 1075 (71%) of the 1512 patients with bottle labels possessed by the Applicant, were from sites in Russia, which enrolled only 9.9% of the ITT patient population. This suggests that the population with collected bottle labels may not be representative of the study as whole.*

Figure 3 Number of Legible Bottle Labels by Country



Reviewer's analysis:\med_error\crf800_lbl030, sponsor's data: lbl030

b. The Applicant suggests that some the errors found by comparing the trial database bottle dispensing and return data with the IVRS database amounts to mere "transcription errors," not errors in dispensing. Their evidence for this comes from the fact that among bottles that did not match IVRS assigned bottles, a high percentage of the bottle numbers of bottles returned and a lower percentage of the bottles dispensed do not

correspond to bottles that were sent to the relevant site (determined from the drug shipment database). For example, in the apixaban arm, a total of 859 medication errors were identified by comparing the dispensing and return data to the IVRS. Of these 712 (83%) involved the identification of mismatched bottles that were entered into the study database as being dispensed or returned that should not have been at the site if the shipping records are correct. In warfarin arm patients, the analogous data are 645 of 795 mismatched bottles (82.3%) that should have been shipped to other sites. Most of the bottles that should not have been at the site were identified as bottles that were returned (but never dispensed). The Applicant suggests that these represent transcription errors. Another interpretation is that the shipping records are suspect, and that the bottle dispensed was not correctly recorded – perhaps the investigator simply entered the number given to him or her by the IVRS.

In addition, The Applicant compared the bottle label data to the clinical trial database information on medication errors. Although the clinical trial database suggested that the per-bottle error rate was more than 3X the per-bottle error rate derived from the bottle label database, only eight of the 44 errors (18%) found in the bottle-label database (see **Table 5**) were also found to be errors in the clinical trial database, meaning that 36 errors (82%) were not picked up in the clinical trial database. Note that the monitors were supposed to compare the bottle labels to the information in the clinical trial database. The Applicant suggests that the reason that some of these errors were not picked up by the monitors is that they do not represent dispensing errors, but mere errors of record-keeping at the site: the bottles were correctly dispensed, but the bottle labels were placed in the wrong patient file. The evidence for this contention is the fact for some bottle label errors, the bottle represented by the label was recorded as being dispensed to a different patient at the same site in the study database and the medication log kept at the site.

Reviewer Comment: This last claim by the Applicant is difficult to evaluate. If site staff simply used their notes from the IVRS call or the IVRS fax or email as the source for the information in the study database and medication log, those entries would match the IVRS database, but the bottle could have been dispensed to someone else at the site. Regardless of the identity of the patient that actually received the bottle in question, the discrepancy between the bottle label data and the study database or medication log might have been picked up by good monitoring and possibly resolved.

Table 5 Comparison of Bottle Label Information to Clinical Trial Database

	Apixaban Arm (18,001 labels)	Warfarin Arm (17,858 labels)	Total (35,859 labels)
Errors found bottle label database	**18**	**26**	**44**
Errors above also found in clinical trial database, n (%)	4 (22)	4 (15)	8 (18)
Errors above not found in clinical trial database, n (%)	14 (78)	22 (85)	36 (82)

Denominator for percentages is n in first data row.

Reviewer Comment: The Applicant may well be right that some of the errors in the clinical database may represent only transcription errors, and the patient received the correct bottle (or another bottle of the same type), but the number of a bottle of a different type was entered into the database. The fact that many of the errors involve bottles that were recorded as being sent to other sites supports that notion that these errors may simply be errors in data entry.

A very different picture results from analysis of the bottle label data. Of the more than 35,000 bottles for which the sponsor has bottle labels, our analysis indicates that every bottle was sent to the site where it was supposedly dispensed.

This reviewer draws the following inferences from these data:

- *The clinical database information on bottles dispensed and returned is probably not reliable,*
- *study monitoring did not appear to have found inconsistencies between the bottle labels and the eCRF (the source of information in the clinical database) and did not lead to effective preventive measures such as modifying the error-prone dispensing system adopted by the Applicant,*
- *and, most important, the best evidence of what was actually dispensed to the patients probably is the bottle label information. Only 8% of the labels were collected by the Applicant; the remainder is presumably still at the investigational sites.*

c. The Applicant argues that the good TTR observed in ARISTOTLE (an overall mean of 62% and median of 66% indicates that warfarin control was not affected much by medication errors. However, TTR is not a sensitive measure of the effect of sporadic medication errors. The most important medication error that would affect TTR in warfarin arm patients would be substitution of placebo warfarin for active warfarin. If one bottle were substituted, the patient might receive placebo for the days he or she took tablets from the bottle, which might be less than one month. By the time of the next INR visit, the patient might have resumed taking active warfarin, and perhaps returned to near his or her usual INR level. Even if the patient was still under-treated, the amount of imputed time out of range would be some fraction of the number of days in between the previous in-range value and the next in-range value in range. To gauge the effect of a placebo-for-active warfarin medication error, we assumed that a patient was in the study for the warfarin arm median of 88 weeks with 65% of days in range (very near the trial median). This patient would have 57.2 weeks in range. If a substitution of placebo for active warfarin reduced the time in range by 4 weeks to 53.2 weeks, the patient's TTR would fall to 60.5% - a reduction of 4.5%. If this occurred in 1/10 of warfarin arm patients in the trial (a rate considerably higher than that suggested by the Applicant), and this patient was representative, the overall TTR would fall by less than ½% due to the medication errors.

   However, stopping warfarin or apixaban therapy for 28 days due to a medication error could increase the rate of stroke (see Sec. **6.1.10.2**. Taking two active drugs for a month could increase the rate of bleeding. Thus, the fact that overall TTR in ARISTOTLE was reasonably good does not rule out the possibility that medication errors affected the trial outcome.

d. In its initial submissions regarding dispensing errors, the Applicant only counted a dispensing event as an error if the medication was of the "wrong type", i.e., the patient received a medication he or she should not have received. On March 22 and then again on April 12, we asked the sponsor to perform analyses of dispensing errors that included instances where the patient received a bottle with an erroneous bottle number, even if the medication was of the right type. Because the site was supposed to obtain the bottle number from the IVRS system at the time of the patient visit or just before the visit, we also asked the Applicant to count a dispensing event as an error of the IVRS assigned a patient a bottle, but the patient received the bottle before the assignment or more than 7 or 30 days after the assignment, because this would suggest an error at the site.

   Table 6 provides container level data. About 99.3% or more of bottles dispensed had been assigned to the patient who received the bottle

(according to the study database, based on the CRFs) were assigned by the IVRS from 0 to 7 days prior to the indicated dispensing event. The remainder of bottles dispensed was dispensed to a subject either before or more than 7 days after days after the bottle was assigned by the IVRS to that patient. For <0.1% of bottles dispensed, that particular bottle was never assigned to the patient who received it. This would include bottles of the correct type (containing the medication the patient was supposed to receive, ignoring errors involving the wrong dose of apixaban) and also bottles of the wrong type.

Table 7 and **Table 8** provide patient level data analogous to the bottle-level data in **Table 6**. In **Table 7** data for the two types of medication each patient was supposed to receive are disaggregated, while in **Table 8**, they are combined. **Table 8** indicates that about 7.5% and 7.9% of patients in the apixaban and warfarin arms respectively were dispensed at least one bottle that was not assigned to the patient by the IVRS or was assigned after it was dispensed or more than 7 days prior to the dispensing date. About 4.9% and 5.3% of patients in the apixaban and warfarin arms, respectively received at least one bottle that was not assigned to the patient by the IVRS or was assigned after it was dispensed or more than 30 days prior to the dispensing date.

Reviewer Comment: The Applicant explained to us verbally that they believe that bottles indicated as being dispensed before being assigned represent errors in the assignment date. They also commented that many bottles that were dispensed late represent deliberate acts by the site when more than one bottle of a type was ordered by the IVRS. In such cases, the investigator might dispense one bottle while keeping the others for dispensing in the future. This claim is difficult to evaluate.

*While the error rates in **Table 6** through **Table 8** are not problematic, it should be noted that the dispensing field information would be inaccurate if the investigator merely copied the bottle number provide to the site by fax or email into the CRF, but dispensed a bottle with a different number. When errors in the "verified" field (for bottles brought back to the site for a tablet count and then taken home again by the patient) and the "returned" field (i.e., the final return of a bottle) are counted, the per patient error rate increases to over 30% in each arm. However, the data that include the verified and returned field information are confounded, by the fact that in many cases where a bottle indicated as verified or returned was not dispensed to the patient and the bottle was never shipped to the patient's site. However, some modest percentage of the errors involving bottle numbers in the verified and returned fields may represent true errors.*

Because of the uncertainty around these data, the reviewers believe that the bottle label information is the most accurate way to estimate the medication error rate.

Table 6 Container Level – Dispensing Dates vs. IVRS Call Dates

Container Type	Prior to IVRS Call	0-7 days after IVRS Call	8-14 days after IVRS Call	15-30 days after IVRS Call	> 30 days after IVRS Call	> 1 year after IVRS Call	Dispensed but not assigned by IVRS	
APIXABAN ARM	Number (%) of Bottles Dispensed Relative to Date of IVRS Call and Assignment							
Active Apixaban	23 (0.03)	66540 (99.37)	113 (0.17)	91 (0.14)	154 (0.23)	2 (<0.01)	23 (0.03)	
Placebo Warfarin	60 (0.04)	150944 (99.37)	290 (0.19)	258 (0.17)	281 (0.18)	5 (<0.01)	26 (0.02)	
Active Warfarin	N/A	N/A	N/A	N/A	N/A	N/A	36 (0.02)	
Placebo Apixaban	N/A	N/A	N/A	N/A	N/A	N/A	24 (0.04)	
WARFARIN ARM	Number (%) of Bottles Dispensed Relative to Date of IVRS Call and Assignment							
Active Warfarin	53 (0.04)	140263 (99.38)	237 (0.17)	216 (0.15)	271 (0.19)	2 (<0.01)	47 (0.03)	
Placebo Apixaban	32 (0.05)	65141 (99.29)	122 (0.19)	88 (0.13)	181 (0.28)	2 (<0.01)	28 (0.04)	
Active Apixaban	N/A	N/A	N/A	N/A	N/A	N/A	25 (0.04)	
Placebo Warfarin	N/A	N/A	N/A	N/A	N/A	N/A	35 (0.02)	

Denominator for all percentages is number of bottles dispensed as the indicated product, per the study database, Apixaban Arm, N=9088. Bottles dispensed per database: Active Apixaban – 66964; Placebo Warfarin – 151899 Warfarin Arm, N=9052. Bottles dispensed per database: Active Warfarin – 141133; Placebo Apixaban – 65606

Table 7 Patient Level – Dispensing Dates vs. IVRS Call Dates

Container Type	Prior to IVRS Call	0-7 days after IVRS Call	8-14 days after IVRS Call	15-30 days after IVRS Call	> 30 days after IVRS Call	> 1 year after IVRS Call	Dispensed but not assigned by IVRS	
APIXABAN ARM	Number (%) of Patients Receiving Bottles Dispensed in Date Range Relative to Date of IVRS Call and Assignment							
Active Apixaban	23 (0.25)	9087 (99.99)	80 (0.88)	84 (0.92)	142 (1.56)	2 (0.02)	23 (0.25)	
Placebo Warfarin	58 (0.64)	9088 (100.0)	138 (1.52)	196 (2.16)	213 (2.34)	5 (0.06)	25 (0.28)	
Active Warfarin	N/A	N/A	N/A	N/A	N/A	N/A	35 (0.39)	
Placebo Apixaban	N/A	N/A	N/A	N/A	N/A	N/A	23 (0.25)	
WARFARIN ARM	Number (%) of Patients Receiving Bottles Dispensed in Date Range Relative to Date of IVRS Call and Assignment							
Active Warfarin	52 (0.57)	9047 (99.94)	123 (1.36)	176 (1.94)	218 (2.41)	2 (0.02)	44 (0.49)	
Placebo Apixaban	30 (0.33)	9046 (99.93)	88 (0.97)	87 (0.96)	167 (1.84)	2 (0.02)	28 (0.31)	
Active Apixaban	N/A	N/A	N/A	N/A	N/A	N/A	25 (0.28)	
Placebo Warfarin	N/A	N/A	N/A	N/A	N/A	N/A	35 (0.39)	

Denominator for all percentages is number of bottles dispensed as the indicated product (regardless of whether active or placebo), per the study database,
Apixaban Arm, N=9088. Bottles dispensed per database: Active Apixaban – 66964; Placebo Warfarin – 151899
Warfarin Arm, N=9052. Bottles dispensed per database: Active Warfarin – 141133; Placebo Apixaban – 65606

Table 8 Patient Level – Dispensing Dates vs. IVRS Call Dates
Number of Subjects with Dispensing Dates in Time Periods Relative to IVRS Call and Assignment

Container Type	Prior to IVRS Call	>7 days after IVRS Call	>30 days after IVRS Call	Prior to or >7 days after IVRS Call	Prior to or >30 days after IVRS Call	Dispensed but not assigned by IVRS	Prior to or >7 days after IVRS Call or Dispensed but not assigned by IVRS	Prior to or >30 days after IVRS Call or Dispensed but not assigned by IVRS
Apixaban Arm N=9088	72 (0.79)	557 (6.13)	296 (3.26)	612 (6.73)	361 (3.97)	98 (1.08)	685 (7.54)	443 (4.87)
Warfarin Arm N=9052	72 (0.80)	571 (6.31)	316 (3.49)	629 (6.95)	379 (4.19)	119 (1.31)	714 (7.89)	481 (5.31)

Denominator for all percentages is number of bottles dispensed as the indicated product, per the study database,
Apixaban Arm, N=9088. Bottles dispensed per database: Active Apixaban – 66964; Placebo Warfarin – 151899
Warfarin Arm, N=9052. Bottles dispensed per database: Active Warfarin – 141133; Placebo Apixaban – 65606

e. The Applicant performed a series of assumption-based modeling analyses of key endpoint results of ARISTOTLE. The Applicant assumed a series of per-bottle error rates higher than the observed error rate of 0.38%. Assumed error rates assumed ranged from 0.5% to 1% at increments of 0.1%, then 1% to 5% at increments of 1%, and finally, an assumed error rate of 10%. The Applicant described the methodology as follows:

> "In the simulations, for each assumed rate of treatment assignment errors, e.g., 0.5%, correct bottles were randomly selected and assigned to an incorrect treatment type. These simulated cases were combined with the 1654 known cases to achieve the overall error rate in treatment dispensations at the specified level; for each assumed error rate, 100 replications were generated and, for each replicate, sensitivity analyses on primary efficacy and ISTH major bleeding were performed. The sensitivity analyses used the observed endpoints in the study and followed the methodology used in the CSR analyses but excluding endpoints on or after a subject first received an incorrect bottle type and censoring subjects who did not have an endpoint prior to first receiving an incorrect bottle type."
>
> *Reviewer Comment: The Applicant presented mean results (event rates, hazard, ratio, 95% CI, and superiority p) for each run of 100 replications. Also, all-cause death comparisons were not estimated in these models.*

For the primary endpoint, superiority of apixaban over warfarin (i.e., $p<0.05$) was maintained for error rates up to and including 5%, with $p=0.036$ for the 5% modeling run. For ISTH major bleeding, superiority of apixaban was maintained up to including an error rate of 10%, with $p=0.0013$ for the 10% modeling run. For each endpoint, the general trend was a very small reduction in the hazard ratio (i.e., a change favoring apixaban) as the error rate approached 10%, with widening of the confidence intervals as progressively more patients were censored for assumed errors (**Table 6**). In **Table 9** the Applicant argues that this suggests that even with these very high error rates, the superiority results for the primary endpoint and ISTH major bleeding are robust.

> *Reviewer comment: This reviewer has concerns with the Applicant's approach. The Applicant's modeling results are based on means of 100 replications of the study results with censoring based on an assumed error rate. Rather than attempting to model the effects of the observed medication errors, the Applicant's analysis merely tells us how many patients can be removed at random times from the key study analyses without affecting the*

finding of superiority of apixaban over warfarin. A priori, in this type of analysis, the hazard ratio is unlikely to change very much. Instead, the confidence interval of the HR would be expected to widen as patients and data were excluded from the analysis. Eventually the CI would cross zero. This is exactly what occurred. However, the actual impact of medication errors might be extreme in either direction (i.e., towards making apixaban look better or worse than it does in the protocol-specified analyses of the relevant endpoints).

Thus, the most sensible approach to understanding the impact of medication errors on the study results is to obtain the best possible data regarding medication errors – i.e., the bottle labels, including those at BMS and those still at the sites – and to use the data from all the labels and the IVRS system to determine what errors occurred and when they occurred. Patients should be censored on the basis of actual data, and the results based on those data, not the means of 100 modeling runs. There is no sense in relying on modeling when the actual data are available and could be obtained by the Applicant.

Table 9 Applicant's Modeling of Medication Errors - Mean Hazard Ratios and Confidence Intervals

Values represent means of 100 runs at each rate

Assumed Error Rate, %	Primary Endpoint (Stroke/SE) HR (95% CI)	ISTH Major Bleeding HR (95% CI)
0.5	0.78 (0.65, 0.94)	0.68 (0.59, 0.79)
0.6	0.78 (0.64, 0.94)	0.68 (0.59, 0.79)
0.7	0.78 (0.65, 0.94)	0.68 (0.59, 0.79)
0.8	0.78 (0.64, 0.94)	0.68 (0.59, 0.79)
0.9	0.78 (0.64, 0.94)	0.68 (0.59, 0.79)
1	0.78 (0.64, 0.94)	0.68 (0.58, 0.79)
2	0.77 (0.63, 0.95)	0.67 (0.57, 0.79)
3	0.77 (0.62, 0.95)	0.67 (0.56, 0.79)
4	0.76 (0.60, 0.95)	0.66 (0.55, 0.79)
5	0.75 (0.59, 0.95)	0.65 (0.54, 0.79)
10	0.74 (0.55, 0.98)	0.62 (0.49, 0.79)

FDA also performed sensitivity tests to determine the potential effects of medication errors on the trial outcome. The results of all such testing suggested that the study results showing superiority to warfarin for the primary efficacy and safety endpoints

would be negated only of one assumed high error rates, and that non-inferiority for these outcomes would not be threatened. These analyses are summarized below:

1. The Medical Officer who performed the efficacy review modeled the results for the primary efficacy endpoint (composite of stroke and systemic embolism). The model examined the as follows:

- The baseline case was the ITT analysis.
- The model assumed the worst case: that medication errors occurred and they biased the study only in favor of apixaban, so that when patients with errors were removed from the analysis, the study results would become less favorable for apixaban. The model assumed one type of error: substitution of warfarin placebo for active warfarin in the warfarin arm, which according to the clinical database was the most common error that would have biased the study in favor of apixaban for the primary endpoint.
- Various error rates were assumed
- There were two assumptions on the magnitude of the effects of errors on total stroke, which comprised most of the primary endpoint events.
- The duration of the effect of the error on the event rate was assumed to be 38 days, the mean number of days that tablets from a warfarin bottle were taken,
- No more than one error per patient was assumed
- The number of strokes added by the various error scenarios was compared to the number of events, if subtracted from the warfarin arm, would negate the findings of superiority (13) or non-inferiority (78) of apixaban to warfarin in the ITT analysis.

Table 10 shows the number of additional strokes that would be expected to result from various error rates as described in previous paragraph. The first 3 data rows assume per-patient error rates of 10%, 40% and 100%, with all errors resulting in 38 days of an event rate 3x the primary efficacy endpoint event rate in the warfarin arm in the ITT analysis. Note that the actual error rate was about 8% based on the clinical database and about 3% based on the bottle label database.

Dr. Bai's analysis indicated that it would take 13 fewer events in the warfarin arm to negate superiority of apixaban to warfarin, and 78 fewer events to negate non-inferiority. A 10% error rate would be expected to account for 3 extra events in the warfarin arm. The limit of 13 events to negate superiority is not closely approached until the error rate reaches 40%, when 12 errors would be expected. If all warfarin arm patients had experienced an error, 30 additional events would be expected. In this extremely unlikely scenario, superiority would be negated, but non-inferiority would not be lost when the error-induced events were eliminated from the analysis.

If the effect of an error was an event rate 5x the ITT rate, (somewhat higher than one would expect) it would require a 20% error to add 12 events. An error rate of 100% would yield 60 extra events, less than what would be required to negate non-inferiority.

Note the number of added events in **Table 10** is a point estimate; the CI was not calculated. If the upper end of the 95% CI was used instead of the point estimate as the metric of interest, the error rates needed to negate superiority and non-inferiority might be considerably lower than suggested here.

Table 10 Modeling of Effects of Medication Errors[1] on Strokes in ARISTOTLE
(Near-worst-case assumptions for apixaban)

Warfarin arm patients with errors, % N=9052 [2]	Error effect on primary endpoint rate, multiple of ITT rate [3]	Error-induced additional events, per 100 pt-yr [4]	Additional events
10%	3x	3.2	**3.0**
40%	3x	3.2	**12.1**
100%	3x	3.2	**30.2**
20%	5x	6.4	**12.1**
100%	5x	6.4	**60.3**

1. All errors were assumed to be substitution of placebo warfarin for active warfarin in warfarin arm patients (see text).
2. Treated patients
3 The primary endpoint occurred at a rate of 1.6 events per 100 patient-years in the warfarin arm in the ITT analysis.
4. Effect of error was assumed to last for 38 days before return to the baseline rate in each case

3.1.2 Fraud in China

In third quarter 2011 the Division requested that DSI inspect ARISTOTLE site 1200 in Shanghai, China (PI, Dr. Shiyao Wu) because of "excess warfarin arm primary endpoint events and deaths". In January 2012, the Applicant informed the Agency of "suspected misconduct" at site 1200. Monitoring staff alleged that BMS senior clinical site manager (initials (b) (6) responsible for all sites in China, altered source documents to eliminate evidence of GCP violations in preparation for the scheduled FDA inspection of the site. This site randomized 35 subjects.

Allegations and background
Sites in China were monitored by PPD (a CRO), with oversight from BMS site managers. In November 2011, DSI informed BMS of their intentions to inspect sites 1200 and 1178 in China. One of the PPD monitors for site 1200 (initials (b) (6) visited the site during the study and also on December 6-7, 2011 to prepare for the DSI inspection.

Following the pre-inspection visit, she alleged that source documents had been altered at the site prior to her pre-inspection visit. Specifically, (b) (6) stated that:

- Outpatient records she had monitored 2 years earlier were different (hand written comments and clarifications from 2 years earlier were no longer present).
- She had described her concerns to a second PPD monitor of the site (initials (b) (6) She stated that (b) (6) told her that the documents had been changed at the direction of the BMS site manager ((b) (6) (b) (6) states that (b) (6) told her that he had personally altered records. ((b) (6) later "refused to confirm or deny" this.)
- (b) (6) told (b) (6) that he had a USB drive with patient data that (b) (6) "understood to have been altered." (b) (6) made a copy of the drive.
- During (b) (6) pre-inspection visit, she had seen a nurse writing on a patient chart in pencil. (b) (6) understood that the penciled entries would later be rewritten and the penciled markings erased. The nurse allegedly told (b) (6) that she "had been told to do this by BMS."

BMS subsequently undertook a series of actions to investigate (b) (6) allegations, including phone and in person interviews and several visits at the site. These site visits included two multi-day visits by internal BMS auditors. Persons interviewed included the two PPD monitors ((b) (6) and (b) (6) the BMS site manager ((b) (6) and several supervisory staff from BMS. Internal BMS communications were also reviewed. Interviews with "key subinvestigators" at site 1200 were conducted with translators. Apparently, the PI was not interviewed.

Key Findings by BMS

BMS found that there was evidence to support the allegation by (b) (6) that documents were altered prior to her pre-inspection visit.

- For example, the electronic files for subject 17557 show modifications (not described by BMS) were made on 3 days in late November 2011. BMS staff ((b) (6) were on-site for only one of the 3 days when alterations were made.
- The electronic files on the USB drive indicate that outpatient records were modified in late November and early December 2011. However, there was no audit trail, and the documents do not indicate who made the alterations.
- The BMS site manger, (b) (6) identified significant GCP issues during a site visit on Nov. 14-18, 2011. These were provided to her manager (initials (b) (6) the "China Hub" unit manager) and to (b) (6) a BMS Associate Monitor Manager. However, some key issues were withheld from the Global apixaban team.
- GCP issue identified by (b) (6) included:
    - Drug accountability issues such as missing bottles of warfarin from the last visit for 5-6 subjects, differences in drug return dates in the eCRF and drug log, and dosing compliance issues for "some subjects."

- Subject diaries were missing dates or subject number. There were differences between diaries and source documents (not explained).
- Informed consents were signed by "different people (subject or relative)"; handwriting appeared inconsistent in some cases.
- There were potential unreported SAEs (all non-endpoint events for 4 subjects, potential late reporting of SAEs (not noted to be endpoint events for 3 subjects).
- **There were also subjects with potential unreported endpoint events:**
  - Subject 6838 (warfarin arm, potential bleed): possibly a minor anal bleed during a hospitalization for HF
  - Subject 11384 (apixaban 5 mg arm, potential bleed x 2): During a hospitalization for HF in May 2010, the subject reported blood in stool, a test showed "weak positive." No AE was reported. An outpatient record indicates that the investigator believed this was not a bleed. In December 2010, the patient was rehospitalized for HF, and had "blood in sputum"; there are no further details. This was not reported as a bleed.
  - Subject 9094 (apixaban 5 mg arm, potential MI): The patient was hospitalized for unstated reason and duration. During the hospitalization, the patient had severe lesion of "R branch of coronary artery; moderate lesion of L branch of coronary artery," and had PCI. No further details were provided. No endpoint event was reported; it is not stated what, if anything was reported.
  - Subject 5985 (apixaban 5 mg arm, potential stroke): It is not stated why a stroke was suspected. The letter from BMS states that "[BMS] auditors reviewed hospital records for subject 5985. No evidence of a stroke was found in the records provided."

    Reviewer Comment: None of this establishes that an endpoint was missed.

- "Most of the adverse events in the inpatient charts … were not reported; some adverse events found in the outpatient chart were not reported."

  Reviewer Comment: This may be a comment of general applicability. It does not refer to any specific patient.

- Source documents were missing patient number.
- Physical exams were not done at the end of treatment visit.

- ECGs were not done at the screening visit (it's not clear how often this happened**).**
- Concomitant medications were not recorded.

Key findings at Site 1178 (Dr. Shulin Wu, Guangdong, China)

"Examples" of findings made by the BMS site manager, (b) (6) were confirmed by the BMS auditors, including these potentially unreported or late-reported endpoints:

- Subject 20235 (apixaban 2.5 mg arm) - Unreported AE of blood in stool
- Subject 19579 (warfarin arm) – Late reporting of SAE of R cervical artery occlusion (adjudicated as "No Event", but there was also an adjudication of ischemic stroke occurring with onset 2 days later)

Reviewer Comment: There were no allegations of modification of documents at this site.

Actions Taken by BMS or involved staff
BMS states the following actions have occurred/are planned:

- The BMS site manager who was alleged to have told the PPD monitor (b) (6) to alter records resigned from BMS in January 2012
- A disciplinary process was instituted for the 2 other BMS employees who had knowledge of the irregularities at site 1200 and withheld information from the global team, (b) (6) and (b) (6) Both these employees were terminated in January 2012.
- Re-training of BMS clinical research staff in China on scientific misconduct and "The Standards of Business Conduct and Ethics" will be completed by the end of 2Q 2012
- PPD will conduct an investigation of the activities of their staff at site 1200 and provide refresher training to "all PPD operational staff in China."
- An audit (by an unidentified party) will be conducted of PPD China and will include an assessment of monitor selection and training, a review of the PPD investigation, and verification of the PPD refresher training.

Summary of Conclusions by BMS:

BMS made the following conclusions:
A local BMS site manager identified CGP deficiencies during pre-inspection visits at site 1200. She withheld information about her findings from the global product team.

In response to allegations about activities at site 1200, an investigation was launched. There is evidence suggesting that documents were altered at the site to influence the outcome of the inspection.

BMS believes that the problems identified at site 1200 are "site-specific." The GCP violations identified at site 1200 "would not affect the interpretation of the study results from the site, but BMS is prepared to provide revised statistical analysis of [ARISTOTLE] that exclude the data from site 1200." Disciplinary actions with respect to the involved employees are being undertaken and the Company has instituted measures "to ensure that these issues are not repeated."

> *Reviewer Comment: The situation in China is complex and can be viewed in several ways. OSI has made a recommendation that excludes 24 specific sites where (b) (6) or Mr. (b) (6) worked from key analyses.[h] However, since (b) (6) was responsible for all 36 sites in China, one could reasonably exclude all Chinese sites from important analyses* ***(Table 11).***
>
> *The OSI inspection of site 1200 in China reveals major problems with study medication accountability, and that monitoring documentation at this site, including but not limited to documentation of errors in study drug accountability was inadequate. OSI's inspection of site 1178 in China did not reveal these problems, however they recommend excluding this site from key analyses.*
>
> *While there is no compelling evidence that key endpoint events from site 1200 were not reported, GCP violations, including under-reporting of adverse events, abounded at this site. Monitoring during the study seems to have been ineffective in finding and preventing the GCP errors. Given the plethora of problems at this site, it would not be surprising if key endpoint events were indeed not reported. The findings at site 1200 are consistent with our global concerns about medication errors and lax monitoring in this study.*
>
> *These problems suggest a pattern of sloppy execution in ARISTOTLE. Notably, ARISTOTLE was planned as a non-inferiority study. In this type of the study, "noise" resulting from sloppiness in execution tends to favor success by minimizing apparent differences between the experimental agent and the control. This is the opposite of the effects of noise in superiority studies, where noise tends to obscure the differences that one hopes to find. Thus, the surprising number of serious execution errors in ARISTOTLE may have helped the Applicant achieve its primary study goal of demonstrating the non-inferiority of*

h *OSI recommends that data from the Chinese sites where either (b) (6) or Mr. (b) (6) worked be excluded from the study analysis. These are Sites 1168, 1178, 1180, 1182, 1198, 1199, 1200, 1206, 1207, 1221, 1223, 1244, 1246, 1247, 1266, 1287, 1547, 1548, 1549, 1550, 1552, 1555, 1556, and 1558. The listed sites constitute 24 of the 36 sites in China in the ARISTOTLE study.*

apixaban to warfarin for preventing stroke and systemic embolism in patients with nonvalvular atrial fibrillation.

The reviewers performed sensitivity analyses that (1) excluded Site 1200, (2) excluded the sites recommended by OSI for exclusion, and (3) excluded all sites in China for key study analyses (the primary efficacy endpoint (stroke and systemic embolism), its components, MI, all-cause death, the primary safety endpoint (ISTH major bleeding), and GUSTO severe bleeding. Results for the primary efficacy endpoint and all cause death are shown in **Table 11**. *Excluding Site 1200 disfavors apixaban slightly in the analyses of both the primary endpoint and all-cause death, but superiority of apixaban is lost only for all-cause death. Excluding 24 sites or all 36 sites in China favors apixaban in both analyses, and does not change the conclusions of the overall analyses.*

Table 11 Sensitivity Analyses Relating to Fraud in China

Event	Apixaban n/N	Rate*	Warfarin n/N	Rate*	A vs W Hazard Ratio (95% CI)	p value
Efficacy event rates and hazard ratios (ITT analysis – all sites (N=18,201))						
Primary Endpoint (Stroke/SE)	212 / 9120	1.27	265 / 9081	1.60	0.79 (0.66, 0.95)	0.0114
All-Cause Death	603 / 9120	3.52	669 / 9081	3.94	0.89 (0.80, 1.00)	0.0465
Efficacy event rates and hazard ratios (ITT analysis without Site 1200 (N=18,166))						
Primary Endpoint (Stroke/SE)	212 / 9103	1.27	263 / 9063	1.59	0.80 (0.67, 0.96)	0.0145
All-Cause Death	603 / 9103	3.53	666 / 9063	3.93	0.90 (0.80, 1.00)	0.0565
Efficacy event rates and hazard ratios (ITT analysis without OSI specified Chinese sites (N=17,567))						
Primary Endpoint (Stroke/SE)	192/8804	1.19	242/8763	1.51	0.78 (0.65, 0.95)	0.0122
All-Cause Death	587/8804	3.55	655/8763	3.99	0.89 (0.79, 0.99)	0.0379
Efficacy event rates and hazard ratios (ITT analysis without all Chinese sites (N=17,358))						
Primary Endpoint (Stroke/SE)	184 / 8698	1.15	239 / 8660	1.51	0.76 (0.63, 0.92)	0.0053
All-Cause Death	578 / 8698	3.53	652 / 8660	4.02	0.88 (0.79, 0.98)	0.0239

Reviewer's analysis:\china\china, sponsor's data: adefl, adbs2, adbl2

When the analyses for the primary safety endpoint (ISTH major bleeding) and GUSTO severe bleeding are compared in an analogous fashion, sensitivity analyses had no material effect on the overall results (data not shown).

3.1.3 Dataset Quality

We are concerned that the trial datasets do not match the information in the CRF. In our review of four observations of data out of over one million in your medication error dataset, smed.xpt (used for most of your medication error analyses), we found an observation with a valid date in the CRF that was misrepresented by a period in the dataset, indicating that a valid date was missing. This error in the dataset has the possible effect of reducing the reported medication error rate since the observation was subsequently excluded from the count of medication errors because of the missing date in the dataset.

The raw dataset smed.xpt, containing information on drug dispensing, was used to create the analysis dataset, adinctrt.xpt. In smed.xpt the apixaban/placebo bottle verification date (variable name: smedd) for subject CV185030-1106-2032 contains a period(.), implying that the date was missing, invalid or partial. However, this subject's eCRF has an apixaban bottle verification date of "24OCT2008". Notably, the warfarin/placebo bottle medication return date is also "24OCT2008" in the eCRF, and it correctly appears in the dataset (variable name: smrtd). The SDTM dataset da.xpt (contains similar information as smed.xpt) is also incorrect since it appears that the Applicant used smed.xpt to create da.xpt. (reviewer's analysis: med_err\adinctrt check, sponsor's dataset smed, da) These dates are important because they were used to determine the subject's actual treatment in relation to the primary safety and efficacy endpoints. For this particular bottle, the Applicant removed it from the analysis dataset, adinctrt.xpt which was used for the majority of the medication error analyses.

The identified date problem, found with little effort, is worrisome because the primary analyses are time to event analyses. The Applicant should assure us that the datasets for these analyses are accurate and describe why they believe so. If their data cleaning processes were different for these datasets than they were for the medication error datasets, then they should apply similar processes to cleaning their medication error datasets. They should explain how their new process differs from that used for this NDA submission. There is certainly a concern that there may be other aspects of the datasets that do not match the CRF that the reviewer has not identified.

3.2 Compliance with Good Clinical Practices

3.2.1 Unblinding

ARISTOTLE was a blinded study. BMS' Clinical Supplies Operations had access to individual subject treatment assignments. In the event of a medical emergency or

pregnancy, in which knowledge of the investigational product ([(b) (6)] was critical to the subject's management, the blind for that subject could be broken. Before breaking the blind, the investigator had to determine that the information was necessary (i.e., that it could alter the subject's immediate management). In addition to the pre-specified unblinding for DMC members and the DMC reporting statistician, 78 subjects (30 on apixaban, 48 on warfarin) were unblinded during the study. Reasons given for unblinding were SAEs for all subjects except one on apixaban (reason unknown, site 837, country Austria). OSI's report suggests that there may be cases where unblinding was not needed to inform medical management of the subject.

3.3 Financial Disclosures

The only trial providing substantial support for approval is ARISTOTLE, performed globally at more than 1000 centers. No single investigator provided a meaningful fraction of the safety data. Out of thousands of principal and sub-investigators in ARISTOTLE, only 6 at sites that enrolled patients disclosed a financial interest, which consisted of either a substantial equity interest or substantial payments. These 6 investigators were at 6 sites that randomized a total o[(b) (6)] subjects (range, [(b) (6)] subjects per site). The subjects at these sites accounted for [(b) (6)]% of the 18,201 randomized subjects in ARISTOTLE, an ostensibly blinded study.

In addition to these 6 investigators, there were 3 investigators with disclosed interests who were at sites that failed to enroll or randomize any patients.

Other disclosure issues involving ARISTOTLE were (payments are in US $ unless otherwise noted):

- One physician from [(b) (6)] who was a member of the [(b) (6)] and a member of the [(b) (6)]" disclosed substantial payments (his salary was supported by grants from BMS and Pfizer, and he received <10K in consulting fees from BMS).
- Another [(b) (6)] member disclosed substantial payments (about 25K as an unrestricted research grant for work as Co-PI on the [(b) (6)] study (company not named).
- Another [(b) (6)] member disclosed substantial payments (consultant and/or advisory board member for BMS).
- The [(b) (6)] disclosed substantial payments (research grant support (BMS), consultant, (BMS & Pfizer), [(b) (6)] member and "spokesman" (BMS).
- One of the 6 investigators with a disclosed interest noted above (substantial payments -- his [(b) (6)] receives a 16K UK pound grant based on his research) was also a member of the [(b) (6)]. His site randomized [(b) (6)].

The Applicant notes that the investigators and patients were blinded to treatment assignment and the study had many sites. While study sites may have been unblinded for reasons discussed elsewhere in this review, we have no reason to believe that members of the various study committees were unblinded, except for those (b) (6) .

Reviewer Comment: Because less than 1% of subjects were at sites where investigators disclosed interests and because all key analyses were based on events adjudicated by persons who should have had no access to the randomization code or to samples of study drug, we believe that bias by the potentially conflicted investigators or committee members was very unlikely to have affected the outcome of the study in a meaningful way.

4 Significant Efficacy/Safety Issues Related to Other Review Disciplines

4.1 Chemistry, Manufacturing and Controls

The CMC review team identified no issues that affect the interpretation of the clinical trial data. Based on the material they have reviewed, they believe the application may be approved. However, they noted that "GMP inspections are pending for the proposed tablet manufacturing sites in the U.S. and Puerto Rico, and for the proposed contract packaging site in Italy."

For a discussion of the issue of visual and tactile differences in thickness between the 5 mg active apixaban tablet used in the study and its placebo, see page **74.**

4.2 Clinical Microbiology

Not applicable to this submission – no clinical microbiology data were submitted.

4.3 Preclinical Pharmacology/Toxicology

This brief summary is based on Dr. Patricia Harlow's Pharmacology/toxicology (PT) final review dated 21Feb 2012 and addendum dated 13Apr2012. Please see her review for details. Apixaban is approvable from a PT perspective. Most of the toxicities were attributable to the pharmacodynamic effect of apixaban. The safety margins relative to human therapeutic exposures were satisfactory.

Apixaban did not significantly affect the central nervous, cardiovascular, respiratory, or renal systems after repeat dose studies in rats and dogs. There were no significant

effects on QTc in conscious dog studies; in vitro hERG studies were negative for both apixaban and O-desmethyl apixaban sulfate (its major metabolite).

Tissue distribution studies in rats showed that the GI tract (stomach, small intestine, large intestine, cecum), thyroid, urinary bladder, adrenal glands, liver and kidneys were exposed to the highest concentrations of apixaban. The brain, heart, and bone marrow were exposed to the lowest concentrations. Radioactivity was still present in the eyes at 168 hours after dosing. In contrast, radioactivity was last measureable in the blood and plasma at 24 hours post dose. The presence of radioactivity in the eye is noteworthy because the intraocular bleeding was higher on apixaban than on warfarin (see **Sec 7.3.2.1**). Elimination half-life estimates for apixaban-equivalents were less than 5 hours for adrenal glands, blood, plasma, and testes, but greater than 60 hours for the eyes, bone marrow and cecum. Apixaban showed no phototoxic potential in an acceptable mouse fibroblasts study.

The toxicology findings in chronic (6 months in rats and 12 months in dogs) repeat dose toxicology studies were related to the pharmacologic activity of apixaban (prolongation of coagulation parameters, evidence of bleeding, and effects on red cell parameters). There were transient effects on serum potassium. The NOAELS (no observable adverse effect level) correspond to exposure ratios of ~4 times (in rats) and 20-28 times (in dogs) the human exposure of unbound apixaban 5 mg BID.

Activated charcoal reduced the clearance of apixaban, with the largest effect at 3 hours after the dose in dogs. Hemodialysis (for 4 hours) had no effect on oral apixaban AUC_{0-24} in male beagle dogs; it did reduce the AUC_{0-4} by 6.5%. Mean Cmax concentrations for oral apixaban were reduced by 12% during dialysis.

Acceptable animal studies indicate that apixaban and its metabolites are not carcinogenic, mutagenic, clastogenic, or genotoxic. In the carcinogenicity studies CD-1 male and female mice and Sprague Dawley rats did not have a statistically significant increase in tumors at exposure ratios of ~9, ~20, and 2-4 times, respectively, the human exposure dose of unbound apixaban 5 mg BID.

There were no drug-related effects on estrous cycling, mating or fertility in male or female rats; at the highest dose of 600 mg/kg body weight gain decreased slightly in male rats. The NOAEL doses were 200 mg/kg in males and 600 mg/kg in females, corresponding to exposure ratios that are 2.7 fold and 4.2 fold, respectively, the human exposure dose of 5 mg BID.

In an acceptable embryo-fetal development study in rats, the NOAEL for maternal toxicity was 1000 mg/kg/day because of a decrease in mean body weight gain and increased vaginal bleeding at 3000 mg/kg/day. The NOAEL corresponds to an exposure ratio of 4.2 times the human exposure dose of unbound apixaban 5 mg BID.

Reviewer comment: Note that the exposure ratios for fetal development and fertility were correctly calculated as 4.2 times the human exposure despite the difference in drug dose (600 mg/kg vs. 1000 mg/kg).

A prenatal/postnatal development study in rats showed that dosing in the main study did not need to be interrupted during parturition. Bleeding and prolonged PTs were observed with 200 and 1000 mg/kg. There was no drug related mortality of F0 dams. At the NOAEL dosages of 200 and 1500 mg/kg, the exposure ratios were 4.9 and 5.4 times, respectively, the human exposure dose of unbound apixaban 5 mg BID.

Apixaban is excreted into the milk of rats primarily as apixaban. Concentrations of apixaban in milk were higher than concentrations in blood and plasma at all time points. Approximately 12 % of the maternal dose is excreted into the milk over a 24 hour period.

The main juvenile toxicology study in rats had a potential drug-related AE of an increased incidence of unilateral or bilateral degeneration of the testicular seminiferous tubules with 600 mg/kg/day in the Set 1 males that were necropsied at the end of dosing. The incidence of testicular degeneration was 33% (above the maximum of 26.7% in the historical control data). This finding appeared reversible in the recovery group; however the PT reviewer believes that this finding was drug related because of the severity of the testicular finding and the correlation with hypospermia in the epididymides.

The pharmacology/toxicology reviewer has the following recommendations:

1. Warn women of child-bearing potential of the high risks for bleeding during labor and delivery (there were no deaths during parturition in the pre/postnatal development study)
2. Conduct an additional juvenile animal study to determine if a critical period can be identified because of the severity of the testicular degeneration in the juvenile study. This should be done prior to any clinical pediatric studies.

4.4 Clinical Pharmacology

Most of this section is based on the final Clinical Pharmacology review by Drs. Lai and Menon-Anderson (Clinical Pharmacology) and Drs. McDowell and Marathe (Pharmacometrics) dated 15Feb2012.

4.4.1 Mechanism of Action

Activation of Factor X is the initial step in the final common coagulation pathway. FXa cleaves prothrombin to generate thrombin, which triggers the conversion of fibrinogen to fibrin, the fibrous protein that polymerizes to form a clot in conjunction with platelets. The activity of FXa is greatly increased when it is complexed with activated co-factor V

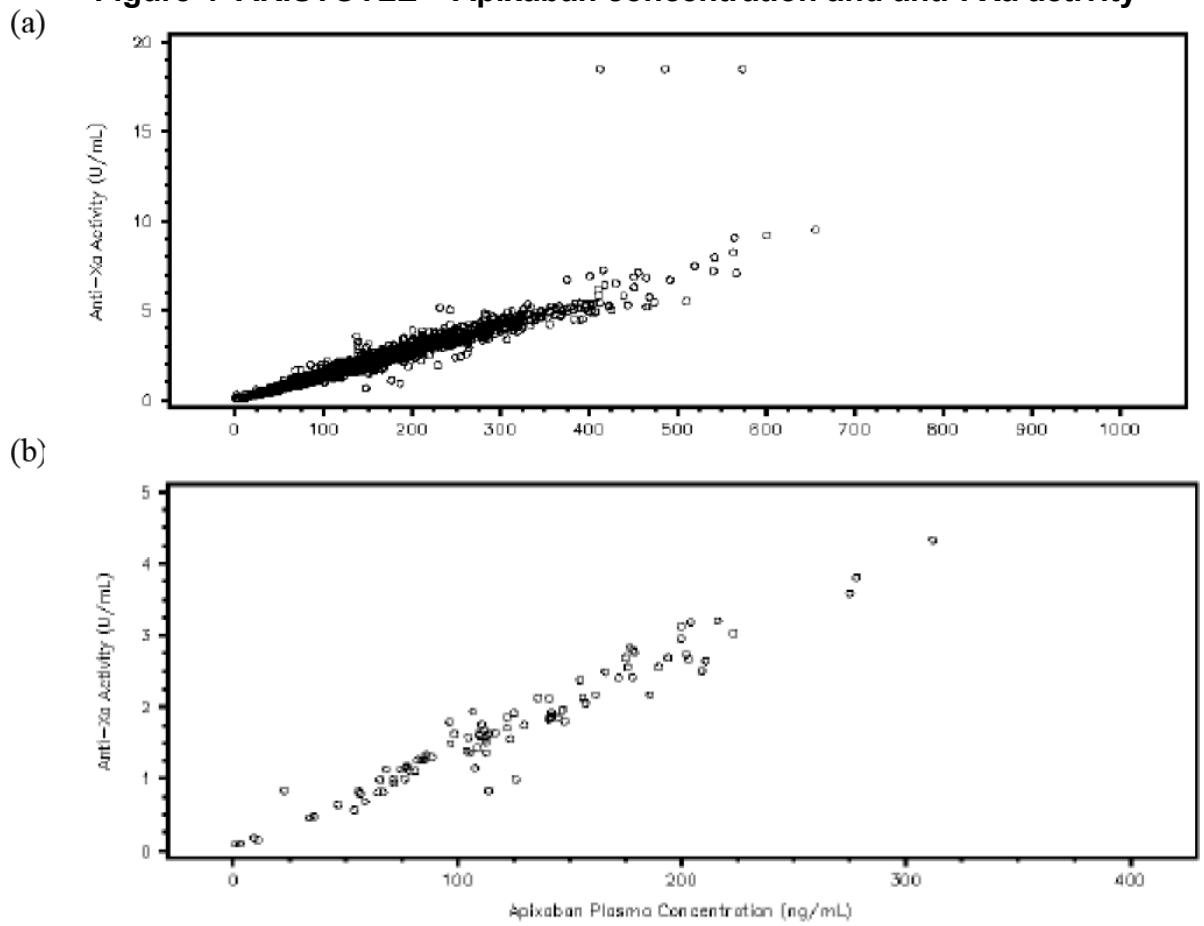
in the prothrombinase complex. By inhibiting FXa, apixaban inhibits the formation of thrombin from prothrombin and the downstream formation of fibrin and blood clots. Because of the functional location of FXa at the top of the final common coagulation pathway, apixaban affects clotting induced through both the intrinsic and extrinsic clotting cascades. Studies of the FXa inhibitory action of apixaban are discussed in Section. **4.4.2.**

4.4.2 Pharmacodynamics

Apixaban 2.5 mg and 5 mg did not affect PT or aPTT.

There is a linear relationship between apixaban plasma concentration and anti-FXa activity (**Figure 4**). Anti-FXa activity was ~ 15-20% less in the apixaban 2.5 mg group compared to the 5 mg group; this corresponds with the 25% lower apixaban concentration in the 2.5 mg group.

Figure 4 ARISTOTLE – Apixaban concentration and anti-FXa activity



Source: ARISTOTLE CSR, Figure 10.2.A, (a) is 5 mg BID, (b) is 2.5 mg BID
Data shown are for a small subset of ARISTOTLE subjects with a single random PK sample (n=3231) and/or PD sample (n=3125) collected at month 2.

Specific drug-drug interactions studies examining the PD effect on platelet aggregation and bleeding time did not show clinically relevant changes when studied with aspirin, clopidogrel, or aspirin and clopidogrel. Coadminstration with naproxen (P-gp inhibitor) resulted in a 1.5-fold increase in apixaban concentration and increased clotting time, but no changes were observed on arachidonic acid-induced platelet aggregation or clinically relevant prolongation of bleeding time. Coadministration with enoxaparin has an additive effect on anti-FXa activity (~50%). There was no PK effect and there was no clinically relevant bleeding.

4.4.3 Pharmacokinetics

The pharmacokinetic and biopharmaceutic properties of apixaban are described in **Table 12**.

Table 12 Apixaban pharmacokinetic and biopharmaceutic properties

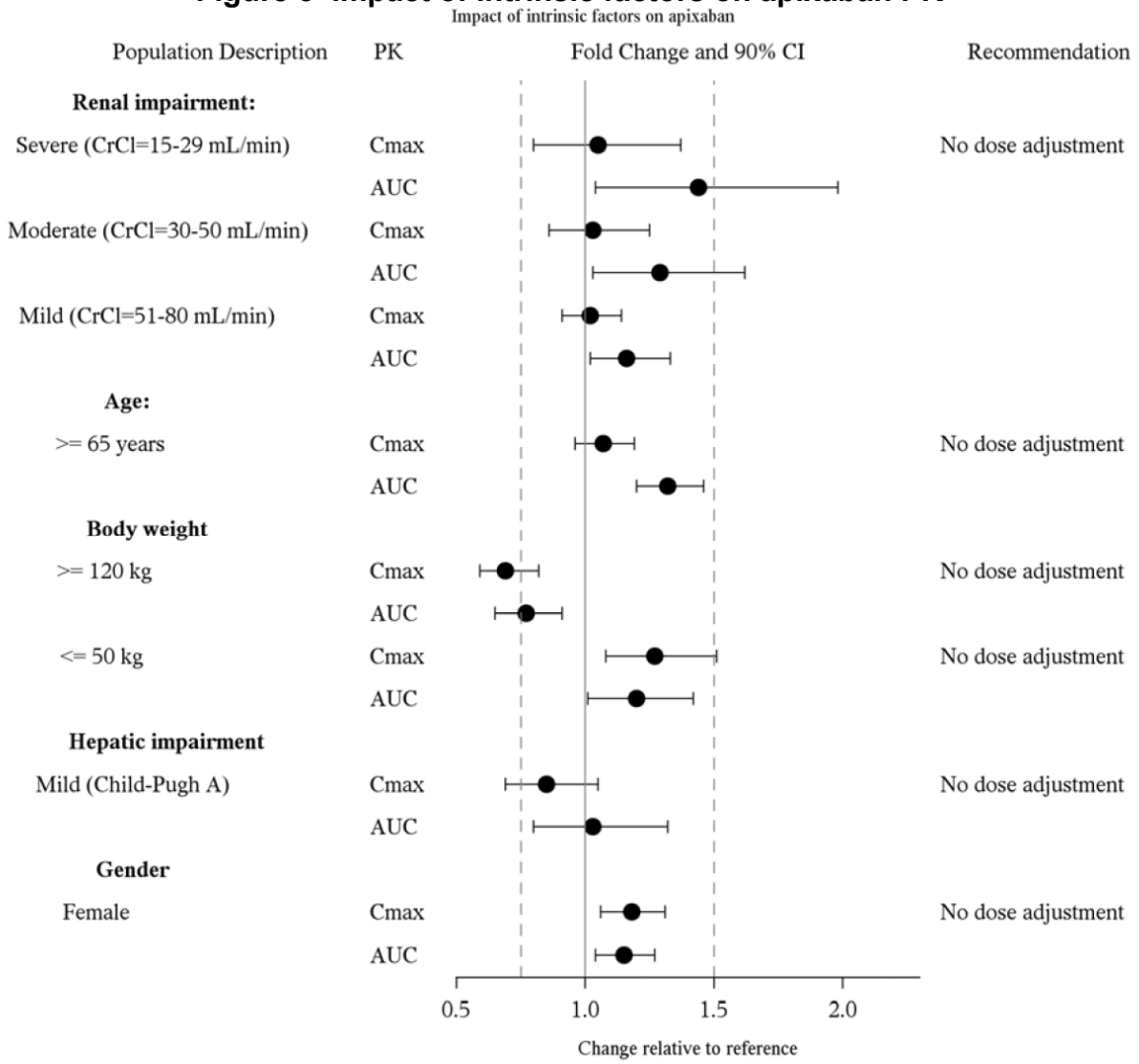
Absorption	50% of oral tablet dose
Tmax	~3-4 hours after an oral tablet dose
Distribution	Apparent volume of distribution (Vss/F) ~ 50 L
	87% bound to plasma proteins (so not highly protein bound like warfarin) No difference in plasma protein binding between healthy subjects and subjects with mild to moderate hepatic impairment
Metabolism	Not extensively metabolized; ~20% metabolized, primarily by CYP3A4 (minor contributions by CYP1A2 and CYP2J2)
Metabolites	Irrelevant towards pharmacological activity
Excretion	Multiple elimination pathways, but primarily eliminated by excretion (as unchanged drug in urine).[1] There is also a small component of hepatic clearance (small amount in bile and small amount metabolized by CYP3A4).
Half-life	12 hours
Dose proportionality	Linear over dose range of 2.5 to 10 mg QD. Less than dose proportional increase in CMax and AUC at doses greater than 10 mg QD
Accumulation ratio	<2 after BID dosing for 7 days
Food effect	Apixaban may be taken without regard to food.
BCS Class	III (high solubility, low permeability)

BCS= biopharmaceutical classification system

1. In a mass balance study, ~78% of the dose was recovered in 9 days, of which 25% was eliminated in urine (21% unchanged), ~2.4% in bile, and 56% in feces (34% unchanged)

The PK effect of various intrinsic factors alone in dedicated PK studies (**Figure 5**) and in population PK analyses indicate that no dose adjustment is needed in these populations when considered by itself. Dosing recommendations for hepatic impairment greater than mild cannot be provided.

Figure 5 Impact of intrinsic factors on apixaban PK



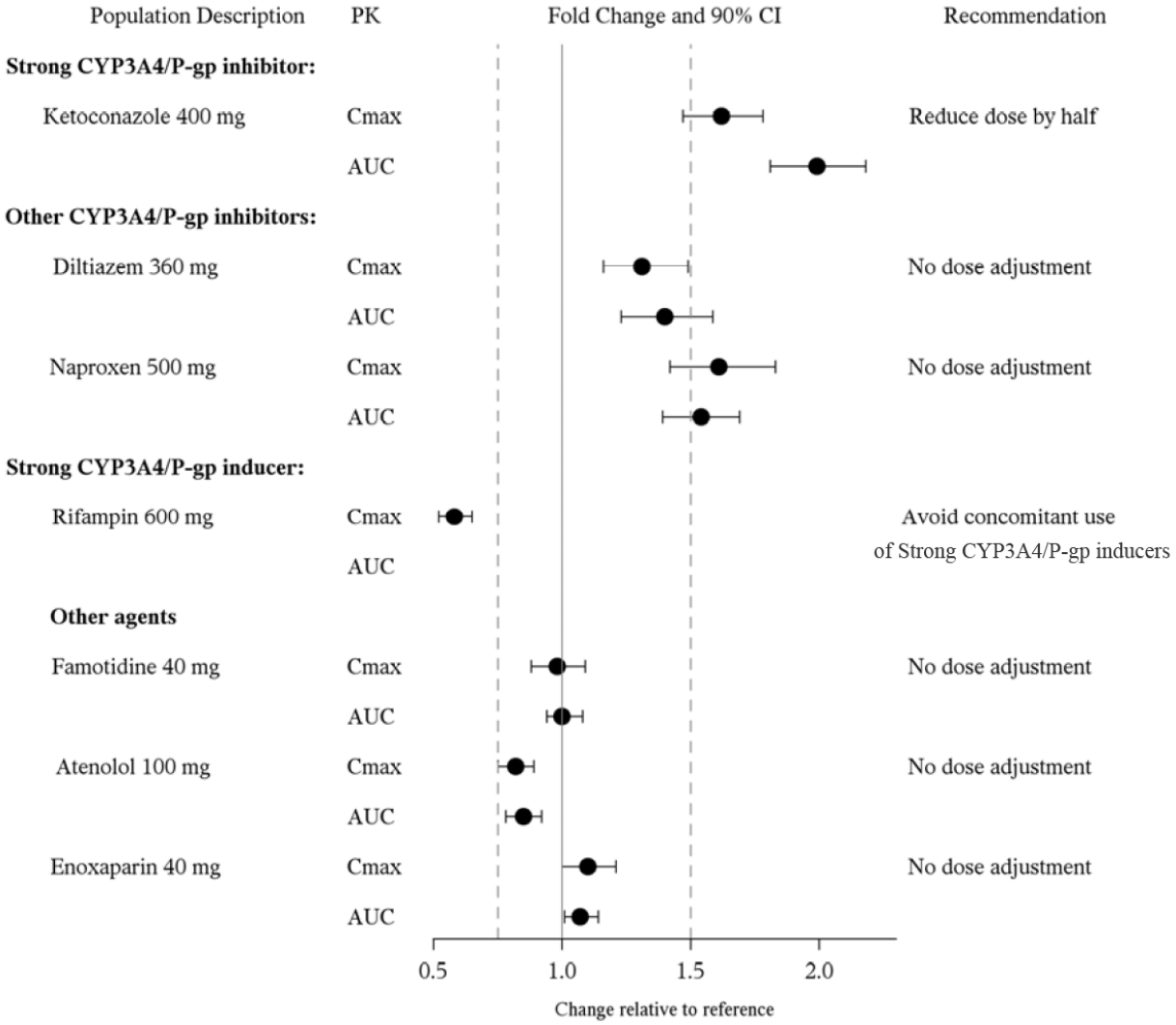
Source: Clinical pharmacology review, page 24

As already stated in **Table 12,** apixaban is not extensively metabolized. In addition to CYP3A4, apixaban is also a substrate for the drug efflux transporter proteins, p-glycoprotein (P-gp) and breast cancer resistance protein (BCRP). Coadministration with a strong inducer (rifampin) decreased apixaban concentrations by ~50% (**Figure 6**). Coadministration with strong inhibitors of CYP3A4 and P-gp (ketoconazole) increases apixaban AUC by 100% (or 2-fold).

Reviewer comment: The reviewer agrees with the Clinical Pharmacology reviewers recommendations in the figure.

Pertinent negatives are that apixaban is not a substrate for the key transporters, MRP, OATP1B1, OATP1B3, OATP2B1, OAT1, and OAT3. Apixaban is unlikely to inhibit CYP enzymes nor P-gp. It is not an inducer of CYP enzymes either.

Figure 6 Effect of drug interaction on apixaban PK



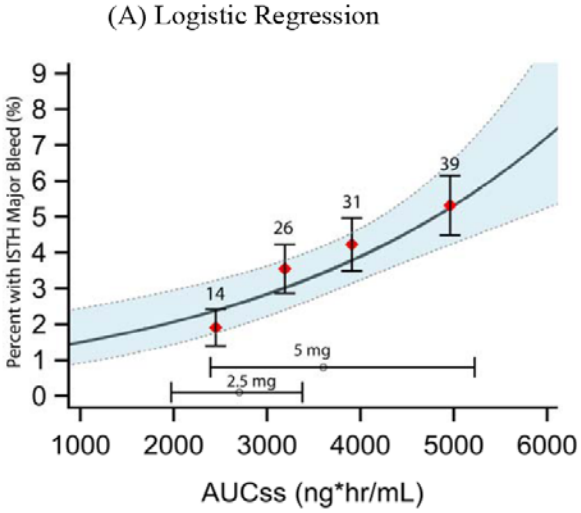
Source: Clinical pharmacology review, p. 27

4.4.4 Exposure-Response Modeling

The clinical pharmacology reviewers modeled the relationship between apixaban exposure and response, specifically ischemic stroke and ISTH major bleed. The probability of ischemic stroke was independent of apixaban concentration (Source: Clin Pharm review, Section 4.4 of pharmacometric review, p. 180-181). These analyses (and subsequent interpretation) are limited because of the small number of subjects with ischemic strokes (n=27) and apixaban concentration.

The clinical pharmacology reviewer's exposure response (ER) models for ISTH major bleed show that there is a direct relationship between exposure and ISTH major bleed (**Figure 7** shows logistic regression). Their Cox proportional hazards model (Figure 3B, p. 12 of Clin Pharm review) did not adjust for the covariates in each dose group, so the figure is not shown.

Figure 7 Probability of ISTH major bleed and apixaban concentration



Source: Clinical Pharmacology review, Figure 3, p. 12. The solid line represents the predicted probability from the logistic regression; shaded blue area is the 95%CI. The red diamonds represent the observed probability at the median AUC for a given quartile. The number above each quartile is the number of subjects with an ISTH major bleed. Apixaban treated subjects that contributed to the concentration data=2932; of these 110 had an ISTH major bleed (7 were hemorrhagic strokes).

The Cox proportional hazards model used to create the predictions for risk of ISTH major bleed in one year did adjust for the covariates in each dose group. The model predictions were close to the actual annual event rates in ARISTOTLE (This is discussed in more detail in **Sec 7.3.2.1.6 Apixaban 2.5 mg dose**).

5 Sources of Clinical Data

5.1 Tables of Studies/Clinical Trials

Table 13 Clinical Trials Supporting Safety and Efficacy of Apixaban for use in Patients with AFib

STUDY	Indication	Goal	Phase
STUDIES SUPPORTING EFFICACY			
CV185030 (ARISTOTLE)	AFib – stroke/SE prevention	To demonstrate NI of apixaban to warfarin in preventing stroke and SE in patients with AFib	3
CV185048 (AVERROES)	AFib - stroke/SE prevention	To demonstrate superiority of apixaban to aspirin in preventing stroke and SE in patients with AFib	3
STUDIES SUPPORTING SAFETY			
ARISTOTLE	(see above)	(see above)	3
AVERROES	(see above)	(see above)	3
CV185068 (APPRAISE 2)	ACS	To determine if apixaban is superior to placebo in patients with recent ACS receiving SOC in preventing MACE	3
STUDIES SUPPORTING DOSING REGIMEN			
CV185010	VTE prevention post knee replacement	To determine dose response of apixaban in preventing thrombotic events after knee replacement`	2
CV185017	Treatment of VTE	To assess safety, efficacy vs. conventional treatment and optimal dose for Phase 3 of apixaban in the treatment of DVT	2

5.2 Review Strategy

The clinical review is split between one reviewer focusing on efficacy and one reviewer focusing on safety. The reviews are combined in this document.

The efficacy review focuses primarily on the ARISTOTLE trial, the only Phase 3 warfarin-controlled trial performed by the Applicant intended to evaluate the clinical efficacy of apixaban in preventing stroke and SE in patients with non-valvular AFib or AFl. Besides this trial, the Applicant conducted the AVERROES trial. This is an aspirin-controlled trial of apixaban in patients with AFib in patients who failed or were deemed unsuitable for VKA therapy. This trial is merely supportive of efficacy; its results will be summarized and selectively used in support of the ARISTOTLE trial. The design features of both these trials are described in Section **5.3.2**; this section also includes the efficacy results of AVERROES.

The results of ARISTOTLE and AVEROES were not pooled by the Applicant for the ISE or ISS. The efficacy results of ARISTOTLE stand alone in Section **6**. The data supporting the dose of apixaban used in ARISTOTLE and AVERROES, which come primarily from a DVT prevention dose ranging trial, are also discussed in Section **6**.

The safety review also focuses primarily on ARISTOTLE, and those data alone permit a substantive review. For rare SAEs, the reviewer also analyzed the data in AVERROES, and for information on bleeding while on concomitant antiplatelet therapy, the reviewer analyzed the data in APPRAISE-2, a study in ACS patients that was stopped early because the benefit did not outweigh the risk of TIMI major bleeding.

5.3 Discussion of Individual Studies/Clinical Trials

The evidence for the efficacy of apixaban in the prevention of strokes and SE in patients with non-valvular AFib comes primarily from the Applicant's global study No. CV185030, "A Phase 3, Active (Warfarin)-Controlled, Randomized, Double-blind, Parallel Arm Study to Evaluate Efficacy and Safety of Apixaban in Preventing Stroke and Systemic Embolism (SE) in Subjects with Non-valvular Atrial Fibrillation (AF)" (ARISTOTLE).

5.3.1 ARISTOTLE

5.3.1.1 Study Design and Objectives

ARISTOTLE was a randomized, parallel-group, active-controlled, double-blind, multicenter, event-driven non-inferiority trial comparing warfarin titrated to the target INR range (2.0 to 3.0) vs. fixed dose apixaban given twice daily, using a classic double-dummy design to maintain the blind. The primary objective was to determine whether the efficacy of apixaban is non-inferior to that of dose-adjusted warfarin for the prevention of the composite of stroke and SE in subjects with non-valvular atrial fibrillation.

5.3.1.2 Geographic Scope

ARISTOTLE was conducted at 1053 sites in 40 countries. There were 1034 sites that randomized at least one patient, including sites on each of the 6 continents with permanent residents (i.e., all continents except Antarctica).

For administrative purposes and for many analyses, the countries where the trial was conducted were organized into 4 regions – North America, Latin America, Europe, and Asia Pacific. For a few analyses, Europe was split into Eastern and Western regions. For a breakdown of countries and regions, see **Attachment 1**.

5.3.1.3 Study Duration/Dates

The protocol anticipated that patients who survived and did not drop out would be treated and followed for as long as 60 months, based on the estimated time to reach the event target with 18,000 randomized patients. The study's dates of first and last patient

visits were 19 December 2006 and 25 May 2011. The database was locked on 10 June 2011. The study report date was 25 August 2011.

The study was planned to end after the event target of 448 adjudicated primary endpoint events was reached. Late in the study, it was estimated that the target would be reached on January 30, 2011. This date was set as the end of the "intended treatment period," which marked the cutoff for the primary efficacy endpoint analysis. At this time, the sites were notified to schedule end of study visits. There was no open-label roll-over period.

5.3.1.4 Patients

Patients who met each of the inclusion criteria below could enroll:

1. Men or women aged ≥18 years with non-valvular AFib or AFl not due to a reversible cause
2. AFib or AFl was to be documented by ECG evidence –
    a. at the time of enrollment (i.e., screening) OR
    b. on two occasions at least 2 weeks apart in the 12 months prior to enrollment
3. At least one of the following risk factors for stroke was present –
    a. Age 75 years or older
    b. Prior stroke, transient ischemic attack (TIA) or SE
    c. Either symptomatic congestive heart failure within 3 months or left ventricular (LV) dysfunction with an LV ejection fraction (LVEF) ≤40% by echocardiography, radionuclide study or contrast angiography
    d. Diabetes mellitus
    e. Hypertension requiring pharmacological treatment.
4. Patient provided signed written informed consent

Reviewer Comment: These risk criteria would lead to a population with all subjects having a $CHADS_2$ score of at least 1. The risk criteria are similar to those of RE-LY (a study of dabigatran) but less stringent than those of ROCKET (a study of rivaroxaban), in which most patients had a $CHADS_2$ score of 3 or more.

Patients who met any one or more of the following criteria were excluded:

- AF due to reversible causes (e.g., thyrotoxicosis, pericarditis)
- Clinically significant (moderate or severe) mitral stenosis
- Increased bleeding risk that was believed to be a contraindication to oral anticoagulation (e.g., previous intracranial hemorrhage)
- Conditions other than atrial fibrillation that require chronic anticoagulation (e.g. prosthetic mechanical heart valve)

- Persistent, uncontrolled hypertension (systolic blood pressure [SBP] >180 mm Hg, or diastolic BP [DBP] >100 mm Hg)
- Active infective endocarditis
- Planned major surgery
- Planned AF or flutter ablation procedure to be performed
- Use of an unapproved, investigational drug or device within the past 30 days
- Required treatment with aspirin > 165 mg/day
- Simultaneous treatment with both aspirin and a thienopyridine (e.g., clopidogrel, ticlopidine)
- Severe comorbid condition with life expectancy of ≤ 1 year
- Active alcohol or drug abuse, or psychosocial reasons that make study participation impractical
- Recent ischemic stroke (within 7 days)
- Severe renal insufficiency (serum creatinine > 2.5 mg/dL or a calculated creatinine clearance < 25 mL/min)
- ALT or AST > 2X ULN or a Total Bilirubin ≧ 1.5X ULN (unless an alternative causative factor [e.g., Gilbert's syndrome] is identified
- Platelet count ≤ 100,000/ mm^3
- Hemoglobin < 9 g/dL
- Inability to comply with INR monitoring
- Prior randomization into an apixaban clinical study
- Prisoners or subjects who are involuntarily incarcerated
- Subjects who are compulsorily detained for treatment of either a psychiatric or physical (e.g., infectious disease) illness
- Women of child bearing potential (WOCBP) unwilling or unable to use an acceptable method (defined in the protocol) to avoid pregnancy, or those who were pregnant or breastfeeding at enrollment, or became pregnant prior to the first dose of study drug. :

Reviewer comment: The pivotal studies of dabigatran and rivaroxaban did not enroll patients with atrial flutter, so this trial is unique among trials recently reviewed by the Division for the target indication. However, current guidelines for treatment of atrial fibrillation suggest that patients with atrial flutter should receive similar treatment to prevent thrombotic events as those with AFib.[3] For efficacy results in patients with atrial flutter, see **Table 53** *and related discussion.*

5.3.1.5 Randomization and Treatments

After meeting the study enrollment criteria, eligible subjects were randomized to treatment with apixaban or warfarin in a 1:1 ratio. A telephonic IVRS system (available at all hours world-wide) was used for randomization, which was stratified by site and prior VKA status (naïve or experienced). IVRS inputs included site no., patient name and number, age, prior VKA status, serum creatinine, and weight. VKA naïve patients

were defined as those who had not previously received warfarin or another VKA or had received ≤ 30 consecutive days of treatment with warfarin or another VKA in the past.. Other patients were defined as experienced. There was a target of at least 40% VKA naïve patients at each site, which was enforced through the IVRS system. If the number of VKA-experienced patients was more than 60% of the number randomized at any site +2 patients, another VKA experienced patient could not be randomized by the IVRS system. This rule was implemented during the first year of enrollment and maintained until the end of enrollment.

In patients taking VKA at study entry, VKA therapy was discontinued prior to randomization. Randomization and study drug initiation occurred when the INR was < 2.0.

A classic double dummy design was employed. Subjects in the apixaban arm received placebo for warfarin, and subjects in the warfarin arm received placebo for apixaban. Apixaban or matching placebo was provided as 2.5 or 5 mg plain red-brown oval tablets; the two strengths noticeably differed in size. Warfarin or matching placebo was supplied as 2 mg lavender round scored tablets only. However, as discussed in the comment below, this reviewer believes there was a blinding issue.

> *Reviewer Comment: We were provided with samples of study drug, including bottles of active and placebo for apixaban 5 mg and warfarin 2 mg. (We only received one bottle of apixaban 2.5 mg tablets, so no comparison between drug and placebo was possible) For apixaban 5 mg, the placebo tablets were somewhat thicker than the actives (3 people agreed). For warfarin the active tablets were slightly thicker than placebo (3 people agreed); the difference was not as marked as for apixaban 5 mg. We did not know if we would have gotten the same results with other bottles of tablets.*
>
> *Following a discussion of this issue with the Applicant, we received a submission that downplayed the importance of the difference in thickness between the placebo and active 5 mg apixaban tablets, but acknowledged its existence. The Applicant argued that individual patients would be exposed only to active or placebo (assuming no dispensing errors at the site) and would ordinarily never have the opportunity to appreciate the difference. While that might be true, site personnel handled active drug and placebo tablets on a routine basis when subjects returned their unused tablets to be counted. The Applicant did not address the issue of unblinding of site personnel. Site personnel could have become aware of a difference in thickness of apixaban 5 mg tablets. If for some reason the results of an open INR measurement then became known, the site might become completely unblinded to the assignment of all patients at the site.*
>
> *The Applicant supplied additional information on tablet thickness from various lots, including samples, and this information was reviewed by the CMC team. They concluded: "Based on evaluation of the samples, there was a slight*

difference in thickness of placebos vs. actives but it was within operating parameters of the Applicant's manufacturing process, and adequate care was taken to match the placebos with the corresponding active tablets."Because of this conclusion, and because we have no evidence that the differences in thickness between active and placebo tablets actually led to unblinding, we do not believe this is an issue of regulatory importance.

The usual apixaban dose was 5 mg po twice daily. However, there was a dose modification, applicable only at the time of randomization, for risk factors for bleeding. Any person who had "any two" of the three defined risk factors for bleeding was randomized to apixaban (or matching placebo) at a dose of 2.5 mg po bid. The risk factors for bleeding were:

- Age ≥ 80 years
- Body weight ≤ 60 kg
- Serum creatinine ≥ 1.5 mg/dL

Reviewer Comment: The Applicant has confirmed that "any two" risk factors was intended to mean any two or more risk factors. This was communicated to the sites at investigator meetings.

The dose of apixaban was otherwise fixed and not dependent on any measured coagulation parameters or changes in bleeding risk factor status after randomization.

Warfarin dose was titrated to a target INR of 2.0 to 3.0, as discussed below.

5.3.1.5.1 Warfarin Dosing Based on INR Measurements

During the study INR was to be measured using a point-of-care (POC) device provided to the site. The device and associated procedures were designed to minimize the likelihood of unblinding based on INR data. After analyzing a blood sample, this device displayed a coded INR value. The site then provided the coded INR to the IVRS system by phone along with other patient information, including patient number and the last 3 warfarin doses. The IVRS system decoded the INR input and provided a response to the site. For warfarin arm patients, the true INR value was provided to the site, ranging from 0.8 to 9.9, the limits of detection of the device. For apixaban arm patients, the site was provided with the true INR if the true INR was > 4.1. Otherwise, a sham INR was provided. The sham value was in part dependent on the last 3 warfarin/placebo doses entered into the system to seem realistic. True INR values <0.8 or > 9.9 for patients in either arm generated requests for additional blood samples to be sent to the central laboratory for testing. INR results were provided to the site by phone (via computer-generated speech) and also sent within 24 hours via the site's preferred written contact mode (fax and/or email).

Reviewer Comment: The rationale for providing true INR values >4.1 for apixaban arm patients is not stated. However, it might have protected patients somewhat from medication errors (i.e., dispensing of active warfarin instead of placebo for warfarin), which were not rare in this study.

The study documents included dosing guidelines for the initiation of warfarin therapy. The protocol recommended the following --

- For subjects who taking warfarin at baseline, the recommendation was to begin study therapy at the patient's previous dose, if known.
- For those who were not taking warfarin at baseline or those whose warfarin dose was not know, the following was recommended:
  - In subjects < 80 years old, initiate warfarin at a dose of up to 6 mg/day, with INR testing on day 3 or 4.
  - In subjects ≥ 80 years old, initiate warfarin at a dose of up to 4 mg/day, with INR testing on day 3 or 4.

The sites were also provided with a "Guidance for the Use and Dosing of Warfarin for Sites and Investigators" that was prepared for this trial (**Attachment 2**). A warfarin initiation nomogram was provided in this guidance, but not one for maintenance dosing. The guidance recommendations are both considerably more detailed and not entirely congruent with those in the protocol. The investigators were told that the protocol should be the preferred source of recommendations.

In addition to the guidance document, the sites received a slide rule-like device that provided guidance on warfarin maintenance therapy. This device was a special edition of the commercially available "Anticoagulation-Advisor®" (see **www.anticoauglation-advisor.com**.) The input to the device is the current weekly warfarin dose. The outputs are 10% increases and decreases from the current daily dose and a recommended dose for each day of the week to achieve a given total weekly dose. A nomogram printed on the front of the device provides advice on adjustments to the total weekly warfarin dose, based on the most recent INR value (see **Table 14**). Except in the case of an INR value in the range of 4.1 to 5, no recommendations are provided with regard to when the next INR should be drawn.

Sites could use other warfarin dosing nomograms at their discretion. We were informed verbally that one dosing algorithm that was used by some sites was an online algorithm at www.warfarin.dosing.org.

Table 14 Maintenance Dosing Nomogram of ARISTOTLE "Slide Rule"

INR Value	Recommendation
< 2	Increase weekly warfarin dose ~ 10%
2 - 3	No change in weekly dose
3.1 - 4	Decrease weekly dose ~ 10%
4.1 – 5	Hold warfarin until INR ~ 3 Recheck INR in 1-3 days Decrease dose 10-20%
> 5	Consult protocol and Warfarin Guidance Manual

The protocol recommends that INR should be obtained twice a week for two weeks (on Days 4, 7, 10 and 14), then once a week for two weeks, and monthly thereafter.

The protocol, guidance and "slide-rule" were advisory with respect to warfarin dose; the investigator retained discretion as to what warfarin dose to use and whether additional INRs should be obtained following changes in warfarin dose. The Applicant monitored frequency of INR measurements and had discussions about appropriate time of INRs in cases where out of range patients were not being brought back for INRs in a reasonable time. However, no sites were dropped for failure to manage INR appropriately.

With the exception of a screening INR in VKA experienced patients, INR data was not captured in the case record. However, study INR data were downloaded electronically from the IVRS provider to BMS and became part of the study database. Both the sham INR that were reported to the sites for apixaban arm patients and true INR data (for patients in both arms) were captured.

5.3.1.5.2 Duration of Treatment

Except has provided below, treatment with blinded study drug was to continue until the end of the study, which was to occur following attainment of the target number of endpoint events. Patients could withdraw from treatment at their discretion, but would have been followed up as described in Sec. **5.3.1.7** unless they specifically withdrew from follow-up. For information on the last study visit and follow-up of patients ending blinded study drug, see Section **5.3.1.7.1**.

The protocol indicated that double-blind treatment was to be discontinued for the following reasons (reasons that do not involve the discretion of the investigator or patient are underlined):

- Withdrawal of informed consent (subject's decision to withdraw for any reason)

- Any clinical AE, laboratory abnormality or intercurrent illness which, in the opinion of the Investigator, indicates that continued participation in the study is not in the best interest of the subject
- Clinical jaundice is present for a subject at any time
- If ALT ≥ 5 x ULN on any two consecutive occasions
- Total bilirubin ≥ 2.0 x ULN on any two consecutive occasions in the absence of an alternative causative factor [e.g., Gilbert's syndrome] is identified
- Pregnancy
- Termination of the study by Applicant
- Loss of ability to freely provide consent through imprisonment or involuntary incarceration for treatment of either a psychiatric or physical illness (e.g., infectious disease)

There were to be continued attempts throughout the duration of the trial, if clinically appropriate, to resume study medication for any subject who had study drug discontinued. If, in the judgment of the investigator, the subject could not continue to receive study treatment, or if the subject withdrew consent, then continued follow-up was to be pursued with the subject, the subject's family or designated representative to ascertain the subject's vital status.

A number of concomitant medications were prohibited during study treatment, including ASA > 165 mg/day, "potent" CYP314 inhibitors (examples: itraconazole, ketoconazole, clarithromycin, telithomycin, ritonavir, indinavir, nelfinavir, atazanavir, saquinavir, and nefazadone), other antithrombotics (e.g., UFH or LMWH (unless used as part of a bridging strategy), direct thrombin inhibitors, or fondaparinux), and GP IB/IIIa inhibitors (e.g., abciximab, etifiatide, tirofiban. If treatment with any of these was required, study drug should have been discontinued, but should have been restarted as soon as possible following discontinuation of the prohibited therapy.

The administration of the following agents in subjects on study drug should be done cautiously given the increased risk of bleeding. In such cases, consideration of interruption of the study drug may be warranted; this decision should be made after a careful assessment of the risks and potential benefits.

- Concomitant (simultaneous) use of both aspirin (≤ 165 mg/day) and a thienopyridine (e.g., clopidogrel, ticlopidine)
- Chronic (> 3 months) daily NSAIDs
- Cytotoxic/myelosuppressive therapy

In addition, if a subject received an agent that is a potent inducer of CYP3A4 (e.g., rifampin), the investigator was to carefully evaluate that subject's risk of thromboembolism, as the plasma concentration of apixaban may be lower than that in subjects not receiving a potent inducer of CYP3A4. No further instructions were provided.

Reviewer Comment: Increase in the dose of study drug was not allowed. If there was concern about thromboembolism in the setting described above, the logical course would have been to d/c study drug and begin open label anticoagulation, perhaps with LMWH or some other fast-acting agent.

5.3.1.5.3 Special Dosing Procedures

See **Attachment 4** for information on dosing instructions for such events as invasive procedures and cardioversion.

5.3.1.5.4 Switching from blinded study drug to open label standard of care

The following procedures were advised:

The subject was to take the morning dose of blinded apixaban (apixaban/placebo), on the morning of the final treatment visit (FTV). The blinded warfarin study medication containers were collected at the FTV, and the bottle of blinded apixaban study drug was redispensed to the subject with four tablets of the blinded apixaban study drug inside. The investigator prescribed open label warfarin (or VKA). If the subject had a stable INR during the months prior to FTV, it was suggested that it would be reasonable to consider the recent dosing schedule of blinded warfarin as a starting point for the open label warfarin dose; but the final decision regarding warfarin (or VKA) dose rested with the investigator.

Switching is accomplished using an "apixaban bridge" by the following schedule, (Switch Day 1 is the day of the FTV):

- Switch Day 1 (PM): Subject takes open label warfarin (or VKA) dose AND one blinded apixaban tablet.
- Switch Day 2 (AM): Subject takes one blinded apixaban tablet.
- Switch Day 2 (PM): Subject takes open label warfarin (or VKA) dose AND one blinded apixaban tablet.
- Switch Day 3 (AM): Subject takes one blinded apixaban tablet.
- Switch Day 3 (PM): Subject takes open label warfarin (or VKA) dose.

While the above procedure was recommended, the protocol states that "Investigators, at their discretion, may choose to switch subjects to open label warfarin (or a VKA) without overlapping with blinded apixaban, or they may choose to use low molecular weight heparin or unfractionated heparin bridging instead in appropriate cases, or they may chose to switch the subject to another antithrombotic drug (e.g. aspirin or a novel oral anticoagulant) that is approved for this use in their country based on local standards of care.

5.3.1.6 Blinding

See Sec. **5.3.1.5** for a discussion of the IVRS randomization procedure.

The investigator was not openly provided with randomization codes. However, the identity of the blinded study medication (active drug or placebo) was provided on a scratch-off panel on each bottle of study drug. The scratch-off portion of the label was to be separated from the bottle when it was given to the patient and affixed to a sheet that was specific for each patient but not part of the case record. These sheets were ordinarily not collected by the Applicant. The bottles were provided to the patient and then returned each month. The bottles were openly labeled as either BMS-562247-01(apixaban) 2.5 mg or placebo, BMS-562247-01 5mg or placebo, or warfarin 2mg or placebo.

We inspected bottle samples. The codes could not be read without peeling and scratching the label; if scratched, the identity of the investigational product would have been obvious to any observer. However, the sites were not asked to provide the returned bottles to study monitors.

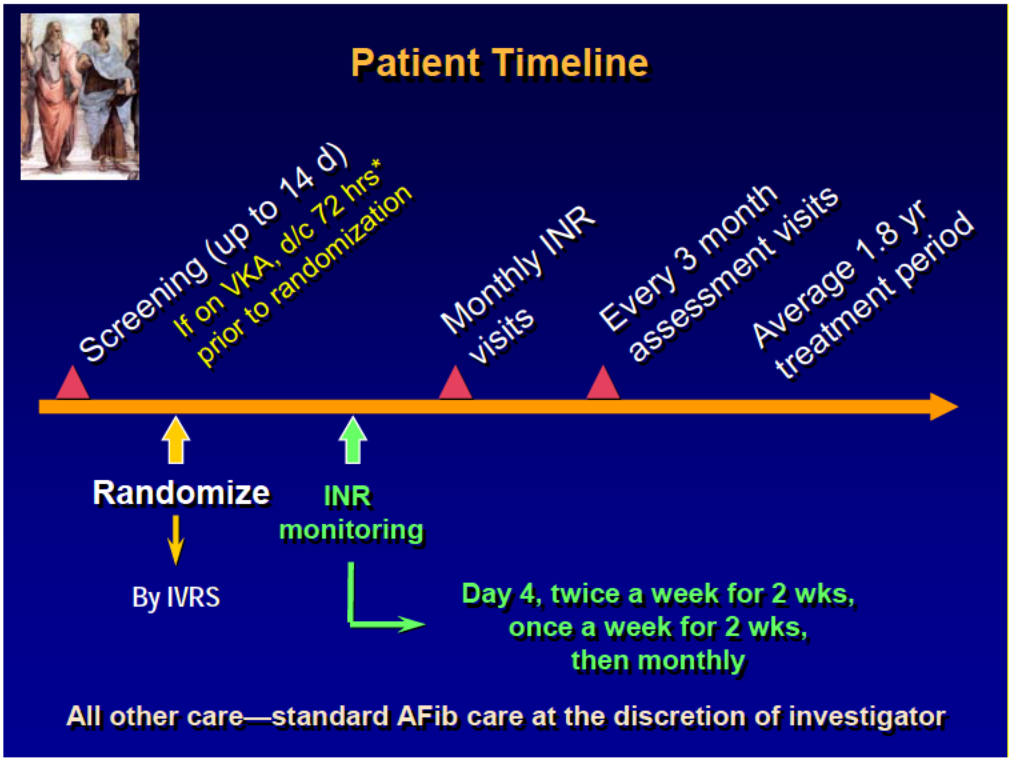
> *Reviewer Comment: Because the bottles were not provided to the monitors, undocumented code breaking could have occurred.*

5.3.1.7 Study Plan and Procedures

The study was divided into a screening period, a (double-blind) treatment period that closed with the final treatment visit (FTP), and a post-treatment observation period. At the FTP, subjects could be transitioned from study drug to an open-label VKA or other appropriate therapy. At the end of the post-treatment observation period, a follow-up visit occurred. For patients who completed the study on treatment, this was a telephonic visit and was planned as the last contact with the subject. **Figure 8** (based on an Investigator Meeting slide) is a simple schematic of the trial plan. Note that some patients were in the trial for over 4 years.

All randomized subjects were to be followed until the study end trigger (the occurrence of 448 adjudicated primary endpoint events) and the subsequent procedures, even if they did not ever take study drug or prematurely discontinued study drug. Efforts were to be made to contact any subjects lost to follow-up and collect information on the occurrence of efficacy endpoint events and the reason for discontinuation.

Figure 8 Study flow diagram



5.3.1.7.1 Study Visits and Information Collected

Screening procedures were to be performed after obtaining informed consent and within 30 days of randomization. Patients determined to be eligible for the study on the basis of screening procedures were asked to return for the Baseline (Day 1) visit, when randomization and dispensing of study drug were to occur. However, if the patient was taking a VKA at baseline, the relevant procedures in Section **5.3.1.5** were followed prior to randomization.

In general, during the double-blind treatment period, there were 2 types of visits: INR visits and assessment visits. The time points for these visits are detailed in **Table 15**. These visits included, but were not limited to, the following assessments.

INR visits (see **Figure 8** for timing):

- Perform POC testing for INR
- Obtain pharmacogenomics sample (should be drawn at the Month 2 visit, however, it could be drawn at any scheduled lab collection visit after randomization.)
- Obtain laboratory tests for assessment of LFT and CK (at Months 1 and 2 visits only)

- Obtain urine pregnancy test (at all monthly visits)
- Assess for outcomes (death, stroke, systemic embolism, myocardial infarction, bleeding)
- Assess changes in concomitant medication use (at Month 1 visit only)
- Assess for AEs
- Collect Study Medication
- Assess Study Medication use
- Dispense Study Medication

Quarterly Assessment Visits (Months 3, 6, 9, 15, 18, 21, 27, 30, 33, 39, 42, 45, 51, 54, and 57):

- Obtain vital signs
- Obtain clinical laboratory tests including Hematology and Chem 7 panels
- Obtain clinical laboratory tests for assessment of LFT and CK (at the Months 6, 18, 30, 42 and 54 visits only)
- Obtain urine pregnancy tests
- Perform POC testing for INR
- Assess for fractures (at the Months 6, 18, 30, 42 and 54 visits)
- Assess for AEs
- Assess for outcomes (death, stroke, systemic embolism, myocardial infarction, bleeding)
- Assess changes in concomitant medication use
- Collect Study Medication
- Assess Study Medication use
- Dispense Study Medication

Annual Assessment Visits (Months 12, 24, 36 and 48):

- Obtain same information as Quarterly Assessment Visits, and -
- Obtain 12 Lead ECG
- Obtain physical measurements including weight and hip and waist circumference only

Table 15 Time and Events Schedule

Procedure	Screening	Day 1	Day 4, Week 1, Day 10, Week 2, Week 3	Months 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23, 25, 26, 28, 29, 31, 32, 34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52, 53, 55, 56, 58, 59 (INR)	Months 3, 6, 9, 15, 18, 21, 27, 30, 33, 39, 42, 45, 51, 54, 57	Months 12, 24, 36, 48	Final Treatment Visit [a]	Follow-up [b]	Protocol Section
Eligibility Assessments									
Informed Consent	X								3.3
Inclusion/Exclusion Criteria	X	X							4.2
Medical History	X								6.1.3
Safety Assessments									
Physical Examination	X								6.3.5
Height, Weight, Hip, Waist	X					X[c]			6.3.6
Vital Signs: BP, HR	X	X			X	X	X		6.3.3
12-Lead Electrocardiogram	X					X	X[d]		6.3.4
Fracture Assessment	X				X[e]	X	X		6.3.7
Adverse Events Assessment		X	X	X[f]	X	X	X	X[g]	7
Outcome Assessment		X	X	X	X	X	X	X	6.3.1, 6.4.1, 6.4.2
Hematology[h]	X	X			X	X	X		6.3.2
Chemistry Panel[h]	X[i]	X[i]			X[j]	X[i]	X[i]		6.3.2
LFT and CK Only[h]				X[k]	X[k]				6.3.2
Biomarkers		X							6.6
INR (Central Lab)[h]	X								6.3.2

Table 15 continued

Procedure	Screening	Day 1	Day 4, Week 1, Day 10, Week 2, Week 3	Months 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23, 25, 26, 28, 29, 31, 32, 34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52, 53, 55, 56, 58, 59 (INR)	Months 3, 6, 9, 15, 18, 21, 27, 30, 33, 39, 42, 45, 51, 54, 57	Months 12, 24, 36, 48	Final Treatment Visit [a]	Follow-up [b]	Protocol Section
INR (POC)[l]			X	X	X	X			4.1
Urinalysis[h]	X	X				X	X		6.3.2
Urine Pregnancy	X	X		X	X	X	X		6.3.2.1
Genomic Blood Draw				X[m]					6.7
Clinical Drug Supplies									
Randomization		X							5.2, 6.1.4
Dispense Study Treatment		X	X	X	X	X	X[n]		5.3
Assessment of Study Medication Use			X	X	X	X	X		6.1.5
Assessment of Concomitant Med Use	X	X		X[f]	X	X	X		5.5.1, 6.3.8
Collect Study Medication			X	X	X	X	X[n]		6.1.5

Procedures for Discontinuation of Study Drug

Procedures were specified for discontinuation of study drug at the end of the study as well as for early discontinuation of study drug.

The final treatment visit (FTV) was to be scheduled as soon as possible and no later than five weeks after the end of treatment period date (EOTP date). The latter was the estimated date of attainment of the event target of 448 primary endpoint events, which was January 30, 2011. The FTV was the final in the double-blind treatment period for subjects on study drug at that time.

Investigators were encouraged, but not required, to transition patients to open-label anticoagulant therapy at the FTV and to use an apixaban bridging strategy (see Sec. **5.3.1.5.4**).

Following the FTV, visit there was to be a 30 day observation period to follow subjects after transition from study drug to open-label VKA or other appropriate therapy. In addition to ad hoc return visits to assess INR control (scheduled at the investigators' discretion), completing subjects had a telephonic "follow-up visit" approximately 30 days after the permanent discontinuation of study drug. For subjects who completed the scheduled double-blind treatment period, this was the final subject contact.

Subjects who had prematurely discontinued study drug (or who were planning to discontinue) were to have the FTV, and if possible, continue taking study drug up to the FTV. At this visit, they were to be started on open-label VKA treatment or other appropriate therapy using the same procedures as those patients who completed the study, including visits for INR measurements. After the ESMDV, patients were to have a final in person site visit 30 days later. They were then followed up by phone every 12 weeks for the occurrence of efficacy endpoints until the EOTP date.

The following table summarizes planned study drug discontinuation and end of study procedures.

Table 16 ARISTOTLE – Early Termination and End of Study Procedures

Patients with Early Termination of Study Drug	**Patients who Completed the Study**
Decision by subject or investigator to terminate study drug ↓ ↓	Sponsor notifies sites that the target number of primary adjudicated primary endpoints have occurred (end of treatment period (EOTP) date), triggering the end-of-study procedures ↓
If possible, continue study drug to Final Treatment Visit (FTV) ↓	Site schedules FTV to occur within 5 weeks after EOTP date. Subjects are to continue study drug to FTV. ↓
FTV Transition to standard of care medication or other treatment as appropriate. Apixaban bridge may be used to transition to VKA (but protocol amendment describing bridge occurred late in study). ↓	FTV Transition to standard of care medication as appropriate. Apixaban bridge may be used to transition to VKA. ↓
Other discretionary visits for monitoring anti-coagulation therapy. Follow up site visit in 30 days (intended as the last in-person site visit). ↓	Other discretionary visits for monitoring of anti-coagulation therapy. Follow up telephone call in 30 days (intended as the final planned contact). ↓
Phone contacts q 12 weeks until the end of the study. Upon "EOTP date" (see event No. 1 in the next column), a final phone contact is made. ↓	**Efficacy endpoint information to be collected through final contact**
Efficacy endpoint information to be collected through final contact	

5.3.1.8 Efficacy Endpoints

5.3.1.8.1 Primary Endpoint

The primary efficacy outcome was time to the first occurrence of the composite of stroke (all types) and non-CNS systemic embolism. Adjudicated results were to be used. .

5.3.1.8.2 Secondary Endpoints

Secondary efficacy endpoints included to the first occurrence of

- ischemic stroke or stroke of unspecified type
- hemorrhagic stroke
- systemic embolism
- all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism, major bleeding
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism, all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism, major bleeding, all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism, myocardial infarction, all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism and major bleeding in warfarin naive subjects

All these analyses were to use adjudicated results.

5.3.1.9 Safety Endpoints

5.3.1.9.1 Primary Safety Endpoint

The primary safety endpoint was time to first occurrence of confirmed major bleeding.

5.3.1.9.2 Secondary Safety Endpoints

The secondary safety outcome for this trial is a composite of confirmed major bleeding and confirmed clinically significant non-major bleeding. Other safety outcome measures will also be assessed, and will include minor bleeds, fractures and other AEs as well as abnormal standard clinical laboratory test results.

All major bleeding and clinically relevant non-major bleeding outcomes will be adjudicated by the CEC.

5.3.1.9.3 Safety Procedures

Adverse Events

AE and SAE definitions were generally consistent with the IND safety reporting provisions of 21 CFR Sec. 312.32. However, BMS practice is to exclude certain hospitalizations from the definition of SAE, including those that are –

1) admissions as per protocol for a planned medical/surgical procedure
2) routine health assessment requiring admission for baseline/trending of health status (e.g., routine colonoscopy)
3) medical/surgical admission for purpose other than remedying ill health state and was planned prior to entry into the study. Appropriate documentation is required in these cases
4) admission encountered for another life circumstance that carries no bearing on health status and requires no medical/surgical intervention (e.g., lack of housing, economic inadequacy, care-giver respite, family circumstances, administrative).

Also, all overdoses with study drug were to be considered as SAEs.

Reviewer Comment: These variations from the regulatory definition of SAE are also not problematic.

Adverse events were to be spontaneously reported by the subject or elicited through "open-ended" questioning, but no scripted question(s) was included in the protocol.

Serious adverse events were to be immediately reported (within 24 hours of the investigator's awareness). When required, and according to applicable local law and regulations, serious adverse events were reported to the IRB or Ethics Committee and Regulatory Authorities.

All SAE reports were reviewed by (b) (4) with a primary focus on subjects who experienced serious adverse events of special interest: bleeding events, liver-related events, pancreatitis, hypersensitivity reactions and other potential safety issues (e.g., organ toxicity, renal toxicity).

AE reporting of study endpoints

In ARISTOTLE, the clinical efficacy endpoint events of myocardial infarction, stroke (any type) and non-CNS systemic embolism were to be considered adverse events or serious adverse events. All bleeding events (including CNS bleeds) were to be reported as adverse events or serious adverse events, as appropriate. **However, none of the events mentioned in the previous two sentences were to be reported as Expedited Safety Reports (ESRs) to IRBs, investigators, or FDA, provided that they were non-fatal.** Fatal events were reported as ESRs, as well as SAEs.

5.3.1.9.4 Monitoring and Evaluation of Liver Function

Liver function tests (LFTs) were drawn at screening, baseline, and then monthly through the final treatment visit. These tests included ALT, AST, total and direct bilirubin, AP, and GGT.

The following management instructions were provided in the protocol:

If at any time during the treatment period a subject's LFTs results show:

- An isolated elevation of either ALT ≥ 3 x ULN **OR** a total bilirubin ≥ 2 x ULN, obtain the following laboratories: ALT, AST, total bilirubin, direct bilirubin, alkaline phosphatase, GGT, CK within one week
- An elevation of **BOTH** ALT ≥ 3 x ULN **AND** total bilirubin ≥ 2 x ULN, obtain the following laboratories: ALT, AST, total bilirubin, direct bilirubin, alkaline
- phosphatase, GGT, CK as soon as possible (i.e., within ≤ 3 days)

If the repeat tests indicate:

- ALT < 3 x ULN and total bilirubin < 2 x ULN, study medication may continue
- ALT ≥ 3x ULN but < 5 x ULN and total bilirubin < 2 x ULN study medication may continue but repeat LFTs weekly until ALT < 1.5 x ULN or to baseline if subjects entered the study with an ALT ≥ 1.5 x ULN
- If the repeat ALT ≥ 3x ULN AND the total bilirubin is ≥ 2 x ULN, study medication must be discontinued unless, in consultation with the BMS Medical Monitor/Trial Helpline, an alternative causative factor (e.g., Gilbert's syndrome) is identified.

Study medication must be discontinued if:

- Clinical jaundice is present for a subject at any time OR
- If ALT ≥ 5 x ULN on any two consecutive occasions OR
- Total bilirubin ≥ 2.0 x ULN on any two consecutive occasions in the absence of an alternative causative factor [e.g., Gilbert's syndrome] is identified

In addition, all subjects with an ALT ≥3x ULN or total bilirubin ≥ 2x ULN will be followed weekly until ALT and total bilirubin return to < 1.5x ULN or to baseline if subjects entered the study with an ALT ≥ 1.5 xULN.

If study medication is discontinued due to elevated ALT or bilirubin, as defined above, inform the Medical Monitor and perform the following:

- Hepatitis screen (anti-HAV, HbsAg, anti-HBc, anti-HBs and anti-HCV)
- Abdominal ultrasound, including liver and hepatobiliary system

5.3.1.9.5 Additional data to be collected

Other type of data were collected in optional substudies at interested sites:

- pharmacokinetics,

- pharmacodynamics (including assessment of hemostasis, inflammation, platelet activation, endothelial dysfunction, and/or cardiovascular disease risk markers at baseline, followed by coagulation system parameters on treatment (including but not limited to, D-dimer, hs-CRP, sCD40L, ADMA, and NTproBNP), and .
- pharmacogenomics, and
- health care economics

5.3.1.10 Endpoint Definitions

The following definitions in the Clinical Endpoint Committee (CEC) charter were used in assessing endpoints. Any event suspected as being one of events described below and all deaths were to be referred for adjudication by the CEC.

Stroke was defined as "the non-traumatic abrupt onset of a focal neurological deficit and lasting at least 24 hours in duration." A retinal ischemic event (embolism, infarction) was to be considered a stroke. It was "strongly recommended" that a CT scan or MRI should be performed for all suspected strokes. Strokes were to be classified as -

1) ischemic,
2) ischemic with hemorrhagic transformation,
3) (primary) hemorrhagic , or
4) uncertain.

Hemorrhagic strokes were to be sub classified as subdural, subarachnoid, or intraparenchymal. Criteria for making the distinctions between the various types of stroke were not provided.

Hemorrhagic strokes, either primary hemorrhage or infarction with hemorrhagic conversion, were considered as major bleeds as well as stroke endpoints. However, the site was not required to complete a bleeding event form.

Transient ischemic attack (TIA) was defined as a non-traumatic abrupt onset of a focal neurological deficit and lasting less than 24 hours in duration.

Stroke and TIA were to be further sub-classified based on whether or not there was imaging evidence of a new cerebral infarction that correlates with the clinical presentation of the subject.

Systemic embolism was defined as a clinical history consistent with an acute loss of blood flow to a peripheral artery (or arteries), which is supported by evidence of embolism from surgical specimens, autopsy, angiography, vascular imaging, or other objective testing.

Myocardial infarction (MI) had a complex definition that took into account clinical context:

. ,

1) For patients without a recent myocardial infarction, an MI was defined by -
   a) Elevation of CK-MB or Troponin T or I ≥ 2 x ULN, or
   b) If no CK-MB or Troponin is available, elevation of total CK ≥ 2 x ULN, or
   c) New, significant (≥ 0.04 s) Q waves in ≥ 2 contiguous leads.

2) For patients who had a recent myocardial infarction, an MI was defined by
   a) Elevation of CK-MB or Troponin T or I to ≥ 2 x ULN (if prior level was normal), or
   b) Re-elevation of CK-MB or Troponin T or I to ≥ 2 x ULN and > 50% above the prior level (if prior level was above normal), or
   c) Re-elevation of total CK ≥ 2 x ULN and > 25% above the prior level (if CK-MB is unavailable), or
   d) New, significant (≥ 0.04 s) Q waves in ≥ 2 contiguous leads and discrete from the prior MI.

3) For patients who had undergone revascularization, an MI was defined by:
   a) Peri-PCI: CK-MB or Troponin T or I (or total CK, if CK-MB and Troponin is unavailable) ≥ 3 x ULN and increased by at least 50% from the level before the procedure or new significant (≥ 0.04 s) Q waves in ≥ 2 contiguous ECG leads.
   b) Peri-CABG: CK-MB (or total CK, if CK-MB is unavailable) ≥ 10 x ULN and increased by at least 50% from the level before the procedure or CK-MB ≥ 5 x ULN and increased by at least 50% from level before the procedure with new significant (≥ 0.04 s) Q waves in ≥ 2 contiguous ECG leads.

Reviewer Comment: "Recent" MI was not defined. Note that definitions in paragraphs 1) and 2) are not consistent with the Universal definition of MI.

Bleeding definitions were extensive. For the primary endpoint, definitions adapted from the ISTH criteria were used.

Major bleeding was defined as a bleeding event that is:

1) Acute, clinically overt bleeding that is accompanied by one or more of the following:
   a) A decrease in hemoglobin (Hgb) of 2 g/dL or more
   b) A transfusion of 2 or more units (U) of packed red blood cells (PRBC)
   c) Bleeding that occurs in at least one of the following critical sites:
      i) Intracranial
      ii) Intra-spinal
      iii) Intraocular (within the corpus of the eye; a conjunctival bleed is not an intraocular bleed)
      iv) Pericardial
      v) Intra-articular
      vi) Intramuscular with compartment syndrome

vii) Retroperitoneal.

2) Bleeding that is fatal.

Clinically relevant non-major bleeding was defined as a bleeding event that is:

Acute or subacute clinically overt bleeding that does not satisfy the criteria for major bleeding and that leads to either:

1) Hospital admission for bleeding
2) Physician guided medical or surgical treatment for bleeding
3) A change in antithrombotic therapy (including study drug) for bleeding

Minor bleeding: All acute clinically overt bleeding events not meeting the criteria for either major bleeding or clinically relevant non-major bleeding were to be classified as minor bleeding.

Bleeding events were also to be classified by the TIMI criteria and GUSTO criteria:

TIMI Bleeding Criteria

TIMI major bleeding:

1) Intracranial bleeding (ICH)
2) Clinically overt bleeding (including bleeding evident on imaging studies) associated with a ≥ 5 gm/dL fall in hemoglobin or a 15% fall in hematocrit from baseline, accounting for the effect of transfusions (1 unit packed red blood cells = 1 gm/dL Hgb = 3% hematocrit Hct]).

TIMI minor bleeding

Clinically overt bleeding (including bleeding evident on imaging studies) associated with a ≥ 3 gm/dL fall in hemoglobin or a 9% fall in hematocrit from baseline, accounting for the effect of transfusions (1 unit packed red blood cells = 1 gm/dL Hgb = 3% hematocrit Hct]).

TIMI minimal bleeding

Clinically overt bleeding (including bleeding evident on imaging studies) not meeting criteria for TIMI minor bleeding.

GUSTO Bleeding Criteria

GUSTO severe bleeding

1) Intracranial bleeding (ICH)
2) Bleeding resulting in hemodynamic compromise requiring treatment

GUSTO moderate bleeding

Bleeding resulting in the need for transfusion.

GUSTO mild bleeding

Bleeding that does not require transfusion or cause hemodynamic compromise

Note: All overt bleeding episodes not meeting the criteria for either major bleeding o clinically relevant non-major bleeding were to be classified as minor bleeding. Minor bleeding was not be adjudicated.

.

Death

Cause of death was to be classified as either cardiovascular or non-cardiovascular. All deaths that were unobserved were to be assumed to be cardiovascular in nature unless a non-cardiovascular cause can be clearly provided.

1) Cardiovascular

Cardiovascular deaths were classified as deaths due to –

(1) ischemic stroke,
(2) hemorrhagic stroke,
(3) systemic embolism,
(4) myocardial infarction,
(5) sudden death,
(6) heart failure, and
(7) other cardiovascular and unobserved deaths.

2) Non-cardiovascular

Non-cardiovascular deaths include deaths due to a clearly documented non-cardiovascular cause. Non-cardiovascular deaths were to be further classified into the categories -:

(1) bleeding,
(2) study drug toxicity other than bleeding,
(3) malignancy,
(4) infection,
(5) trauma, and
(6) pulmonary causes of death.

3) Unknown death

Observed deaths of unknown cause were to be classified as non-cardiovascular deaths.

5.3.1.11 Adjudication of Endpoints by the Clinical Endpoint Committee

An independent Clinical Endpoint Committee (CEC), which operated under a charter, was created to adjudicate the endpoints described below. The CEC was comprised of board certified/eligible physicians from the Duke Clinical Research Institute (DCRI), the Uppsala Clinical Research Institute (UCR, Uppsala, Sweden) and the Brazilian Clinical

Research Institute (BCRI, San Paulo, Brazil).Duke University. CEC physicians were chosen for their clinical expertise in relevant fields, clinical trial experience, availability for the duration of the trial and commitment. Physicians from outside of the Duke community could also be selected for membership. The adjudicated endpoints were:

- Stroke (or TIA)
- Systemic embolism
- Cause of death
- Myocardial infarction
- Major bleeding event
- Non-major clinically significant bleeding event

5.3.1.11.1 CEC structure and responsibilities

The CEC Chair was responsible for presiding over CEC meetings and conference calls, the finalization and dissemination of endpoint criteria, the assurance of quality of the adjudication process through ongoing QC reviews, and participation in the adjudication process.

The CEC Coordinator, from DCRI, played a central in the adjudication process. Among other responsibilities, the Coordinator was to:

- collaborate with the Applicant CEC chair, and others in developing CEC processes and the CEC Charter,
- collaborate with the Applicant in providing the sites with the necessary tools and training to provide the CEC with complete data required for event adjudication,
- train and oversee the day-to-day work of the CEC team members,
- organize and participate in the CEC meetings,
- manage CEC workflow, including the collection of necessary source documents

5.3.1.11.2 Ascertainment of events for adjudication

Events brought for adjudication were identified by review of specified data fields on the eCRF determined to be CEC-critical variables. These variables were described in a CEC "Triggers" document (see **Attachment 5**).

Reviewer Comment: The triggers for stroke included such items as completion of the "suspected stroke" endpoint CRF page or indication that a stroke or TIA had occurred between study visits on the "clinical event assessment" pages for each visit. However, the performance of a head CT or MI, without one of the other named triggers, would not have trigger referral for adjudication of stroke. Likewise, the performance of a leg angiogram, by itself, would not have triggered referral for

adjudication of systemic embolism. Some studies have used such procedural triggers. However, it seems likely that omission of referrals on this basis would be rare because it is very likely that patients with a suspected stroke or SE would have some trigger besides the procedural trigger. Also, there is no evidence that that such missed referrals, if they exist at all, would be biased in favor of apixaban.

Once all eCRF data fields necessary for CEC review were query-clean, the case was ready for adjudication. It was the responsibility of the Coordinator to ensure that records were complete enough for adjudication.

In addition, the CEC could adjudicate events ("CEC identified events") that it discovered during its review of events identified by the computer program.

Both pages from the CRF and source documents were typically reviewed for each type of event.

5.3.1.11.3 Adjudication procedures

Adjudication off all events of the previously named classes that occurred after randomization was performed in "phases". Phase I adjudication for deaths and MIs involved adjudication by one physician (always a cardiologist in the case of Mis). These cases were resolved in Phase I. For stroke, systemic emboli, and bleeding events, Phase I involved two independent physicians, at least one of which was a neurologist for stroke cases. If the two agreed, the case was resolved. If not, the case went to Phase II. Phase II reviews were performed by a committee of at least 3 physicians which operated by consensus. For stroke cases, there was at least one neurologist. However, stroke and systemic embolic events occurring within one month of randomization went directly to Phase II.

A random sample of events underwent QC review by the CEC Phase II committee.

Reviewer comment: These processes seem adequate on their face. While it might have been preferable to adjudicate all hospitalizations, to find efficacy and safety events, the algorithms used in ARISTOTLE appear to be unbiased and sufficiently inclusive.

However, of the adjudication "packages" provided to us for stroke adjudications that were reviewed by the efficacy reviewer had evidence of review by only one physician, and no notes were provided regarding the reasoning underlying the final adjudication. We understand that single document represents the final sign-off. It would have been helpful to have the individual adjudication documents and the rationale for the final decision.

5.3.1.12 Statistical Plan

There were two versions of the statistical plan, Ver.. 1.0 and 2.0. The discussion below describes the final version, Ver. 2.0. For a history of the statistical plan, see Sec. 5.3.1.12.4

5.3.1.12.1 Sample Size

This study was planned to stand alone as support of efficacy of apixaban for the target indication, with a primary analysis involving non-inferiority (NI) to warfarin for the time to the primary endpoint. The statistical assumptions were:

- True risk ratio RR) for apixaban vs. warfarin of 1.0
- 1% loss to follow-up
- Stoke rate of 1.20%/year, with 2.1 years follow-up, and one of the following two statistical requirements -

- NI margin of 1.38 (with 95% confidence interval (CI) of RR) (statistical requirement A)
- (b) (4)) (statistical requirement B)

Note that while FDA was satisfied with requirement A, some regulatory agencies imposed requirement B, which was more stringent due to the wider CI, even though the NI margin for B is higher.

The Applicant calculated that 448 primary endpoint events would provide (b) (4) power to demonstrate non-inferiority under requirement B, with higher power under requirement A.. This was estimated to require randomization of enrollment of about 18,000 subjects.

Reviewer Comment: The precise power under requirement A was not stated in the protocol or statistical plan.

The 1% lost to follow-up rate seems very low, and might have jeopardized the success of the trial due to inadequate power. However, the assumed RR of 1.0 was conservative, and in retrospect, the trial was adequately powered.

5.3.1.12.2 Analysis Plan – Efficacy Variables

The primary time to first event analysis (using adjudicated primary endpoint events of stroke or systemic embolism) was performed in the all randomized patients population followed to the end of the "intended treatment period", .i.e., the expected date of attainment of the event target (the "ITT/ITP" analysis).

For the US, the primary endpoint analysis was based on scenario A described in the previous section: a NI margin of 1.38 for the 95% CI of the RR (apixaban vs. warfarin) for the primary endpoint analysis of time to first event. This was calculated using a Cox proportional hazards model with treatment arm as a covariate, with stratification by investigative site (pooled at the geographic region level) and prior VKA status (experienced or naïve). A sensitivity Cox model analysis was performed that was stratified by investigative site and prior VKA status, with covariates of treatment arm, history of stroke, TIA or non-systemic embolus, history of DM, treatment for hypertension, history of MI, and history of "coronary heart failure."

There were various analyses of efficacy result in the "Evaluable" (per protocol) patient population, using event windows starting with the first dose of study drug and ending 2, 7, or 30 days after the last dose of study drug. The Evaluable population excluded randomized patients who were not dosed or who had or protocol violations that might have affected the outcome of the primary endpoint, including:

- Compliance with apixaban/apixaban placebo < 80%
- Error in treatment assignment resulting in a subject being dosed with an incorrect treatment

In the case of an error or errors in treatment assignment, the patient was censored at the time of the first such error for analyses of the Evaluable patient population.

Using the same methods as the primary analysis, a series of secondary endpoints would be analyzed. These included:

- ischemic stroke or stroke of unspecified type
- hemorrhagic stroke
- systemic embolism
- all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic
- embolism and major bleeding in warfarin naive subjects
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic
- embolism, major bleeding
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), major bleeding,
- all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), systemic embolism, all cause death
- composite of stroke (ischemic, hemorrhagic, or of unspecified type), myocardial infarction, all cause death.

As noted earlier, there was prespecified hierarchy of endpoints:

1) *non-inferiority for the primary endpoint of time to stroke/systemic embolism (with margin of 1.38 for the 95% CI for the US)*
2) *superiority for time to stroke/systemic embolism*
3) *superiority for time to ISTH major bleeding event*
4) *superiority for time to all cause mortality*

There was a formal interim analysis of the primary endpoint planned when half the target number of primary endpoint events had occurred. The DMC could recommend stopping the study if the one-sided p for the primary endpoint analysis was < 0.0001. Termination for futility was also possible. The impact of the interim look on the final p for success in the superiority analysis for the primary endpoint would be a reduction of the one sided p to 0.02499 (from 0.025). The Applicant claims that the impact of the interim look on the non-inferiority analysis was calculated and is essentially 0, and accordingly no adjustment was made to this analysis. The interim analysis was performed and the DMC recommended that the study should continue.

Reviewer Comment: The Applicant's lack of adjustment of the NI analysis for the interim analysis was discussed with Dr. S. Bai of Biometrics, who indicated his agreement with that aspect of the statistical plan.

(b) (4)

owever, during the review the Applicant agreed to repeat several of the evaluable patient analyses using the larger treated population instead; these are reviewed in Sec. **6.1.4** and **6.1.5.**

5.3.1.12.3 Analysis Plan – Safety Variables

The primary safety endpoint was ISTH major bleeding during the Treatment Period, and the specified analysis was superiority for time to the first event using the Cox proportional hazards model including treatment group as a covariate and stratified by investigative site and prior warfarin/VKA status.

Secondary analyses included the time from first dose of study drug to first occurrence of:

- composite of confirmed (per ISTH guidelines) major bleeding and clinically relevant non-major bleeding during the Treatment Period
- all bleeding events reported by the Investigator during the Treatment Period.

A Cox model including treatment group as a covariate and stratified by investigative site and, when applicable, prior warfarin/VKA status was to be used for these analyses.

5.3.1.12.4 History of the Statistical Plan

There were two versions of the statistical plan. Dates of these plans, along with major study milestones, are provided in **Table 17**.

Table 17 ARISTOTLE – History of the SAP and Other Relevant Events

SAP Version or Relevant Event	Date	Comments
First patient randomized	12/19/2006	-
Original SAP – Ver. 1.0	08/01/2007	See text
1st DSMB meeting with data review	11/30/2007	-
Last patient randomized	02/04/2010	-
AVERROES interim analysis performed	02/19/2010	Stopping rule met for efficacy; AVERROES was stopped
SAP Ver. 2.0	05/11/2010	See text
ARISTOTLE interim analysis performed	07/08/2010	DSMB recommended that the study should continue as planned
Database lock	06/10/2011	--

Version 1.0 of the SAP was completed more than 8 months after the first study patient was randomized.

Version 2.0, the final version, was completed almost 3 ½ years after the first patient was randomized, and about two months prior to the only interim analysis. Although there were many changes between Ver. 10 and 2.0, there were few changes with more than a trivial likelihood of affecting approval or labeling. These included:

- Addition of a justification for the sample size increase effected by Protocol amendment 7, which increased sample size from 15,000 to 18,000 to account for the unexpectedly low overall event rate for the primary endpoint.
- Clarification of the definition of the treatment period.
- Changes in the testing strategy to address comments from FDA.
- Changes in provisions for handling stratification by site in the Cox proportional model analysis of the primary endpoint: sites were to be pooled at the level of Geographic Region.
- Addition of BMI, weight, level of renal impairment, and $CHADS_2$ score as baseline subgroups of clinical relevance that will be summarized.
- Adopted Rosendaal method for INR imputation.
- Added last dose + 2 day analysis for per protocol analysis of NI.
- Added formal testing for superiority for major bleeding and all-cause death.
- Specified analysis period for each laboratory value analysis.
- Aligned interim analysis stopping rule with stopping rule developed by DMC (prior to the DMC examining any data).

Reviewer Comment: The SAP was changed to add formal testing for superiority for major bleeding and all-cause death after the trial was fully enrolled with 18,000 patients. These analyses were both successful and each is cited by the Applicant to provide support for a label claim. The timing of this SAP amendment should be considered in the evaluation of the relevant label claims.

5.3.1.13 Study Committees

The study protocol described the following committee structure:

Executive Committee (EC): The EC consisted of members of the academic leadership of the study and one member from each sponsoring company. It was chaired by Lars Wallentin (Uppsala U.) and Christopher Granger (DCRI). The EC collaborated with the sponsor to oversee the design and execution of the study, as well at its statistical analysis. The two chairs were to oversee analytic confirmation of the main study results at their respective institutions.

Operations Committee (OC): The OC was a small group of EC members, including representatives of the sponsor and CRO, tasked with "ensuring that study execution and management were of the highest quality." It met every 2 weeks to discuss and report on the conduct of the study.

Steering Committee (SC): The SC consisted of EC members and the National and/or Regional Coordinating investigators of the trial. They were responsible for the operational aspects of the trial in their geographic areas and advised and assisted the EC. .

Independent Data Monitoring Committee (IDMC): The IDMC was established pursuant to a charter to monitor the progress of the study and ensure that the safety of subjects. The DMC was chaired by Marc Pfeffer, MD. Members included Stuart Pocock as the statistician and 4 other clinicians. The IDMC received unblinded data from an independent statistical contractor. The sponsor provided blinded data to the contractor. to this group.

Reviews of unblinded data were to be conducted on an ongoing basis, at least twice per year. The IDMC sent its recommendations to the SC chair.

Data reviewed periodically included:

- Summary of bleeding events
- Summary of clinical outcomes
    - Strokes and systemic emboli
    - Death / Cause of death

- Myocardial ischemia/MI
- Serious Adverse Events
- Permanent discontinuation of double-blind study drug
- Marked Laboratory abnormalities
- Events associated with cardioversion

Independent Central Adjudication Committee , also known as the Clinical Endpoint Committee (CEC): The composition and functions of the CEC are described above in Section **5.3.1.11.**

The protocol also provided for consultation with blinded external hepatologists (there were 3 in total) who were experts in drug-induced liver injury. Cases referred to these experts included ALT elevations > 3 X ULN and total bilirubin elevations > 2 X ULN (it is not clear if both had to present for referral) or SAEs of hepatitis, hepatic failure and jaundice. The goal was to have two types of blinded hepatologist assessments: (1) individual hepatologist review of specific liver-related cases; (2) periodic collective review during which the hepatologists discuss the cases with each other and produce a consensus assessment of each case (see Section 1.2). Additionally, the hepatologists could be consulted for advice regarding management of ongoing individual liver-related patient situations.

Also, SAEs with preferred terms included in the MedDRA high level terms of "acute polyneuropathies" (acute polyneuropathy, critical illness polyneuropathy, Guillain-Barre syndrome, Miller Fisher syndrome, polyneuropathy), and preferred term (PT) amyotrophic lateral sclerosis (ALS) were further assessed by 3 independent, blinded, external neurologists, who are recognized experts regarding peripheral nerve disease.

5.3.1.14 Protocol Amendments

There were a series of protocol amendments. Note that the discussion above describes the final protocol (Ver. 4, dated 8/4/2010) Amendments applicable to the US are described below.

Table 18 ARISTOTLE Protocol Amendments Applicable to the US

Amend-ment No.	Date	Description
1	11/4/2006	Provided for obtaining blood samples at month 2 for voluntary pharmacogenomic testing (the genomic markers to be studied were not specified). Patients were to give specific written consent.
2	7/30/2007	Multiple modest modifications to protocol language, notably including revisions to the AF inclusion criterion, the definition of minor bleed, allowing flexibility in the timing of stopping open VKA therapy prior to randomization (but not changing the INR limit), narrowing the exclusion for prior stroke from 30 days to 7 days, adding language encouraging resumption of study drug, if appropriate, in those who discontinued., adding a requirement for at least 40% of patients to be VKA naïve, adding language about when to unblind to Sec. 5.4., changing dosing recommendations for elective procedures, and emergency procedures, added suggested bridging strategy for procedures, changed provisions for prohibited concomitant medications, added section on management of cardioversion, added new Sec. 7.1.3 on "Events of Special Interest" – thrombocytopenia, elevated LFTs, neuropathy, added plan for 5 year "passive" follow-up after end of study, added new Sec. 8.7 indicating the EC would manage publications. .
5	11/7/2007	Provided for biomarker sub-study, including markers of cardiac necrosis, endothelial function, inflammation, platelet activity, coagulation, renal function, and lipoproteins.
7	8/5/2009	Provided for increase in sample size to from 15K to 18K due to lower than expected overall event rate (a related provision to increase the average expected duration of follow-up from 1.8 to 2.1 years was mistakenly omitted; this was added in Amendment 10); added provision that patients could discontinue study drug but continue follow-up; clarified that compliance data would be based on apixaban/placebo tablet counts only.
10	5/11/2010	Change to satisfy unnamed regulatory agencies in statistical analysis of net clinical benefit endpoint – superiority in efficacy and bleeding would be assessed separately. Added secondary endpoints superiority in time to major bleeding and time to all cause mortality, resulting in the final 4 step hierarchical analysis plan (see Sec. **5.3.1.12.2)**.
11	8/4/2010	Clarified that early discontinuation patients should be followed up with an by phone quarterly and have in person final follow-up visit within 30 days of reaching the study event target. Informed consent form was clarified to indicate Sponsor expectations on collecting vital status and occurrence of endpoints in early discontinuation patients. SAEs to be captured up to 30 days after last dose of study drug. Added transition regimen for end of study switch to open-label VKA.

Reviewer Comment: The late changes in the analysis plan (Amendment 10) to add superiority for major bleeding and all-cause death to the hierarchical analysis plan is problematic, as the study was fully enrolled and much data had been generated and passed along to the contractor preparing data for the IDSMB. The timing of this protocol amendment should be considered in our deliberations regarding potential label claims based on the last two superiority analyses in the hierarchical analysis plan (ISTH major bleeding and all-cause mortality).

In addition, to the 6 amendments above and other 5 amendment affecting sites in one or more foreign nations but not the US, there were a series of "administrative letters" that communicated minor changes in the protocol, such as changes in contact information, spelling errors, style, or administrative issues that were deemed to "not significantly affect the safety of subjects, study scope, or scientific quality…" of the study. These letters were reviewed and seemed to meet the definition of an "administrative letter" as quoted in the previous sentence.

Reviewer Comment: The administrative letters provided seem unobjectionable.

Note that not every amendment led to a new version of the protocol. The final protocol (Ver. 4), included all amendments affecting the US as well as changes created by 8 o the 9 administrative letters included with the protocol. The last such letter dated 18 February 2011 (after the sites had been notified to end the study) provided advice to US sites regarding follow-up issues related to the revised consent form of Sept. 10, 2010 (not included with the letter).

5.3.2 Supporting Study: AVERROES

AVERROES refers to a global study entitled, "(Apixaban Versus Acetylsalicylic Acid [ASA] to Prevent Stroke in Atrial Fibrillation Patients Who Have Failed or Are Unsuitable for Vitamin K Antagonist Treatment)" conducted by the Applicant. The trial was stopped early following a recommendation to terminate by its IDSMB because of the results favoring apixaban for the primary endpoint (time to first stroke or systemic embolism), with a p for superiority of 0.000002 (z = 4.76).

Because of the nature of the control used in this study, aspirin at a dose of 81 to 324 mg/day, the study can only be considered as supportive. Aspirin is not approved to prevent stroke or thrombotic event in patients with non-valvular atrial fibrillation, although it has been extensively studied. Its effects for the apixaban target indication are substantially less than those of warfarin. A modern meta-analysis indicates that warfarin is significantly superior to aspirin in terms of stroke prevention in patients with AF (risk reduction of 38% (95% CI, 18 to 52%). [1] The results of 6700-patient ACTIVE W trial in patients with atrial fibrillation, which compared warfarin with aspirin + clopidogrel, favored warfarin for stroke prevention, with a HR of 0.58 (p=0.001);.[8] this trial was included in the Hart meta-analysis. However, a meta-analysis of the effects of aspirin vs

placebo on stroke in patients with AF did show a modest effect (risk reduction of 22% (95% CI, 2 to 39%).

The current (2006) AHA/ACC/ESC consensus guidelines on the management of patients with atrial fibrillation recommend aspirin as an alternative in some patients with AFib:

> "6. Aspirin, 81–325 mg daily, is recommended as an alternative to vitamin K antagonists in low-risk patients or in those with contraindications to oral anticoagulation. *(Level of Evidence: A)"* (emphasis added) [3]

The guidelines recommend warfarin therapy for patients at moderate risk of stroke (those with one moderate risk factor (age ≥ 75 years, impaired LV function, diabetes mellitus, or hypertension), and also for patients at high risk of stroke (those with more than 1 moderate risk factor or with a high risk factor (prior history of stroke, TIA or systemic embolism)[i] Thus warfarin is preferred in patients with $CHADS_2$ scores of ≥ 2; those with lower scores have the alternative of aspirin. A recent update to the ESC guidelines is quite similar to the 2006 guidelines.[9]

5.3.2.1 Design of AVERROES and Contrasts with ARISTOTLE

Similarities and differences between the ARISTOTLE protocol and the AVERROES in terms of design features, enrollment data, and several key patient baseline characteristics affecting stroke risk, are described in the following table.

i Rheumatic mitral stenosis is also listed as a high risk factor, but the proposed indication for apixaban excludes patients with valvular AFib; such patients were excluded from both ARISOTLE and AVERROES.

Table 19 Features of ARISTOTLE and AVERROES

	ARISTOTLE	AVERROES
Basic design	Randomized, prospective, double-blind (double dummy) warfarin-controlled, event-driven, parallel trial	Same, except that the trial was aspirin-controlled
Primary objective	Demonstrate non-inferiority of apixaban to warfarin in terms of prevention of primary endpoint events (stroke, SE)	Demonstrate superiority of apixaban to aspirin in terms of prevention of primary endpoint events (stroke SE)
Geographic scope	6 continents	6 continents
Patients	Adults (≥18 yrs) with atrial fibrillation with ≥ 1 of the 5 named stroke risk factors (similar to the CHADS2 risk factors)	Adults ≥ 50 y with ≥ 1 of 6 named stroke risk factors (the 5 in ARISTOTLE + PAD) and NOT taking VKA at enrollment because it was demonstrated as unsuitable or expected to be unsuitable for the patient
Planned sample size	About 18,000	About 5,600
Enrolled	18,201	5,598
Event target	448	226
Apixaban dose	5 mg po once daily (2.5 mg for those with ≥ 2 of 3 named risk factors for bleeding vs. warfarin tablets once daily	Same
Control agent and dose	Warfarin titrated to INR target of 2.5 (range, 2.0-3.0) for all ages; blinded INR results obtained from point-of-care device	Aspirin 81 mg tablets, 1 to 4 daily (max of 324 mg) at discretion of investigator.
Follow up of completers	30 days after end-of-study (EOS) visit	Same, except some patients entered a long term open-label extension (LTOLE)
Follow-up of those with premature discontinuation	30 days after EOS visit, then phone follow-up q 12 weeks until overall end of study	Subjects who do not enter LTOLE have a follow-up visit 30 days after last dose of study drug
Anticoagulation required after study drug d/c'ed?	No – Institution of anticoagulation was at the investigator's discretion	Same
Primary efficacy endpoint analysis	Non-inferiority to warfarin for time to first stroke or SE in ITT population; patients followed to fixed date	Superiority to aspirin for time to first stroke or SE in ITT population, otherwise the same.
Primary safety endpoint analysis	Non-inferiority to warfarin for time to first major or non-major clinically relevant bleeding event in safety population on treatment	Compare to aspirin for time to first major or non-major clinically relevant bleeding event in safety population on treatment (no specific goal stated)
Important endpoints adjudicated?	Yes	Yes

	ARISTOTLE	AVERROES
PK/PD data collected?	Yes	Yes
First patient entered	Dec 19, 2006	Aug 31,2007
Last patient entered	4 Feb 2010	Dec 23, 2009
Last patient visit	May 25, 2011	Sept 20, 2010
Median F/U	1.8 yr	1.1 yr
Mean Baseline CHADS2	2.1	2.0
Prior Stroke/TIA/SE	19.4%	13.6%
Prior Use of VKA	57.1%	40.0%

5.3.2.2 Efficacy Results of AVERROES

Efficacy results of AVERROES are provided in this section. The reader desiring to understand the primary data supporting the efficacy of apixaban may elect to proceed directly to Section **6**, which contains the results of the single definitive study, ARISTOTLE, and then return here to review the abbreviated efficacy results of AVERROES. Safety results of AVERROES and ARISTOTLE are discussed in Section **7.**

5.3.2.2.1 Demographics

Demographic and risk factor data for the ITT population (N=5598), with 2807 and 2791 subjects in the apixaban and aspirin arms, respectively, are provided here.

In general, the treatment arms in the ITT population were quite well balanced at baseline. Each arm had a mean age of 70 years, with 32% and 35% with age ≥ 75 years in the apixaban and aspirin arms, respectively. Women comprised 41% and 42% of the apixaban and aspirin arms, respectively. Race was comparable in the two arms: 79% and 78% of patients were classified as white in the apixaban and aspirin arms, respectively, and 19% in each arm were Asian. Mean weight was 79 kg and mean BMI was 28 in each arm.

Table 20 is a display of relevant medical history at baseline in the treatment arms. There were no notable differences.

Table 20 AVERROES – Baseline Medical History – ITT Population

Condition, risk factor, treatment, or substance use	Apixaban n = 2807 N (%)	Warfarin N = 2791 N (%)
Heart failure (NYHA class≥2) or LVEF≤35%	961 (34.2)	926 (33.2)
Hypertension with pharmacological therapy	2408 (85.8)	2429 (87.0)
Diabetes mellitus	536 (19.1)	559 (20.0)
Peripheral arterial disease	66 (2.4)	87 (3.1)
A Fib on Screening ECG	1900 (67.7)	1866 (66.9)
Persistent AFib	430 (15.3)	401 (14.4)
Permanent AFib	1463 (52.1)	1461 (52.5)
Prior use of aspirin	2135 (76.3)	2085 (75.0)
Prior use of clopidogrel	91 (3.3)	92 (3.3)
Current tobacco user	206 (7.3)	201 (7.2)
Alcohol ≥ 3 days/week	354 (12.6))	345 (12.4)

Table 21 and **Table 22** provide information on baseline renal function and CHADS2 score, respectively. Both tables show similarity between the groups. Notably, more than 60% of patients in each arm had a CHADS2 score of at least 2, making them potential candidates for warfarin therapy. About 38% in each arm had a CHADS2 score of exactly 1, which by the consensus guidelines for the management of AFib, made them candidates for either aspirin or warfarin.

Table 21 AVERROES – Baseline Renal Function – ITT Population

Level of Renal Impairment (CrCl)*	**Apixaban N=2807 N (%)**	**Warfarin N=2891 N (%)**
Severe (<= 30 ml/min)	55 (2.0)	61 (2.2)
Moderate (> 30 - <= 50 ml/min)	490 (17.5)	478 (17.1)
Mild (.50 - <= 80 ml/min)	1068 (38.2)	1072 (38.6)
Normal (> 80 ml/min)	955 (34.0)	923 (33.1)
Not available	233 (8.3)	254 (9.1)

*CrCl= creatinine clearance

Table 22 AVERROES – Baseline $CHADS_2$ Score – ITT Population

$CHADS_2$ Score at Enrollment	Apixaban N=2807 N (%)	Aspirin N=2891 N (%)
0	10 (0.4)	9 (0.3)
1	1056 (37.6)	1067 (38.2)
2	1037 (36.9)	936 (33.5)
3	443 (15.8)	491 (17.6)
4	187 (6.7)	208 (7.5)
5	63 (2.2)	74 (2.7)
6	11 (0.4)	6 (0.2)
<=1	1066 (38.0)	1076 (38.6)
>=3	704 (25.1)	779 (27.9)
Mean (SD)	2.0 (1.04)	2.0 (1.07)
Median	2.0	2.0

There were 182 (6.5%) and 184 (6.6%) subjects in the apixaban and aspirin arms, respectively, who were candidates for the reduced (2.5 mg bid) dose of apixaban/placebo because they met at least two of the relevant criteria (age ≥80 years, weight ≤60 kg or serum creatinine ≥1.5 mg/mL) for increased bleeding risk. Six and 4 subjects in the apixaban and aspirin arms, respectively had all three risks; the remainder of candidates for the lower dose (176 and 180 respectively) had 2 risks.

The treatment arms were also similar with respect to the reasons why VKA therapy was deemed to be unsuitable for the patient. For about 40% of patients in each arm, VKA therapy had been demonstrated to be unsuitable. For the remainder, it was expected to be unsuitable.

For patients with prior VKA use, the most common reason for unsuitability was “unable to maintain VKA in therapeutic range (about 16.5% in each arm). Only about 3% in each arm had “significant bleeding while on Coumadin ” as the reason for unsuitability.

“Subject refused treatment” was only the reason for unsuitability in 15% and 14% of patients in the apixaban and aspirin arms, respectively. “$CHADS_2$ score =1 and physician does not recommend VKA use” was the only reason for unsuitability in about 11% of patients in each arm.

5.3.2.2.2 Subject Disposition

Information regarding patients who discontinued treatment during double-blind therapy is provided in **Table 23**.

More patients in aspirin arm failed to complete treatment (19.9% vs. 23.3%). Death and non-fatal AEs in general (including stroke, SE, and bleeding) were less commonly associated with withdrawal from treatment in the apixaban arm, while MI (10 vs. 6 subjects) was more common in the apixaban arm. Of note, "other" as a cause for withdrawal was more common in the apixaban arm (4.7% vs. 3.7%).

Note that AVERROES had a long-term open label extension (LTOLE) that is ongoing.

Table 23 AVERROES - Reasons For Failure To Complete Double-Blind Treatment
ITT Population

	APIXABAN N=2807 n (%)	**Aspirin N=2791 n (%)**
Total who failed to complete	558 (19.9)	649 (23.3)
Death	31 (1.1)	56 (2.0)
Non-fatal adverse event	174 (6.2)	260 (9.3)
Stroke	18 (0.6)	73 (2.6)
Systemic embolism	0	12 (0.4)
Myocardial infarction	10 (0.4)	6 (0.2)
Bleeding	35 (1.2)	20 (0.7)
Other	112 (4.0)	149 (5.3)
Subject requested to withdraw from treatment	156 (5.6)	171 (6.1)
Subject withdrew consent	92 (3.3)	98 (3.5)
Lost to follow-up	26 (0.9)	29 (1.0)
Poor or non-compliance	17 (0.6)	22 (0.8)
Subject no longer meets study criteria	18 (0.6)	18 (0.6)
Other	132 (4.7)	102 (3.7)

Table 24, Subject Disposition, indicates that in general, follow-up during the treatment phase was good in this study, with a low lost-to-follow-up rate during treatment (≤0.3% in each arm). However, 23% and 22% of subject in the apixaban and aspirin arms, entered the follow-up phase (i.e., they stopped treatment) but did not enter the LTOLE or complete follow-up for some reason that was not provided. However the Applicant notes that "many" of these subjects failed to enter the LTOLEs due to delays in approval of the protocol amendment that provided for the LTOLE at IRBs or regulatory authorities.

Table 24 AVERROES – Subject Disposition
(At end of follow-up period, ITT Population)

	APIXABAN N=2807 n (%)	Aspirin N=2791 n (%)
Subjects entered LTOLE	1133 (40.4)	1065 (38.2)
Subjects completed follow-up and did not enter LTOLE	887 (31.6)	933 (33.4)
Subjects did not enter LTOLE and did not complete follow-up	707 (25.2)	700 (25.1)
Death	64 (2.3)	91 (3.3)
Other	643 (22.9)	609 (21.8)
Subjects did not enter follow-up	80 (2.9)	93 (3.3)
Death	52 (1.9)	63 (2.3)
Withdrew consent	20 (0.7)	24 (0.9)
Lost to follow-up	8 (0.3)	6 (0.2)

LTOLE= Long-term open-label extension (ongoing)

Table 25 is a display of the number of subjects in the various study populations used in the efficacy analyses described under the next heading.

Table 25 AVERROES - Analysis Populations

Population	Apixaban	Aspirin	Total
ITT	2807	2791	5598
Treated	2798	2780	5578

About 6.4% and 6.5% of subjects in the apixaban and aspirin arms, respectively qualified for the reduced dose of apixaban, under criteria identical to those in ARISTOTLE.

Median exposure to double-blind study drug (first dose to last dose) was 58.5 weeks vs. 58.6 weeks in the apixaban and aspirin arms, respectively. Maximum exposure was 140 weeks in the apixaban arm and 151 weeks in the aspirin arm.

5.3.2.2.3 ***Analysis of Efficacy Endpoints***

AVERROES was stopped early at the recommendation of its DSMB when the first scheduled interim analysis, which was scheduled to take place after 113 adjudicated events (50% of the event target), met the prespecified criteria for early termination due to superiority of apixaban for the primary endpoint (i.e., a critical value based on a Haybittle-Peto boundary of at least 4 SD (one-sided $p < 0.00003$). The following sequence of events occurred:

- The results of the first interim analysis (data cut-off, January 31, 2010, 104 events, 46% of target) were z = 4.9 SD in favor of apixaban)
- This was followed by 2 confirmatory analyses with cut-offs of –
  - April 30, 2010 (140 events, 62%, z = 4.6 SD)
  - May 27, 2010 (145 events, 64%, z = 4.76 SD, p=0.00002).
- The DSMB met on May 28, 2010, considered all the data above plus subgroup data and safety data and recommended termination of the study.
- The recommendation was accepted by the Sponsor and Executive Committee.
- The key ITT analyses are based on a data cut-off of May 28, 2010.

The primary endpoint in AVERROES was the time to an adjudicated stroke or SE in the ITT population, during the Intended Treatment Period (ITP), which was defined in the same way as in ARISTOTLE.

Results for the analysis of the protocol-specified primary efficacy endpoint are displayed in the first data row in **Table 26.** Other data rows show results for the two major secondary endpoints, the composite of stroke/SE/MI/vascular death, and all-cause death.

Table 26 AVERROES – Key Efficacy Results
ITT Population, ITP

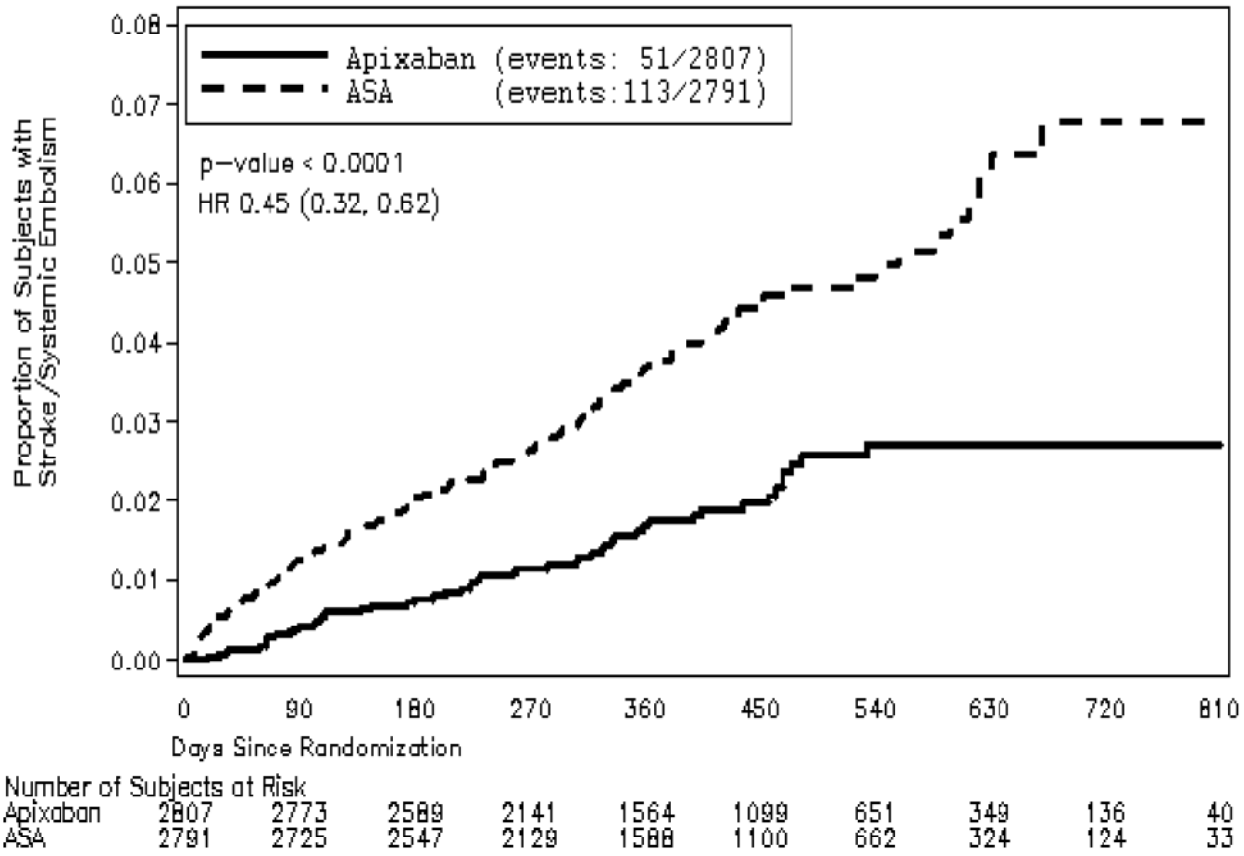
	Apixaban N=2807		**Aspirin N=2791**		**A vs ASA HR (95% CI)**	**p (Super-iority)**
Parameter	n (%)	Events/ 100 pt-yr	n (%)	Events/ 100 pt-yr		
Stoke/SE	51 (1.82)	1.62	113 (4.05)	3.63	0.45 (0.32, 0.62)	<0.00001
Stroke/SE/MI/ Vasc. Death	132 (4.70)	4.21	197 (7.06)	6.35	0.66 (0.53, 0.83)	0.00026
All-cause Death	111 (3.95)	3.55	140 (5.02)	4.42	0.79 (0.62, 1.02)	0.068

Highly significant results in favor of apixaban were obtained for the primary endpoint and the composite of stroke/SE/MI/vascular death. There was a trend in favor of apixaban for all-cause death (p=0.07). Note that a hierarchical analysis was not prespecified.

Information on rates of the individual components of the primary endpoint and other secondary endpoints are discussed below.

The Kaplan-Meier curve for time to first primary efficacy event in the protocol-specified primary efficacy analysis is shown in **Figure 9** .

Figure 9 AVERROES -- Time To First Primary Efficacy Endpoint Event
(ITT Population, ITP)



Additional Endpoints

Rates for the occurrence of secondary and tertiary endpoints in the ITT Population during the ITP are displayed in **Table 27**. These endpoints include the individual components of the primary endpoint (types of stoke, systemic embolism), as well as MI, vascular death, and non-vascular death. The data indicate the most of the advantage of

apixaban over aspirin is due prevention of ischemic/unspecified stroke, but there is an a beneficial effect of the rate of SE as well. Both vascular and non-vascular death were numerically reduced by apixaban compared to aspirin. Hemorrhagic stroke rates were low and similar.

Table 27 AVERROES – Rates Of Secondary Efficacy Endpoints
(ITT Population during ITP)

	Apixaban N=2807		**Aspirin N=2791**		**A vs ASA HR (95% CI)**
Parameter	n (%)	Events/ 100 pt-yr	n (%)	Events/ 100 pt-yr	
FIRST EVENT					
Ischemic/unspecified stroke	38 (1.35)		94 (3.37)		
Hemorrhagic stroke	5 (0.18)		6 (0.21)		
SE	2 (0.07)		11 (0.39)		
MI	21 (0.75)		23 (0.82)		
Vascular death	84 (2.99)	2.65	96 (3.44)	3.03	0.87 (0.65, 1.17)
Non-vascular death	27 (0.96)	0.85	44 (1.58)	1.39	0.62 (0.38, 1.00)

The Applicant also claims in the study report and in labeling that the following analyses significantly favored apixaban:

- Time to stroke with modified Rankin score of 0-2 (non disabling stroke), HR 0.51, 95% CI (0.29, 0.91), p=0.02 (see **Attachment 3** for Rankin scale)
- Time to stroke with modified Rankin score of 3-6 (disabling or fatal stroke) HR 0.43, 95% CI (0.28, 0.65), p<0.0001
- Incidence of CV hospitalization, HR 0.79, 95% CI (0.69, 0.91), p=0.0009

Reviewer Comment: None of these analyses was prespecified in the SAP. The study DSMB did request analyses of strokes prevented "according to Rankin scores at various cut points." Nonetheless, these analyses represent post-hoc looks at the data.

5.3.2.2.4 Subgroups

Analyses of the primary endpoint in major subgroups supported the overall findings (see **Attachment 7**).

The primary endpoint results of AVERROES in the subgroup of patients with a $CHADS_2$ score of 1 at baseline is of interest. These patients comprised about 38% of the ITT population. The AHA/ACC consensus guidelines on management of AF recommend aspirin (81-325 mg daily) "…as an alternative to vitamin K antagonists in low-risk patients or in those with contraindications to oral anticoagulation." (Class IA).[3] "Low-risk" as defined in the guidelines is concordant with a CHADS2 score of 0-1. There were very few patients in the study with $CHADS_2$ =0 (and they had no events); consequently, this analysis will focus on those patients with a score of 1.

Data for the primary endpoint results in the subgroup of patients in AVERROES with a $CHADS_2$ score of 1 at baseline are shown in **Table 28.** The first row of data shows results for AVERROES. The HR is 0.63, slightly worse than the overall HR of 0.45. This result is consistent with results in other studies of stroke prevention in AFib patients shown better results (i.e., lower hazard ratios) for effective interventions in patients at higher risk, who constitute over 60% of AVERROES. The second row of data shows the comparison of apixaban to warfarin in ARISTOTLE in the analogous and considerably larger population of patient with $CHADS_2$ = 1 in that study. Again the HR, 0.85, is slightly larger than the overall HR of 0.79 in ARISTOTLE. However, both studies suggest than the beneficial effects of in apixaban are preserved in patients with a $CHADS_2$ score of 1. The AVERROES data indicate that apixaban may be a better choice than aspirin in many patients at risk for stroke, including those with $CHADS_2$ =1.

Table 28 Primary Endpoint Results in Patients with Baseline $CHADS_2$ Score =1

	Apixaban N=2807		Aspirin N=2791		A vs ASA HR (95% CI)
	n/N (%)	Events/ 100 pt-yr	n/N (%)	Events/ 100 pt-yr	
AVERROES ($CHADS_2$ = 1)	12/1056 (1.14)	1.01	19/1067 (1.78)	1.60	0.63 (0.31, 1.30)
ARISTOTLE ($CHADS_2$ = 1)	44/3046 (1.44)	0.75	51/3025 (1.69)	0.88	0.85 (0.57, 1.27)

Reviewer comment: The results of AVERROES provide no direct information about the comparison of apixaban to warfarin, which is the critical comparison with respect to approval. However, if apixaban is approved key results from AVERROES, including results in the subset of low risk patients, should be included in apixaban labeling in Sec.14 to permit promotional claims of superiority to aspirin in stroke/SE prevention to encourage physicians to use this more effective product as an alternative to aspirin.

For a discussion of primary efficacy endpoint events that occurred after the last dose of study drug in AVERROES, see Sec. **6.1.10.2.2**.

6 Review of Efficacy

Efficacy Summary

The primary support for the proposed indication is the warfarin-controlled ARISTOTLE trial. In addition, the Applicant conducted the aspirin-controlled AVERROES trial in patients who had failed or were considered likely to be unsuitable for VKA treatment. Because warfarin has been convincingly demonstrated to be a superior to antiplatelet agents (including aspirin alone or in combination with clopidogrel)[1;8] in preventing thrombotic events in patients with nonvalvular AFib, the AVERROES trail should be considered as merely supportive.

ARISTOTLE primary endpoint results:

ARISTOTLE was a large (>18,000 subjects), randomized, double blind (double dummy), event-driven, warfarin-controlled, non-inferiority trial in adults with non-valvular atrial fibrillation (AFib) or atrial flutter (AFl) with at least one additional risk factor for thrombotic events. The dose of apixaban was 5 mg po twice daily (2.5 mg twice daily for those with pre-specified risk factors for bleeding); warfarin was to be titrated to a target INR range of 2.0 to 3.0. The primary endpoint was time to a composite of stroke and systemic embolism (SE). The Applicant's designated primary endpoint analysis was for non-inferiority (with a margin of 1.38 for the HR) in the ITT (all randomized patient) population. This analysis included events occurring during the "intended treatment period" (ITP), which extended from each patient's date of randomization to January 30, 2011, the estimated date of attainment of the event target of 448 primary endpoint events. This analysis will be referred to as the "ITT/ITP" analysis. As the initial analysis in a 4-step hierarchical analysis plan, it was used to evaluate (1) non-inferiority to and then (2) superiority to warfarin if non-inferiority was achieved. Steps 3 and 4 in the hierarchy are discussed below. The ITT/ITP analysis of the primary endpoint yielded a hazard ratio (apixaban vs. warfarin) of 0.79, with a 95% CI of 0.66 to 0.95, p (superiority) = 0.0114, thus supporting both non-inferiority and superiority. However, if the standard for a superiority claim based on a single study is a $p \leq 0.01$, one could arguably deny a superiority claim. Additional analyses of the primary efficacy endpoint included ones of "evaluable" patients (i.e., the per-protocol population),

counting events occurring during the on-treatment period (first dose of study drug to last dose + 2 days), and two other analyses based on the on-treatment period but with event windows extending to 7 or 30 days after the last dose of study drug. All these analyses of the primary endpoint results supported superiority of apixaban to warfarin, with p values ranging from 0.011 in the worst case (the ITT/ITP analysis) to < 0.001in the best case (for both the on-treatment and last dose + 7 days analyses in the per-protocol population). Additional analyses of the primary endpoint in all treated patients were likewise supportive of superiority.

These efficacy findings appeared to be preserved across major subgroups of patients in the ITT/ITP analysis, including each gender, the elderly, subjects previously treated with a VKA, subjects with a prior history of stroke, TIA or systemic embolism, subjects in each of the 4 specified geographic regions, those who qualified for the lower dose, and those enrolled from US sites.

Secondary endpoints

The primary endpoint findings were also supported by numerical imbalances for most important secondary efficacy endpoints that favored apixaban over warfarin in the ITT/ITP analysis. These endpoints included the rates of strokes (all types combined), hemorrhagic strokes, fatal strokes, systemic emboli, vascular deaths, and non-vascular deaths. The results for myocardial infarction also favored apixaban, unlike in the RE-LY trial of dabigatran. The results for death also support the primary endpoint findings, and are discussed below.

There was a small imbalance of pure ischemic strokes in favor of warfarin in the ITT population during the ITP (140 vs. 136 events occurring during the ITP). However, when ischemic strokes with hemorrhagic conversion are also included, the results favored apixaban (152 vs. 156), and the additional inclusion of strokes of uncertain type also favored apixaban (166 vs. 177 for the combined categories). There were few systemic emboli and the difference in rate between the treatment arms was small. Thus the primary endpoint results favoring apixaban were driven mostly by an excess of hemorrhagic strokes in the warfarin arm (40 vs. 78). While all strokes were counted as efficacy events, hemorrhagic stroke is a risk of anticoagulation, not something that is prevented by anticoagulation. Thus, if apixaban is superior to warfarin in terms of stroke, it is superior because it causes less hemorrhagic stroke than warfarin. There is a modest (non-significant) difference between apixaban and warfarin in terms of reducing the rate of ischemic stroke, the primary reason for giving anticoagulants to patients with atrial fibrillation.

In the pre-specified, 4-step hierarchical analysis plan, step 3 was a superiority analysis for time to ISTH major bleeding, the primary safety endpoint, in the safety population on treatment. This analysis was robustly successful and is discussed further in the safety summary. Step 4 in the hierarchy was an analysis of superiority of apixaban for time to all-cause death, conducted in a manner similar to that of the primary endpoint ITT/ITP

analysis. This was just barely successful: the HR was 0.89, with a 95% CI of 0.80 to 1.00 (p=0.0465). Hazard ratios for CV death and non-CV death differed little from each other and from the all-cause death HR.

The following issues are relevant to the interpretation of the efficacy results of the trial:

Medication errors and other trial conduct issues in ARISTOTLE:

The clinical reviewers are concerned that study medication errors and deficiencies in monitoring and the data quality assurance process may have affected outcomes in ARISTOTLE. This complex issue is summarized in Sec. 1 **above** and described in depth in Sections **3.1.1** and **3.2**.

(b) (4)

(b) (4), there is explicit language in Sec. 14 of the Applicant's proposed PI regarding superiority to warfarin.

The primary support for (b) (4) is the just-barely significant reduction in time to all-cause death in the ITT/ITP analysis (HR=0.89, 95% CI, 0.80 to 1.00, p=0.0465). This analysis is part of a hierarchy that conserves overall alpha error at the 0.05 level. However, one additional death in the apixaban arm or one fewer death in the warfarin would negate the statistical significance of this finding. An analysis with a variable cutoff date that includes a 30 day follow-up period for treated patients after the last dose of study drug (rather than the ITT/ITP analysis, in which all completing patients (about 75% of the total) were still on treatment at the analysis cutoff date) was slightly less favorable for apixaban, with p = 0.08 an upper limit of the HR of 1.01.

In addition, 590 patients (3.2% of those randomized) discontinued follow-up alive during the study and had no information on vital status at the cut-off date for the ITT/ITP analysis of death; they were censored on the date of their last contact prior to the cut-off date. As discussed in Sec. **3**, there were systemic blinding issues in this study that might have lead to unblinding of individuals or all patients at a site. If such unblinding occurred (we don't know how often that happened), a site might have made lesser efforts to track the vital status for some patients than for others, thus potentially biasing the mortality results. Note that only one additional death in the apixaban arm (or one fewer in the warfarin arm) might have negated the superiority finding for all-cause death in the ITT/ITP analysis.

Also, the results of analyses based on site-specific INR control in the warfarin arm suggest that unlike the primary endpoint results, better INR control was associated with less favorable results for apixaban for all cause death. The Applicant's analyses show

at sites above the median in time in therapeutic range of INR (TTR), the HR for all cause was 0.93, and sites in the top quartile of TTR, it was 1.2 (see **6.1.10.1.3**).

Finally, there was fraud committed at site 1200 in China that was committed by the BMS site supervisor (see Sec. **3.1.2**). In a sensitivity analysis that excludes this site, superiority of apixaban for all-cause analysis is lost). However, when all Chinese sites are excluded from the analysis of mortality, the results become more favorable for apixaban (**Table 11**).

(b) (4)

. However, these concerns might not affect the absolute claim for a reduction in mortality. Warfarin itself may reduce mortality in patients with AFib, as evidenced by the results of published meta-analysis by Hart. et. al. of the placebo-controlled studies of warfarin therapy in patients with atrial fibrillation.[1] Thus, one might conclude that apixaban was non-inferior to warfarin in preventing all-cause death on the basis of an appropriately designed and conducted non-inferiority study. However, the Hart analysis for mortality would yield a NI margin of 1.015, if the same procedures for calculating the NI margin were used as those used to calculate the NI margin of 1.38 for the primary endpoint. A margin of 1.015 is trivially different from 1.0, which means that this NI analysis would have nearly the same results as a test of superiority to warfarin. (see p. **141**). If there are reasons to be concerned about the use of the superiority findings of the ITP/ITT analysis to support a claim of superiority, the same reasons would probably apply to use of a NI analysis with a margin of 1.015 to support a claim of NI to warfarin for all-cause death.

(b) (4)

.

Adequacy of anticoagulation in the warfarin treatment arm and constancy assumption issues:

Although the primary efficacy analysis in this study involves non-inferiority to warfarin, satisfaction of the constancy assumption would not be a critical issue if we accept the nominal superiority of apixaban over warfarin in the prespecified primary endpoint ITT analysis. Nonetheless, patients in ARISTOTLE were reasonably similar to those in the RE-LY study of dabigatran, which satisfied the constancy assumption with respect to the 6 placebo-controlled studies of warfarin. Treatment goals, endpoints, and other critical design features were similar to or more rigorous than those in RE-LY. In addition, other key characteristics of the patient population and study design of ARISTOTLE were reasonably similar those of the 6 historical placebo-controlled studies

of warfarin, and the constancy assumption is satisfied for ARISTOTLE (see Sec. **6.1.10.1**).

Quality of INR control in the warfarin arm in ARISTOTLE was in between that obtained in the ROCKET study of rivaroxaban (another study satisfying the constancy assumption with respect to use of warfarin) and the better results obtained in RE-LY, but was closer to RE-LY than to ROCKET in this regard. Thus, anticoagulation in ARISTOTLE was good enough so that the overall findings of non-inferiority or superiority to warfarin should not be rejected due to inadequate dosing in the control arm. However, for subjects at sites with TTR above the median subjects, and in the best quartile of TTR, results for all-cause death suggest lack of superiority over warfarin, which might be a labeling issue (see text **above** and Sec. **6.1.10.1.3**).

Efficacy events occurring after discontinuation of study drug in completers:

Approximately 3/4 of patients in ARISTOTLE in each arm continued taking study drug until the end of this event-driven study. In these patients, blinded study medication was stopped, and the investigator was to transition patients to alternative anticoagulant therapy, usually a vitamin K antagonist such as warfarin. Like in several other recent trials of novel anticoagulants in AFib patients (the Sportif V trial of ximelagatran and the RE-LY trial of dabigatran), investigators were urged to implement a short period of dual therapy with study drug and open-label warfarin (which was to continue long-term) for patients in the apixaban arm to continue effective anticoagulation during the lag period of INR control at the start of warfarin therapy. In ARISTOTLE, this was done using blinded apixaban study drug, which was to be administered for four additional (twice-daily) doses. Blinding of study drug was maintained, so that warfarin arm patients received placebo for their transition medicine. About 60% of completing subjects in each arm received this transition regimen; about 84% received a VKA during the 30 days after the end of the double blind treatment period.

However, as in the ROCKET study of rivaroxaban, in the 30 days after the last dose of blinded study drug, there were significantly more primary endpoint events (mostly ischemic strokes) in the apixaban than in the warfarin arm, with a HR of about 4 (21 vs 5 events). Events were distributed throughout the 30 day period. In the apixaban arm, there were 3 hemorrhagic strokes, all in patient who received open-label VKA treatment, and all occurring in the second half of the 30 day post-dose period. These events were probably related to warfarin use. Unfortunately, INR information during this period was not routinely collected, and was absent for the vast majority of subjects. However, less than 50% of subjects were receiving a VKA at the start of the study, and it is seems reasonable to believe that INR control was suboptimal following the end of the study in many subjects.

There was an analogous finding (a significantly increased rate of stroke/SE in the 30 days after the last dose of study drug in the apixaban arm compared to the warfarin arm) in completing patients in AVERROES, the aspirin controlled study of apixaban in

patients with AFib who had failed or were deemed unsuitable for VKA therapy. In that study, no completing patient was known to have received VKA treatment, so this explanation for the difference in stroke rate is not applicable (see **Table 29** for primary endpoint results in completers during the 30 days after the last dose of study drug in ARISTOTLE, AVERROES and the ROCKET study of rivaroxaban).

Table 29 Primary Efficacy Endpoint Events Following Completion of Study Treatment in Completing Patients in Phase 3 Studies

Study		Apixaban N=6810		Warfarin N=6588		
	Days after last dose	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr	HR (95% CI)
ARISTOTLE	1-30	21 / 6791	4.02	5 / 6569	0.99	4.07 (1.54, 10.81)
AVERROES	1-30	9 / 1472	10.55	1 / 1421	1.20	8.78 (1.11, 69.34)
ROCKET (rivaroxaban)	3-30	22 / 4587	6.42	6 / 4652	1.73	3.72 (1.51, 9.16)

Reviewer Comment: The Applicant did not collect information on the concentration and function of clotting system constituents after cessation of long term treatment with apixaban, and nothing is known about the pharmacodynamics of warfarin that is initiated in this setting. Given the findings in ARISTOTLE and AVERROES, it would be desirable for the Applicant to perform a study to collect such information.

There was a modest excess of primary endpoint events and deaths in the period from day 1 to 30 days after the last dose of study drug in the warfarin arm in patients who discontinued study drug early. However, the differences between the treatment arms were not statistically significant.

For more information regarding the rate of events after discontinuation of study drug in ARISTOTLE see Sec. **6.1.10.2.2**, and for information on the Applicant's proposed instructions for the transition from apixaban to warfarin see Sec. **5.3.1.5.4**.

Dosing regimen:

The Applicant evaluated one dosing regimen in its pivotal trial: 5 mg of apixaban bid for most patients, 2.5 mg bid for patients with at least 2 of 3 specified risk factors for bleeding. The Applicant established that this regimen overall was superior to warfarin

for the primary endpoint and ISTH major bleeding. Only 5% of subjects qualified for the 2.5 mg bid dose. Safety and efficacy results in this subgroup of patients were slightly more favorable for apixaban than in the much larger subgroups that qualified for the 5 mg bid dose, so the dose reduction criteria seem reasonable. However, criteria for dose reduction other than those selected by the Applicant may also be reasonable. The same regimen was used in AVERROES, an aspirin controlled trial of apixaban in patients with AFib who were not candidates for warfarin therapy.

This dosing regime was based primarily on the results of a parallel-arm dose finding study in post-operative orthopedic surgery patients at risk for VTE. For EOP2 meeting for the AFib program, we reviewed the relevant data and agreed to the use of this regimen in both ARISTOTLE and AVERROES prospectively. (Sec. **6.1.8**).

6.1 Indication

The Applicant's proposed indication is:

> ELIQUIS (apixaban) is indicated to reduce the risk of stroke, systemic embolism, (b) (4) in patients with nonvalvular atrial fibrillation (b) (4)

6.1.1 Methods

The Applicant provided an ISE, but did not pool the results of ARISTOTLE and AVERROES. The designs of the two studies have already been described. Because the results of AVERROES are generally not useful to shape efficacy labeling in the US, only the results of ARISTOTLE are discussed here. The efficacy results of AVERROES are found at the end of the preceding section starting on page **114**.

6.1.2 Demographics

Baseline data for demographic and disease-related parameters are displayed in **Error! Reference source not found.** for the ITT population, N=18,201).

As expected in a study of this size, the treatment arms were well balanced for all important demographic and prevalent disease specific features. About 65% of subjects in each arm were male. The mean age in both arms was 69 years. About 57% in each arm were VKA experienced at baseline, as defined in the protocol. The mean $CHADS_2$ score (a widely used scale for stroke risk stratification of A Fib patients that ranges from 0 to 6, with higher scores associated with greater risk) in each arm was 2.1, and 19-20% of subjects had a prior history of stroke/TIA/systemic embolism, generally considered the most important risk factor for subsequent stroke.

Table 30 ARISTOTLE – Baseline Demographics and Disease-Related Parameters
(ITT population)

Characteristic	Apixaban N=9120	Warfarin N=9081	Total N=18,201
Male, N (%)	5886 (64.5)	5899 (65.0)	11785 (64.7)
Age, years, N, %			
Mean (SD)	69.1 (9.6)	69.0 (9.7)	69.1 (9.7)
18 to <65	2731 (29.9)	2740 (30.2)	5471 (30.1)
65≤ to <75	3539 (38.8)	3513 (38.7)	7052 (38.7)
≥75	2850 (31.3)	2828 (31.1)	5678 (31.2)
Race, N (%)			
White	7536 (82.6)	7493 (82.5)	15029 (82.6)
Black	125 (1.4)	102 (1.1)	227 (1.2)
Asian	1310 (14.4)	1332 (14.7)	2642 (14.5)
American Indian / Alaska Native	26 (0.3)	24 (0.3)	50 (0.3)
Other	123 (1.3)	129 (1.4)	248 (1.4)
Ethnicity, N (%)			
Hispanic or Latino	1808 (19.8)	1803 (19.9)	3611 (19.8)
Body metrics, Mean (SD)			
Weight (kg)	83.9 (20.8)	84.1 (20.6)	84.0 (20.7)
Height (cm)	168.7 (10.7)	168.7 (10.7)	168.7 (10.7)
BMI	29.3 (5.9)	29.4 (6.1)	29.4 (6.0)
Creatinine clearance strata (mL/min), N (%)			
≤30,	137 (1.5)	133 (1.5)	270 (1.5)
>30 to ≤50	1365 (15.0)	1382 (15.2)	2747 (15.1)
>50 to ≤ 80	3817 (41.9)	3770 (41.5)	7587 (41.7)
>80	3761 (41.2)	3757 (41.4)	7518 (41.3)
Prior Stroke/TIA/Non-CNS Systemic Embolism, N (%)			
Yes	1748(19.2)	1790 (19.7)	3538 (19.4)
Hypertension with pharm. treatment, N (%)			
Yes	7962 (87.3)	7954 (87.6)	15916 (87.4)
Diabetes mellitus, N (%)			
Yes	2284(25.0)	2263 (24.9)	4547 (25.0)

Symptomatic HF or LVEF ≤40%, N (%)			
Yes	3235 (35.5)	3216 (35.4)	6451 (35.4)
Prior MI, N (%)			
Yes	1319 (14.5)	1266 (13.9)	2585 (14.2)
Baseline $CHADS_2$ score			
Mean (SD)	2.1 (1.10)	2.1 (1.11)	2.1 (1.10)
Median	2.0	2.0	2.0
Score -- N, (%)			
0	54 (0.6)	58 (0.6)	112 (0.6)
1	3046 (33.4)	3025 (33.3)	6071 (33.4)
2	3262 (35.8)	3254 (35.8)	6516 (35.8)
3	1681 (18.4)	1598 (17.6)	3279 (18.0)
4	767 (8.4)	814 (9.0)	1581 (8.7)
5	273 (3.0)	289 (3.2)	562 (3.1)
6	37 (0.4)	43 (0.5)	80 (0.4)
Apixaban/placebo dose reduction factors and dose assigned, N (%)	(If ≥2 factors present, dose of apixaban/placebo was to be 2.5 mg bid; otherwise, it was to be 5.0 mg bid)		
Age ≥ 80 years	1225 (13.4)	1211 (13.3)	2436 (13.4)
Weight ≤ 60 kg	1018 (11.2)	967 (10.6)	1985 (10.9)
Serum creatinine ≥ 1.5 mg/dL	626 (6.9)	615 (6.8)	1241 (6.8)
0 Factors present	6675(73.2)	6681 (73.6)	13356 (73.4)
1 Factor present	2032 (22.3)	2014 (22.2)	4046 (22.2)
2 Factors present	402 (4.4)	379 (4.2)	781(4.3)
3 Factors present	11 (0.1)	7 (<0.1)	18 (<0.1)
2.5 mg bid dose assigned	428 (4.7)	403 (4.4)	831 (4.6)
5 mg bid dose assigned	8692 (95.3)	8678 (95.6)	17370 (95.4)
VKA use, N (%)	"'VKA Experienced" defined as > 30 consecutive days of VKA use at any time in the past		
"VKA Experienced"	5208 (57.1)	5193 (57.2)	10401 (57.1)
VKA use at screening	5070 (55.6)	5117 (56.3)	10187 (56.0)
Aspirin use at randomization, N (%)			
Yes	2846 (31.2)	2762 (30.4)	5608 (30.8)

Table 31 provides information regarding the percentage of patients with selected risk factors for stroke, as well as the mean $CHADS_2$ score, in ARISTOTLE and two other recent studies of approved anticoagulants published since 2010, the RE-LY study of dabigatran and the ROCKET study of rivaroxaban. In terms of stroke risk, patients in ARISTOTLE were more like those in RE-LY than in ROCKET.

Table 31 Selected Baseline Stroke Risk Factors in Recent Trials of Novel Anticoagulants

	ARISTOTLE (apixaban)	RE-LY (dabigatran)	ROCKET (rivaroxaban)
No VKA use or VKA "inexperienced", (%)	44	39	38
Prior stroke/TIA/systemic embolism, (%)	19	21	55
Female, (%)	35	37	40
Mean $CHADS_2$ score	2.1	2.1	3.5

Table 32 provides information on enrollment by region. Note the region of "Europe" included parts of Asia and Africa. Subregions of "Eastern Europe" and "Western Europe" were used for a limited number of analyses. Enrollment in the two arms was similar in each region; only the aggregate enrollment in both arms is provided. For additional geographic information on enrollment, see **Attachment 1**

Table 32 Enrollment by Region
ITT Population

Region	N	%
North America	4474	25
Latin America	3468	19
Europe [1]	7343	40
Eastern Europe [2]	3964	22
Western Europe [3]	2940	16
Asia-Pacific	2916	16
Total	**18,201**	**100**

1 Includes Europe (with Asiatic Turkey and Russia), Israel and South Africa
2 Includes participating Eastern European countries (except Turkey)
3 Includes participating Western European countries

Unlike ROCKET and RE-LY, ARISTOTLE allowed patients with atrial flutter (AFl) to enter the study as well as those with AFib. In addition, patients without either of these arrhythmia at entry could be randomized with a demonstrated history of either AFib or AFl (see Sec **5.3.1.4)**. The consensus guidelines on the management of patients with AFib states that patients with AFl should receive anticoagulant therapy like those with AFib (level of evidence: C). [3]

Table 33 provides information on heart rhythm at the baseline ECG as well as other study visits during the trial for convenience. The study arms were well balanced with regard to rhythm; only the overall data are shown.

About 81% of subjects has AFib at baseline, while 4% had AFl. Thus, 15% had neither AFib nor AFl at baseline, and about 13% were in sinus rhythm at baseline.

About 88% had at least one study visit ECG with AFib; 7% had at least one ECG with AFl. About 70% had AFib on all study ECGs. Only 1% had AFl on all ECGs, consistent with the generally unstable nature of AFl. About 10% of subjects had not a single study ECG with either AFib or AFl, suggesting that these subjects might have been be at lower risk of thromboembolism than others.

> *Reviewer Comment: The last finding described above may not affect our acceptance of positive study results. To the extent that patients did not have an atrial arrhythmia that predisposed them to thrombotic events during the study, it would be more difficult to show superiority of one antithrombotic over another, although it might be easier to show non-inferiority. Thus, if superiority is shown, the fact that 10% of patients may not have been at increased risk for thrombotic events is not problematic in interpreting the study results.*

Table 33 ARISTOTLE Heart Rhythm Data
(ITT Population)

Rhythm (Relevant Study Visit(s) with ECG)	Total N=18,201 n (%)
Atrial Fibrillation (Screening)	14,753 (81)
Atrial Fibrillation (At least one)	15,959 (88)
Atrial Fibrillation (All)	12,673 (70)
Atrial Flutter (Screening)	802 (4)
Atrial Flutter (At least one)	1192 (7)
Atrial Flutter (All)	187 (1)
Neither AFib nor AFl (Screening)	2712 (15)
Neither AFib nor AFl (All)	1845 (10)
Sinus Rhythm (Screening)	2292 (13)

An analysis of medications received prior to baseline reveals no imbalances between the groups in the use of any of the classes of medications expected to be used by the enrolled patients, many of whom had hypertension and heart failure. The most commonly used medication classes (> 30% of subjects) were beta blockers, diuretics, ACE inhibitors, antiarrhythmics (including digoxin), lipid lowering agents, oral antiplatelet agents, and calcium channel blockers (**Table 34**).

Oral anticoagulants (OACs) were being taken by 3% of subjects, despite the fact that subjects were to discontinue oral anticoagulants before randomization. It is possible that these subjects simply had not yet discontinued their oral anticoagulant prior to the baseline visit. In any event, patients taking oral anticoagulants at baseline were well balanced between the treatment arms. CYP3A4 inhibitors (discouraged but not forbidden by the protocol) were taken by about 0.4% of subjects.

Table 34 ARISTOTLE – Concomitant Medications Received at Randomization
ITT population

Medication Class N (%)	Apixaban N=9120	Warfarin N=9081	Total N=18,201
Beta Blocker	5797 (63.6)	5685 (62.6)	11482 (63.1)
Diuretic	4823 (52.9)	4880 (53.7)	9703 (53.3
ACE Inhibitor	4446 (48.8)	4439 (48.9)	8885 (48.8)
Antiarrhythmics*	4208 (46.1)	4238 (46.7)	8446 (46.4)
Lipid Lowering	4104 (45.0)	4095 (45.1)	8199 (45.0)
Oral Antiplatelet	3037 (33.3)	2953 (32.5)	5990 (32.9)
Calcium Channel Blocker	2744 (30.1)	2823 (31.1)	5567 (30.6)
ARB	2204 (24.2)	2108 (23.2)	4312 (23.7)
Diabetes Medications	1774 (19.5)	1784 (19.6)	3558 (19.5)
Acid Suppressing Drugs	1683 (18.5)	1667 (18.4)	3350 (18.4)
NSAIDs	752 (8.2)	768 (8.5)	1520 (8.4)
Alpha Blocker	672 (7.4)	681 (7.5)	1353 (7.4)
Nitrates	646 (7.1)	603 (6.6)	1249 (6.9)
Antidepressants	514 (5.6)	561 (6.2)	1075 (5.9)
Oral Anticoagulants	268 (2.9)	287 (3.2)	555 (3.0)
CYP3A4 Inhibitors	34 (0.4)	37 (0.4)	71 (0.4)
Other	5537 (60.7)	5480 (60.3)	11017 (60.5)

*Includes digoxin

6.1.3 Subject Disposition and Compliance with Study Drug

6.1.3.1 Disposition

There were 20,998 subjects screened for entry into ARISTOTLE; 18,201 were randomized, with 2797 screen failures and a screen failure rate of 13.3%.

Reasons for screen failure in at least 50 subjects, in decreasing order of frequency, included violation of the one of the inclusion/exclusion criteria (N=1575, 7.5 % of all subjects screened), withdrawal of consent (N=951, 4.5%), lost to follow-up during screening period and adverse event occurring during the screening period (each with N=59, 0.3%), and compliance issues and "administrative reason by Applicant" (each with N=56, 0.3%). Case records for 3 subjects (<0.1%) included no reason for why they were not randomized.

Table 35 provides information on subjects who discontinued follow-up prior to the end of the ITP. Such subjects who left the study alive did not have follow-up for most study endpoints, but some may have information on time of death from death databases or contact with a subject's family. There were 440 and 458 subjects who discontinued

follow-up alive in the apixaban and warfarin arms, respectively. The majority of these subjects withdrew consent.

Table 35 ARISTOTLE - Early Discontinuation of Follow-up
Before end of ITP, ITT Pop.

	APIXABAN N=9120 n (%)	WARFARIN N=9081 n (%)
COMPLETED ITP	**8105 (88.9)**	**8000 (88.1)**
DID NOT COMPLETE ITP	**1015 (11.1)**	**1081 (11.9)**
DEATH	575 (6.3)	643 (7.1)
DISCONTINUED ALIVE	440 (4.8)	458 (4.8)
WITHDREW CONSENT	260 (2.9)	259 (2.9)
LOST TO FOLLOW-UP	180 (2.0)	179 (2.0)

shows vital status at "end of study," using all information available before database lock, including from death registries and contacts with a subject's family. It is my understanding from conversations with Applicant representatives that if the last known "contact" (information from any source) indicated a patient was alive, the patient would be classified as alive, even if the contact occurred before the end of the ITP. Note that 656 and 718 subjects in the apixaban and warfarin arms, respectively were known to be dead; 180 and 200 subjects, respectively, had unknown vital status.

We requested analogous information for vital status at the end of the ITP to help us interpret the Applicant's (b) (4) the requested data are shown in **Table 37**. Compared to **Table 36,** in **Table 37** there are fewer known deaths in the apixaban arm (656 and 603, respectively), and considerably more subjects with unknown vital status (180 and 288, respectively). Data for the warfarin arm show the same pattern, with slightly more deaths

> *Reviewer Comment: The fact that that there are more warfarin patients than apixaban arm patients with unknown vital status subjects at the end of the ITP is somewhat reassuring. However, possible unblinding may have affected the diligence of site staff in seeking out vital status information on some subjects who discontinued follow-up. This issue is discussed further in* ***Sec. 1.2.***

Table 36 ARISTOTLE – Vital Status at "End of Study"
ITT Population

	APIXABAN N=9120 n (%)	WARFARIN N=9081 n (%)
ALIVE	8284 (90.8)	8163 (89.9)
DEAD	656 (7.2)	718 (7.9)
UNKNOWN	180 (2.0)	200 (2.2)
WITHDREW CONSENT	92 (1.0)	107 (1.2)
LOST TO FOLLOW-UP	35 (0.4)	34 (0.4)
OTHER	53 (0.6)	59 (0.6)

Source: CSR Table S.2.1D

Table 37 ARISTOTLE – Vital Status at End of Intended Treatment Period
ITT Population

	APIXABAN N=9120 n (%)	WARFARIN N=9081 n (%)
ALIVE	8229 (90.2)	8110 (89.3)
DEAD	603 (6.6)	669 (7.4)
UNKNOWN	**288 (3.2)**	**302 (3.3)**
SUBJECT WITHDREW CONSENT	150 (1.6)	155 (1.7)
LOST TO FOLLOW-UP	138 (1.5)	147 (1.6)

Table 38 provides information on subjects who discontinued study drug prior to the end of the ITP. There were fewer such subjects, 2310 (25.3%) vs.2493 (27.5%), in the apixaban arm compared to the warfarin arm. Bleeding was the most common reason for discontinuation other than death.

Table 38 ARISTOTLE - Early Discontinuation of Treatment
(Before end of ITP, ITT Pop.)

	APIXABAN N=9120 n (%)	**WARFARIN N=9081 n (%)**
DISCONTINUED TREATMENT	**2310 (25.3)**	**2493 (27.5)**
DEATH	331 (3.6)	349 (3.8)
ADVERSE EVENT	679 (7.4)	738 (8.1)
STROKE	75 (0.8)	108 (1.2)
SYSTEMIC EMBOLISM	14 (0.2)	8 (<0.1)
MYOCARDIAL INFARCTION	24 (0.3)	15 (0.2)
BLEEDING	154 (1.7)	190 (2.1)
OTHER ADVERSE EVENT	424 (4.6)	438 (4.8)
NOT REPORTED	1 (<0.1)	0
SUBJCT REQUEST TO DISCONTINUE	921 (10.1)	989 (10.9)
INCONVENIENCE	310 (3.4)	336 (3.7)
INR MONITORING ISSUES	116 (1.3)	136 (1.5)
PERCEIVED SIDE EFFECTS*	109 (1.2)	141 (1.6)
OTHER	448 (4.9)	435 (4.8)
NOT REPORTED	3 (<0.1)	1 (<0.1)
LOST TO FOLLOW-UP	51 (0.6)	39 (0.4)
POOR/NON-COMPLIANCE	57 (0.6)	77 (0.8)
PREGNANCY	1 (<0.1)	0
SUBJECT NO LONGER MEETS STUDY CRITERIA	87 (1.0)	100 (1.1)
PHYSICIAN REFUSED TO CONTINUE TREATMENT	81 (0.9)	89 (1.0)
PERCEIVED RISK	21 (0.2)	31 (0.3)
DESIRE TO HAVE SUBJECT ON OPEN-LABEL WARFARIN	38 (0.4)	32 (0.4)
OTHER	22 (0.2)	29 (0.3)
NOT REPORTED	1 (<0.1)	0
OTHER	80 (0.9)	92 (1.0)
NOT REPORTED	11 (0.1)	12 (0.1)

* Not reported as adverse event by investigator

The duration of double-blind treatment was similar in each arm at 88 weeks (1.7 years), with a range of 1 day to 215 weeks (4.1 years).

6.1.3.2 Compliance with Study Drug

Table 39 is a display of compliance with study drug as determined by counts of returned tablets (apixaban active for apixaban arm patients and apixaban placebo for warfarin arm patients). Note that mean compliance in both arms was over 100%, suggesting possible irregularities in ascertainment. The Applicant did not comment on this finding

Table 39 Compliance with Study Drug by Counts of Returned Tablets

	Apixaban	Apixaban Placebo (surrogate for Warfarin)
All Patients	N=9088	N=9052
Percent Compliance, N (%)		
<80%	568 (6.3)	572 (6.3)
80%-120%	8355 (91.9)	8292 (91.6)
>120%	65 (1.8)	188 (2.1)
Mean (SD)	106.4 (379.55)	123.2 (673.39)
Median	97.7	97.5
MIN, MAX	(0.0 , >200.0)	(0.0 , >200.0)
VKA Naïve Patients	N=3892	N=3872
Mean (SD)	114.7 (561.07)	139.1 (860.45)
Median	97.7	97.4
MIN, MAX	(0.0 , >200.0)	(0.0 , >200.0)
VKA Experienced Patients	N=5196	N=5180
Mean (SD)	100.3 (126.91)	111.3 (488.61)
Median	97.7	97.5
MIN, MAX	(0.0 , >200.0)	(0.0 , >200.0)

6.1.3.3 Analysis Populations

Analysis populations for the ARISTOTLE efficacy analyses are shown in Table 40.

Table 40 ARISTOTLE – Efficacy Analysis Populations

Population [1]	Apixaban [2]	Warfarin [2]	Total [2]
ITT (All randomized)	9120	9081	18, 201
Treated	9088 (99.65%)	9052 (99.68%)	18,140 (99.66%)
Evaluable	8518 (93.40%)	8475 (93.33%)	16,993 (93.36%)

1 ITT Population – All randomized patients ; Treated Population – Randomized patients who took at least one dose of study drug; this is also the safety population; Evaluable Population – ITT population minus patients important protocol violations that could affect efficacy (see criteria in text below).
2 In each column, % is calculated using the ITT population as the denominator

Patients were excluded from the evaluable population for the following reasons:

1. They received no study treatment
2. They received the wrong study treatment at first visit.
3. Study drug compliance was less than 80%.

A patient who received the correct study drug initially and then was given the wrong study drug (through an error at the site) was included in the evaluable population but was censored on the day that the wrong drug was first dispensed.

Table 41 Exclusions and Censoring in the Evaluable Population
(Base population is ITT population)

	Apixaban	Warfarin	Total
ITT (All randomized)	9120	9081	18,201
REASONS FOR EXCLUSION			
Not dosed	32	29	61
Incorrect study drug at Visit 1	2	5	7
Compliance < 80%	568	572	1140
TOTAL EXCLUDED	602	606	12,080
Censored for receiving wrong study drug after Visit 1	662	104	766
Total who received wrong study drug	664	109	773

*Reviewer Comment: This reviewer believes that the information in the Applicant's dataset regarding the number of medication errors is not reliable, and analyses of endpoint data based on such information should not be relied upon. For further information on this issue, see Sec. **3.1**.*

6.1.4 Analysis of Primary Endpoint

The primary efficacy endpoint was the composite of stroke (ischemic, hemorrhagic or of unknown type) or non-CNS systemic embolism. The primary efficacy analysis was the

time to the first occurrence of a primary endpoint event in the ITT Population during the intended treatment period (ITP, defined as the time from randomization up to January 30, 2011, which was estimated late in the study as the day the primary efficacy event target would be reached). The Applicant's intent (for US regulatory purposes) was to establish that apixaban is non-inferior to warfarin, using a non-inferiority (NI) margin of 1.38 for the hazard ratio and a 95% CI.

The results for the primary endpoint analysis are shown below in **Table 42** and **Figure 10.** Unless otherwise specified, all efficacy analyses pool data from the roughly 5% of patients who received apixaban 2.5 mg bid (the dose in those with at least 2 of the 3 specified risk criteria for bleeding) with the remainder of patients who received 5 mg bid.

There were 212 and 265 first primary efficacy events in the apixaban and warfarin arms, respectively, yielding respective event rates of 1.27 and 1.60 events per 100 patient-years and a hazard ratio of 0.79 (95% CI 0.66 to 0.95). The p for non-inferiority was <0.0001 with a NI margin of 1.38. The p for superiority was also significant, with a value of 0.0114. Note the in the pre-specified event hierarchy, if non-inferiority is established, the Applicant may examine superiority of subsequent events at a one-sided p of 0.02499, which takes into account one interim analysis.

Table 42 ARISTOTLE -- Primary Efficacy Endpoint Results
Stroke or non-CNS systemic embolism (Adjudicated data, ITT Population, ITP)

	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
Time to First Event						
Any Primary Efficacy Endpoint Event	**212 (2.32)**	**1.27**	**265 (2.92)**	**1.60**	**0.79 (0.66, 0.95)**	**0.0114**
Ischemic or unspecified stroke	159 (1.74)	-	173 (1.91)	-	-	-
Hemorrhagic stroke	38 (0.42)	-	76 (0.84)	-	-	-
Systemic embolism	15 (0.16)	-	16 (0.18)	-	-	-
Event at Any Time During ITP						
Any stroke	199 (2.18)		250 (2.75)			
Ischemic stroke	140 (1.54)		136 (1.50)			
Ischemic stroke with hem-orrhagic conversion	12 (0.13)		20 (0.22)			
Hemorrhagic stroke	40 (0.44)		78 (0.86)			
Subarachnoid hemorrhage	2 (0.02)		4 (0.04)			
Subdural hematoma	5 (0.05)		6 (0.07)			
Intraparenchymal hemorrhage	31 (0.34)		66 (0.73)			
Stroke of uncertain type	14 (0.15)		21 (0.23)			
Systemic embolism	15 (0.16)		17 (0.19)			

Figure 10 ARISTOTLE - Time to First Primary Efficacy Endpoint Event
Time to first event – stroke or non-CNS systemic embolism (Adjudicated data, ITT Population, ITP)

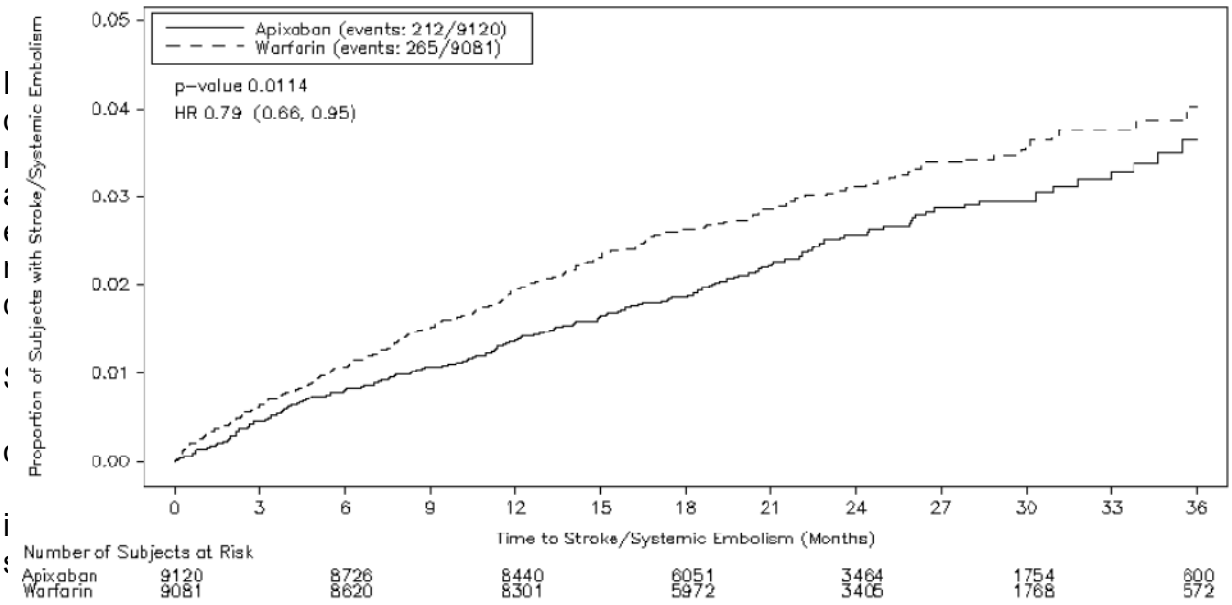


Table 43 ARISTOTLE – Additional Analyses of the Primary Endpoint Results
Time to first event – stroke or systemic embolism (Adjudicated data, Treated Population, various observation periods)

Observation Period - from first dose to:	Apixaban (N=9088)		Warfarin (N=9052)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
Last dose + 2 days	176 (1.94)	1.14	225 (2.49)	1.49	0.77 (0.63, 0.93)	0.0080
Last dose + 7 days	184 (2.02)	1.18	236 (2.61)	1.55	0.76 (0.63, 0.93)	0.0060
Last dose + 30 days	218 (2.40)	1.36	255 (2.82)	1.62	0.84 (0.70, 1.00)	0.0526

Table 44 is a display of analogous results in the somewhat smaller Evaluable Population. Note that the most common reason for Treated Population patients to be excluded from the Evaluable Population was compliance with study medication dosing <80%, based on returned tablet counts. The same pattern of increasing hazard ratios over the period from last dose +7 days to last dose + 30 days is observed as in the

Treated Population, although, as one might expect, the hazard ratios are more favorable to apixaban than those for the Safety Population.

Table 44 ARISTOTLE – Additional Analyses of the Primary Endpoint Results
Time to first event – stroke or systemic embolism (Adjudicated data, Evaluable Population, various observation periods)

Observation Period - from first dose to:	Apixaban (N=8518)		Warfarin (N=8475)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
Last dose + 2 days	138 (1.62)	0.96	200 (2.36)	1.39	0.69 (0.56, 0.86)	0.0009
Last dose + 7 days	144 (1.69)	1.00	210 (2.48)	1.45	0.69 (0.56, 0.85)	0.0006
Last dose + 30 days	174 (2.04)	1.17	228 (2.69)	1.52	0.77(0.63, 0.94)	0.0086

Reviewer comment: The tables above support the superiority of apixaban to warfarin during treatment. However, the various time cuts for events in the Treated and Evaluable populations show more events in the apixaban arm compared to the warfarin arm in the interval between the end of the last dose + 7 days analysis and the last dose + 30 days analysis. For example, in the Treated Population, there were 34 vs.19 events that occurred during this 23 day period in the apixaban and warfarin arms, respectively. This finding is explored further in Section **6.1.10.2**.

FDA performed an analysis of how many additional primary endpoint events would be required in the apixaban arm to negate the findings of non-inferiority and superiority in the Applicant's primary efficacy analysis. The analysis was done using two different methods. The first method calculated odds ratios for the number of events, without taking time into effect. Events were simply added to the apixaban arm results or subtracted from the warfarin arm results (but not both in the same analysis), an exact test was used to calculate statistical significance. The second used the hazard ratios for time to event. For adjustments to in the apixaban arm results , one day was subtracted (if possible) from the time to event for patients with the shortest time to event, one at a time. For adjustments to the warfarin arm results, one day was added to the time to event for patients with the shortest time to event, one patient at a time.

The results are displayed in **Table 45**. The non-inferiority result are quite robust: using the odds ratio approach, it would require 98 additional events in the apixaban arm or 78 fewer events in the warfarin arm to negate the finding of non-inferiority with a margin of 1.38. Using the hazard ratio approach, the non-inferiority results are even more robust.

The superiority findings are considerably less robust. Using either approach, it would take 12 additional events in the apixaban arm or 13 fewer event in the warfarin arm to negate superiority.

Table 45 Sensitivity Analyses of Non-Inferiority and Superiority Findings for the Primary Efficacy Endpoint
(Number of additional events in the apixaban arm or fewer events in the warfarin arm needed to negate finding)

	By Odds Ratio Change in No. of Events		By Hazard Ratio Change in No. of Events	
	Apixaban	Warfarin	Apixaban	Warfarin
Non-inferiority [1]	212→ 310 =↑**98**	265 → 187 =↓**78**	212 → 314 =↑**102**	265→185=↓**80**
Superiority	212→224 =↑**12**	265 → 252=↓**13**	212 → 224=↑**12**	265→252=↓**13**

[1] Based on NI margin of 1.38

6.1.5 Other Efficacy Endpoints

6.1.5.1 Death

Table 46 is a display of event rates, hazard ratios, and p-values (superiority) in the ITT population during the Intended Treatment Period for secondary endpoint data, including the components of the primary endpoint, various categories of stroke, all-cause death and several categories of cause-specific death, and myocardial infarction.

All-cause death was the fourth step in hierarchical analysis plan. The first 3 steps in the hierarchy (1, non-inferiority for the primary efficacy endpoint in the ITT population during the intended treatment period; 2, superiority for the primary efficacy endpoint in the ITT population during the intended treatment period; and 3, superiority for the primary safety endpoint (ISTH major bleeding) in the Treated Population during the treatment period), all met their respective, prespecified criteria for statistical significance (see Sec. 7 for a discussion of bleeding). Thus, observed significant finding for superiority of apixaban for all-cause mortality (b) (4). However, nominal success of any analysis in the hierarchy means that there is no increase in alpha error inherent in moving down to the next analysis in the hierarchy. It does not necessarily imply regulatory recognition of the finding for the purposes of labeling.

Table 46 ARISTOTLE – Secondary Endpoints - Deaths
Adjudicated endpoints, ITT Pop, ITP

Endpoint	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
All-cause death	603 (6.61)	3.52	669 (7.37)	3.94	0.89 (0.80, 1.00)	0.0465
CV death (Caused by ↓)	308 (3.38)	1.80	344 (3.79)	2.02	0.89 (0.76, 1.04)	-
Stroke	38 (0.42)	-	65 (0.72)	-	-	-
Systemic embolism	1 (0.01)	-	2 (0.02)	-	-	-
MI	21 (0.23)	-	17 (0.19)	-	-	-
Sudden death	126 (1.38)	-	129 (1.42)	-	-	-
Heart failure	76 (0.83)	-	92 (1.01)	-	-	-
Other CV cause	23 (0.25)	-	22 (0.24)	-	-	-
Unobserved death	23 (0.25)	-	17 (0.19)	-	-	-
Non-CV death (Caused by ↓)	196 (2.15)	1.14	208 (2.29)	1.22	0.93 (0.77, 1.13)	-
Bleeding	15 (0.16)	-	17 (0.19)	-	-	-
Malignancy	60 (0.66)	-	66 (0.73)	-	-	-
Infection	67 (0.73)	-	52 (0.57)	-	-	-
Trauma	7 (0.08)	-	13 (0.14)	-	-	-
Respiratory failure	19 (0.21)	-	35 (0.39)	-	-	-
Other non-CV cause	28 (0.31)	-	25 (0.28)	-	-	-
Unknown cause of death	99 (1.09)	0.58	117 (1.29)	0.69	0.84 (0.64, 1.09)	-

The overall results for all-cause death significantly favored apixaban (HR 0.89, 95% CI, 0.80 to 1.00, p=0.465). As one would expect, most of the advantage of apixaban over warfarin (66 fewer deaths) was due a reduced number of CV deaths (36 fewer than in the warfarin arm), but there were numerical differences favoring apixaban for non-CV death and death of unknown cause. Note that unwitnessed deaths of unknown cause were counted as CV deaths.

One potential shortcoming of the ITT analysis is that it includes deaths occurring months after the last dose of study drug in patients who discontinued early. Such deaths seem unlikely to be due to the effects of study treatment. Accordingly, we asked the Applicant to perform three analyses of treated patients, from first dose of study to: last dose + 2 days, last dose + 7 days, and last dose + 30 days (**Table 47**).

Table 47 ARISTOTLE – Additional Analyses of Adjudicated All-cause Death
Safety Population, Various Event Windows

Observation Period - from first dose to:	Apixaban (N=9088)		Warfarin (N=9052)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
Last dose + 2 d	265 (2.92)	1.70	296 (3.27)	1.94	0.87 (0.74, 1.03)	0.113
Last dose + 7 d	330 (3.63)	2.10	372 (4.11)	2.42	0.87 (0.75, 1.00)	0.056
Last dose + 30 d	429 (4.72)	2.65	471 (5.20)	2.97	0.89 (0.78, 1.01)	0.076

Unlike the primary endpoint analyses, where the last dose +2 days and last dose +7 days analyses were both more favorable for apixaban than the ITT analysis, here all 3 analyses off all-cause death in the treated population failed to show a significant difference between apixaban and placebo and were less favorable (i.e., had a higher p value) than the ITT analysis of death.

Reviewer Comment: These results tend to disfavor a claim of superiority of apixaban over warfarin, but they could be viewed as supporting a claims of non-inferiority to warfarin in terms of death (using a NI of 1.015) if one accepts the death finding in the Hart meta-analysis.

Table 48 Sensitivity Analyses of Superiority Finding for All-cause Death
(Number of additional events in the apixaban arm or fewer events in the warfarin arm needed to negate finding)

Finding	By Odds Ratio Change in No. of Events		By Hazard Ratio Change in No. of Events	
	Apixaban	Warfarin	Apixaban	Warfarin
Superiority	603→ 605=↑2	669→667=↓2	603→ 604=↑1	669→668=↓1

6.1.5.2 Other Secondary Endpoints

Other secondary endpoints included various stroke subtypes, systemic emboli, and MI. Results for these endpoints are displayed in **Table 49.**

Table 49 ARISTOTLE – Secondary Endpoints – Stroke, Systemic Emboli and MI
Adjudicated endpoints, ITT Pop, ITP

Endpoint	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR (95% CI)	2-sided p (super-iority)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr		
Ischemic/unspecified stroke	162 (1.78)	0.97	175 (1.93)	1.05	0.92 (0.74, 1.13)	0.4220
Hemorrhagic stroke	40 (0.44)	0.24	78 (0.86)	0.47	0.51 (0.35, 0.75)	0.0006
Systemic embolism	15 (0.16)	0.09	17 (0.19)	0.10	0.87 (0.44, 1.75)	0.7020
MI	90 (0.99)	0.53	102 (1.12)	0.61	0.88 (0.66, 1.17)	0.3720

As in the studies of rivaroxaban and dabigatran, the rate of hemorrhagic stroke was substantially lower the in apixaban arm; this finding was statistically significant. The rate of ischemic/unspecified death numerically favored apixaban, but the results were not statistically significant. Thus, the pattern of benefit in ARISTOTLE is not like that of the RE-LY study of dabigatran, where both hemorrhagic and ischemic/unknown stroke rates were significantly reduced by dabigatran.

Table 50 is a display of stroke rates by severity strata defined by the Modified Rankin Score: non-disabling score of 0-2), disabling (3-6) and fatal (6). This was a post-hoc analysis. Results in all strata numerically favor apixaban. Confidence intervals are provided for interest only.

Reviewer Comment: These data are post-hoc, and labeling should not suggest that differences in any stratum are statistically significant. However, they are reassuring in that they suggest that the reduction in stroke rate with apixaban is clinically meaningful.

Table 50 ARISTOTLE – Stroke Severity
ITT Pop, ITP (Based on Modified Rankin Score)

Rankin Score*	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR (95% CI)
	n (%)	Events / 100 pt-yr	n (%)	Events / 100 pt-yr	
0 - 2	63 (0.69)	0.37	70 (0.77)	0.42	0.89 (0.64, 1.26)
3 – 6	85 (0.93)	0.50	118 (1.30)	0.71	0.71 (0.54, 0.94)
6 (Fatal)	32 (0.35)	0.19	54 (0.59)	0.32	0.59 (0.38, 0.91)
Rankin missing	59 (0.65)	0.35	62 (0.68)	0.37	0.95 (0.66, 1.35)

* Includes all subtypes of strokes.

A sensitivity analysis was performed to determine the number of additional events in the apixaban arm or fewer event in the warfarin arm would be to negate the finding of superiority for all-cause mortality, which on its face would appear to be quite fragile. The methodology described in connection with the sensitivity analysis for the primary endpoint results (see the paragraph immediately above **Table 45** for details) was used. Using the odds ratio approach, which does not take time into account, either two more events in the apixaban arm or two fewer events in the warfarin arm would negate superiority. Using the hazard ratio approach, which takes time into account, only one more event in the apixaban arm or one fewer event in the warfarin arm would be required to negate superiority (see **Table 48)**.

This result is not surprising given that the upper limit of the hazard ratio for all-cause death is 1.00. As noted in Sec.**3**, there were issues regarding integrity of the study blind which may have affected event ascertainment, and there were GCP issue at a site 1200 in China, where there were 3 deaths in the warfarin arm, compared to 2 deaths in the apixaban arm. If, for example, Site 1200 were eliminated from the analysis, statistical significance for superiority for all-cause death probably would be lost. There are also ongoing analyses of medication errors that might have affected study outcomes (Sec. **3.1.1)**.

There are additional reasons for and against accepting the Applicant's proposed language indicating that apixaban is superior to warfarin for all-cause death. The reasons in favor of apixaban include:

1) The ITT analysis, usually considered the most conservative, supports superiority
2) On treatment analyses in the safety population and per protocol population support superiority

3) There were significantly fewer strokes (including significantly disabling and fatal strokes) in the apixaban arm. One would expect a drug that reduced stroke risk to show an effect on all cause death.
4) The results of AVERROES (which was terminated early because of superiority of apixaban for stroke/SE prevention) show a strong trend in favor of apixaban for all-cause death vs. aspirin, which ought be at least as good as placebo.

Reasons against accepting the Applicant's proposed language include:

5) There were 590 patients who left the study early and had unknown vital status at the end of the ITP. Such patients were censored at the last known contact prior to the end of the ITP, when they were known to be alive. If ascertainment of vital status was affected by unblinding (i.e., less effort was made to uncover the vital status of apixaban arm patients), the results could be biased. However, there is no evidence that any such bias occurred.
6) The effect of apixaban on all-cause death is most apparent at sites below the median TTR. At sites above the median, the HR > 0.9, and is 1.2 in the Applicant's analysis of the best quartile of TTR (**Table 68**) and is 1.0 in FDA's analysis of the best quartile of TTR (**Table 69**). Note that this objection is most relevant to the comparative claim, not an absolute claim of reduction in all-cause death.

Also, if one accepts that warfarin prevents death, one might show that apixaban is NI to warfarin even of not superior. Then (b) (4) might be appropriate, even though a comparative claim would not. There are several problems with this approach:

1) The Hart meta-analysis showing the benefit of warfarin on all cause death has not been vetted by FDA, although such vetting could be undertaken.
2) Even if one accepts the results of the Hart meta-analysis, the resulting NI margin of 1.015 is so close to 1.0 that nearly any analysis that fails to support superiority would also fail to support non-inferiority.

6.1.6 Other Endpoints

See Sec. **7** for safety information.

6.1.7 Subpopulations

6.1.7.1 Subpopulations of the global study population

Results for the primary efficacy endpoint were analyzed in various subgroups of patients, based on geographic region, demographic factors, disease-related factors, and prior medication use. With one exception (the subgroup of persons with age < 65 years at entry, the point estimate for the HR was < 1 in all major subgroups (**Table 51**).

Table 51 Primary Efficacy Endpoint Results by Subgroup
ITT Population, ITP

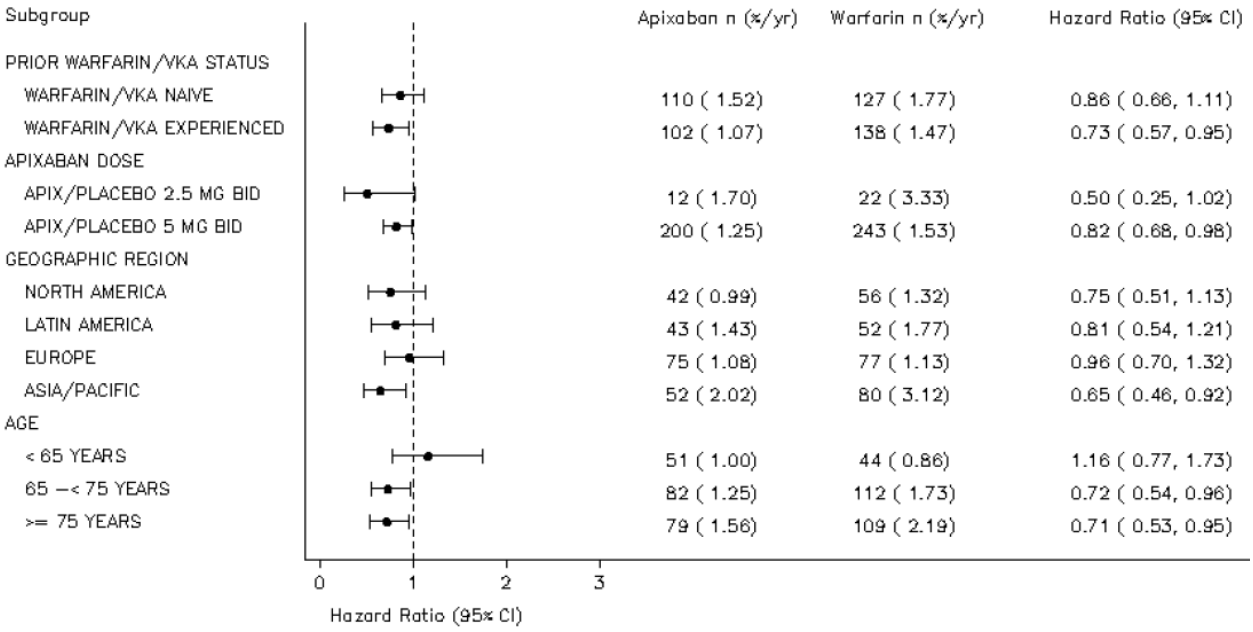
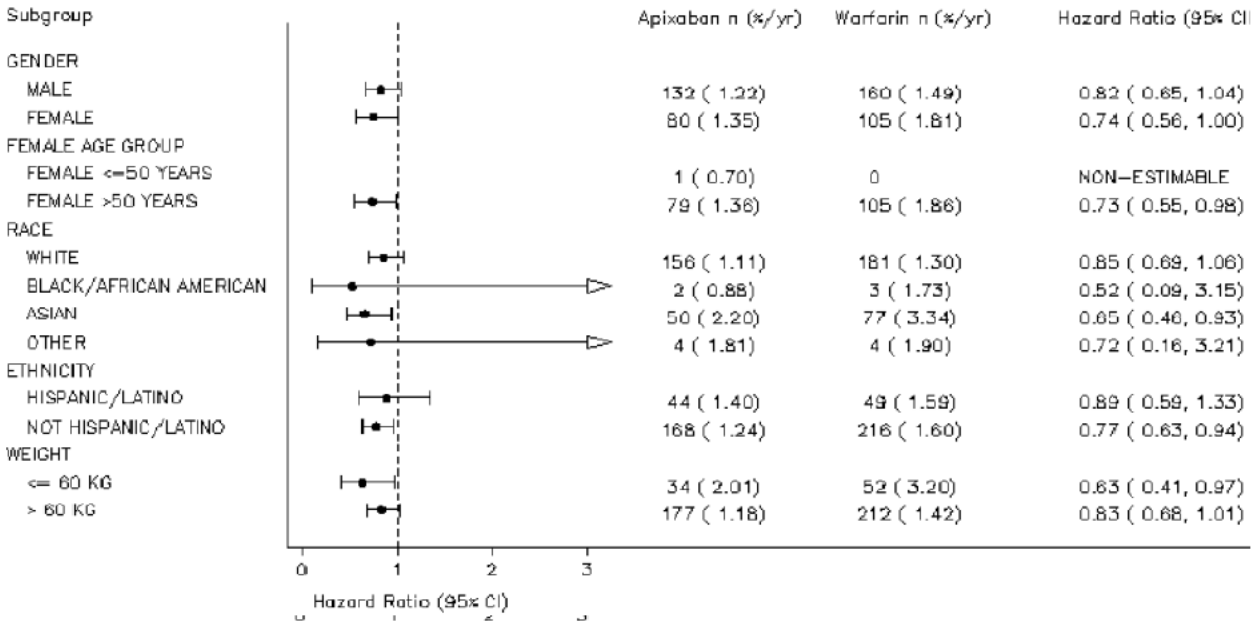


Table 51 Continued



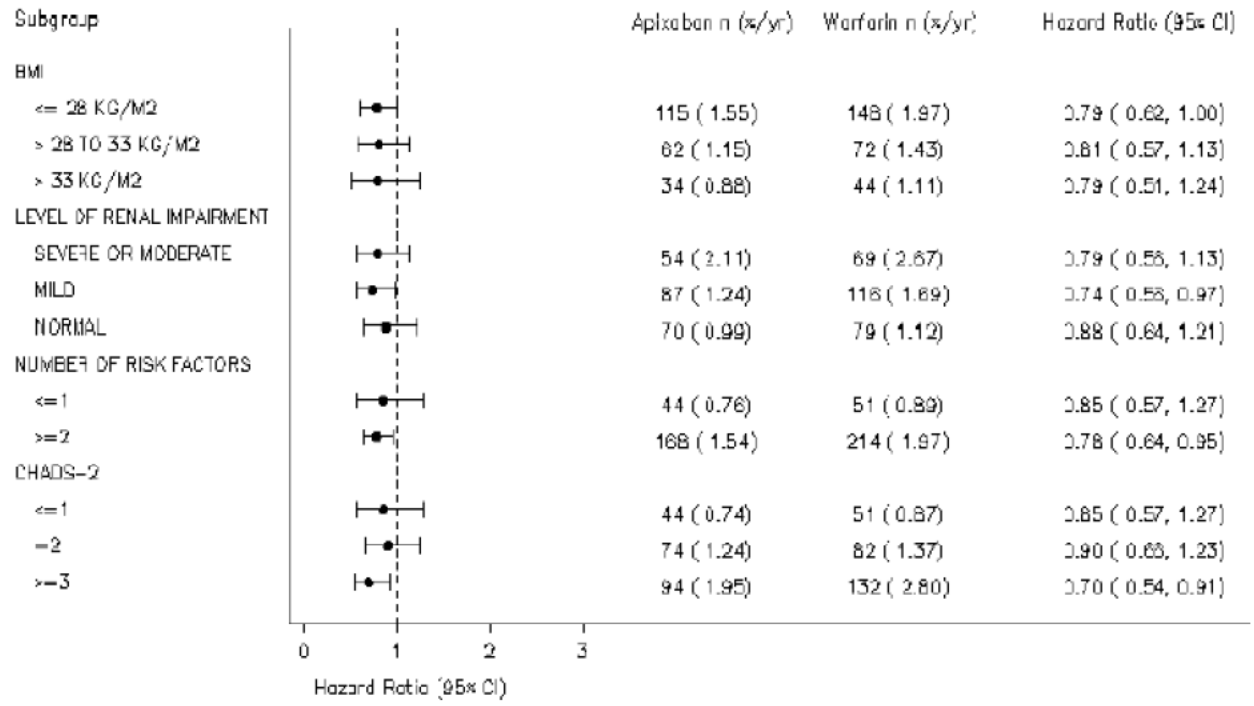


Table 51 Continued

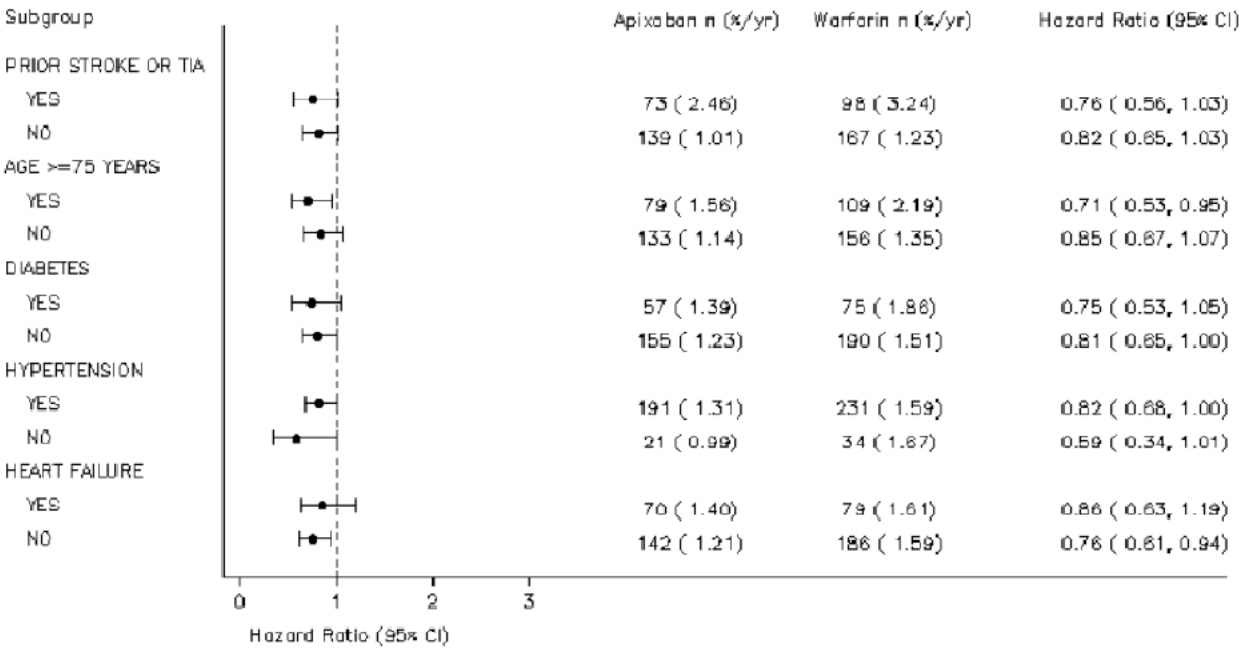


Table 52 provides treatment by subgroup interaction p values for the various subgroups analyzed in **Table 51** No p value was less than 0.05, and only one interaction term, for age (p=0.12) was less than 0.15. Examination of the results for age, which was divided into 3 subgroups, reveals that the results in the subgroup with age < 65 years slightly favored warfarin, while the results in the 65 - <75 and ≥ 75 years subgroups both favored apixaban and were quite similar in terms of HR (see page **142**). Note that age < 75 and age ≥ 75 subgroups were analyzed for interaction in another analysis, and this interaction term was 0.36. Unexpectedly, the HR for VKA experienced patients was more favorable for apixaban than the one for VKA naïve patients (see page **142**).

Table 52 Treatment by Subgroup Interaction Terms for Primary Endpoint Results

Characteristic	Interaction p	Characteristic	Interaction p
Prior VKA status	0.3879	Number of risk factors	0.7002
Apixaban dose	0.2165	$CHADS_2$	0.4457
Geographic region	0.4356	Prior stroke or TIA	0.7090
Age	0.1156	Age>=75	0.3602
Gender	0.5957	Diabetes	0.7050
Female age group	0.9760	Hypertension	0.2629
Race	0.6143	Heart failure	0.4998
Ethnicity	0.5608	Aspirin at randomization	0.4415
Weight	0.2558	Clopidogrel at randomization	0.8570
BMI	0.9908	Type of AF	0.7047
Level of renal impairment	0.7185		

Table 53 provides information on the primary endpoint results in subgroups based on heart rhythm at study entry and at other times during the study. Note that ECGs were obtained at screening, yearly during treatment, and at the final visit. Study patients included those with a screening ECG showing AFib or atrial flutter (AFl), and also those with either rhythm documented on 2 occasions at least 2 weeks apart in the 12 months prior to enrollment. The data indicate that while about 4% of subjects had AFl at screening and 7% had AFl on at least one study ECG, only 1% of subjects had this rhythm on all study ECGs, consistent with usual description of AFl as an unstable rhythm. There were too few events in patients in any of the AFl subgroups mentioned above to determine whether apixaban therapy was beneficial compared to warfarin, although hazard ratios favor apixaban in each of the 3 AFl subgroups.

Table 53 Primary Endpoint Results in Subgroups Based on Heart Rhythm
(ITT population, ITP)

Rhythm (% of total)	Apixaban (N=9120)		Warfarin (N=9081)		A vs. W HR (95% CI)
	n / N	Events / 100 pt-yr	n / N	Events / 100 pt-yr	
AFib at Screening (81)	183/7384	1.36	230/7369	1.72	0.79 (0.65, 0.96)
AFib on any study ECG (88)	188/7974	1.28	241/7985	1.65	0.78 (0.64, 0.94)
AFib on all ECGs (70)	167/6346	1.46	211/6327	1.87	0.78 (0.64, 0.96)
AFib on no study ECG (12)	24/1139	1.17	24/1088	1.21	0.95 (0.54, 1.67)
AFl at Screening(4)	8/420	1.02	12/382	1.70	0.59 (0.24, 1.45)
AFl on any study ECG (7)	10/609	0.86	16/583	1.42	0.60 (0.27, 1.31)
AFl on all ECGs (1)	3/99	1.91	4/88	3.12	0.58 (0.13, 2.66)
Neither* at Screening (15)	23/1348	0.91	24/1364	0.93	0.97 (0.55, 1.72)
Neither* on all ECGs (10)	18/932	1.06	18/913	1.07	0.98 (0.51, 1.88)

*Neither = neither AFib nor AFl

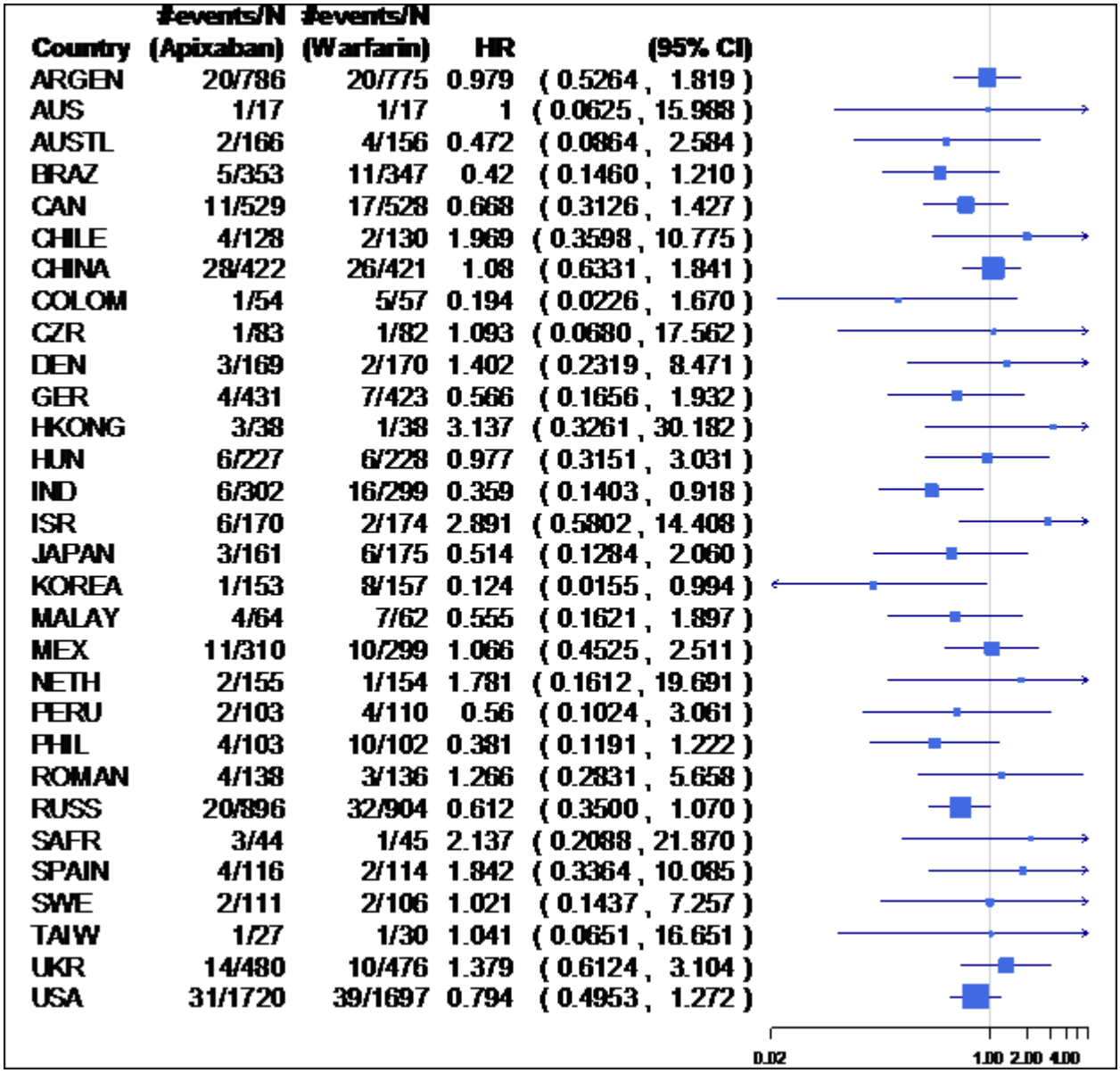
Patients with neither AFib nor AFl at screening and those with neither rhythm on any study ECG were not uncommon in this study (about 15% of the total for the former subgroup and 10% for the latter). Both subgroups had low event rates in each arm (about 1 event per 100 pt-years) and hazard ratios near 1.

On the other hand, each of the AFib subgroups (those with AFib at screening, on any ECG, and on all study ECGs) had fairly similar event rates in the apixaban arm, and similar (but higher) event rates in the warfarin arm, with hazard ratios very close to the overall study HR of 0.79. Those with AFib on no study ECG had low event rates in each arm (about 1.2 events per 100 patient-years) and a hazard ratio near 1.

> *Reviewer Comment: The fact that 10% of subjects had neither AFib nor AFl on all study ECGs might have confounded the interpretation of the NI analysis if nominal superiority had not been demonstrated. However, the evidence that apixaban was superior to warfarin for prevention of primary endpoint events in the overall ITT/ITP analysis, on treatment, and in the subsets of patients with AFib at some time or all times during the study is reassuring that apixaban does have a beneficial effect in AFib patients.*

Figure 11 is a forest plot of primary endpoint results by country. With the exception of China, Mexico and Ukraine, all countries that contributed more than 8 events had an apixaban vs. warfarin hazard ratio < 1.0. Additional information on the results at US sites is presentenced in Sec. **6.1.7.1.1**.

Figure 11 ARISTOTLE – Forest Plot of Primary Endpoint Results by Country
ITT population



Country	#events/N (Apixaban)	#events/N (Warfarin)	HR	(95% CI)
ARGEN	20/786	20/775	0.979	(0.5264 , 1.819)
AUS	1/17	1/17	1	(0.0625 , 15.988)
AUSTL	2/166	4/156	0.472	(0.0864 , 2.584)
BRAZ	5/353	11/347	0.42	(0.1460 , 1.210)
CAN	11/529	17/528	0.668	(0.3126 , 1.427)
CHILE	4/128	2/130	1.969	(0.3598 , 10.775)
CHINA	28/422	26/421	1.08	(0.6331 , 1.841)
COLOM	1/54	5/57	0.194	(0.0226 , 1.670)
CZR	1/83	1/82	1.093	(0.0680 , 17.562)
DEN	3/169	2/170	1.402	(0.2319 , 8.471)
GER	4/431	7/423	0.566	(0.1656 , 1.932)
HKONG	3/38	1/38	3.137	(0.3261 , 30.182)
HUN	6/227	6/228	0.977	(0.3151 , 3.031)
IND	6/302	16/299	0.359	(0.1403 , 0.918)
ISR	6/170	2/174	2.891	(0.5802 , 14.408)
JAPAN	3/161	6/175	0.514	(0.1284 , 2.060)
KOREA	1/153	8/157	0.124	(0.0155 , 0.994)
MALAY	4/64	7/62	0.555	(0.1621 , 1.897)
MEX	11/310	10/299	1.066	(0.4525 , 2.511)
NETH	2/155	1/154	1.781	(0.1612 , 19.691)
PERU	2/103	4/110	0.56	(0.1024 , 3.061)
PHIL	4/103	10/102	0.381	(0.1191 , 1.222)
ROMAN	4/138	3/136	1.266	(0.2831 , 5.658)
RUSS	20/896	32/904	0.612	(0.3500 , 1.070)
SAFR	3/44	1/45	2.137	(0.2088 , 21.870)
SPAIN	4/116	2/114	1.842	(0.3364 , 10.085)
SWE	2/111	2/106	1.021	(0.1437 , 7.257)
TAIW	1/27	1/30	1.041	(0.0651 , 16.651)
UKR	14/480	10/476	1.379	(0.6124 , 3.104)
USA	31/1720	39/1697	0.794	(0.4953 , 1.272)

Table 54 Primary Endpoint Results by Apixaban Dose

Apixaban Dose	Apixaban (N=9088)*		Warfarin (N=9052)*		A vs. W HR (95% CI)
	n/N (%)	Events / 100 pt-yr	n/N (%)	Events / 100 pt-yr	
2.5 mg bid	12 / 424	1.70	22 / 402	3.33	0.50 (0.25, 1.02)
5 mg bid	200 / 8664	1.25	243 / 8650	1.53	0.82 (0.68, 0.98)

* Numerators are from ITT population, while the denominators and total N are from the Treated Population, reflecting available data. The difference in the populations is 61 patients, about 1/3% of the ITT population.

The number of patients who were randomized to 2.5 mg apixaban was small, but the data suggest that the benefit of apixaban for the prevention of stroke/SE was maintained in that subgroup.

6.1.7.2 US patients only

6.1.7.2.1 Demographics and Disposition

Study centers in the US randomized 3417 subjects, 18.8% of patients in the global ITT population. The study arms were similar for important characteristics that might have affected outcome (see Attachment 6.) The mean $CHADS_2$ score was 2.1 in each arm, similar to the global data.

Disposition data in the US were similar in the two arms in terms of dropouts from treatment or follow-up. Discontinuation of treatment occurred in about 30% of subjects in each arm, about 3-5% more than in the global population. Discontinuation of follow-up for reasons other than death occurred in about 8% of subject sin each arm (see Attachment 6.)

6.1.7.2.2 Efficacy results

The US data for control of INR were better than the global results. Mean overall (imputed) INR in the warfarin arm was 64%, slightly higher than the global mean. . About 21% of days on warfarin were associated with INR values < 2.0, and about 15% of days were associated with INR values > 3.0.

Key efficacy results for the US population are shown in **Table 55**. Hazard ratios for the primary efficacy endpoint and all-cause mortality results were very similar to the global

results; the latter were 0.79 and 0.89, respectively. Event rates in both arms were lower in the US than in the respective arms globally for both these endpoints.

Table 55 US Patients – Key Efficacy Results
ITT Population, ITP

	Apixaban N=1720		Warfarin N=1697		HR (95% CI)
Analysis	**N (%)**	**Events / 100 pt-yr**	**N (%)**	**Events / 100 pt-yr [1]**	
Primary Efficacy Endpoint	31 (1.8)	0.95	39 (2.3)	1.20	0.79 (0.49,1.27)
All-cause Mortality	103 (6.0)	3.05	114 (6.7)	3.39	0.90 (0.69,1.18)

In summary, the US results for the primary efficacy endpoint favored apixaban to the same extend as the global results. Results for all-cause death were likewise similar to the global results.

6.1.8 Analysis of Clinical Information Relevant to Dosing Recommendations

Only one dosing regimen of apixaban was evaluated in ARISTOTLE, the primary study supporting efficacy for the proposed indication. The regimen was 5mg apixaban bid for most patients and 2.5 mg bid for patients with at least 2 of 3 prespecified risk factors for bleeding at baseline (see page **75**). As is typical, phase 2 studies in patients with AFib were not performed prior to Phase 3. Instead, the Phase 3 dosing regimen was based on the results of Phase 2 trials for another thrombotic indication.

The Applicant states that two dose-ranging Phase 2 venous thromboembolism (VTE) studies, CV185010 and CV185017, support the ARISTOTLE dosing regimen.

CV185010 was a multicenter, randomized, partially blinded trial assessing DVT prevention in trial in patients undergoing elective knee replacement, who have a high rate of post-operative DVT. There were 8 arms – each dosed for 12 ± 2 d, with 97-111 subjects per arm:

- 3 once daily doses of apixaban - 5, 10, 20 mg
- 3 bid doses of apixaban – 2.5, 5, 10 mg bid
- Enoxaparin 30 mg q 12 h
- Warfarin 5 mg QD initially, titrated to INR 1.8 to 3.0

Apixaban and enoxaparin dosing was blinded using a double dummy technique; warfarin was given open-label.

The primary endpoint was a composite of adjudicated VTE events (asymptomatic and symptomatic DVT, non-fatal PE) and all-cause death during the Evaluation period (up to the last dose of study drug + 2 days.). The primary safety endpoint was adjudicated major bleeding events.

Table 56 summarized the efficacy results of this study. The top 3 rows show the number of events, event rate in %, and the 95% CI for the event rate for the primary endpoint. Numerically, the best efficacy results were observed with 5 mg bid. The results for this dose and apixaban 10 mg bid doses were significantly superior to enoxaparin (b) (5) .

Table 56 DVT Prevention Study CV185010 – Efficacy Results

Apixaban
BMS-562247

CV185010
Clinical Study Report

Comparison of Incidences of VTE/All-cause Death during the Evaluation Period

	BMS QD 5mg (N= 97)	BMS QD 10mg (N= 105)	BMS QD 20mg (N= 110)	BMS BID 2.5mg (N= 111)	BMS BID 5mg (N= 105)	BMS BID 10mg (N= 110)	Enoxaparin (N = 109)	Warfarin (N = 109)
VTE / All-cause Death, n	11	13	9	11	5	6	17	29
Event rate (%)	11.3	12.4	8.2	9.9	4.8	5.5	15.6	26.6
95% CI	(5.8, 19.4)	(6.8, 20.2)	(3.8, 15.0)	(5.1, 17.0)	(1.6, 10.8)	(2.0, 11.5)	(9.4, 23.8)	(18.6, 35.9)
Individual Components								
DVT, n	11	13	8	10	5	5	15	29
Symptomatic PE, n	0	0	1	0	0	1	2	0
Death, n	0	0	0	1	0	0	0	0
Comparisons to BMS Low Dose								
Diff (%) (High - Low)	N/A	1.0	-3.2	N/A	-5.1	-4.5	N/A	N/A
95% CI		(-8.3, 10.4)	(-12.0, 5.3)		(-12.8, 2.0)	(-12.2, 2.9)		
Comparisons to Enoxaparin								
Ratio (%) (BMS/Enox)	0.73	0.79	0.52	0.64	0.31	0.35	N/A	N/A
95% CI	(0.33, 1.49)	(0.38, 1.56)	(0.23, 1.11)	(0.29, 1.31)	(0.09, 0.77)	(0.11, 0.82)		
Diff (%) (BMS-Enox)	-4.3	-3.2	-7.4	-5.7	-10.8	-10.1		
95% CI	(-14.0, 5.4)	(-13.0, 6.4)	(-16.5, 1.3)	(-14.9, 3.3)	(-19.5, -2.7)	(-18.9, -1.9)		

Table 57 is a display of the safety data from the same study. The results for major bleeding are displayed in the top 3 rows of data. The bottom rows show results for major bleeding/potentially significant bleeding/minor bleeding. While the and 10 mg od and 2.5 mg bid doses were associated with less major bleeding than 5mg bid, the latter dose was associated with less total bleeding than the 10 mg dose, the second-most efficacious dose. The Applicant selected 10 mg as the dose to be used in the Phase 3 trials in the apixaban AFib program. FDA concurred with this decision in our minutes of the EOP2 meeting held on October 2, 2006.

Table 57 DVT Prevention Study CV185010 – Safety Results

Summary of Incidence of Adjudicated Bleeding Events during Evaluation Period, by Treatment Group - Treated Subjects

	BMS QD 5mg (N= 151)	BMS QD 10mg (N= 155)	BMS QD 20mg (N= 151)	BMS BID 2.5mg (N= 154)	BMS BID 5mg (N= 153)	BMS BID 10mg (N= 153)	Enoxaparin (N = 149)	Warfarin (N = 151)
Major Bleeding, n	4	1	5	0	4	4	0	0
Event rate (%)	2.6	0.6	3.3	0.0	2.6	2.6	0.0	0.0
95% CI	(0.7, 6.6)	(0.0, 3.5)	(1.1, 7.6)	(0.0, 2.4)	(0.7, 6.6)	(0.7, 6.6)	(0.0, 2.4)	(0.0, 2.4)
Minor Bleeding, n	1	9	10	6	6	11	6	8
Event rate (%)	0.7	5.8	6.6	3.9	3.9	7.2	4.0	5.3
95% CI	(0.0, 3.6)	(2.7, 10.7)	(3.2, 11.8)	(1.4, 8.3)	(1.5, 8.3)	(3.6, 12.5)	(1.5, 8.6)	(2.3, 10.2)
Potentially Significant Bleed (PSB), n	0	1	0	0	0	0	2	0
Event rate (%)	0	0.6	0	0	0	0	1.3	0
95% CI	(0.0, 2.4)	(0.0, 3.5)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)	(0.2, 4.8)	(0.0, 2.4)
Thrombocytopenia	0	1	0	0	0	0	0	0
Event rate (%)	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0
95% CI	(0.0, 2.4)	(0.0, 3.5)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)	(0.0, 2.4)
Major Bleeding / PSB	4	2	5	0	4	4	2	0
Event rate (%)	2.6	1.3	3.3	0.0	2.6	2.6	1.3	0.0
95% CI	(0.7, 6.6)	(0.2, 4.6)	(1.1, 7.6)	(0.0, 2.4)	(0.7, 6.6)	(0.7, 6.6)	(0.2, 4.8)	(0.0, 2.4)
Major Bleeding / PSB / Minor Bleed	5	11	15	6	10	15	8	8
Event rate (%)	3.3	7.1	9.9	3.9	6.5	9.8	5.4	5.3
95% CI	(1.1, 7.6)	(3.6, 12.3)	(5.7, 15.9)	(1.4, 8.3)	(3.2, 11.7)	(5.6, 15.7)	(2.3, 10.3)	(2.3, 10.2)

In the NDA, the Applicant also submitted summary results of Study 185017 as support of the proposed dosing regimen for apixaban. This was a partially blinded RCT with 130 patients/arm with established DVT. They were randomized to 3 weeks of treatment with:

- Apixaban 5 mg bid,
- Apixaban 10 mg bid;
- Apixaban 20 mg qd,
- or open-label standard of care, which could be LMWH (enoxaparin or tinzaparin) + VKA or fondaparinux +h VKA (but all patients in this arm received LMWH + VKA)

The primary endpoint was change at week 12 in composite of adjudicated symptomatic recurrent VTE (recurrent DVT, PE (fatal or not) or increase of thrombotic burden by bilateral venous compression ultrasound and perfusion lung scan. The primary safety endpoint was adjudicated major bleeding + clinically relevant non-major bleeding.

Efficacy and safety results of Study 185017 are displayed in **Table 58** and **Table 59,** respectively. The efficacy results indicate the event rate (6%) in the apixaban 5 mg bid arm was roughly double the rate in the apixaban 20 mg qd arm, and also higher than

the rate in the SOC arm. However, there was one fatal PE in the 20 mg QD arm and a non-fatal PE in the SOC arm; the 5 mg bid arm had no PEs.

The safety results fail to show a dose response for bleeding across the apixaban arms. Bleeding with apixaban was slightly less than with SOC.

While the results of this study fail to make a strong case that the 5 mg bid dose appeared to be best among the apixaban doses tested, it does not clearly support the choice of another dose.

Table 58 DVT Treatment Study CV185017 – Efficacy Results

Table 2: Summary of Symptomatic Recurrent VTE/Deterioration during the Treatment Period – Primary Subjects

	APIX BID 5mg (N=117)	APIX BID 10mg (N=125)	APIX QD 20mg (N=116)	Any APIX N=358)	LMWH/Fond and VKA (N=118)
Symptomatic Recurrent VTE / Deterioration, n	7	7	3	17	5
Event rate (%)	6.0	5.6	2.6	4.7	4.2
95% CI	(2.4, 11.9)	(2.3, 11.2)	(0.5, 7.4)	(2.8, 7.5)	(1.4, 9.6)
Individual Components *					
Fatal PE, n	0	0	1	1	0
Non-Fatal PE, n	0	0	0	0	1
Symptomatic Recurrent DVT, n	3	4	1	8	2
Deterioration, n	4	3	1	8	2
Comparisons to Control Group					
Difference (%) (APIX - LMWH)	1.7	1.4	-1.7		
95% CI	(-4.4, 8.2)	(-4.6, 7.5)	(-7.3, 3.6)		
No Change, n **	21	16	22	59	18
Event rate (%)	17.9	12.8	19.0	16.5	15.3
95% CI	(11.5, 26.1)	(7.5, 20.0)	(12.3, 27.3)	(12.8, 20.7)	(9.3, 23.0)
Improvement, n ***	89	102	91	282	95
Event rate (%)	76.1	81.6	78.4	78.8	80.5
95% CI	(67.3, 83.5)	(73.7, 88.0)	(69.9, 85.5)	(74.2, 82.9)	(72.2, 87.2)

*. Intent-to-treat analysis. If a subject has multiple events, only the most severe one will be counted. Individual components are presented in decreasing order of severity.
**. If a subject does not have a recurrent VTE, and the results from his/her ultrasound tests are normal, and perfusion lung scan results are normal or no relevant change, then the subject is categorized as 'no change' on the primary endpoint
***. If a subject does not have a recurrent VTE, and the results from his/her perfusion lung scan and ultrasound tests have at least one improvement and no deterioration, then the subject is categorized as 'Improvement' on the primary endpoint.

Table 59 DVT Treatment Study CV185017 – Safety Results

Table 3: Summary of Bleeding Events Confirmed by Adjudication during the Treatment Period – All Treated Subjects

	APIX BID 5mg (N= 128)	APIX BID 10mg (N= 133)	APIX QD 20mg (N= 124)	Any APIX (N= 385)	LMWH/Fond and VKA (N= 126)
Major Bleeding or Clinically Relevant Non-major Bleeding, n	11	6	9	26	10
Event Rate (%)	8.6	4.5	7.3	6.8	7.9
95% CI	(4.4, 14.9)	(1.7, 9.6)	(3.4, 13.3)	(4.5, 9.7)	(3.9, 14.1)
Major Bleeding*, n	1	0	1	2	0
Event Rate (%)	0.8	0.0	0.8	0.5	0.0
95% CI	(0.0, 4.3)	(0.0, 2.7)	(0.0, 4.4)	(0.1, 1.9)	(0.0, 2.9)
Clinically Relevant Non-major Bleeding*, n	10	6	8	24	10
Trivial Bleeding*, n	3	11	4	18	10
All Bleeding**, n	14	17	13	44	20
Event Rate (%)	10.9	12.8	10.5	11.4	15.9
95% CI	(6.1, 17.7)	(7.6, 19.7)	(5.7, 17.3)	(8.4, 15.0)	(10.0, 23.4)

*. If a subject has multiple events, only the most severe one will be counted.
**. All bleeding is a composite of major bleeding, clinically relevant non-major bleeding, and trivial bleeding.

Reviewer Comment: The 010 study seems easier to interpret than the 017 study. The former supports the proposed apixaban primary dose (5 mg bid). The division agreed to this choice, which seems rationale.

As noted in Sec. **6.1.7.1.1**, choice of the 2.5 mg apixaban dose for patients with at least 2 of the 3 pre-specified risk factors for bleeding seems rational as well from an efficacy standpoint. This dose also performed well in AVERROES. Safety information for this dose is discussed in Sec. **7**

Dose in renal failure patients and other subgroups: Because of the limited renal excretion of apixaban OCP is not recommending dosing modifications based on renal function. OCP recommends a 50% reduction in dose when apixaban is used together with strong CYP3A4+P-gp inhibitors, and avoidance of concomitant use with strong 3A4 and P-gp inducers.

Japanese Phase 2B study in patients with non-valvular AFib: A small Phase 2 study was performed in Japan in 2008-2009 to assess the effects of 12 weeks of

therapy with 2 doses of apixaban (2.5 mg BID and 5.0 mg BID) versus warfarin on the composite endpoint of major and clinically relevant non-major bleeding events.

This study, CV18506 was a partially blinded, multicenter trial in adults with AFib and one additional risk factor for stoke. Warfarin therapy was open label and titrated to an INR of 2.0-3.0. Apixaban 2.5 mg bid or 5 mg bid was blinded using a double dummy technique. Patients received treatment for 12 weeks. The primary efficacy endpoint was the rate of the composite of stroke and systemic embolism and the primary safety endpoint was the rate of the composite of ISTH major bleeds and clinically relevant non-major bleeding events. Efficacy and safety endpoints were assessed during study treatment.

Notably, this study was started after ARISTOTLE and AVERROES had begun and was not relied upon for choice of dose in the Phase 3 program.

A total of 22 subjects were randomized into each arm, and all but 4 were treated. Patients in each arm were similar at baseline (data not shown). Results for efficacy and safety are shown in **Table 60.**

Table 60 Safety and Efficacy Summary of Study CV185067

	Warfarin N (%)	Apixaban 2.5 mg bid N (%)	Apixaban 5 mg bid N (%)
EFFICACY			
N Randomized	**74 (100)**	**74 (100)**	**74 (100)**
Stroke/SE	3 (4.1)	0	0
All-cause death	0	0	0
MI	0	0	0
SAFETY			
N Treated [2]	**75 (100)**	**72 (100)**	**71 (100)**
Major/CRNM bleeds	4 (5.3)	1 (1.4)	1 (1.4)
Major bleeds	1 (1.3)	0	0
CRNM bleeds	3 (4.0)	1 (1.4)	1 (1.4)
All bleeds	13 (17.3)	9 (12.5)	17 (23.9)
Discontinued for AE	4 (5.3)	4 (5.6)	4 (5.6)

[1] 2 ischemic strokes, 1 hemorrhagic stroke.
[2] One apixaban arm patient received warfarin erroneously through the entire treatment period and is included in the warfarin arm for safety analyses.

In this small and short-duration study, there were no strokes, SEs, deaths, or Mis in either the apixaban 2.5 mg bid or 5 mg bid arm. Each apixaban arm had one CRNM bleed. The higher dose had a higher rate of overall bleeding. Although data suggest that the lower dose may have preferable from the standpoint of bleeding risk, the study

was much too small to provide a reliable comparison of the effects of the 2 apixaban doses on efficacy.

Patients in both ARISTOTLE and AVERROES who met 2 of 3 specified criteria for increased risk of bleeding (age ≥ 80 years, body weight ≤ 60 kg, or serum creatinine ≥ 1.5 mg/dL) were eligible for treatment with a reduced dose of apixaban, 2.5 mg bid, if randomized to the apixaban arm. The rationale for this dose was not based on explicit exposure matching criteria. However, in the AVERROES study report (sec. 3.4.4) the Applicant indicates that results of the influence of intrinsic factors such as age, weight, and renal function indicated that patients with combinations of these factors could result in greater exposure. Thus, for patients at higher risk of bleeding ("e.g., an elderly subject with small stature or renal impairment"), a lower dose of apixaban was selected.

As noted in the primary endpoint analysis in patients receiving the lower dose (**Table 54**) and the analysis of bleeding data in this subgroup (**Table 94**), the ARISTOTLE data support the use of this reduced dose. This issue is also explored in the Clinical Pharmacology Review in Sec. 2.2.6.

6.1.9 Discussion of Persistence of Efficacy and/or Tolerance Effects

The Kaplan Meier curve for time to the primary efficacy endpoint in ARISTOTLE suggests that efficacy is maintained with continued treatment for over 3 years (**Figure 10**). While separation of the survival curves is maintained, the curves approach each other from 30 to 36 months of treatment but don't meet. However, there was an excess of events in the apixaban arm when study treatment was discontinued (see Section **6.1.10.2)**.

In addition, the Applicant provided ARISTOTLE primary endpoint event rate analyses covering 3 specified periods following the date of randomization. The data are summarized in the table below, and suggest that the HR for stroke/SE was reasonably similar for the periods up to and after 180 days following randomization (**Table 61**). However, the HR in the first 30 days after randomization was quite favorable for apixaban, as expected.

Table 61 ARISTOTLE - Primary Endpoint Results by Days from Randomization
ITT Population

	Apixaban N=9120		Warfarin N=9081		HR (95% CI)
Event Window (days after randomization)	N (%)	Events / 100 pt-yr	N (%)	Events / 100 pt-yr [1]	
0 – 30	12 / 9120	1.62	23 / 9081	3.11	0.52 (0.26, 1.04)
0 – 180	70 /9120	1.60	93 / 9081	2.14	0.75 (0.55, 1.02)
181 – End of ITP	142 / 8733	1.15	172 / 8627	1.41	0.82 (0.65, 1.02)

6.1.10 Additional Efficacy Issues/Analyses

6.1.10.1 Adequacy of Comparator

ARISTOTLE provides reasonably robust evidence of the superiority of apixaban to warfarin in preventing primary endpoint events. Because warfarin therapy is very unlikely to be worse than placebo, adequacy of warfarin therapy and in particular satisfaction of the "constancy assumption" j is not critical. However, it may be useful to assess how the design and patient population of ARISTOLE compare to those of the placebo controlled studies supporting the efficacy of warfarin, as well as the warfarin-controlled trials of other novel anticoagulants evaluated in patients with AFib.

Table 62 provides information on the demographics, control of INR in the warfarin arm, and results of the six published placebo controlled trials of warfarin therapy in atrial fibrillation patients at risk for stroke and systemic embolism. These studies were conducted in the 1990s. In five of these studies, most (> 90%) of patients did not have a prior history of stroke. In the sixth, the EAFT study, 100% of patients had a prior history of recent stroke or TIA; 76% of these had had a stroke. No $CHADS_2$ score data are available for the historical studies, but EAFT must have had patients at higher risk of stroke than the other studies.

Table 62 indicates that the historical studies utilized a broad range of INR targets. In the US studies, INR had not yet been adopted widely, and the INR target (and its attainment) was back-calculated from the PT target and the assumed ISI of the thromboplastin used in the PT assay. The INR target range of ARISTOTLE, 2.0 to 3.0, falls within the range of INR targets for the placebo-controlled trials. Similarly, the mean

j Satisfaction of the constancy assumption is generally critical to the validity a non-inferiority trial. This involves an assessment of the likelihood that the effect of the active control in the NI study of interest is similar to its effect in the trials that established the efficacy of the active control. It requires an assessment of dose, patient characteristics, concomitant therapies, definition and ascertainment of endpoints, and other study design and execution features.

time in therapeutic range, 62%, falls within in the range of mean TTR or % of INRs in range in the placebo controlled trials. Thus, it seems that constancy holds for the issue of control of anticoagulation as an isolated question.

Table 62 ARISTOTLE vs. Six Placebo-Controlled Warfarin Trials
Selected Parameters

	ARISTOTLE (apixaban vs. W) [1]	**5 Primary Prevention Studies** (W vs. placebo)	**EAFT** (W vs. placebo)
N (ITT)	18,201	2461	439
% female	36	0-47	43
% with h/o stroke/TIA/SE	19	**6**	**100**
Mean $CHADS_2$ Score	2.1	-	-
Target INR (range)	(2.0-3.0)	(1.4-2.8 to 2.0-4.5)	(2.5-4.0)
Mean TTR or % in range*	**62**	**42-83**	**59***
Endpoint	Stroke + SE	Ischemic stroke to Stroke + TIA + SE	Stroke
Event Rate Warfarin	1.60	0.62 - 3.08	4
Event Rate Apixaban or Placebo	1.27	2.99 – 8.20	12
HR (95% CI)	0.79 (0.66, 0.95))	0.21 – 0.65	0.34 (0.20, 0.57)
FDA meta-analysis of 6 placebo-controlled studies (random effects model) --		HR for W vs. Placebo = 0.36 (0.24, 0.53)	

Table 63 indicates that the design and patient population of ARISTOTLE do not deviate notably from the other warfarin-controlled trials of novel anticoagulants in patients with AFib. Control of INR falls between that achieved in ROCKET and RE-LY, but is closer to the latter than the former.

Table 63 ARISTOTLE vs. Warfarin-controlled Trials of other OAC
Selected Parameters

	ARISTOTLE (apixaban vs W)	**ROCKET** (rivaroxaban vs W)	**RE-LY** (dabigatran 150 mg vs W)	**SPORTIF III** (ximelaga-tran vs W)	**SPORTIF V** (ximelaga-tran vs W)
N (ITT)	18,201	14,171	12,098	3397	3922
Blinding	Double dummy (DD)	DD	Open-label	Open-label	DD
% female	36	40	37	30	31
% with h/o stroke/TIA//SE	19	55	22	29	22
Mean $CHADS_2$ Score	2.1	3.5	2.1	-	-
% w prior VKA therapy	56	62	61	73	85
Target INR (range)	(2.0 – 3.0)	2.5 (2.0-3.0)	(2.0-3.0)	(2.0 – 3.0)	(2.0-3.0)
Mean TTR (%)	62	56	64	66	68
Primary endpoint	Stroke/SE	Stroke/SE	Stroke/SE	Stroke/SE	Stroke/SE
Event rate warfarin	1.60	2.4	1.71	2.29	1.16
Event rate test agent	1.27	2.1	1.11	1.64	1.61
HR or Δ (95% CI)	0.79 (0.66, 0.95)	0.88 (0.74, 1.03)	0.65 (0.52, 0.81)	-0.66%/yr (-1.45, 0.13)	0.45%/yr (-0.13, 1.03)

Rates are expressed in events per 100 patient-years.

6.1.10.1.1 *Importance of the quality of warfarin management*

The consensus guidelines for the management of patients with AFib recognize that the efficacy of warfarin in preventing thrombotic events is dependent on the quality of INR control,[3] a conclusion also reached in other publications describing the inverse relationship between center TTR and event rate in warfarin-treated patients in studies of stroke prevention in AFib patients.[6, 7] The recommended target range in the guidelines for patients with non-valvular AFib in need of anticoagulation with a VKA is 2.0 to 3.0.

Quality of INR control is usually assessed by calculating imputed percentage time in therapeutic range (TTR)[10], is a critical factor in interpreting a trial in which warfarin is used as an active control, and data for this parameter has been often in reports of clinical trials of warfarin performed in recent years.

6.1.10.1.2 *INR in ARISTOTLE*

In ARISTOTLE, the target therapeutic range of INR was 2 – 3, consistent with the recommendation in the consensus guidelines. Data on overall TTR in the warfarin arm of ARISTOTLE is displayed in **Table 64**.

Table 64 ARISTOTLE – Percent Time In INR Range In Warfarin Arm

INR Range	Mean	Median
INR < 2.0	24.73	19.01
2.0 <= INR <= 3.0	62.19	66.02
INR > 3.0	13.06	10.83
INR < 0.8	0.01	0.00
0.8 <= INR < 1.8	15.48	9.03
1.8 <= INR < 2.0	9.24	8.18
2.0 <= INR <= 3.0	62.19	66.02
3.0 < INR <= 3.2	4.58	3.86
3.2 < INR <= 5.0	8.09	6.09
5.0 < INR <= 12.0	0.38	0.00
INR > 12.0	0.00	0.00

INR was calculated using method of Rosendaal, excluding the first week of treatment and treatment interruptions.

INR was in the target range of 2 to 3 about 62% of days in ARISTOTLE. Of the 38% of time spent outside the therapeutic range, about 25% was spent below range (meaning that there was an increased risk of ischemic stroke over the risk when in range) and the remainder, about 13%, was spend above the therapeutic range (meaning that there was an increased risk of hemorrhagic stroke over the risk when in range).

6.1.10.1.3 Analyses of site-based TTR in ARISTOTLE

In other applications, we have examined the results for the primary endpoint in various subsets of the study based on site level TTR. Use of site-level data preserves the effects of randomization and is less prone to bias than simply comparing all patients in the apixaban arm to those with various levels of TTR in the warfarin arm. The latter type of comparison could be greatly confounded by the effects of nationality, region, demography, and general quality of care, which could differ greatly in patients with poor vs. good warfarin control.

The Applicant performed an analysis of the primary endpoint results in quartiles of site-based TTR, as well as above and below the median TTR. The Applicant used the method of Connolly, et. al. to calculate mean site TTR, as is customary.[11] This method calculates site mean TTR as the mean of all the individual TTR values of patients at the site, without weighting for duration of treatment. However, rather than basing the inter-quartile points on site TTR as is customary, the Applicant used data based on individual TTR, which produced a larger inter-quartile range. The Applicant's data are displayed in **Table** 65. Note that the first and fourth quartiles contained relatively few patients. FDA's analysis, performed using the same method as the Applicant to calculate site TTR, but using inter-quartile points based on site TTR (with equal numbers of sites in each quartile), is shown in **Table 66**. Note that number of patients in the various quartiles varies less in FDA's analysis, and the inter-quartile range is smaller.

Table 65 ARISTOTLE – Applicant's Analysis of Primary Endpoint By Site TTR
ITT Population, during ITP

Site TTR (%)	Apixaban		Warfarin		A vs W - Hazard Ratio (95% CI)
	N=9120 n / J	Events / 100 pt-yr	N=9081 n / J	Events / 100 pt-yr	
> 52.35	26 / 1196	1.23	52 / 1178	2.58	0.48 (0.30, 0.77)
52.35 - < 65.99	99 / 3453	1.58	107 / 3474	1.71	0.93 (0.71, 1.22)
65.99 - < 76.50	70 / 3357	1.11	85 / 3407	1.34	0.83 (0.60, 1.13)
≥ 76.50	16 / 1064	0.80	21 / 1015	1.08	0.73 (0.38, 1.40)
< 65.99	125 / 4649	1.49	159 / 4652	1.92	0.78 (0.62, 0.98)
≥ 65.99	86 / 4421	1.04	106 / 4422	1.28	0.81 (0.61, 1.08)

1 J = number of patients in subgroup

Table 66 ARISTOTLE – FDA's Analysis Of Primary Endpoint By Site TTR
ITT Population, during ITP

Site TTR (%)	Apixaban		Warfarin		A vs W - Hazard Ratio (95% CI)
	N=9120 n / J	Events / 100 pt-yr	N=9081 n / J	Events / 100 pt-yr	
≤ 55.3	69 / 2210	1.77	88 / 2189	2.34	0.76 (0.55, 1.04)
>55.3 - ≤ 64.6	72 / 2829	1.40	86 / 2854	1.65	0.85 (0.62, 1.17)
>64.6 - ≤ 72.7	41 / 2398	0.90	55 / 2423	1.20	0.75 (0.50, 1.12)
> 72.7	29 / 1633	0.95	36 / 1608	1.19	0.79 (0.49, 1.29)
≤ 64.6	141 / 5039	1.56	174 / 5043	1.94	0.80 (0.64, 1.00)
> 64.6	70 / 4031	0.92	91 / 4031	1.20	0.77 (0.56, 1.05)

1 J = number of patients in subgroup

In the Applicant's analysis, the event rates in both treatment arms tend to decrease as site TTR increases. However, the hazard ratios do not show a coherent pattern. The lowest HR is in the first (worst quartile of TTR), while the highest HR is in the second

quartile, the spread in values from lowest to highest is 0.45. FDA's analysis shows a similar pattern as the Applicant's analysis with respect to the relationship between event rates and TTR, except that the results in the 3rd and 4th quartiles are not notably different from each other. However, the HR values among the 4 quartiles in FDA's analysis vary much less than in the Applicant's analysis– no quartile has a HR markedly different from the overall HR of 0.79, and the spread from lowest to highest HR is 0.10.

Reviewer Comment: FDA's analysis suggests that the HR for the primary endpoint favors apixaban in a reasonably consistent across a broad range of warfarin control, including quite good control in the 4th quartile (site TTR > 72.7%).

The Applicant also provided primary endpoint results in quartiles of site based time below therapeutic range (TBTR) in the warfarin arm. [k] TTR does not distinguish between out of range high values of INR and out of range low values of INR. The former are associated with a rate of hemorrhagic stroke that increases slowly as INR rises above 3, with little change in the rate of ischemic stroke. The latter are associated with a steeply increasing rate of ischemic stroke (which is more common than hemorrhagic stroke in the therapeutic range) with little change in the rate of hemorrhagic stroke. Thus, a metric that focuses on out of range low values of INR might result in a larger spread of warfarin arm event rates that one that treats high and low values, and possibly a larger spread in hazard ratios among the quartiles. .

FDA also analyzed the primary endpoint results by site TBTR. Again, the quartiles varied less in size in the FDA analysis. The FDA analysis is shown below in **Table 67**. Note that the best warfarin control is represented in the first quartile and the worst control is represented in the fourth quartile. Event rates both arms tend to increase as TBTR increases; this relationship is stronger in the warfarin arm than in the apixaban arm. However, there is no clear relationship between TBTR and the HR.

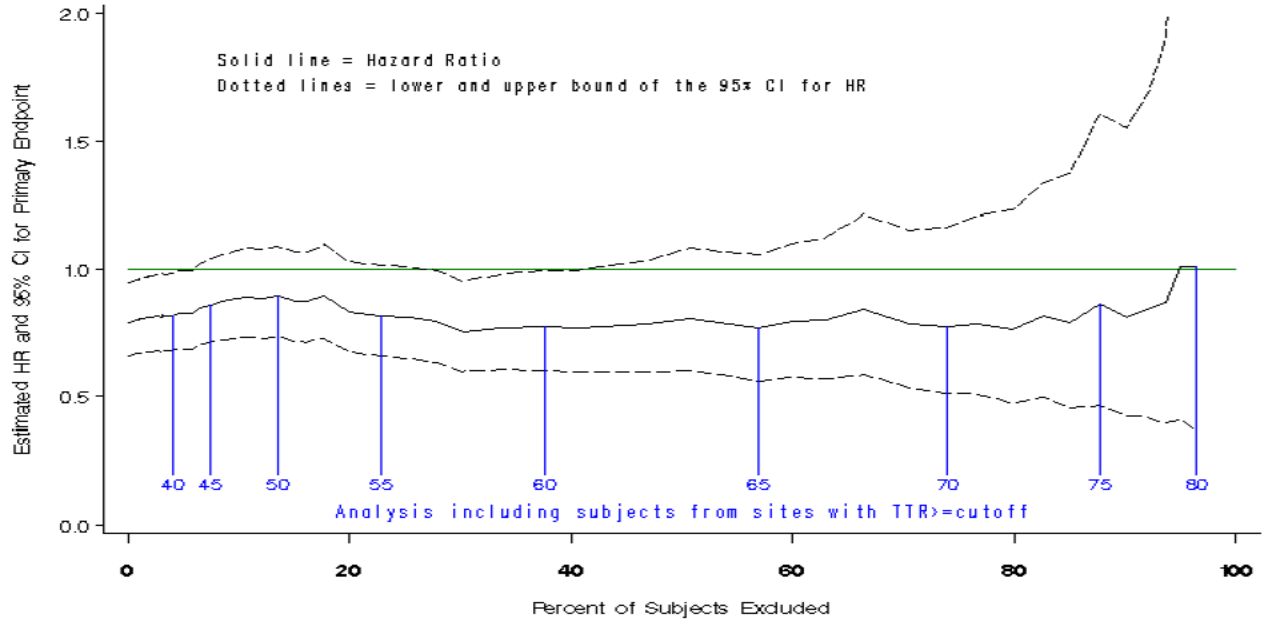
k We hypothesized that quartiles or other subsets based on this parameter might better distinguish centers in terms of primary event rates than a conventional TTR analysis. The underlying rationale is based on the fact that most primary endpoint events are ischemic strokes. The risk of ischemic stroke increases sharply as INR falls below 2. On the other hand, ischemic stroke risk is little affected by INR > 3 compared to INR in the therapeutic range of 2 – 3. The risk of hemorrhagic stroke does increase as INR increases over 3, but the rate of increase is modest, and such strokes are decidedly less common than ischemic strokes in studies in atrial fibrillation patients. Accordingly, while INRs above the therapeutic range count against TTR as it is usually measured, they have only modest effects on primary endpoint rates. This would tend to blunt the power of a primary endpoint analysis that takes into account such INRs to distinguish between subsets based on INR control. Accordingly an analysis that considers only time below the therapeutic range might better distinguish among subgroups of centers with different levels of INR control. However, INRs above therapeutic range would be relevant in an analysis of bleeding risk, and the conventional TTR analysis (or an analysis that considers only time above range) would be expected to be useful in assessing the affects of differences in INR control on bleeding events.

Table 67 ARISTOTLE - FDA's Analysis Of Primary Endpoint Results by Site TBTR
ITT Population, during ITP

Site TB TR (%)	Apixaban		Warfarin		A vs W - Hazard Ratio (95% CI)
	N=9120 n / J	Events / 100 pt-yr	N=9081 n / J	Events / 100 pt-yr	
≤13.3	28 / 1682	0.89	35 / 1660	1.13	0.79 (0.48, 1.30)
>13.3 - ≤ 20.5	34 / 2225	0.82	50 / 2233	1.21	0.67 (0.44, 1.04)
>20.5 - ≤29.9	81 / 2796	1.58	90 / 2818	1.73	0.91 (0.68, 1.23)
>29.9	68 / 2367	1.62	90 / 2363	2.17	0.74 (0.54, 1.01

To further examine the relationship between control of INR and efficacy, we asked the Applicant to create a graphical analysis of the primary endpoint (ITT, ITP) in which the x axis is site TTR ranging from 0% to 100% and the y axis is the HR for apixaban vs. warfarin (see **Figure 12**). The figure is a plot of Y=f(x) where f(x) is the point estimate for the HR for primary endpoint for apixaban vs. warfarin at all centers where TTR was in the interval from x to 100%. Thus, for x=0%, the HR corresponded to the HR for the entire study, and for x=K%, the HR was the HR for the centers with TTR ranging from K% to 100%. As K increases, the number of patients in the analysis decreases, and the CI becomes wider. We also plotted the 5th and 95th percentile for the HR. Note that the HR point estimate curve (the center curve) is fairly flat from X=0% to about X>~75% and goes up slightly. It approaches1.0 only when top several percent of patients at the sites with highest TTR are included in the analysis.

Figure 12 Hazard Ratio and 95% CI for the Primary Efficacy Endpoint Analysis as a Function of Center TTR
ITT Population, ITP



Note: INR measurements on or after Day 8 of randomization until the earliest of (last dose of warfarin, last INR measurement), after excluding INR measurements during warfarin interruptions, are included in TTR computation.
Program source: c:\pgm\ApixabanFDATTR01A-v01.sas Run Date: 03JAN2012:09:20:17

Table 68 and **Table 69**, respectively, are the Applicant's and FDA's analyses of the results for all-cause death by site TTR. As before, the variability of the number of patients in each site TTR quartile is lower in the FDA's analysis, and the spread of hazard ratios is not as large. In the Applicant's analysis, the hazard ratios range from 0.72 in the first (worst) quartile of site TTR to 1.23 in the fourth (best) quartile of site TTR. In the FDA analysis, the analogous hazard ratios are 0.79 and 1.00, respectively. Note that the HR for the overall analysis is 0.89.

Reviewer Comment: Unlike the case for the analysis of the primary endpoint by TTR, both the analyses of the Applicant and FDA here suggest a relationship between quartiles of site TTR and the apixaban vs warfarin HR for all-cause death: higher (better) TTR is associated with a higher (more favorable to warfarin) HR. An inverse relationship of TTR to event rates was suggested in both treatment arms; this relationship was more prominent in the warfarin arm.

The above observation regarding possible differences in the relationship between TTR and the HR for stroke vs. death is consistent with a recent publication of a retrospective look at outcomes in a Swedish database of > 19,000 patients with atrial fibrillation managed with warfarin.[12] For individuals, the association

between TTR and various outcomes (i.e., the decrease in risk of an event associated with a 1 SD increase in TTR) was greater for all-cause mortality than for stroke, bleeding, or hospital admission.

Table 68 ARISTOTLE – Applicant's Analysis of All-cause Death by Site TTR
ITT Population, during ITP

Site TTR (%)	Apixaban		Warfarin		A vs W - Hazard Ratio (95% CI)
	N=9120 n / J	Events / 100 pt-yr	N=9081 n / J	Events / 100 pt-yr	
< 52.35	80 / 1196	3.66	107 / 1178	5.05	0.72 (0.54, 0.96)
52.35 -- <65.99	245 / 3453	3.82	267 / 3474	4.14	0.92 (0.78, 1.10)
65.99 -- <76.50	214 / 3357	3.34	245 / 3407	3.80	0.87 (0.73, 1.05)
≥ 76.50	62 / 1064	3.07	50 / 1015	2.54	1.23 (0.84, 1.78)
< 65.99	325 / 4649	3.78	374 / 4652	4.37	0.86 (0.74, 1.00)
≥ 65.99	276 / 4421	3.27	295 / 4422	3.50	0.93 (0.79, 1.10)

Table 69 ARISTOTLE – FDA's Analysis of All-cause Death by Site TTR
ITT Population, during ITP

Site TTR (%)	Apixaban		Warfarin		A vs W - Hazard Ratio (95% CI)
	N=9120 n / J	Events / 100 pt-yr	N=9081 n / J	Events / 100 pt-yr	
≤ 55.3	156 / 2210	3.87	193 / 2189	4.91	0.79 (0.64, 0.97)
>55.3 -- ≤64.6	215 / 2829	4.09	235 / 2854	4.41	0.93 (0.77, 1.11)
>64.6 -- ≤72.7	142 / 2398	3.06	155 / 2423	3.32	0.92 (0.73, 1.16)
> 72.7	88 / 1633	2.83	86 / 1608	2.81	1.00 (0.74, 1.34)
≤ 64.6	371 / 5039	3.99	428 / 5043	4.62	0.86 (0.75, 0.99)
> 64.6	230 / 4031	2.37	241 / 4031	3.12	0.95 (0.79, 1.14)

6.1.10.1.4 Summary of data

Reviewer's conclusion regarding the adequacy of comparator:

- *The quality of INR control in ARISTOTLE was in the range achieved in other studies of novel anticoagulants with warfarin controls.*
- *There was no apparent relationship between quartile in site INR control and the results for the primary endpoint HR.*
- *There was a relationship between quartile of site INR control and the results for all-cause death. Most of the advantage of apixaban was apparent at sites with TTR below the median.*
- *The constancy assumption was met.*

6.1.10.2 Endpoint Events Occurring After Discontinuation of Study Drug

6.1.10.2.1 ARISTOTLE

In ARISTOTLE, the cutoff for the primary efficacy endpoint analysis occurred at a fixed time based on attainment of the event target. At this time, about 75% of randomized patients were still taking study drug, which was then discontinued at the final treatment visit, with a subsequent follow-up visit 30 days later. Post-discontinuation events in these completing patients were not included in the primary endpoint analysis or the secondary endpoint of all-cause death. However, post-discontinuation events occurring up to the cutoff date were included for the roughly 25% of patients who discontinued study drug early.

This section will examine the rate of post-discontinuation events (both primary endpoint events and all-cause deaths, considered separately) occurring in the first 30 days after discontinuation of study drug. The two populations, those who discontinued study drug early and those who reached the data cutoff while on study drug, will be examined separately, as they tend to be quite different in terms of their health status and differ substantially in their rates of the events of interest.

Table 70 is a display of the rate of primary endpoint events (composite of stroke and systemic embolism) in early discontinuation patients occurring from 1-30 days after the last dose of study drug.

Table 70 Early Discontinuation Patients – Primary Efficacy Endpoint Events after Discontinuation of Study Drug
Treated Population

	Apixaban		Warfarin		A vs. W HR (95% CI)
Event Window – Days after Last Dose	n/N (%)	Event Rate / 100 pt- yr	n/N (%)	Event Rate / 100 pt- yr)	
1 – 30 days	**52 / 1841 (2.82)**	**39.86**	**67 / 2028 (3.30)**	**46.99**	**0.86 (0.60, 1.24)**
1 – 2 days	29 / 1839 (1.58)	299.47	39 / 2026 (1.92)	366.00	-
3 – 7 days	4 / 1700 (0.24)	17.49	11 / 1868 (0.59)	43.87	-
8 – 14 days	8 / 1636 (0.49)	25.82	8 / 1798 (0.44)	23.48	-
15 – 30 days	11 / 1604 (0.69)	16.03	9 / 1759 (0.51)	11.94	-

Note that primary endpoint events were numerically more frequent in the warfarin arm during the first 30 days after the last dose of study drug in this population, although the difference was not statistically significant. Event rates were about 40 vs 47 per hundred patient-years in the apixaban and warfarin arms, respectively. These event rates are over 10 fold greater than the analogous rates during treatment. Over half these events occurred in the first two days after the last dose in each arm, with progressively decreasing event rates in the subsequent segments of the 30 day period. As noted previously, most if not all of these events were captured in the Applicant's primary efficacy endpoint analysis.

Note that death rates were similar in the two study arms and quite high in this population, with about 13.5% of patients dying in each arm over 30 days. Death rates were 188-189 per 100 patient-years over this period, roughly 50x the rate during treatment. Nearly half of deaths occurred in the first 2 days after the last dose of study drug, with death rates falling progressively in subsequent segments of the 30 day period. Again, most if not all of these deaths were captured in the Applicant's ITT analysis of death during the intended treatment period.

Reviewer Comment: In ARISTOTLE, as in the ROCKET study of rivaroxaban, rates of stroke/SE and death were over the 30 days after the last dose of study drug in patients who discontinued study, but rates in the two study arms did not differ significantly.

Table 71 shows rates of stroke and SE over the 30 days after the last dose of study drug in patients who completed the study. The overall (Day 1 to 30) event rate was significantly higher in the apixaban arm, with a hazard ratio of about 4.1. The 21 events in the apixaban arm were distributed fairly evenly throughout the 30 day period. There were 5 events in the warfarin arm. Note that while the stroke/SE rate in the rivaroxaban arm in the post-treatment period is considerably higher than the rate of stroke/SE during treatment (1.14 events/100 pt-yr), the rate in the warfarin arm less than the rate during treatment (1.60 events/100 pt-year).

Reviewer Comment: Unless otherwise specified, the "last dose of study drug" refers to the last dose of active study drug, which might have been active apixaban administered during the 4-dose transition regimen received by 60% of completers in the apixaban arm.

Table 71 ARISTOTLE – Primary Efficacy Endpoint Events Following Completion of ITP

	Apixaban N=6810		**Warfarin N=6588**		
Days after last dose	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr	HR (95% CI)
1-30	21 / 6791	4.02	5 / 6569	0.99	4.07 (1.54, 10.81)
1-2	1 / 6791	2.69	1 / 6569	2.78	-
3-7	4 / 6787	4.31	0 / 6566	0	-
8-14	5 / 6780	3.85	1 / 6559	0.80	-
15-30	11 / 6771	4.18	3 / 6548	1.18	-

Reviewer Comment: These results are reminiscent of the results of an analogous analysis in the ROCKET study of rivaroxaban, where the rate of

stroke/SE in completers was significantly higher in the rivaroxaban arm than in the warfarin arm , with a point estimate for the HR of about 4. The reviewer (MR) agreed with the rivaroxaban Applicant that the increased rate of events in the rivaroxaban arm after the last dose of study drug was probably related to inadequate control of anticoagulation, but that induction of a hypercoagulable state by long term treatment with rivaroxaban had not been ruled out. Note that the PI for rivaroxaban includes a boxed warning about the increased risk of stroke in patients who discontinue rivaroxaban and recommends that when rivaroxaban is must be discontinued for a reason other than pathological bleeding, the health care provider should "…consider administering another anticoagulant."

Table 72 shows the event types and timing of events in the study arms for the events represented in **Table 71**. In the apixaban arm, the 21 events included 13 ischemic strokes, 3 hemorrhagic strokes, 4 strokes of uncertain type, and 2 SE. Note that all 3 hemorrhagic strokes occurred in the 15 to 30 day segment (all in subjects given VKA in the 1-30 day period), and 3 of the 4 strokes of uncertain type occurred during this same segment. The 5 events in the warfarin arm included 3 ischemic strokes and 2 SE.

Table 72 Primary Endpoint Events following Completion of ITP – Event Types

Apixaban Arm					
	Event Type				
Days after last dose	Ischemic Stroke	Hemorr Stroke	Uncertain Stroke	SE	**All**
1-2	1				**1**
3-7	4				**4**
8-14	4		1		**5**
15-30	4	3	3	1	**11**
Total	**13**	**3**	**4**	**1**	**21**
Warfarin Arm					
	Event Type				
Days after last dose	Ischemic Stroke	Hemorr Stroke	Uncertain Stroke	SE	**All**
1 to 2	1				**1**
3 to 7	0				**0**
8 to 14	1				**1**
15 to 30	1			2	**3**
Total	**3**	**0**	**0**	**2**	**5**
Hemorr = hemorrhagic; Uncertain Stroke = Stroke of uncertain type- SE = systemic embolism					

Table 73 shows primary endpoint event rates over the 30 day period after the last dose in completing patients by use of VKA in the 30 day period. About 84% to 85 of completers received VKA at some time during the 30 day post-treatment period in the apixaban and warfarin arms, respectively. Note that there is almost no information on when VKA were started, what doses were administered, and INR data is almost completely lacking. For both types VKA status (i.e., those who received VKA at some point during the 30 day post-dosing period and those who did not), the primary endpoint event rate was higher in the apixaban arm than in the warfarin arm , and event rates were lower in either arm in patients who did receive VKA compared to those who did not. The apixaban vs. warfarin hazard ratio was somewhat higher in the patients who received VKA.

Table 73 ARISTOTLE – Primary Efficacy Endpoint Events Following Completion of ITP by VKA Status

Subjects who Received VKA Following Treatment Period					
	Apixaban N=5734		**Warfarin N=5585**		
Days after last dose	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr	HR (95% CI)
1-30	14 / 5723	3.19	2 / 5570	0.47	6.83 (1.55, 30.07)
1-2	0 / 5723	0	0 / 5570	0	-
3-7	2 / 5719	2.56	0 / 5568	0	-
8-14	2 / 5715	1.83	0 / 5565	0	-
15-30	10 / 5710	4.54	2 / 5559	0.93	-

Table 73 continued

Subjects who Did Not Receive VKA Following Treatment Period					
	Apixaban N=1075		Warfarin N=1003		
Days after last dose	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr	HR (95% CI)
1-30	7 / 1068	8.40	3 / 999	3.84	2.12 (0.55, 8.21)
1-2	1 / 1068	17.11	1 / 999	18.30	-
3-7	2 / 1068	13.71	0 / 998	0	-
8-14	3 / 1065	14.73	1 / 994	5.26	-
15-30	1 / 1061	2.34	1 / 989	2.49	-

Table 74 shows primary endpoint event rates in completing patients in each arm who did or did not receive the 4 dose transition regimen (TR) that was given to 60% of completing subjects in each arm, and also to a trivial percentage of patients who discontinued early (not shown). This regimen consisted of an additional 4 doses of the patients own apixaban study drug that was given back to the patient at the end of treatment visit. The drug was to be taken in same manner as the rest of the patient's apixaban study medication. Apixaban arm patients thus received active apixaban, while warfarin arm patients received placebo for apixaban

Table 74 ARISTOTLE – Primary Efficacy Endpoint Events Following Completion of ITP by use of Transition Regimen (TR)

	Apixaban N=6810		Warfarin N=6588	
Use of TR	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr
Yes	10 / 3461	3.62	2 / 3377	0.74
No	4 / 2273	2.22	1 / 2208	0.57

"Last dose of study drug" means the last dose for those who did not receive the TR and the last dose before the transition regimen for those who received the TR.

Note that in **Table 74** the "last dose of study drug" has a variable meaning. It refers to (1) the last dose of study drug for those who did not receive the TR and (2) the last dose of study before the start of the TR for those who received the TR. This was done to maximize the likelihood of showing a benefit for the transition regimen. However, there was not a suggestion of benefit. Event rates for apixaban arm subjects, as well as warfarin arm subjects, who received the TR are higher than for those who did not. Apixaban vs. warfarin hazard ratios are similar for those who did and did not receive the TR.

We analyzed the characteristics of the completing patients who had primary endpoint events between day 1 and 30 after the last dose of study drug (**Table 75**). Patients in the warfarin arm were slightly older (mean age 75 vs. 73 years) and had somewhat lower $CHADS_2$ scores (mean of 2.8 vs. 2.2). Note that the overall study $CHADS_2$ mean score was 2.1. Notably, patients in apixaban arm had higher rate of prior history of stroke/TIA: over 43% vs. 0; the overall rate for such a history was 19%.

Table 75 ARISTOTLE - Characteristics of Completing Patients With Primary Endpoint Events 1 To 30 Days after Last Dose of Study Drug

	Apixaban (N=21)	Warfarin (N=5)
Age (mean)	73.3	75.2
Baseline $CHADS_2$ score (mean)	2.8	2.2
score ≥ 3, N (%)	13 (62)	2 (40)
score ≥ 4, N (%)	7 (33)	0
History of stroke/TIA, N, (%)	9 (43)	0

Reviewer Comment: These data suggests that the completers in the apixaban arm who had events in this period simply may simply have been a high risk group with a low tolerance for inadequate anticoagulation. Apixaban arm patients may be been at greater risk than those in the warfarin arm because of the short half-life of apixaban's PD effect compared to warfarin, which could have increased the degree and duration of inadequate anticoagulation in the apixaban arm.

6.1.10.2.1.1 Comparisons of "VKA naïve" patients at the start and after the end of double blind treatment

ARISTOTLE study patients who entered the study VKA naïve and were randomized into the warfarin arm may be similar in terms of their risk of thrombotic events to apixaban arm patients who completed the study and then started warfarin treatment, since both subgroups of patients were started on warfarin during the study after extended period of time without warfarin treatment. Thus, a comparison of event rates in these subgroups of patients might be useful in putting into perspective the excess of strokes in apixaban arm completers in the day 1 to 30 period after the last dose of study drug. Similar event

rates in the two populations, one transitioning from essentially no anticoagulant therapy to warfarin, and the other transitioning from apixaban therapy to open label VKA therapy, might suggest that the elevated event rate observed in the latter group of patients might simply be due to poor INR control, as was observed in the VKA naïve patients at the start of the study.

We asked the Applicant to perform a time to event analysis for primary efficacy endpoint events in warfarin arm patients covering the first 30 days after randomization, looking at all patients and the subgroups of VKA naïve and experienced patients. Data from the VKA naïve patients, along with data for apixaban completers in the day 1 to 30 window after the last dose of study drug are displayed in **Table 76**.

Table 76 Comparison of Primary Endpoint Event Rates in "VKA Naïve" Patients

	Apixaban		Warfarin	
	n/N	Events / 100pt-yr	n/N	Events / 100 pt-yr
Completers treated with VKA after double-blind treatment	14 / 5723	3.19	2 / 5570	0.47
Completers not treated with VKA after double-blind treatment	7 / 1068	8.40	3 / 999	3.84
For contrast – events in first 30 days after randomization in VKA naïve patients	3 / 3912	0.94	17 / 3888	5.39

The primary efficacy endpoint event rate over the 30 days following randomization in warfarin naïve patients who then began blinded VKA therapy was 5.39 events per hundred patient years. At the end of the study, the event rate in the 84% of apixaban arm patients (all presumably warfarin naïve at the time),who then began open label VKA, was 3.19 per hundred patient years, less than the rate for the warfarin naïve patients at the start of the study. These data suggest that the event rate in completing patients who discontinue apixaban and start VKA treatment is no worse than the rate in warfarin naïve patients who begin warfarin treatment at the start of the study. Thus, they are reassuring with regard to our concern that cessation of long-term apixaban therapy might produce a hypercoagulable state or accentuate the pro-thrombotic risk of initiation warfarin therapy

6.1.10.2.1.2 *Events during interruptions of therapy in ARISTOTLE*

Patients with temporary interruptions of therapy might be another subset of study patients at greater risk of thrombotic events. Accordingly, we asked the Applicant to analyze the rate of primary endpoint events during interruptions of double blind treatment > 3 days duration.

Table 77 and **Table 78** provide information on the number of patients in the treatment arms with interruptions of treatment of various durations and the primary endpoint event rates associated with those interruptions, respectively. The two event windows for the analyses shown in **Table 78** are from 1 day after the last dose of study drug prior to the interruption to (1) the day that study drug was resumed and (2) 30 days after the day that study drug was resumed.

Table 77 ARISTOTLE - Interruptions Of Treatment > 3 Days In Duration
ITT Population

Duration of interruption	**Apixaban** (N = 9120) n (%)	**Warfarin** (N = 7082) n (%)
> 3 days	3008 (33.0)	3446 (37.9)
> 7 days	1947 (21.3)	2200 (24.2)
> 14 days	1393 (15.3)	1497 (16.5)
> 28 days	839 (9.2)	823 (9.1)

Table 78 Adjudicated Primary Endpoint Events in Patients with Treatment Interruptions > 3 Days
ITT Population

	Apixaban		Warfarin	
Duration of Interruption	**N= 9120 n/J (%)**	**Event Rate (100 Pt-yr)**	**N= 9081 n/J (%)**	**Event Rate (100 Pt-yr)**
(1) Events though the day that study drug was resumed				
> 3 days	14 / 3008 (0.47)	5.08	12 / 3446 (0.35)	4.12
> 7 days	13 / 1947 (0.67)	5.35	12 / 2200 (0.55)	4.80
> 14 days	10 /1393 (0.72)	4.63	8 / 1497 (0.53)	3.73
> 28 days	8 / 839 (0.95)	4.72	7 / 823 (0.85)	4.36
(2) Events though 30 days after study drug was resumed				
> 3 days	27 / 3008 (0.90)	4.27	25 / 3446 (0.73)	3.49
> 7 days	22 / 1947 (1.13)	4.91	19 / 2200 (0.86)	3.93
> 14 days	15 / 1393 (1.08)	4.25	11 / 1497 (0.73)	3.08
> 28 days	11 / 839 (1.31)	4.46	8 / 823 (0.97)	3.44

Event windows in both (1) and (2) start on the day after the last dose of study drug prior to the interruption.
J= patients in subgroup.

About 33% and 38% of patients in the apixaban and warfarin arms, respectively, had interruptions of treatment > 3 days in duration. Event rates for interruptions of various durations favored warfarin slightly in each comparison, with the maximum difference in any row being 1.17 events/ 100 pt years, with an extra 4 events in the apixaban arm in the row for interruptions > 14 days with event window (2). Event rates in the apixaban arm were not notably different from the event rate for the rate of the primary endpoint in apixaban completers in the 30 days following the discontinuation of study drug, which was about 4% (**Table 71**).

Reviewer Comment: These data do not support an important difference in the risk of thrombotic events in patients during interruptions of therapy with apixaban compared to warfarin.

6.1.10.2.2 Efficacy Endpoint Events after Discontinuation of Study Drug in AVERROES

Limited data from AVERROES, which compared apixaban to placebo, are available regarding post treatment events in the 30 days after discontinuation of study drug. Data from early discontinuation patients are similar to what was seen in ARISTOTLE – elevated rates of the primary endpoint and all-cause death compared to rate in the ITT analysis, with no significant differences between the treatment arms. For both the primary endpoint and for all cause death, the rate was numerically higher in the warfarin arm.

Completers in AVERROES were to receive standard care at the end of the study. There was no provision for a transition regimen. For primary endpoint events in completers, as in ARISTOTLE, there was a significantly higher rate of events in the apixaban arm (N=2243) than in the warfarin arm (N=2139, 9 vs. 1 patients with events, HR = 8.78 (95% CI, 1.1, 69.3, see **Table 79**). Most events in the apixaban arm occurred 8 - 30 days after the last dose of study drug. Note that in this analysis, patients who entered the long term open label extension of AVERROES were censored on the day of such entry.

Table 79 AVERROES – Primary Efficacy Endpoint Events in Patients who Completed the ITP

	Apixaban		Aspirin		Apixaban vs. Aspirin
Days after last dose	n/N	Events / 100 pt-yr	n/N	Events / 100 pt-yr	HR (95% CI)
1-30	9 / 1472	10.55	1 / 1421	1.20	8.78 (1.11, 69.34)
1-2	1 / 1472	13.94	0 / 1421	0	-
3-7	1 / 1142	6.48	1 / 1142	6.67	-
8-14	4 / 1095	19.52	0 / 1095	0	-
15-30	4 / 1034	9.43	0 / 1014	0	-

All events represented in **Table 79** were strokes. Nine patients in the apixaban arm had 10 strokes (only the first for each patient is counted in the first data row). The nine first strokes included 5 ischemic strokes (1 during the first 7 days of the post-treatment, 3 in the segment from day 8 to 14, and 1 on day 15), 3 strokes of uncertain type (1 during the segment from day 8-14 and 2 in the segment from day 15-30) and 1 hemorrhagic stroke that occurred on Day 1 in a 74 yo woman who was $CHADS_2$ at baseline and on no recorded concomitant medications. Two patients who had ischemic strokes were

noted as taking open-label aspirin during the post-treatment period; in both cases, it had been started before the end of study treatment and carried over. One patient received parenteral anticoagulation on an unspecified date during the post-treatment period. None was noted to have received a VKA, consistent with the study's inclusion criteria.

The nine patients in the apixaban arm with events had a mean baseline $CHADS_2$ score of 2.9, compared to 2.0 for the study as a whole. The distribution by score was: 1 patients with a score of 1; 2 with a score of 2, 4 with a score of 3, and 1 each with scores of 4 and 5.

The one patient in the aspirin arm with a stroke was 83 yo man with a baseline $CHADS_2$ score of 4 who had an ischemic stroke on Day 6. No concomitant medications are listed.

The event rates for all cause death during the 30 day post-treatment period in completers were elevated over the rate in the ITT analysis, but were similar in the apixaban and warfarin arms (9 vs. 10 events, 11-12 events per 100 pt-years, data not shown). Events were fairly evenly distributed throughout the 30 day post treatment period in both arms.

> *Reviewer Comment: The data for primary endpoint events in AVERROES completers during the 30 days after the last dose of study drug are qualitatively similar to the analogous data in ARISTOTLE – an elevated rate of stroke in a apixaban arm that is significantly higher than the rate in the warfarin arm. Most events in the apixaban arm occurred after the first week of the post treatment period, and ischemic/unknown strokes predominated. The patients with strokes in this period tended to have a higher baseline risk score than the randomized study population. However, unlike the ARISTOTLE patients, there is no suggestion that initiation of warfarin played a role in these strokes.*
>
> *The anti-platelet effects of aspirin are irreversible with respect to each exposed platelet, but platelet turnover (about 10%/day) results in loss of aspirin effect over time. The duration of the useful antiplatelet effects of aspirin has been estimated to be about 4 days in healthy volunteers.* [13,14] *It is interesting that most events in the apixaban arm occurred more than 8 days after the last dose, vs. no events in the aspirin arm during this period, when aspirin's antiplatelet effects should have dissipated.*
>
> *Thus, the post-treatment event data from completers in AVERROES suggest that patients who stop long term therapy with apixaban may be at higher risk of stroke than those who stop long term therapy with aspirin. While the number of completers who had strokes in the post-treatment period was small compared the number of strokes prevented during treatment, it seems prudent to try to understand this phenomenon in order to prevent it. A PMR should be considered*

if apixaban is approved. Note that there is a long-term open label extension of AVERROES which could be modified to yield useful data on its conclusion.

6.1.10.2.3 Laboratory Data Relating to Hypercoagulability

We asked the Applicant for information relating the potential existence of a hypercoagulable state in patients who came off apixaban therapy that may be contributed to the excess number of events those patients. Specifically, we asked if they had characterized changes in the concentrations of pro- and anti-coagulant proteins and the functional aspects of the coagulation system in patients who had stopped long-term apixaban therapy, as well as the pharmacodynamics of warfarin in such patients. The Applicant indicated that they had not performed such studies.

However, the Applicant makes the following arguments against the existence of a hypercoagulable state in patients who have discontinued long-term therapy with apixaban;

- Apixaban is selective for FXa and its effects on PT and other coagulations studies are reversible, so there is not a strong rationale for evaluating the effect of apixaban on other coagulation parameters.
- There is no early peak in events occurring after discontinuation of apixaban. One would expect an early peak in patients whose hypercoagulable state was masked during therapy with apixaban. The greater number of events on apixaban may be due to a 'catch-up' phenomenon following discontinuation of a very effective anticoagulant in patients at considerable risk for stroke.
- When patients who did not start treatment with open-label VKA are excluded from the analysis of post-treatment events in ARISTOTLE are excluded, the annualized rate of primary endpoint is reduced

  Reviewer Comment: The rate is reduced, but the difference from the rate in the warfarin arm remains substantial: 3.19 vs. 0.47 events per hundred patient-years (14/5723 vs 2/5570 events).

- The patients who had strokes were at high risk of stroke

  Reviewer Comment: This is true for both arms, but the event rate is higher in the apixaban arm. There is no evidence that the treatment arms were different in terms of stroke risk in the completing patients.

- The mean affect of apixaban in ARISTOTLE on INR was 1.16 to 1.19 from day 10 to month 30. This consistent, modest effect suggests that there is

no long term effect on coagulation system.

Reviewer Comment: This reviewer does not agree with the Applicant that there is not a strong rationale for examination of the coagulation system in patients who have discontinued long-term treatment with apixaban. Two studies of apixaban have demonstrated that stroke rates in completing patients were significantly higher in the apixaban arm patients than in the control arm over the 30 days after the last dose of study drug. Similar findings were observed in the ROCKET study of rivaroxaban, another FXa inhibitor. The Applicant's arguments do not rule out an effect on one of the many coagulation system constituents that might not alter the PT during treatment with apixaban. The fact that event rates were elevated compared to aspirin during the post-treatment period in AVERROES suggests that poor warfarin control in warfarin naïve patients cannot explain all the findings.

Thus, the Applicant should be required to conduct a post-approval study to examine the effects of long term therapy with apixaban on the coagulation system. The details of such a study are to be determined.

6.1.10.2.4 Instructions for the transition from apixaban to warfarin

The Applicant has submitted proposed labeling with instructions for transition from apixaban to warfarin based on the optional transition regimen used in ARISTOTLE. The text of the instructions follows:

(b) (4)

Reviewer Comment: However, the available data do not support the value of this transition regimen. Instead, there should be a boxed warning similar to the one in rivaroxaban labeling:

DISCONTINUING ELIQUIS IN PATIENTS WITH NONVALVULAR ATRIAL FIBRILLATION

Discontinuing ELIQUIS places patients at an increased risk of thrombotic events. An increased rate of stroke was observed following ELIQUIS discontinuation in clinical trials in atrial fibrillation patients. If anticoagulation with ELIQUIS must be discontinued for a reason other than pathological bleeding, consider administering another anticoagulant *[see Dosage and Administration (2.1), Warnings and Precautions (5.1), and Clinical Studies (14.1)]*.

6.1.10.2.5 Summary of Data Regarding Post-treatment Events

As noted above, there was an elevated rate of stroke in apixaban arm patients compared to during the 30 days after the last dose of study drug control in both of the Applicant's Phase 3 AFib trials in patients who completed the studies. While the data from ARISTOTLE suggest that poor INR control in patients who discontinued apixaban may have played a role in this phenomenon, this hypothesis cannot explain the similar finding in AVERROES, where no completing patients were expected or documented to have taken VKA in the post treatment period. Thus, it is reasonable to question whether some property of apixaban played a role in these findings. It is notable that the Applicant has not characterized the properties of the coagulation system in patients who have discontinued long-term treatment with apixaban. Such a study should be conducted; the details are yet to be determined.

The 4 tablet transition regimen used in ARISTOTLE did not appear to be beneficial.

6.1.10.3 Data Regarding Adjudication of the Primary Endpoint

The intent of the analysis below is to detect possible signals of bias favoring apixaban in the adjudication of strokes by the CEC. Strokes constituted the vast majority of event s in the primary endpoint analysis (456 of 487, 94%), and an advantage in preventing strokes accounted for the observed superiority of apixaban. Adjudication of strokes will thus be the focus of this section.

The adjudication process by the CEC is described in detail in Sec. **5.3.1.11**. Strokes and TIAs were adjudicated as a group. A total of 964 stroke/TIA events were adjudicated; 452 and 512 from the apixaban and warfarin arms, respectively. The triggers for adjudication of these events in 99% of cases in each arm were stroke/TIA event forms submitted by investigators; other event triggers included adverse events or other information appended to the CRF that suggested that a stroke or TIA might have occurred.

Potential bias will be examined by analyzing two types of discordant results: (1) cases from the apixaban arm where the investigator considered an event to be a stroke but it was adjudicated as not a stroke and (2) cases from the warfarin arm that the investigator considered as not a stroke but the case was adjudicated as a stroke. Either type of case would tend to favor apixaban in the primary endpoint analysis.

Table 80 is a display of the number of adjudications with each type of discordance by arm. Type 1 discordance (investigator's determination was stroke and the adjudication result was no stroke, specifically, an adjudication of TIA or "No Event"), was observed for 109 vs. 137 cases (events) in the apixaban and warfarin arms, respectively. Thus, the net result of this type of discordance substantially favored the warfarin arm in the primary endpoint analysis.

Type 2 discordance, i.e., the investigator's determination was no stroke (specifically, TIA or "stroke/TIA" or in a very few cases, no assessment at all) and the adjudication result was stroke, occurred in 17 and 12 cases in the apixaban and warfarin arms, respectively. Again, the net result of this type of discordance favored the warfarin arm, although the difference between the arms was small. In the few adjudicated cases where the investigator did not make an assessment of stroke or TIA, the case was picked up through review of AE forms or other clinical information that suggested a stroke may have occurred.

Table 80 ARISTOTLE – Discordance in Stroke Adjudication

Discordance Type	Apixaban N=9120 452 Events [1] n	Warfarin N=9081 512 Events [1] n
Type 1: Investigator result: Stroke / Adjudication result: No Stroke	**109**	**137**
Adjud: as TIA	12	14
Adjud. as No Event	97	123
Type 2: Investigator result: No Stroke / Adjudication result: Stroke	**17**	**12**
Inv determin: TIA	11	11
Inv determin: No Event [2]	3	0
Inv determin: "Stroke/TIA"	3	1

"Adjud" = Adjudicated; "Inv determin" = Investigator determination
1. No. of events adjudicated
2. Investigator determination of No Event was recorded in the adjudication database when the adjudication was not triggered by a Stroke/TIA event form
Source: Applicant's adjudication database, adj2.xpt

The Applicant included information from the adjudication packages for all primary endpoint adjudications in the provided eCRFs. This reviewer (MR) examined the adjudication packages for 11 of the Type 1 discordance cases from the apixaban arm and 12 of the Type 2 cases from the warfarin arm. In each case, the adjudication result had a reasonable basis in the medical information provided, although in the majority of cases, the investigator's determination also could have been supported reasonably by the information provided.

Reviewer Comment: The above information provides no evidence of bias favoring apixaban in the adjudication process.

7 Review of Safety

As discussed earlier, there is an ever evolving unknown rate of medication errors and unknown effect of the manual manipulations of the IVRS data on randomization. The apparent lack of the Applicant's knowledge of the issues make the reviewers uneasy about the monitoring and conduct of the trial and the potential implications on important endpoints. This information developed late during the review cycle, and the issues have not been resolved to our satisfaction. The Applicant does not appear to fully understand their data (answers to our questions went through multiple iterations, are not fully explained, and have only instigated more questions), the errors and the impact on safety. Moreover, the Applicant's medication error dataset appears to have errors in it whereby the data do not match the CRF and there are dates after database lock. This carelessness in cleaning the data adds to our skepticism of their responses to our medication error questions and makes us question their study conduct. The timing of the study conduct issues was late in the review cycle and much of the primary safety analyses were complete, so the review presented in this section essential ignores the medication errors. Sensitivity analyses for the fraud in China are presented in **Sec 3.1.2**.

Safety Summary

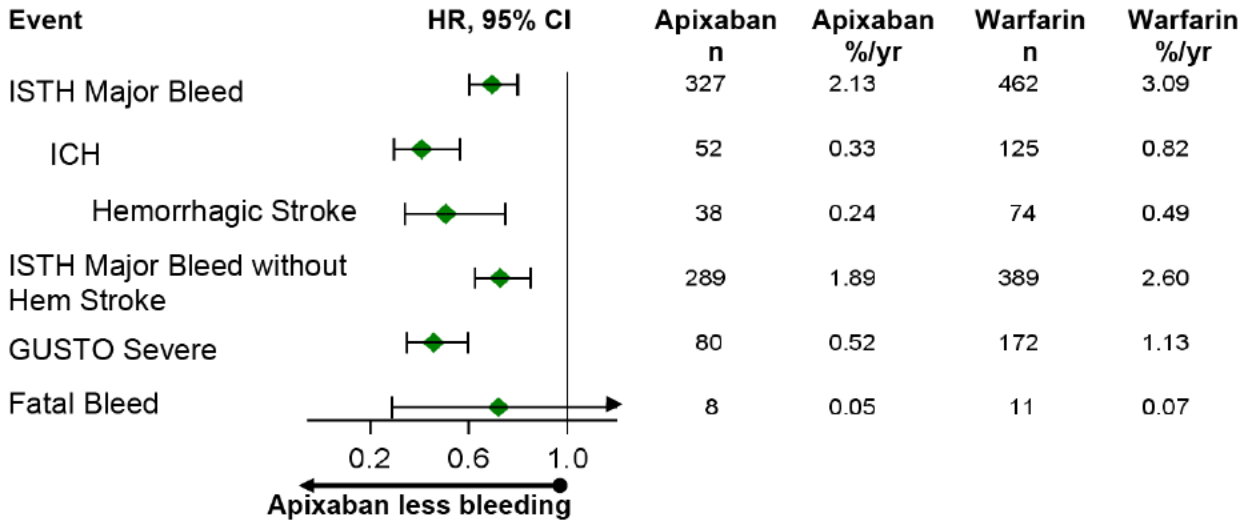
Most of the safety data for apixaban comes from ARISTOTLE (trial discussed in **Sec 5.3.1**). This was the largest randomized trial for the reduction in risk of stroke in AFib to date. As mentioned earlier, over 18,000 subjects with at least one risk factor for stroke were randomized to warfarin or apixaban in this non-inferiority trial; of the subjects randomized to apixaban, >95% of subjects received 5 mg po BID. Apixaban 2.5 mg po BID was for subjects with two out of three risk factors for bleeding at baseline. The mean duration of exposure, ~ 1.7 years, was adequate and similar to that of other large antithrombotic trials. The information from this trial alone is adequate to characterize the safety of apixaban. However, for specific events of interest (drug induced liver injury (DILI), rare serious neurologic adverse events, and concomitant antiplatelet therapy) the reviewer also analyzed the data in other trials.

The primary safety adjudicated endpoint was ISTH major bleed; a bleed that can be insignificant and readily reversible. However, the ISTH definition has historically been used in patients receiving long-term anticoagulation and was the primary safety endpoint in the last two NDAs for this indication. Certainly, it does not rise to the level of GUSTO severe or TIMI major definitions used in thrombolytic trials since the threshold for an ISTH major bleed is rather low (i.e., bleeding leading to a 2 unit transfusion of PRBC).

However, apixaban was consistently superior to warfarin for ISTH major bleed as well as for other serious bleeding (see **Figure 13**). While we do not like to compare across studies, the reviewer did examine across studies to look for consistency (albeit there were some differences between RE-LY, ROCKET-AF and ARISTOTLE). Although the

other antithrombotics were not superior to warfarin on ISTH major bleed, they each had a significant reduction in ICH, similar to apixaban.

Figure 13 ARISTOTLE – Relative and absolute rates of selected major bleeding on apixaban and warfarin



Reviewer's analysis: erate_HR\erateHR runs bleed, Applicant's dataset adbs2, adefl. HR, hazard ratio is apixaban/warfarin, n=number of subjects (first event). Fatal bleed is an adjudicated death due to bleeding during the treatment period.

The superiority findings in ARISTOTLE are robust (see **Table 81** Sensitivity Analyses of Superiority Findings for Bleeding Endpoints). To overturn the superiority findings, an additional 87 ISTH major bleeds in the apixaban arm are needed, or 93 fewer ISTH major bleeds in the warfarin arm are needed.

Table 81 Sensitivity Analyses of Superiority Findings for Bleeding Endpoints
(Number of additional events in the apixaban arm or fewer events in the warfarin arm needed to negate finding)

	By Odds Ratio Change in No. of Events		By Hazard Ratio Change in No. of Events	
	Apixaban	Warfarin	Apixaban	Warfarin
ISTH Major Bleed	327→ 408 =↑**98**	462 → 376 =↓**86**	327 → 414 =↑**87**	462→369=↓**93**
GUSTO Severe Bleeding	80→139 =↑**59**	172 →106 =↓**66**	80 →140 =↑**60**	172→104=↓**68**

Source: FDA Statistical Review, S. Bai

The site of major bleeds was mostly in the gastrointestinal (GI) tract (not a critical organ), followed by intracranial, then intra-ocular (**Table 86**). Similar to dabigatran and rivaroxaban, the site of most major bleeds was also the GI tract. However, in contrast

to both antithrombotics, the annual rate of major GI bleed was lower on apixaban compared to warfarin (by 0.1%) and relative to warfarin, there was no difference in major GI bleed. With the exception of intracranial and intra-ocular bleeds, the site of major bleeds was similar between subjects on either treatment. There were numerically more intra-ocular major bleeds on apixaban (n=33) compared to warfarin (n=22). There are pre-clinical data to support that there may be a pharmacologic basis for this; apixaban radioactivity was still present in the eyes of rats at 168 hours post dose while it was last measureable in plasma at 24 hours. If apixaban is approved, the reviewer recommends that information on major bleeding at this critical site be collected and reported to the Agency in their Safety updates. Attempts should be made to collect adequate information on the subject's medical history as to assess if a particular patient population may be at greater risk for major intra-ocular bleeding.

The reviewer conducted many analyses of all ARISTOTLE bleeding definitions and their relationship with TTR (time in range, time above range, time below range), a measure of warfarin control and safety. The overall mean TTR in ARISTOTLE was 62% (median 66%).[l] The many analyses did not show a consistent relationship between bleeding and TTR, suggesting that using TTR might not be a sensitive marker of individual warfarin control and ultimately bleeding events.

Analysis of countries by TTR quartiles shows that more than half of all US sites were above the median TTR of 66%; the same was true for Canada (the third highest enrolling country) (see **Figure 22**). Russia was the second highest enrolling country, but more than half of the subjects were in the lowest quartile of TTR, suggesting that warfarin control was poor in most subjects treated in Russia. The Ukraine, China, and India were the 9th, 10th and 11th top enrolling countries, but more than 75% of subjects in those countries were below the trial mean TTR of 62%.

Subgroup analyses of ISTH major bleed were generally consistent with the overall findings. A rather novel finding was that females on apixaban had less major bleeding than males and the relative benefit over warfarin was greater in females than in males. This finding is somewhat contrary to the results from a dedicated PK study that showed females have 15-18% higher concentrations than males, so one would not expect less bleeding in females based on PK. This reviewer is cautious about making conclusions based on un-prespecified subgroup analyses since findings can be spurious, and so would not recommend that this finding be included in labeling. Findings from subgroup analyses are hypotheses generating. As age increased, so did the risk of major bleeds, but relative to warfarin there was less major bleeding on apixaban in subjects ≥ 75 years old. While there were 5 subgroups[m] on apixaban that appeared to lose their bleeding advantage over warfarin, for many of these subgroups the effect size was still close to the overall trial effect size **[ISTH major bleed apixaban/warfarin HR (95%CI):**

l This is slightly worse than RE-LY, whose mean TTR was 64.4%.
m The subgroups were subjects < 65 years old, subjects with CHADS2 score ≥4, diabetics, subjects with previous stroke, and African Americans.

0.69 (0.60, 0.80)], so the reason may be due in part to smaller number of subjects. A group worth noting where this might not be the case was in diabetics [HR (95%CI): 0.96 (0.74, 1.25)]. Many of the intra-ocular bleeds were in diabetics and if apixaban is approved, the Applicant should look closer at this patient population. VKA status did not reduce major bleeding on warfarin. Compared to the overall trial, the rates of major bleeding in the US were higher (+0.7%) and the relative benefit of apixaban over warfarin was slightly less, HR 0.75 (95%CI, 0.56, 1.00), nominal p-value=0.0497.

Apixaban 2.5 mg BID, given to subjects having at least 2 of 3 risk factors for bleeding (age ≥ 80 years, weight ≤ 60 kg or serum creatinine ≥1.5 mg/dL) at baseline, was safe and effective. This dose adjustment was not based on PK (exposure in these subjects treated with 2.5 mg was 25% lower than (not equivalent) subjects treated with 5 mg). Most subjects qualified by age and weight. Although <5% of subjects received the lower dose, apixaban 2.5 mg was superior to warfarin on major bleed and stroke/SE (see **Table 94**). Major bleeding was higher in subjects with at least 2 out of 3 risk factors (despite the lower apixaban concentrations) compared to those without 2 out of 3 risk factors, suggesting that bleeding in this subgroup was more due to the population than to apixaban. The dose adjustment translated into a greater relative benefit over warfarin [(HR (95%CI) for apixaban 2.5 mg/warfarin, 0.50 (0.29, 0.86) versus apixaban 5 mg/warfarin, 0.71(0.61, 0.82)]. While the rate of stroke/se was also higher in this population, apixaban 2.5 mg had greater relative benefit over warfarin [(HR (95%CI) for apixaban 2.5 mg/warfarin, 0.40 (0.18, 0.88) versus apixaban 5 mg/warfarin, 0.80 (0.65, 0.98)]. If apixaban is approved, the lower dose should be approved and prescribed as used in ARISTOTLE.

Of the other SAEs there was numerically more syncope on apixaban (n=77) than on warfarin (n=47). The reviewer cannot explain this, but it is worth noting in the label.

Generally, there were no differences in reasons for treatment discontinuation (by System Organ Class) between treatment arms. However in the trial as a whole, more subjects discontinued for an AE in the warfarin arm (8.4% vs. 7.6%, respectively). This was driven in part by a nearly three-fold higher number of warfarin-treated subjects with discontinuations for injury, poisoning, and procedural complications (63 (0.7%) vs. 22 (0.2%) subjects, respectively). The most common reason for treatment discontinuation was in the SOC of nervous system disorders (1.5% of subjects on apixaban vs. 1.7% on warfarin) which consisted mostly of stroke/TIA events, followed by gastrointestinal disorders. There were numerically more major bleeds after apixaban discontinuation (n=44) compared to after warfarin discontinuation (n=29). Sparse data were collected after drug discontinuation. To investigate whether the excess bleeds was due in part from the inability to properly initiate warfarin treatment, the reviewer analyzed the major bleeds after starting treatment. The analyses as well as data in AVERROES do not support this contention. Unfortunately data collection after drug discontinuation was too sparse to know definitively the cause of the excess bleeds.

Apixaban does not appear to cause drug induced liver injury (DILI). There was one fatal case of hepatic failure that occurred on apixaban that independent, blinded, hepatologists judged as possibly related to apixaban or another drug (tianeptine). Otherwise, there were no probable cases and the number of potential Hy's Law cases was balanced between arms.

Apixaban does not appear to cause serious neurologic adverse events. In a P2 dose ranging trial there was 1 case of Guillain-Barre Syndrome (GBS) and 1 case of amyotrophic lateral sclerosis (ALS) in subjects taking apixaban 5 mg QD and 10 mg QD, respectively. Following the two reports, the Applicant enhanced surveillance for neurological events in all Phase 3 studies by use of a supplemental Clinical Safety Plan (CSP) CRF (See Attachment 9, **p.260**). This CRF was used to attain more information to aid in the diagnosis of the neurologic event. For ARISTOTLE, this was instituted about 7 months after the first subject was enrolled. Neurological consultations were also recommended for any SAE that matched a specific list of MedDRA terms. Additionally, in the second quarter of 2010 (near study completion of January 31, 2011) the Applicant instituted external, blinded, independent neurologist assessments[n] of any SAEs with PTs included in the MedDRA high level of terms of "acute polyneuropathies" (acute polyneuropathy, critical illness polyneuropathy, Guillain-Barre Syndrome, Miller Fisher syndrome, and polyneuropathy) and PT amyotrophic lateral. There were a total of 6 GBS cases, 3 ALS cases and 7 cases of acute polyneuropathy identified in the NDA and Safety Update Report (SUR). All but one case were blindly reviewed; of these, all consensus assessments were "unlikely to be drug-related". One subject treated with apixaban did not have a consensus assessment as of late April 2012, but the three individual neurologists' assessments were 2 unlikely, 1 possible. Serious neurologic AEs occurred infrequently and were balanced between treatment arms in ARISTOTLE. Based on the totality of the data, the reviewer believes that apixaban does not cause serious neurologic AEs such as GBS, ALS or acute polyneuropathy.

The most common adverse event was bleeding. Minor bleeding and clinically relevant non major bleeding was lower in the apixaban arm than in the warfarin arm (see **Table 101**). The reviewer was unable to complete the analysis of common adverse events because the sponsor's AE dataset contained errors that were likely created by an investigator filling out two or more CRFs for one unique event, a SAE CRF and a NSAE (non-serious adverse event) CRF. Monitoring did not appear to catch this. There was no systematic pattern for how this error happened, nor was there a way to easily fix the dataset since the AE term was sometimes mapped to different higher MedDRA terms. The Applicant's analysis of common AEs indicates that the frequency of AEs were similar between apixaban and warfarin.

n The external neurologists were Drs. (b) (4), who are recognized experts in peripheral nerve disease

There were no significant laboratory findings. Thrombocytopenia was similar between treatment arms. There were no significant effects on vital signs or ECGs. The Thorough QT study was negative.

Renal elimination does not play a large role in the excretion of apixaban since its elimination is multimodal. So one would not expect a large effect of renal impairment on PK, and certainly there is not a significant one (see **Sec 4.4.3**).[o] However, it is known that subjects with renal impairment are inherently at risk for more adverse events, including bleeds and strokes. Event rates of both major bleeding and stroke/se increase in both treatment arms as level of renal impairment worsens (**Figure 26** and **Figure 27**). Relative to warfarin, apixaban has less major bleeding in subjects with mild-severe renal impairment. For stroke/se, there was no suggestion of worse outcome on apixaban relative to warfarin.

Apixaban is a substrate for CYP3A4 and the drug efflux transporter proteins, p-glycoprotein (P-gp) and breast cancer resistance protein (BCRP). Coadministration with a strong inducer (rifampin) decreased apixaban concentrations by ~50% (**Figure 6**). Coadministration with strong inhibitors of CYP3A4 and P-gp (ketoconazole) increases apixaban AUC by 100% (or 2-fold). The Clinical Pharmacology reviewers recommend, and I agree, to avoid concomitant use with strong CYP3A4/P-gp inducers, and reduce the apixaban dose by half when coadministered with a strong CYP3A44/P-gp inhibitor.

Because the APPRAISE-2 study contains meaningful concomitant antiplatelet information, the reviewer analyzed these data to assess the bleeding risk of apixaban coadministered with antiplatelet drugs. APPRAISE-2 was an ACS trial where randomization was stratified by single or dual antiplatelet therapy after the subject's index qualifying event. The trial was stopped early because the excess in TIMI major bleeding outweighed the benefit of a combined ischemic efficacy endpoint (CV death, MI and ischemic stroke). Single or dual antiplatelet therapy with apixaban had bleeding rates 2-8 times greater than with placebo. The bleeding risk was ~2 times greater in subjects on dual antiplatelet treatment and apixaban compared to single antiplatelet treatment and apixaban. Although the population in APPRAISE-2 is different than those in ARISTOTLE, to put the bleeding in perspective, the rate of ISTH bleeding in APPRAISE-2 on apixaban plus single antiplatelet therapy is almost 3 times greater, and with dual antiplatelet therapy is 6 times greater, than in ARISTOTLE where apixaban was used alone. The rate of bleeding (TIMI major and ISTH major) on apixaban plus single antiplatelet therapy in APPRAISE-2 was similar to the rate of bleeding on warfarin in ARISTOTLE. So in patients that need an anticoagulant and an antiplatelet it may be reasonable to use apixaban with a single antiplatelet.

o The Clinical Pharmacology reviewers recommend no dosage adjustment in moderate or severe renal impairment.

There was not an imbalance in cancers between apixaban and warfarin. The effect of apixaban in pregnancy and lactation has not been studied.

Preclinical data suggests that activated charcoal may be useful for apixaban overdose.

Adjudication of major bleeds appeared balanced, and did not appear to strongly favor apixaban. Adjudication of MI, however, did tend to favor apixaban.

7.1 Methods

The Applicant's summary of clinical safety (SCS) includes information from 23,718 treated subjects with AF (11,886 apixaban, 11,832 warfarin or aspirin) from two large clinical trials, ARISTOTLE and AVERROES, and safety information for 218 subjects with AF who completed a Phase 2 study (CV185067) in Japan. The Applicant did not pool the data from the two Phase 3 trials because of the following differences between the two studies: population (warfarin suitable versus "unsuitable"), comparator arm (warfarin versus aspirin), and treatment duration (1.7 versus 1.1 years). The Applicant also included a safety report for 22,386 subjects in non-AF indications.

Reviewer's Comment: It was reasonable to not pool the data based on the different comparator arms. ARISTOTLE has substantive data for a safety review.

The study design and safety monitoring plans of both trials were similar and appropriate for a large antithrombotic trial. Both ARISTOTLE and AVERROES were event driven trials, so subjects continued study medication until enough subjects reached the primary efficacy endpoint (stroke/SE in both trials). There was a 30 day follow-up period for safety events; however, as is customary for event driven trials, subjects who discontinued study drug were followed for efficacy until enough primary endpoint events were reached. Subjects in AVERROES could enter into a long term open label extension study (LTOLE). Both Phase 3 AF trials had an independent Data Monitoring Committee (DMC), an external independent clinical events adjudication committee (CEC), three expert independent hepatologists who blindly assessed causality of specific hepatic events, and three independent neurologists who blindly assessed subjects with preselected neurologic AE/SAEs.

The Applicant's safety analyses are presented by treatment period (first dose to 2 days post dose) and by the follow-up period (Day 3 post dose to Day 30). For the on treatment analysis, subjects without the event of interest were censored at the earliest of the last date of the study drug + 2 days or the death date. SAEs occurring up to 30 days after the last dose are counted towards the "Treatment period" in the Clinical Study Report.

Reviewer's Comment: The planned oversight of safety data and collection of adverse events appears adequate and typical for a large antithrombotic trial. The proposed time periods for safety assessment are appropriate; however the reviewer's safety analysis

examined ALL data by on treatment as defined by the Applicant and by the follow-up period. Thus, unlike the Applicant's analysis, all SAEs were not lumped into one period since there was precedence for an increase in adverse events after study drug discontinuation (seen with rivaroxaban). The Applicant's coding dictionary for bleeding events, hepatic events and neurologic events appears adequate.

Subjects who did not experience a bleeding endpoint were censored at the earlier of: 2 days after study drug discontinuation, death date, or last contact date when a full assessment was performed (for subjects who withdrew consent to be followed up or were lost to follow-up) at the end of the study. Superiority testing for the primary safety endpoint (ISTH major bleeding) was performed using a Cox proportional hazards model including treatment group as a covariate and stratified by region and prior VKA status.

7.1.1 Studies/Clinical Trials Used to Evaluate Safety

The Applicant's primary safety sources are the AF studies, ARISTOTLE (CV185030) and AVERROES (CV185048). Each trial is described in **Sec 5.3.1 ARISTOTLE** and **Sec 5.3.2 Supporting Study: AVERROES**. The reviewer's safety analysis focused on data in ARISTOTLE (rather than AVERROES). Both trials assessed the same safety endpoint, however ARISTOTLE contained ~3x more subjects and treated subjects on average ~30 weeks longer than in AVERROES. Thus, the data in ARISTOTLE alone are sufficient to allow substantive assessment of the safety of apixaban in an AF population. Additionally, the comparator arm in ARISTOTLE was warfarin, an effective treatment widely used for the prevention of embolic events in AF, but causes serious bleeding. The Applicant asserts that apixaban causes less major bleeding than warfarin, the first drug to have this assertion; thus the reviewer critically examined this safety endpoint. In contrast, the comparator arm in AVERROES was aspirin which offers modest protection against stroke in AF. At the time of the NDA submission it was unclear to the review team the role (if any) that AVERROES would play in the approval decision. For certain safety events (rare neurologic events and DILI, and situations where the trial design was better suited for assessment of a particular safety issue) the reviewer also analyzed data from AVERROES and/or APPRAISE-2.

APPRAISE-2 (CV185068, **Ap**ixaban for **Pr**evention of **A**cute **Is**chemic **E**vents-2) was a phase 3, multi-national, randomized, double-blind study that compared apixaban (5 mg BID) to placebo in patients with recent acute coronary syndrome (ACS) and at least 2 additional risk factors for recurrent ischemic events. The primary endpoint was a composite of CV death, non-fatal MI and ischemic stroke in ACS. The trial started on March 17, 2009; the *DMC stopped the trial early on November 14, 2010 because of an excess of clinically important bleeding that was not offset by clinically meaningful reductions in ischemic events in subjects on apixaban*.

Some key elements of APPRAISE-2 include:

- subjects completed parenteral anticoagulation for index event and were stable
- Randomization was stratified by type of antiplatelet therapy (single or dual) at baseline.
- subjects received standard of care for ACS, including single (ASA or $P2Y_{12}$ antagonist) or dual antiplatelet therapy (left to physician discretion)
- planned 5400 subjects per arm
- estimated 938 patients needed to reach a primary efficacy endpoint (CV death, MI or ischemic stroke)
- decision to stop after 7,392 subjects randomized (3705 on apixaban, 3687 on placebo)
- primary safety endpoint was time to first TIMI major bleed (ISTH major bleed was one of the secondary safety endpoints)
- The dosing regimen was 5 mg BID; subjects with a CrCl < 40 mL/min received apixaban 2.5 mg BID.

The Applicant's cutoff dates for efficacy and safety in the NDA are shown in **Table 82**.

Table 82 Efficacy and safety cutoff dates

	Cutoff date
ARISTOTLE Efficacy (ITT period)	January 31, 2011
ARISTOTLE Database lock and unblinding	June 10, 2011
AVERROES Database lock and unblinding	December 8, 2010
Original NDA Safety Database lock for non-AF indications	April 1, 2011 (with the exception of May 10, 2011 for neurologic cases and July 31, 2011 for hepatic cases)
Safety Update Report (SUR) Database lock	August 15, 2011 (External hepatologists reports received from July 31, 2011 up to September 30, 2011 are included. External neurologists' reports received from May 10, 2011 up to September 30, 2011 are included)

7.1.2 Categorization of Adverse Events

Reviewer's Comment: The Applicant's adverse events analysis dataset (adae.xpt) contains more than one line of observation for the same event for some subjects whereby the same unique event is listed as serious in one line and non-serious in the next line. While it is conceivable that an event initially reported as an AE may become an SAE (indeed this was in the Applicant's protocol to fill out both CRFs if this was the case), it is apparent that the reason for multiple lines of observation for the same unique

event is because site personnel completed both a CRF for non-serious adverse event and SAE. Study monitoring did not catch this. There does not appear to be a consistent pattern to predict when this occurs in the dataset. Moreover, some of the multiple observations (although its one unique event) are subsequently mapped to different MedDRA terms, so cleaning the dataset is labor intensive.[p] An example is highlighted below for Subject CV185030-880-552 who has three lines of observation for an ischemic left foot systemic embolism, summarized as follows:

Table 83 ARISTOTLE - Unique event coded as both SAE and NSAE example

Observation	AE term	Start date	End date	Serious
1	ischemic left foot 2nd to embolism	22Mar2009	21Apr2009[q]	SAE
2	Ischemic left foot	22Mar2009	21Apr2009	NSAE
3	Systemic embolism to left foot	22Mar2009	21Apr2009	NSAE

Dataset adae.xpt, End date variable is AEENDN.

This subject came to the hospital with a chief complaint of "discoloration of her left foot and toes" starting on (b) (6). She was admitted to the hospital and subsequently received urokinase on (b) (6) as part of her care. The event for this subject could be counted as one event (systemic embolism) or two events (ischemic left foot and systemic embolism). Regardless, the event(s) should be listed as an SAE only, not also as an NSAE. Note that the reviewer is not so concerned about this particular case for the AE/SAE analysis because it was an adjudicated event.

The Applicant should clean up the AE analysis dataset and resubmit the revised dataset along with their algorithms for cleaning up the dataset. Additionally the Applicant should provide an analysis and report of significant adverse events as defined in ICH E3 and an analysis of AE with intensity of severe or greater.

The finding of an event being over reported does not appear to affect the SAE analysis. It does affect the "significant AE" and common AE analysis. More discussion of this finding and its implications are discussed in **Sec 3.1.3**. See also **Sec 5.3.1.9.3** for information on AE reporting in ARISTOTLE.

Adverse events were coded and grouped into preferred terms by System Organ Class (SOC) using the Medical Dictionary of Regulatory Activities (MedDRA) version 14.0 for ARISTOTLE and the Applicant's integrated summaries and MedDRA version 13.1 for AVERROES. The Applicant reports that the variable AETERM in the AE dataset contains verbatim terms.

p The reviewer noticed different mapping for the same event up to a level as high as HLGT.
q The initial SAE report has an end date of (b) (6) the date the subject received urokinase. The narrative was written on (b) (6) and estimated the subject's length of stay to be 2-3 more days afterwards. However, in the follow-up SAE report, the subject had never been discharged, and she subsequently had a below the knee amputation; the end date was 21Apr2009.

Reviewer's Comment: The Applicant's categorization of bleeding related AEs, hepatic disorder AEs, and neurologic AEs was reasonable. The list of MedDRA preferred terms were finalized prior to unblinding of the database. [Reviewer's analysis: ae\meddra_coding] Various hepatic SMQs (standardized MedDRA queries) were used, terms for neurologic events appeared reasonable, and the hemorrhages SMQ was used to code bleeding events.

7.1.3 Pooling of Data Across Studies/Clinical Trials to Estimate and Compare Incidence

In general, this was not done for reasons already described in **Methods**. The Applicant did pool the liver function tests, neurologic AEs, and platelet decreases. This was reasonable.

7.2 Adequacy of Safety Assessments

7.2.1 Overall Exposure at Appropriate Doses/Durations and Demographics of Target Populations

Because of two rare AEs of special interest (DILI and serious neurologic AEs), a short discussion of the breadth of apixaban exposure is warranted. As of 01Apr2011 (NDA cutoff) 57,706 subjects received medication in apixaban clinical trials: apixaban 23,961, comparator 22,361, blinded 11,384. The Safety Update Report (SUR) presented safety data for 15,603 treated subjects from three ongoing studies (AVERROES LTOLE, and VTE blinded treatment studies CV185056 and CV185057) and two completed studies, both in non-AFib, CV185070 and CV185036).

For the target population, the mean duration of exposure in ARISTOTLE , ~ 1.7 years, was adequate and similar to that of other large antithrombotic trials. There was not a large difference in exposure between VKA naïve and VKA experienced subjects or between apixaban and warfarin treated subjects (**Table 84**). [Source: Applicant's Table S.4.1B1, 1B2]

Table 84 ARISTOTLE - Exposure

Population	Apixaban	Warfarin
Intent To Treat (n)	9,120	9,081
Subject years	16,741	16,577
As Treated* (n)	9,088	9,052
Not accounting for interruptions		
Mean (SD), weeks	89.2 (42.9)	87.5 (43.7)
Median, weeks	89.9	87.8
Min, max, weeks	0.1, 215.0	0.1, 214.6
Subject years	15,534	15,184

VKA naïve mean, weeks	89.2 (45.2)	86.8 (46.0)
VKA experienced mean, weeks	89.2 (41.1)	88.1 (41.8)
Accounting for interruptions		
Mean (SD), weeks	87.2 (42.8)	85.6 (43.1)
As treated and without major protocol evaluations (Applicant defined as "Evaluable")	8,518	8,475

Applicant's data, source Table 6.1, S.4.1B1

The demographics of the subjects in ARISTOTLE are well balanced between the two treatment arms (**Table 30**) and reflect the type of population with AF fairly well (~70% ≥ 65 years old, ~65% male, with the exception of perhaps only 1.3% of subjects were Black/African American.

7.2.2 Explorations for Dose Response

This is discussed in Sec **6.1.8**.

7.2.3 Special Animal and/or In Vitro Testing

Non-clinical testing was adequate to explore potential adverse reactions. There is a brief summary in **Sec 4.3**. See Dr. Harlow's PT review for more information.

7.2.4 Routine Clinical Testing

At each visit, ascertainment of clinical events for adjudication was done and is described in **Sec 5.3.1.7.1**. **Table 15** describes the time and schedule of procedures in the trial. Notably, LFTs, CKs, and AEs were assessed as often as INR. The safety assessments are appropriate.

7.2.5 Metabolic, Clearance, and Interaction Workup

This was summarized in Section 4.4 Clinical Pharmacology

7.2.6 Evaluation for Potential Adverse Events for Similar Drugs in Drug Class

The identification of AEs in ARISTOTLE appeared to be reasonable. The SAEs specific for this drug class are included in those that were adjudicated and those methods have been discussed. Triggers for suspected events are discussed in **Attachment 5 ARISTOTLE - Adjudication Trigger Document**. Serious adverse events of interests included thrombocytopenia, drug induced liver injury, serious neurologic events, and MI. To identify these events the Applicant monitored AEs, SAEs and laboratory abnormalities. Monitoring reports were run at varying times throughout the trial (as

often as every 2 weeks to on average monthly) depending on the number of subjects enrolled in the trial (personal communication with the Applicant on March 1, 2012). All of the algorithms (except MI and thrombocytopenia) used select MedDRA preferred terms to identify events of interest; hepatobiliary, MI and thrombocytopenia algorithms included laboratory abnormalities. These adverse events of interest were followed-up with a Clinical Safety Plan (CSP) CRF specific for that AE (**Attachment 9**).

The Applicant did a reasonable job identifying potential cases of DILI with the one caveat that the cases submitted for blinded, independent review had the *potential* to exclude subjects with significant total bilirubin elevations that occurred after the transaminase elevation. This is because the selected cases were for those subjects with elevations of ALT>3xULN and total bilirubin>2xULN on the *same* day or for subjects with SAE of jaundice, hepatitis, and hepatic failure. Since hepatocellular injury may take time to impair bilirubin clearance, clinically significant elevations in both transaminase and total bilirubin may not occur simultaneously.

7.3 Major Safety Results

7.3.1 Deaths

Mortality was a primary efficacy endpoint, so it is discussed in **Sec 6.1.5.1**. The safety reviewer did examine cause of death during the ITT period and the treatment period in ARISTOTLE. Other than the numerically higher number of deaths during the treatment period (266 on apixaban, 296 on warfarin) and in the trial overall (656 on apixaban, 717 on warfarin) compared to the ITT period (apixaban 603, warfarin 669), the causes of death were similar between periods. There were some numerical imbalances (e.g., death from infection on treatment apixaban 20, warfarin 12, death from MI post dose apixaban 9, warfarin 6), but the numbers are too small to make general conclusions. The most common cause of death was sudden death. There were more deaths from trauma in the warfarin arm (9) compared to apixaban (4), but the numbers are too small to make any conclusions regarding the ability to stop fatal bleeding due to trauma in subjects on apixaban.

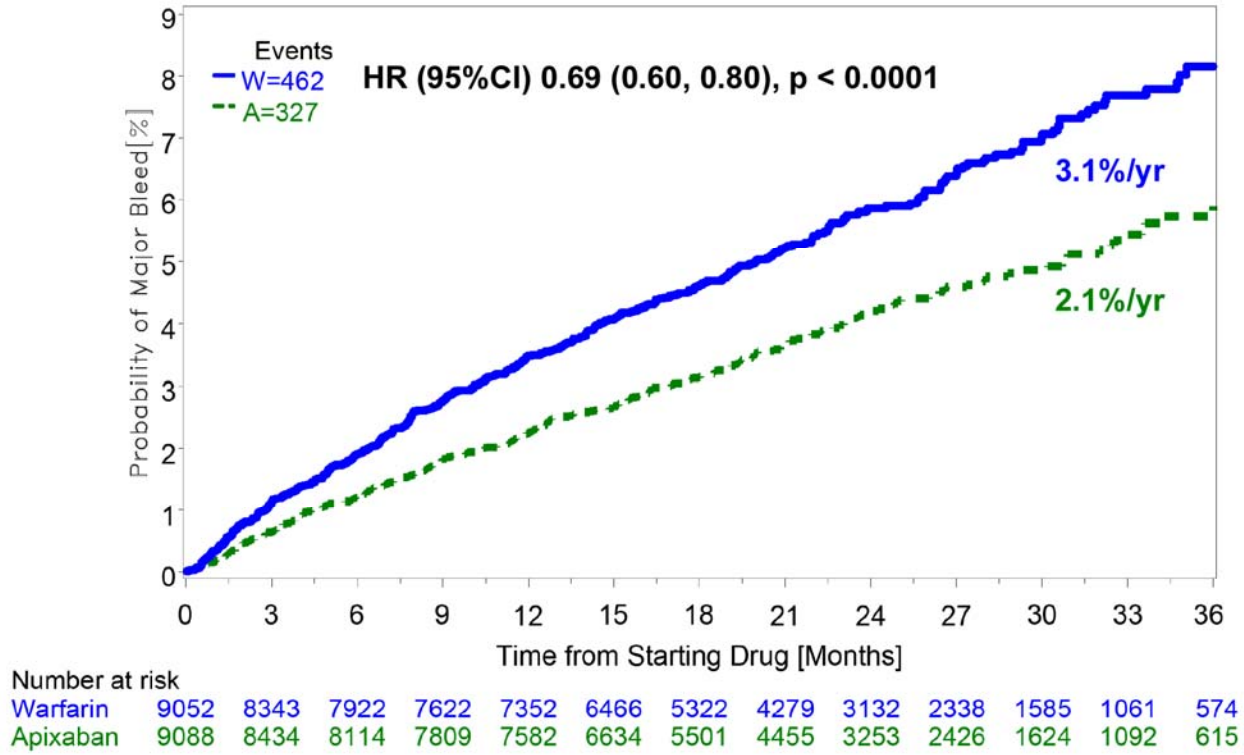
7.3.2 Nonfatal Serious Adverse Events

7.3.2.1 Major Bleeding

Directly related to the pharmacologic activity of apixaban is its primary safety concern and most common adverse event, bleeding. Although this discussion could be placed in Section 7.3.5. Submission specific primary safety concerns, because some major bleeding was serious, the reviewer placed it here. In ARISTOTLE, the time to first ISTH major bleed was the primary safety endpoint (**Figure 14)**. There was an early separation between the two treatments with more events occurring in the warfarin arm; this difference was maintained throughout the trial. The risk of first ISTH major bleed was 31% lower on apixaban as compared to warfarin (p<0.0001). The absolute risk per

100 subject years was 3.1% on warfarin and 2.1% on apixaban; this difference means that treating 100 patients for a year with apixaban (versus warfarin) will avoid 1 major bleed.

Figure 14 ARISTOTLE - Time to first ISTH major bleed



Reviewer's analysis (on treatment): \bleed\tte\tte MB.sas, \erate_HR\erate HR runs bleed, Applicant's datasets adbs2, adbl2. HR, hazard ratio is apixaban/warfarin

While there were more first ISTH major bleeds in the warfarin arm compared to the apixaban arm, the number of rebleeds on treatment was similar between arms (**Table 85**). This just supports that in subjects that bleed, they are likely to bleed again irrespective of treatment.

In contrast to major bleeds during the treatment period, the post treatment period had more major bleeds in the apixaban arm compared to the warfarin arm. The events in the apixaban arm were fairly equal from one week to the next (averaging 11 events per week). In contrast, most of the events (n=16) in the warfarin arm occurred in the first week off drug. The timing of the events in the apixaban arm are not suggestive that they may in part be due to warfarin initiation. Nevertheless to look at this issue further, the reviewer also examined major bleeding by VKA status after drug initiation (**Figure 23**). Although the events are few, during treatment initiation, there were more major bleeds in subjects randomized to *apixaban who were VKA experienced (the most bleeds were not in subjects randomized to warfarin who were VKA naïve)*. Major bleeds in warfarin naïve subjects began to exceed the other subjects after about 15 days of

treatment. Thus, the data do not strongly support that the greater number of major bleeds after apixaban discontinuation is due in part to warfarin initiation.

Table 85 ARISTOTLE - Timing of major bleed

	Treatment	Study Period				
		On treatment	PST 3-30 d	PST 31-60 d	PST >60 d	Total
First bleed	Apixaban	327	44	4	21	397 [a]
	Warfarin	462	29	8	12	511
Total bleeds	Apixaban	355 (+28)	51 (+7)	9 (+5)	26 (+5)	442 (+48)[a]
	Warfarin	493 (+31)	37 (+8)	10 (+2)	13 (+1)	553 (+42)

For total bleeds, if >1 major bleed reported in a day, it was counted as 1 bleed. If a fatal bleed occurred on a different day than the major bleed, it was counted as a separate event.
(n)= the difference in total bleeds compared to first bleed
a. The total does not equal the sum across the row because 1 event occurred during screening.
Reviewer's analysis: bleed\location\totalMB.sas dataset oneperday, Applicant dataset adbl2

One component of the ISTH major bleed definition is clinically overt bleeding at a critical site (defined on **page 90**). These data are shown in **Table 86**. About 30-35% of subjects with a major bleed had bleeding into a critical organ. The site of major bleeds was mostly in the gastrointestinal (GI) tract (not a critical organ), followed by intracranial, then intra-ocular (**Table 86**). With the exception of intracranial and intra-ocular bleeds, the site of major bleeds was similar between subjects on either treatment. Most of the numbers in the table below represent one subject.[r]

r Exceptions are for GI bleeds, intra-ocular bleeds (one subject taking apixaban had two bleeds), and ICH (one subject taking warfarin had two bleeds~ 3 months apart).

Table 86 ARISTOTLE - Site of major bleed[s]

	Treatment	Study Period	
		TRT	PST 3-30 d
Subjects with major bleed into a critical organ (% of subjects)[1]	Apixaban	95 (1.1)	9
	Warfarin	164 (1.8)	5
Site		**Number of bleeds**	
Gastrointestinal	Apixaban	137	16
	Warfarin	146	13
ICH	Apixaban	52	6
	Warfarin	126	2
Intra-ocular	Apixaban	33	1
	Warfarin	22	1
Intra-articular	Apixaban	6	0
	Warfarin	10	1
Retroperitoneal	Apixaban	2	1
	Warfarin	5	0
Intra-spinal	Apixaban	2	0
	Warfarin	2	1
Intramuscular with compartment syndrome	Apixaban	1	0
	Warfarin	1	0
Pericardium	Apixaban	0	1
	Warfarin	0	0

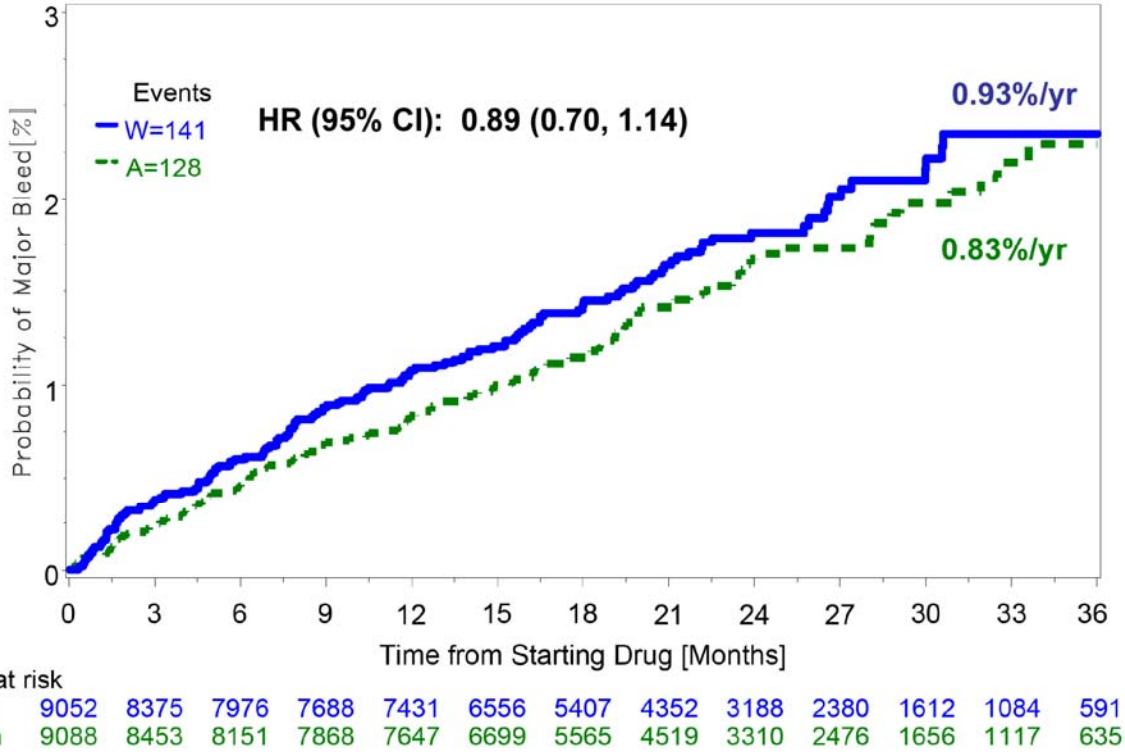
1. Critical site = all sites in table except gastrointestinal.
TRT=on treatment (Day 1 to 2 days after the last dose), PST 3-30 d is the period from 3 to 30 days after the last dose.
Reviewer's analysis: bleed\location\site.sas, Applicant dataset: adblsaf (resubmitted 2/8/12).

Similar to dabigatran and rivaroxaban, the site of most major bleeds was the gastrointestinal tract. However, in contrast to both antithrombotics, the annual rate of major GI bleed on apixaban was lower than that on warfarin (**Figure 15**).[t] The relative

s The reviewer's analysis differs from the sponsor's CSR because after the initial NDA submission the sponsor reassigned major bleed locations marked as "Other", but with information on ICH, GI or intraocular bleeds, to their appropriate location following the reviewer's request.
t The rate of major GI bleed in ROCKET-AF was 2.0%/year on rivaroxaban and 1.2%/year on warfarin and in RE-LY was 1.6%/year on dabigatran 150 mg and 1.1 %/year on warfarin.

difference, however, was not significant.[u] Similar to rivaroxaban, the majority (~60%) of the GI bleeds was in the upper GI tract (reviewer's analysis: location\GI, Applicant's dataset adbl2). So while the site of major bleeding was primarily in the GI tract (upper GI tract to be more specific), relative to warfarin, the major GI bleeding was not different.

Figure 15 ARISTOTLE - Time to first ISTH major gastrointestinal bleed

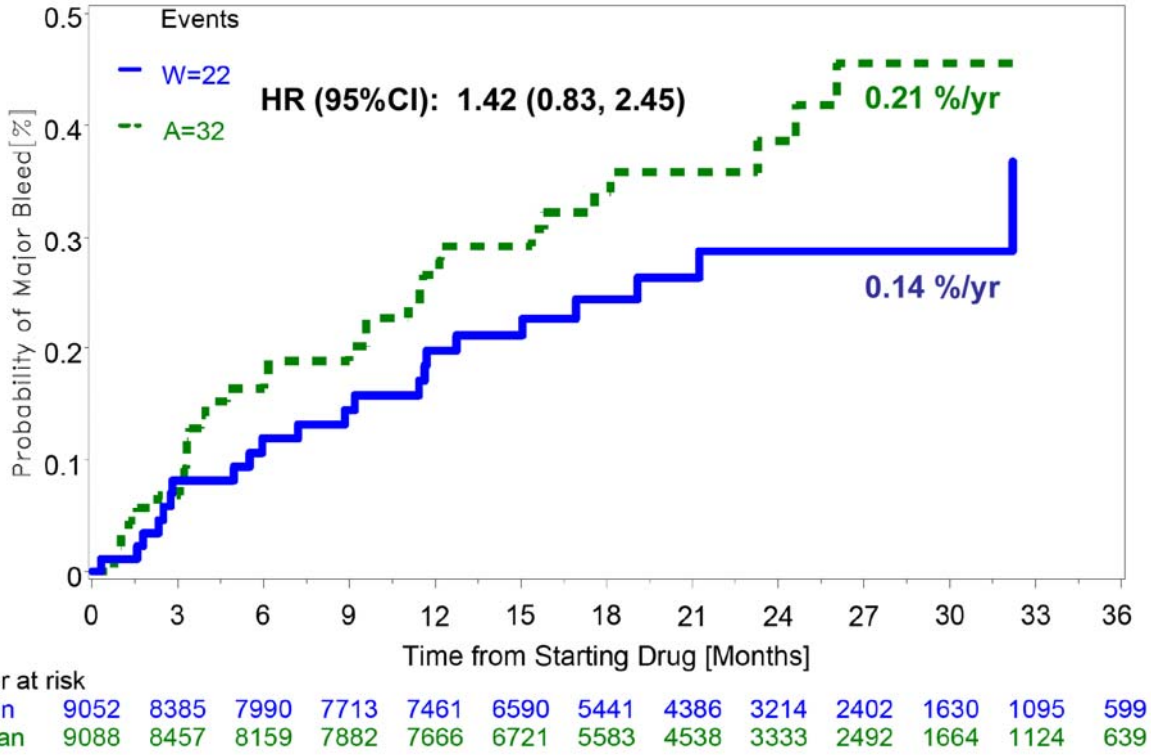


Reviewer's analysis (on treatment): \erate_HR\erate HR runs bleed, \bleed\tte\tte MBGI, Applicant's datasets adbs2, adbl2 (resubmitted 2/8/12). HR=apixabn/warfarin

Although the numbers were small, there were more intra-ocular bleeds on apixaban (0.2%/year) than on warfarin (0.1%/year). The numbers were likely too small to show a relative difference. The pharm/tox review points out that apixaban radioactivity was still measurable in the eye at 168 hours after the dose while the last measureable time point in plasma was 24 hours. Thus, there are pre-clinical data to suggest that the difference in intra-ocular bleeds, albeit small, are likely due to inherent properties of apixaban. If apixaban is approved, the Applicant should monitor and report major bleeding, but special consideration should be given to intra-ocular bleeding as it may warrant a more prominent highlight in the label.

u This is in contrast to RE-LY where the HR (95%CI) of dabigatran 150 mg/warfarin was 1.47 (1.17, 1.85). Cox proportional hazard analysis not found for ROCKET-AF.

Figure 16 ARISTOTLE - Time to first ISTH major intra-ocular bleed



Reviewer's analysis (on treatment): \erate_HR\erate HR runs bleed, bleed\tte\tte intraocular, Applicant's datasets adbs2, adbl2(resubmitted 2/8/12). HR=apixaban/warfarin

The other components of the ISTH major bleed definition, a Hgb reduction of ≥ 2 g/dL . a transfusion of ≥ 2 U of PRBC, and a fatal bleed are shown in **Table 87** along with other components of popular bleeding definitions such as GUSTO severe bleeding. After adjusting for the larger number of total bleeds in warfarin treated subjects, the characteristics of the major bleeds between the two treatment arms were balanced. About 50% of all bleeds required hospitalizations, 30% required medical or surgical intervention, and 13% led to hemodynamic compromise. Of the major bleeds ~8-10% (35/355 on apixaban and 37/493 on warfarin) met the definition solely by hemoglobin drop (reviewer's analysis: bleed\components, Applicant dataset adblsaf). In other words, these subjects did not have major bleeding at a critical site and other criteria for an ISTH major bleed were negative. There was a small number of adjudicated fatal bleeds, with a numerically higher number in warfarin treated subjects.

Table 87 ARISTOTLE - Characteristics of ISTH major bleed

	Treatment	Study Period	
		TRT	PST 3-30 d
Required hospitalization	Apixaban	181	28
	Warfarin	259	12
Hg drop ≥ 2 g/dL	Apixaban	180	24
	Warfarin	228	22
Required medical/surgical intervention	Apixaban	109	22
	Warfarin	160	13
Led to ≥ 2 U transfusion	Apixaban	76	10
	Warfarin	107	8
Hemodynamic compromise	Apixaban	45	12
	Warfarin	65	8
Led to ≥ 4 U transfusion	Apixaban	28	7
	Warfarin	39	5
Reported fatal bleed	Apixaban	15	3
	Warfarin	22	3
Adjudicated fatal bleed	Apixaban	8	5
	Warfarin	11	4

TRT=on treatment (Day 1 to 2 days after the last dose), PST 3-30 d is the period from 3 to 30 days after the last dose.
Reviewer's analysis: bleed\components.sas, Applicant dataset adblsaf

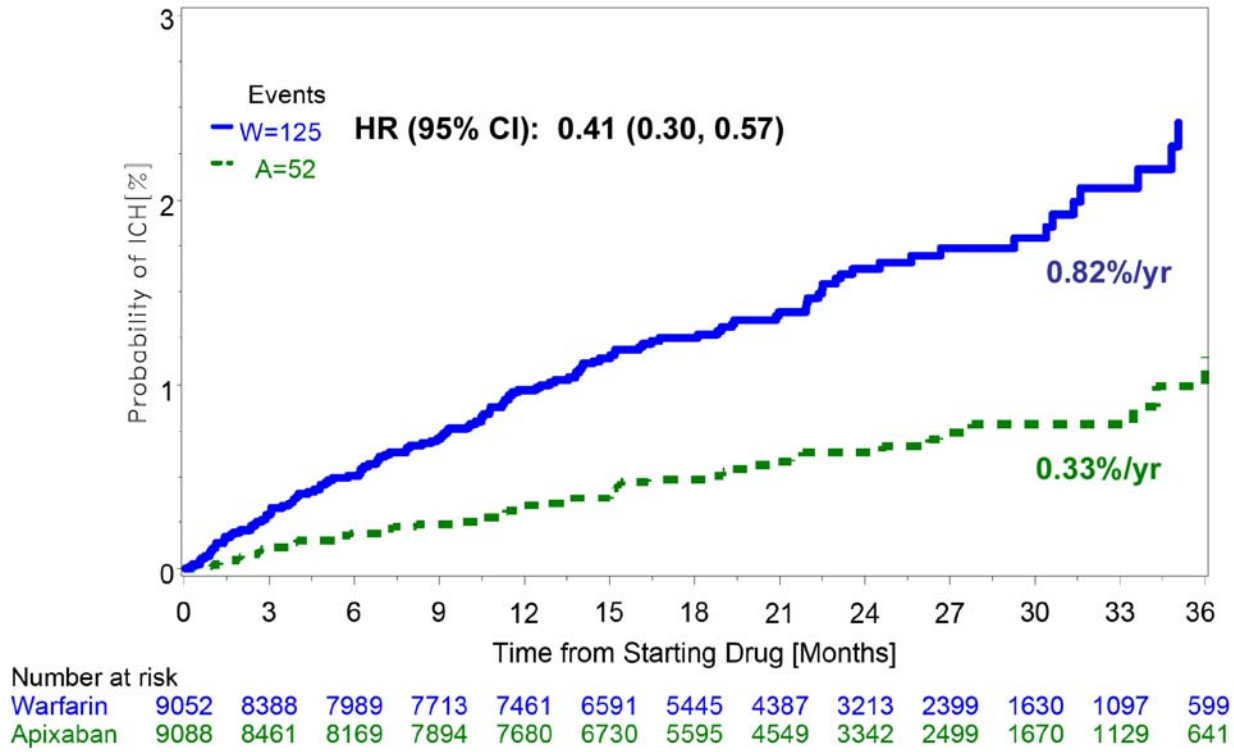
7.3.2.1.1 ICH

Among the most serious types of ISTH major bleeds are ICH and its subset hemorrhagic stroke. Similar to the other approved antithrombotics, there were less intracranial hemorrhages with apixaban relative to warfarin; this benefit was evident soon after starting treatment, continued throughout the trial and was statistically significant **(See Figure 17)**.[v] The relative and absolute benefit compared to warfarin

v. ROCKET-AF rivaroxaban/warfarin HR (95%CI), 0.67 (0.49, 0.91),
RE-LY, dabigatran 150mg/warfarin HR (95%CI), 0.41 (0.28, 0.60). The rate of ICH in ROCKET-AF was 0.47%/year on rivaroxaban and 0.85%year on warfarin and in RE-LY was 0.32%/year on dabigatran 150 mg and 0.76%/year on warfarin.

was strikingly similar to that seen in RE-LY for dabigatran 150 mg. There was a 59% relative reduction in ICH on apixaban as compared to warfarin in ARISTOTLE. The absolute risk per 100 subject years was 0.82% on warfarin and 0.33% on apixaban; this difference means that treating 1000 patients for a year with warfarin results in 5 more ICH compared to apixaban.

Figure 17 ARISTOTLE - Time to first ICH

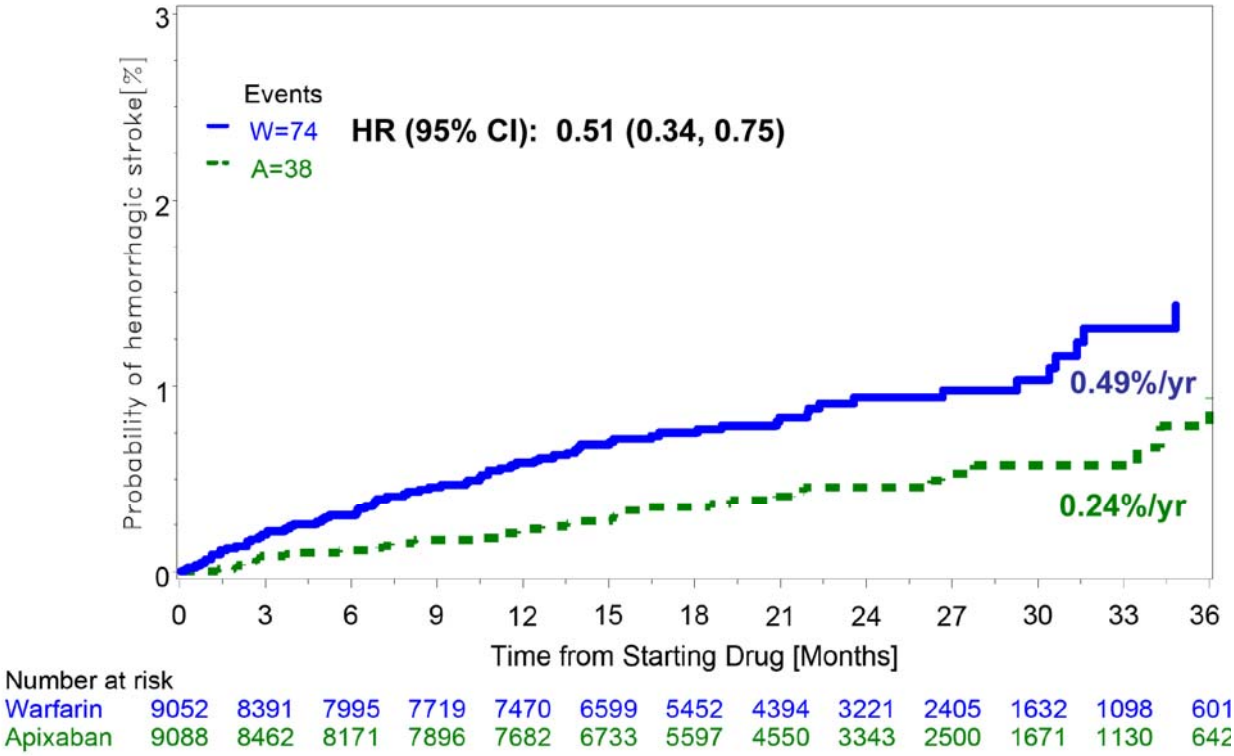


Reviewer's analysis (on treatment): \erate_HR\erate HR runs bleed, \bleed\tte\tte ICH, Applicant's datasets adbs2, adbl2. HR=apixaban/warfarin

7.3.2.1.2 Hemorrhagic Strokes

Treatment with apixaban had a 49% reduction in hemorrhagic strokes compared to treatment with warfarin. The difference in strokes started early and continued to separate during the trial (**Figure 18**). The annualized difference means treating 1000 patients with warfarin for a year will cause 3 additional hemorrhagic strokes versus treatment with apixaban.

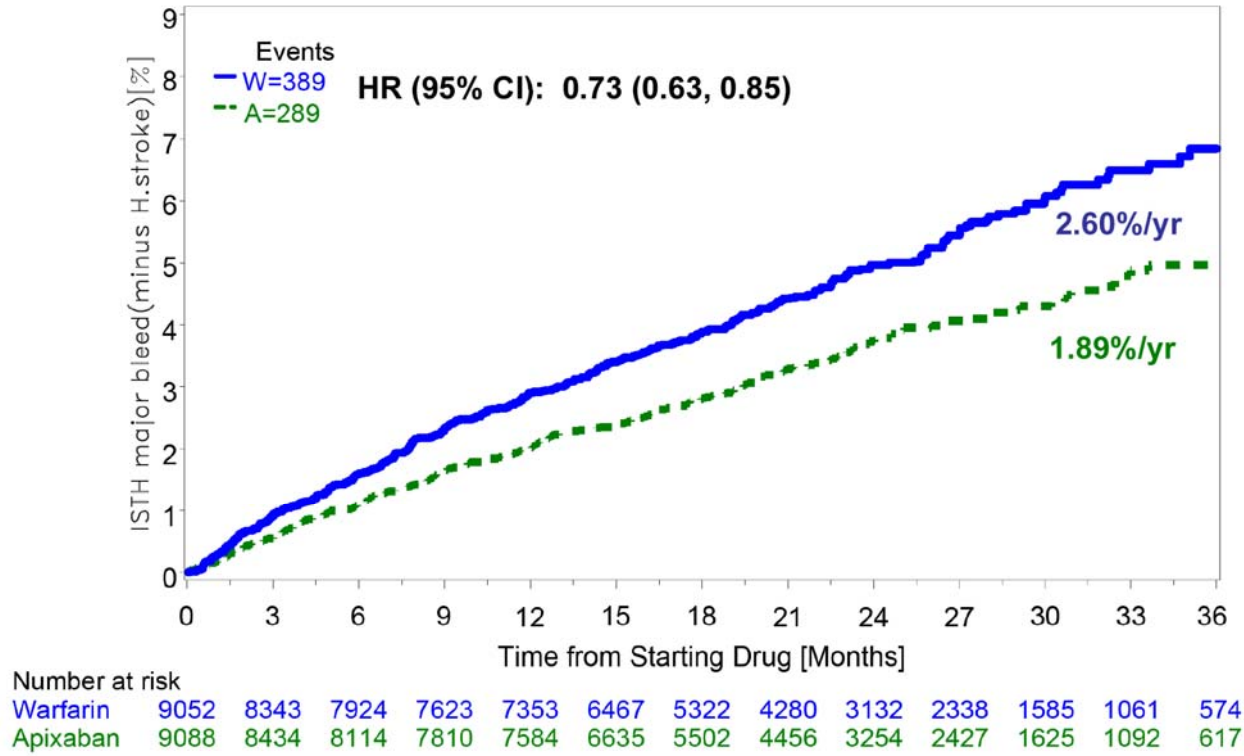
Figure 18 ARISTOTLE - Time to first hemorrhagic stroke



Reviewer's analysis (on treatment): \erate_HR\erate HR runs bleed, \bleed\tte\tte Hstroke, Applicant's datasets adbs2, adbl2. HR=apixaban/warfarin

Since hemorrhagic strokes were also counted in the efficacy analysis, the reviewer removed hemorrhagic strokes from the safety analysis for ISTH major bleeding. After excluding hemorrhagic strokes, the relative benefit of apixaban over warfarin was reduced by 4%, but still statistically significant; there was a 27% relative reduction in ISTH major bleed on apixaban compared to warfarin. Instead of treating 100 patients (as stated in **Sec 7.3.2.1)**, one would have to treat 141 patients to avoid one major bleed on apixaban compared to warfarin.

Figure 19 ARISTOTLE - Time to first ISTH major bleed, excluding hemorrhagic strokes



Reviewer's analysis (on treatment): analysis\erate_HR\erate HR runs bleed, \bleed\tte\tte MB_Hstroke, Applicant's datasets adbs2, adbl2. HR=apixaban/warfarin

7.3.2.1.3 GUSTO and TIMI Bleeding

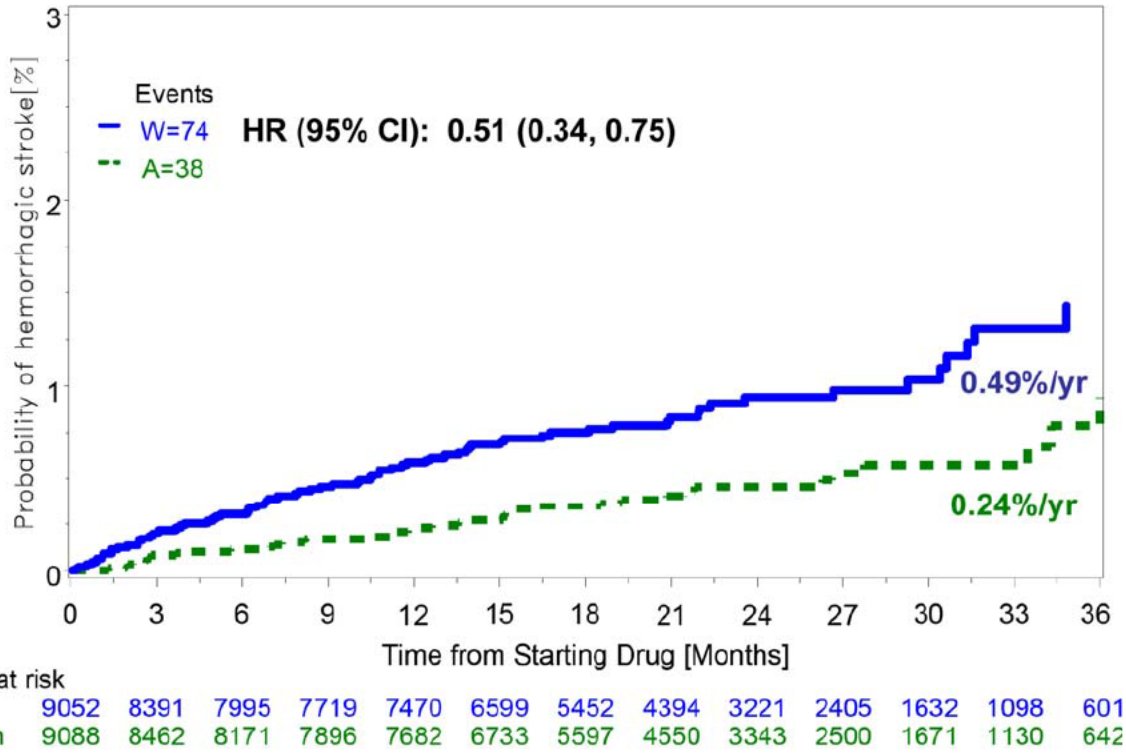
The reviewer also examined other serious definitions of bleeding (GUSTO severe and TIMI major, **see page 91** for definition) used in cardiovascular trials. There remained a significant advantage for apixaban with the serious bleeds GUSTO severe and TIMI major **(See Table 88, Figure 20, Figure 21).**

Table 88 ARISTOTLE -TIMI bleeding

Event	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
	(n)	%/yr	(n)	%/yr	HR	95% CI
TIMI Minor	98	0.63	118	0.78	0.82	(0.62, 1.07)
TIMI Major	148	0.96	256	1.69	0.57	(0.46, 0.70)

Reviewer's analysis (on treatment): erate_HR\erateHR runs bleed, Applicant's dataset adbs2. HR, hazard ratio is apixaban/warfarin, n=number of subjects (first event)

Figure 20 ARISTOTLE - Time to first GUSTO severe bleed

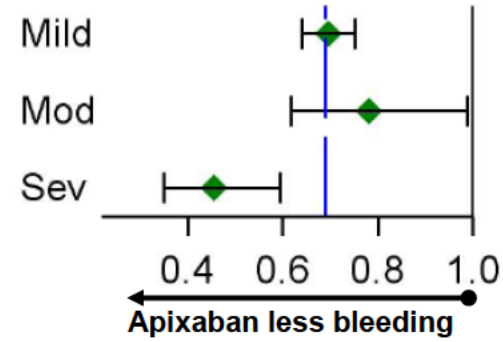


Number at risk

Warfarin	9052	8391	7995	7719	7470	6599	5452	4394	3221	2405	1632	1098	601
Apixaban	9088	8462	8171	7896	7682	6733	5597	4550	3343	2500	1671	1130	642

Reviewer's analysis (on treatment): analysis\erate_HR\erate HR runs bleed, \bleed\tte\tte gustoSEV, Applicant's datasets adbs2, adbl2. HR=apixaban/warfarin

Figure 21 ARISTOTLE - Forest plot of GUSTO bleeding



Type of GUSTO Bleed	HR, 95% CI	Apixaban n/N	Apixaban %/yr	Warfarin n/N	Warfarin %/yr
Mild		1043	7.18	1427	10.4
Mod		124	0.80	157	1.04
Sev		80	0.52	172	1.13

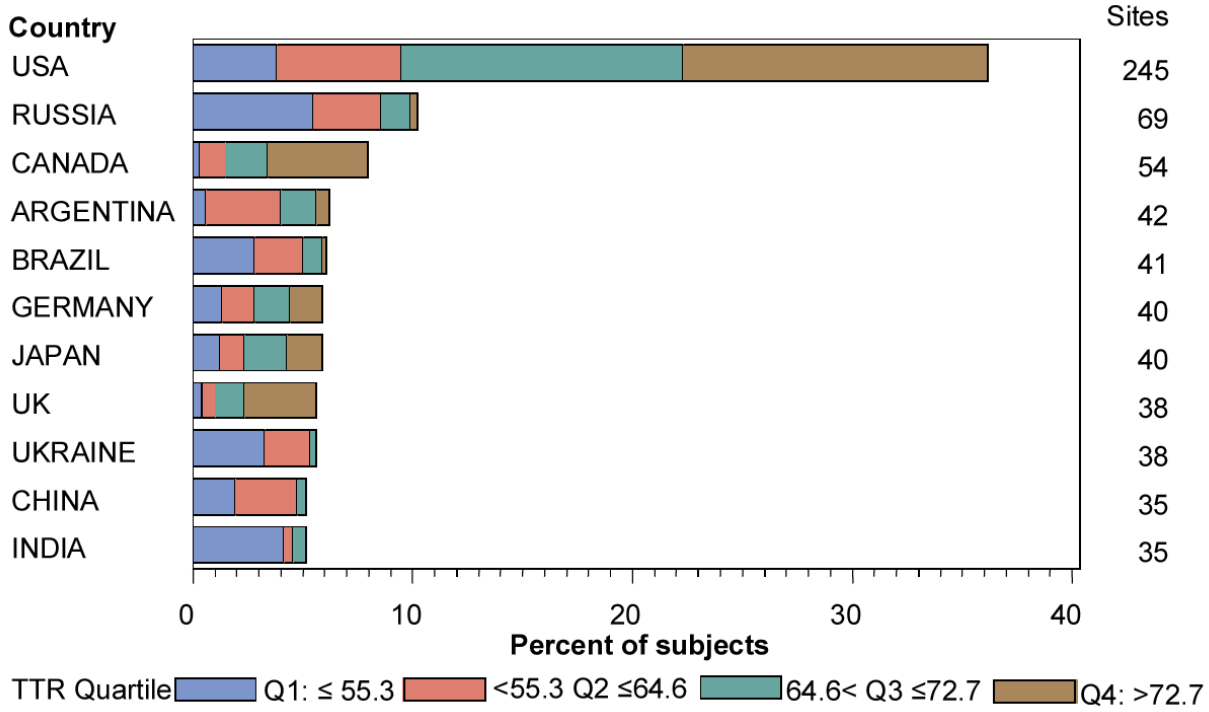
Reviewer's analysis: \bleed\sub\subgroup runs bleed, forestplot_GUSTO, Applicant's datasets adbl2. HR=apixaban/warfarin, n=number of events

7.3.2.1.4 Major bleeding by level of INR control

The reviewer conducted numerous analyses of all bleed definitions in ARISTOTLE and their relationship with time in therapeutic range (TTR), including analyses by quartile and by median for time below therapeutic range, time in therapeutic range, and time above therapeutic range (therapeutic range defined as an INR 2-3). All analyses excluded periods of warfarin interruption. In ARISTOTLE the overall median TTR was 66, the mean was 62 (reviewer's analysis (using individual's TTR): \inr\create quartile, Applicant's data: adinref3). These numbers are in agreement with the Applicant's. In the US the median TTR was 72, the mean was 70. For the reviewer's analyses the quartiles and medians were determined from the mean TTR at each site (not individual data as in the calculation of the overall trial TTR). This differs from the Applicant who determined site quartiles from individual data.

The US was the top enrolling country; both the US and Canada had a large number of sites with very high time in therapeutic range (Quartile 4), where TTR was > 72.7 % (US, 38%, Canada 57% of their sites). Russia was the second highest enrolling country, however more than half of their subjects had low TTR (Q1), suggesting that warfarin control was poor in these subjects.

Figure 22 ARISTOTLE – Quartile time in therapeutic range for top enrolling countries



Reviewer's analysis: inr\create quartile site in_below, Applicant's dataset adinref3. Sites=sites with INR data. Number of sites may be different from total sites per country because some sites may not have INR data.

One might logically hypothesize that subjects who spend more time in range would bleed less. Moreover, the relative advantage of apixaban over warfarin might decrease, such that the hazard ratio moves towards one since warfarin control presumably is better. However, part of the reason why these analyses are done by sites is because sites with good warfarin control might also be sights were quality of care is also better, thus there is an impact in subjects treated with apixaban also. And thus, there would be little change in relative risk. Quartile analysis of time in range did not show a consistent clear relationship between bleeding (not even for all serious bleeds) and time in range. The analysis that showed the most consistent trend was for GUSTO severe bleeding; as time in range improved, the annual event rate of bleeding decreased in both arms. Since both treatment arms had less bleeds as TTR improved, there was little change in relative risk. Some of the analyses are shown below. The site mean time in range was 63%. The site mean time below range was 24% and the mean time above range was 14%. The medians are evident in the tables that follow.

Table 89 ARISTOTLE - Time in therapeutic range bleeding quartile analysis

Type of Bleed	Q	Quartile range	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
			(n)	%/yr	(n)	%/yr	HR	95% CI
Major Bleed	1	≤ 55.3	55	1.55	98	2.95	0.52	(0.38, 0.73)
Major Bleed	2	> 55.3 and ≤ 64.6	93	1.95	138	2.95	0.67	(0.51, 0.87)
Major Bleed	3	> 64.6 and ≤ 72.7	118	2.86	130	3.14	0.91	(0.71, 1.17)
Major Bleed	4	> 72.7	60	2.14	96	3.42	0.63	(0.46, 0.87)
GUSTO Severe	1	≤ 55.3	23	0.64	47	1.40	0.46	(0.28, 0.76)
GUSTO Severe	2	> 55.3 and ≤ 64.6	27	0.56	57	1.20	0.47	(0.30, 0.75)
GUSTO Severe	3	> 64.6 and ≤ 72.7	20	0.47	47	1.12	0.42	(0.25, 0.71)
GUSTO Severe	4	> 72.7	9	0.32	21	0.74	0.42	(0.19, 0.93)
TIMI Major	1	≤ 55.3	32	0.90	57	1.70	0.53	(0.34, 0.81)
TIMI Major	2	> 55.3 and ≤ 64.6	39	0.81	80	1.69	0.49	(0.33, 0.71)
TIMI Major	3	> 64.6 and ≤ 72.7	45	1.07	74	1.77	0.61	(0.42, 0.88)
TIMI Major	4	> 72.7	31	1.10	45	1.58	0.70	(0.44, 1.10)
Less Serious Bleeds								
CRNM	1	≤ 55.3	50	1.41	104	3.16	0.45	(0.32, 0.63)
CRNM	2	> 55.3 and ≤ 64.6	86	1.82	112	2.40	0.76	(0.57, 1.00)
CRNM	3	> 64.6 and ≤ 72.7	105	2.55	133	3.25	0.78	(0.61, 1.01)
CRNM	4	> 72.7	75	2.69	95	3.44	0.78	(0.58, 1.06)
Minor Bleed	1	≤ 55.3	138	3.99	226	7.21	0.56	(0.45, 0.69)
Minor Bleed	2	> 55.3 and ≤ 64.6	191	4.14	298	6.69	0.62	(0.52, 0.75)
Minor Bleed	3	> 64.6 and ≤ 72.7	214	5.32	266	6.77	0.79	(0.66, 0.94)
Minor Bleed	4	> 72.7	154	5.70	199	7.53	0.75	(0.61, 0.92)

Reviewer's analysis: \quartile\erateHRquart runs bleed, Applicant's datasets: adinref3, adblsaf.

Additionally, one might hypothesize that subjects who spend more time above therapeutic range would bleed more. However, there was not a consistent relationship for these analyses either.

Table 90 Time above therapeutic range bleeding quartile analyses - ARISTOTLE

Type of Bleed	Q	Quartile range	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
			(n)	%/yr	(n)	%/yr	HR	95% CI
Major Bleed	1	≤ 8.8	63	1.90	82	2.46	0.76	(0.55, 1.06)
Major Bleed	2	> 8.8 and ≤ 12.4	75	1.79	118	2.86	0.63	(0.47, 0.84)
Major Bleed	3	> 12.4 and ≤ 16.9	80	1.80	156	3.57	0.51	(0.39, 0.66)
Major Bleed	4	> 16.9	108	3.29	106	3.38	0.99	(0.76, 1.30)
GUSTO Severe	1	≤ 8.8	17	0.51	35	1.04	0.49	(0.27, 0.87)
GUSTO Severe	2	> 8.8 and ≤ 12.4	19	0.45	37	0.89	0.51	(0.29, 0.89)
GUSTO Severe	3	> 12.4 and ≤ 16.9	18	0.40	61	1.37	0.29	(0.17, 0.49)
GUSTO Severe	4	> 16.9	25	0.75	39	1.22	0.62	(0.37, 1.02)
TIMI Major	1	≤ 8.8	27	0.81	49	1.46	0.55	(0.34, 0.88)
TIMI Major	2	> 8.8 and ≤ 12.4	43	1.02	57	1.37	0.75	(0.50, 1.11)
TIMI Major	3	> 12.4 and ≤ 16.9	30	0.67	94	2.13	0.31	(0.21, 0.47)
TIMI Major	4	> 16.9	47	1.41	56	1.76	0.83	(0.56, 1.22)

Reviewer's analysis: \quartile\erateHRquart runs bleed, Applicant's datasets: adinref3, adblsaf.

All in all these analyses did not show anything meaningful except to suggest that one should not heavily weight the "time in therapeutic range" calculation as a marker for good INR control and ultimately bleeding events.

7.3.2.1.5 Subgroup Analysis – Demographics

Apixaban was associated with less major bleeding relative to warfarin in most subgroups (**Table 91**). Women had less major bleeds on apixaban; both the absolute and relative difference was better than in women treated with warfarin as well as compared to males.[w] Analyses in various age subgroups show that older subjects bleed more; but the relative safety of apixaban over warfarin is preserved (and statistically significant). Subjects who had a previous stroke also had more bleeding than those who did not have a prior stroke. Relative to warfarin, the effect was similar to the HR in the overall trial, but the confidence intervals were wide, likely because of the small number of subjects. Nevertheless, the differences were not large. Diabetics tended to bleed more, but relative to warfarin there was no difference. In sum, apixaban

w This finding is somewhat contrary to the apixaban concentration data, which in a dedicated gender study (CV185022) showed women to have a 15-18% higher concentration than males, so one would expect more bleeding in women.

is superior to warfarin on major bleeding in most subgroups. In subgroups where the HR is similar to the overall HR but the sample size is small, superiority is lost.

Table 91 ARISTOTLE - Major bleed subgroup analysis

Subgroup	Apixaban		Warfarin		Apixaban vs. Warfarin	
	(n/N)	%/yr	(n/N)	%/yr	HR	95% CI
ISTH Major Bleed	**327/9088**	**2.13**	**462/9052**	**3.09**	**0.69**	**(0.60, 0.80)**
Gender						
Male	225 / 5868	2.26	294 / 5879	2.98	0.76	(0.64, 0.90)
Female	102 / 3220	1.91	168 / 3173	3.29	0.58	(0.45, 0.74)
< 72 years[1]	35 / 1579	1.27	55 / 1580	2.06	0.61	(0.40, 0.94)
≥ 72 years	67 / 1641	2.58	113 / 1593	4.64	0.56	(0.41, 0.76)
Little Old Lady (≥ 80y & ≤60kg)	7 / 189	2.50	11 / 186	4.17	0.63	(0.24, 1.62)
Female < 80yo & >60kg	95 / 3031	1.87	157 / 2987	3.25	0.58	(0.45, 0.75)
Age						
< 65 years	56 / 2723	1.17	72 / 2732	1.51	0.78	(0.55, 1.11)
65 to ≤ 75 years	120 / 3529	1.99	166 / 3501	2.82	0.71	(0.56, 0.89)
≥ 75 years	151 / 2836	3.33	224 / 2819	5.19	0.64	(0.52, 0.79)
CHADS2 Score						
1	75 / 3039	1.38	126 / 3018	2.38	0.58	(0.44, 0.77)
2	125 / 3246	2.30	163 / 3246	3.03	0.76	(0.60, 0.96)
3	68 / 1677	2.51	95 / 1588	3.91	0.65	(0.47, 0.88)
4	39 / 764	3.25	50 / 811	3.98	0.81	(0.53, 1.23)
5	18 / 271	4.36	24 / 288	5.52	0.78	(0.42, 1.44)
6	1 / 37	1.98	4 / 43	7.86	0.28	(0.03, 2.55)
Diabetes						
No	215 / 6812	1.85	348 / 6802	3.08	0.60	(0.51, 0.71)
Yes	112 / 2276	3.01	114 / 2250	3.13	0.96	(0.74, 1.25)

Table continued

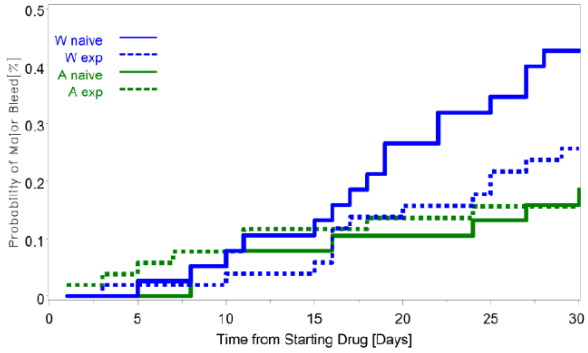
Previous Stroke						
No	272 / 8048	1.99	390 / 7973	2.94	0.68	(0.58, 0.79)
Yes	55 / 1040	3.35	72 / 1079	4.31	0.77	(0.54, 1.10)
Previous Stroke/TIA						
No	250 / 7401	1.98	356 / 7317	2.91	0.68	(0.58, 0.80)
Yes	77 / 1687	2.84	106 / 1735	3.91	0.73	(0.54, 0.98)
VKA[2]						
naive	142 / 3892	2.17	188 / 3872	2.96	0.73	(0.59, 0.91)
experienced	185 / 5196	2.11	274 / 5180	3.18	0.66	(0.55, 0.80)
Race						
White	275 / 7512	2.15	370 / 7469	2.96	0.73	(0.62, 0.85)
Asian	45 / 1302	2.14	81 / 1328	3.83	0.56	(0.39, 0.80)
Black/African Am	4 / 125	2.05	2 / 102	1.30	1.57	(0.29, 8.64)
Other	3 / 149	1.39	9 / 153	4.75	0.20	(0.04, 0.95)

N is the number of subjects in that group. Shaded subgroups indicate where the 95% upper CI crosses 1. The median age of females was 72 years. 2. Analysis did not stratify by VKA status. Reviewer's analysis, Cox proportional hazards model stratified by prior VKA status per the Applicant's SAP: \bleed\sub\subgroup runs bleed, Applicant's dataset: adbs2.

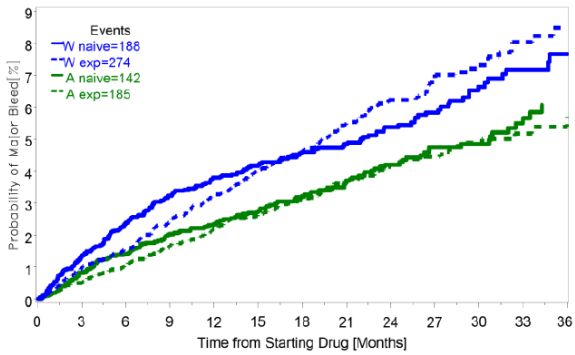
Warfarin-treated subjects bled more than apixaban treated subjects and VKA status did not matter (**Table 91**). One might presume that warfarin-treated subjects that are VKA experienced would have less bleeding than VKA naïve subjects. Indeed, major bleeds in warfarin-treated subjects that were VKA naïve start to separate from the other subjects around Day 15 (**Figure 23A**). However about half way through the trial, warfarin-treated subjects that were VKA experienced crossed over and had more bleeds than VKA naive subjects. The reviewer cannot offer an explanation for this.

Figure 23 ARISTOTLE – Time to first ISTH major bleed by VKA status

A. Major bleeding in first 30 days



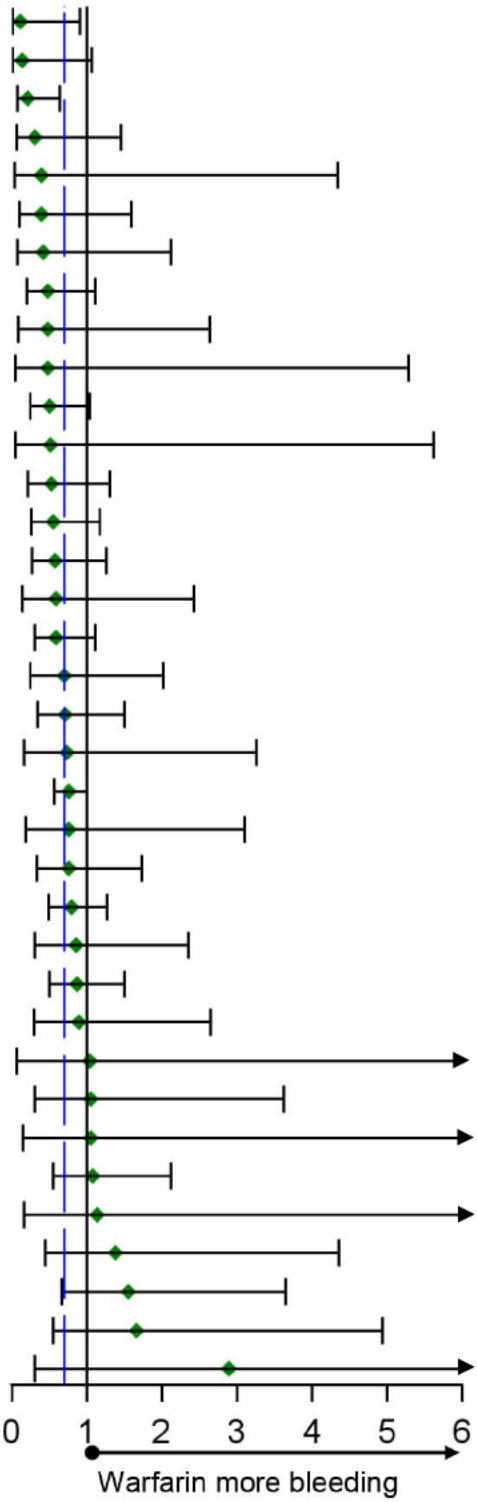
B. Major bleeding during entire trial



Reviewer's analysis: \bleed\tte\tte MB VKA, Applicant's dataset: adbs2, adbl2

Compared to the overall trial results, in the United States the rates of major bleeding on each treatment was higher (+0.7%) and the relative benefit of apixaban over warfarin was slightly less, HR 0.75 (95%CI, 0.56, 1.00), nominal p-value=0.0497.

Figure 24 ARISTOTLE – ISTH Major bleeds by country



Country	HR 95% CI	Apixiban (n/N)	Apixiban %/yr	Warfarin (n/N)	Warfarin %/yr
PHILIPPINES *		1 / 103	0.62	9 / 102	5.93
SOUTH KOREA		1 / 150	0.43	9 / 157	3.76
JAPAN *		4 / 160	1.26	18 / 175	5.99
PERU		2 / 103	1.52	7 / 110	5.10
NORWAY		1 / 47	1.27	2 / 43	3.17
COLOMBIA		3 / 54	3.35	6 / 57	6.99
CHILE		2 / 128	0.93	5 / 130	2.34
AUSTRALIA		8 / 166	2.62	16 / 156	5.86
TAIWAN		2 / 26	4.97	4 / 30	9.80
CZECH REPUBLIC		1 / 83	0.57	2 / 82	1.16
MEXICO		11 / 307	2.21	21 / 298	4.37
POLAND		1 / 156	0.45	2 / 156	0.90
UKRAINE		7 / 480	0.86	13 / 476	1.66
BRAZIL		11 / 353	1.88	18 / 345	3.39
INDIA		10 / 301	2.27	17 / 298	3.91
ROMANIA		3 / 137	1.28	5 / 136	2.12
RUSSIA		15 / 896	0.89	25 / 903	1.54
SPAIN		6 / 116	3.38	8 / 113	5.10
GERMANY		12 / 427	1.72	16 / 420	2.38
NETHERLANDS		3 / 154	1.17	4 / 154	1.65
USA *		83 / 1716	2.84	109 / 1693	3.77
SOUTH AFRICA		4 / 44	4.70	4 / 45	5.55
HUNGARY		10 / 227	2.68	13 / 228	3.61
ARGENTINA		31 / 785	2.34	37 / 774	3.00
BELGIUM		7 / 95	4.57	8 / 98	5.40
CANADA		23 / 528	2.66	28 / 526	3.08
ISRAEL		6 / 167	2.13	7 / 170	2.38
FRANCE		1 / 17	3.42	1 / 18	3.51
MALAYSIA		5 / 64	3.97	5 / 60	4.06
AUSTRIA		2 / 17	9.20	2 / 16	8.55
CHINA		17 / 420	2.62	16 / 420	2.38
ITALY		2 / 84	1.43	2 / 93	1.22
SWEDEN		7 / 110	3.46	5 / 105	2.54
DENMARK		13 / 169	4.41	9 / 170	2.93
UK		9 / 213	2.47	5 / 216	1.50
HONG KONG		3 / 38	4.37	1 / 38	1.38

Reviewer's analysis: \bleed\sub\subgroup runs bleed, Applicant's dataset: adbs2.

7.3.2.1.6 Apixaban 2.5 mg dose

The protocol specified that subjects "at higher risk for bleeding" randomized to apixaban were to receive apixaban 2.5 mg BID. This population was defined as having at least 2 of 3 criteria: age ≥ 80 years, weight ≤ 60 kg or serum creatinine ≥1.5 mg/dL. Subjects continued on this dose even if they no longer met criteria during the study.

The criteria for dose adjustment were not directly based on apixaban pharmacokinetics. Although apixaban PK are linear, the group that received apixaban 2.5 mg had ~25% lower exposure (not equivalent) than the group that received 5 mg (source: Clin Pharm review, Table 9, p. 185).[x] The characteristics of the subjects for each criteria are shown in **Table 92**. Less than 5% of subjects were assigned to apixaban 2.5 mg. Most subjects were aged ≥ 80 years old; most subjects fulfilled the lower dose criteria with age and weight and few subjects had all three criteria.

Table 92 ARISTOTLE - Criteria for subjects that received apixaban 2.5 mg

	Apixaban	Warfarin
Criteria	**n=424 (4.7%)**	**n=402 (4.4%)**
Age ≥ 80 years	389 (4.3%)	373 (4.1%)
Wt ≤ 60 kg	280 (3.1%)	256 (2.8%)
Scr ≥ 1.5 mg/dL	149 (1.6%)	144 (1.6%)
Age & weight	245 (2.7%)	228 (2.5%)
Age &Scr	122 (1.3%)	120 (1.3%)
Weight & Scr	38 (0.4%)	30 (0.3%)
Age, weight & Scr	11 (0.1%)	7 (0.1%)

Note: Two of 3 criteria (shaded lines in table) qualified for lower apixaban dose. Reviewer's analysis: \bleed\sub\low dose, Applicant's dataset: adbs2

It was questionable whether CrCL might be a better marker for the lower dose, since using Scr might exclude some subjects with poor renal function. However, renal elimination does not have a large role in the excretion of apixaban since its elimination is multimodal. While 90% of subjects in the 2.5 mg group had severe or moderate renal impairment (CrCL ≤ 50 mL/min), (the median baseline CrCl was 37 mL/min), only 25% of all subjects with moderate to severe renal impairment received the 2.5 mg dose. Thus, most (75%) subjects with moderate to severe renal impairment received the 5 mg dose, and major bleeding in subjects with renal impairment was less on apixaban relative to warfarin (see **Figure 26** in Renal Impairment Section). In the 5 mg group, 13% of subjects had moderate or severe renal impairment and the median baseline CrCl was 80mL/min. (source: Clin Pharm review, Figure 11, p. 184). Analysis of subjects with moderate or severe renal impairment by dose also shows an absolute and relative advantage for apixaban compared to warfarin for ISTH major bleed **(Table 93)**. This is in spite of apixaban median concentrations being ~40% higher in the 5 mg group with moderate to severe renal impairment.[y]

x Final Clinical Pharmacology review dated 2/15/2012 (Lai, Ju Ping, et al.)
y The median AUC in 5 mg group subjects in ARISTOTLE with concentration data was 4987 ng*hr/mL in

Table 93 ARISTOTLE - ISTH major bleed in subjects with moderate or severe renal impairment by dose group

	Median apixaban AUCss(ng*hr/mL)	Apixaban n/N (%/yr)	Warfarin n/N (%/yr)	Apixaban vs. warfarin HR (95%CI)
≥ 2 risk factors (2.5 mg dose group)	2746	15/382 (2.70)	35/347 (7.44)	0.37 (0.20-0.68)
< 2 risk factors (5 mg dose group)	4987	58/1111 (3.37)	107/1165 (6.16)	0.55 (0.40-0.76)

Source: Adapted from Clin Pharm review, Table 9, p. 184

On treatment analyses of the two doses show that apixaban 2.5 mg was safe (ISTH major bleed) and effective (stroke/se) (**Table 94**). While the rates of major bleeding were higher in the low dose arm compared to the 5 mg arm, the effect relative to warfarin was actually better compared to that in subjects treated with apixaban 5 mg. So it appears that the Applicant's criteria for subjects at risk for bleeding worked; these subjects had more bleeding on warfarin and using a lower dose of apixaban showed a greater relative difference (HR(95%CI): 0.50 (0.29, 0.86). The rates of stroke/SE were also higher in subjects at greater risk of bleeding, but apixaban was effective relative to warfarin. In sum, based on the data in ARISTOTLE, the lower dose of apixaban was safe and effective and should be approved for subjects at high risk for bleeding as defined in the trial.

moderate to severe renal impairment and 3603 ng*hr/mL in all 5 mg treated subjects (so includes all levels of renal impairment). A dedicated renal impairment study (CV185018) showed a 30-40% increase in apixaban concentration in subjects with moderate to severe renal impairment compared to subjects with normal renal function.

Table 94 ARISTOTLE – Safety and efficacy rates on apixaban 2.5 mg and 5 mg

Event	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
	(n/N)	%/yr	(n/N)	%/yr	HR	95% CI
Major bleed						
Apixaban 2.5 mg	20 / 424	3.29	37 / 402	6.71	0.50	(0.29, 0.86)
Apixaban 5 mg	307 / 8664	2.09	425 / 8650	2.95	0.71	(0.61, 0.82)
Stroke/SE						
Apixaban 2.5 mg	9 / 424	1.45	20 / 402	3.57	0.40	(0.18, 0.88)
Apixaban 5 mg	167 / 8664	1.12	205 / 8650	1.41	0.80	(0.65, 0.98)
Hemorrhagic stroke						
Apixaban 2.5 mg	2 / 424	0.32	6 / 402	1.06	0.31	(0.06, 1.53)
Apixaban 5 mg	36 / 8664	0.24	68 / 8650	0.46	0.52	(0.35, 0.78)
Ischemic stroke						
Apixaban 2.5 mg	10/428	1.42	14/403	2.11	0.65	(0.29, 1.47)
Apixaban 5 mg	152/8692	0.95	161/8678	1.01	0.94	(0.75, 1.17)
Death						
Apixaban 2.5 mg	73 / 428	10.0	77 / 403	11.2	0.91	(0.66, 1.26)
Apixaban 5 mg	530 / 8692	3.23	592 / 8678	3.63	0.89	(0.79, 1.00)

Reviewer's analysis (on treatment except for death and ischemic stroke which was ITT).

Additional data that supports approval of apixaban 2.5 mg comes from the Clinical Pharmacology Review. They used a Cox proportional hazards model with adjustments for covariates in each dose group to predict the risk of ISTH major bleed in one year in relation to concentration. The model predictions were close to the actual annual event rates in ARISTOTLE (**Table 95**). Using their ER model, the reviewers also predicted that if the group with at least 2 risk factors for bleeding had received 5 mg instead of 2.5 mg (thus increasing the median AUC by 60%), then these subjects would have an annual rate of major bleeding of 6.3%, almost double.

Table 95 ISTH major bleed - mean model predicted (95%CI) and ARISTOTLE annual event rate

	Predicted annual rate Apixaban (%)	Predicted annual rate Warfarin (%)	ARISTOTLE apixaban event rate (%)	ARISTOTLE warfarin event rate (%)
2.5 mg dose group (≥ 2 risk factors)	3.55 (2.26, 4.82)	7.16 (6.05, 8.26)	3.29	6.71
5 mg dose group	1.79 (1.41, 2.18)	2.87 (2.51, 3.22)	2.09	2.95
2.5 mg dose group treated with 5 mg dose	6.33 (4.43, 8.20)	---	---	---

Source: Adapted from Clinical Pharmacology review, Table 7, p 179

Subgroup analysis of subjects with high body weight (≥120 kg), another population with ~25% lower apixaban concentration, is supportive of the efficacy (**Table 96**). This suggests that the 25% lower apixaban exposure does not result in a loss of benefit.

Table 96 ARISTOTLE – Risk of stroke/se in 2.5 mg dose group and in 5 mg dose group by high body weight

	Median Apixaban AUCss (ng*hr/mL)	Apixaban n/N (% yr)	Warfarin n/N (% yr)	HR (95% CI)
2.5 mg	2703 (n = 128)	12/428 (1.70)	22/403 (3.33)	0.50 (0.20-1.02)
Weight ≥120kg in 5 mg	2690 (n = 179)	4/513 (0.40)	12/519 (1.19)	0.34 (0.11-1.06)
Weight <120kg in 5 mg	3662 (n = 2625)	196/8179 (1.30)	231/8159 (1.55)	0.84 (0.70-1.02)

Source: Clinical Pharmacology review, Table 8, p. 183
High body weight defined as ≥ 120 kg; ITT analysis

*Reviewer's comment: The reviewer is recommending that apixaban 2.5 mg BID be approved for patients at risk for bleeding as defined in ARISTOTLE. The same reviewer recommended not approving the lower dose of dabigatran (110 mg). The factors that led to the differences in decision on the lower dose of each drug are worth elucidating. In RE-LY, subjects were randomized equally to one of three treatment arms. Post hoc comparisons between the high dose versus the low dose were allowed and showed that the higher dose was more effective on stroke/SE than the lower dose (HR 0.72, 95%CI: 0.58, 0.90), albeit there was more bleeding (HR 1.16, 95%CI: 1.00, 1.34). Moreover, the lower dose was associated with more ischemic strokes than either dabigatran 150 mg or warfarin (See **Table 2**). Exploratory analyses of the lower dose did not find any subpopulation where benefit outweighed the risk. Perhaps the largest contrast between ARISTOTLE and RE-LY is that in ARISTOTLE, the study was designed such that a particular subgroup deemed to be at risk for bleeding received the lower dose. And in this subgroup, both safety and effectiveness were preserved and event rates were more favorable on apixaban 2.5 mg.*

7.3.2.2 Other Nonfatal Serious Adverse Events

The reviewer analyzed the SAE data at all MedDRA levels including system organ class (SOC), high level group term (HLGT), high level term (HLT), and preferred term (PT). At each level of analysis, the reviewer examined the data by frequency of occurrence in each arm and by risk difference between arms. The results were compared on treatment and from Day 3-30 days post dose (to look for a possible increase in embolic events post dose).

For the most part, the SAEs were similar between the two treatment arms at any MedDRA level of analysis, both on treatment and Day 3-30 post dose.[z] When there were notable differences, the magnitude of the difference was always very small (< 1%), so it is difficult to draw definitive conclusions on sparse events. The notable SAES are discussed.

The cardiac disorders SOC were 0.6% higher in the apixaban arm compared to the warfarin arm during treatment (12.2% vs. 11.6%, respectively). These consisted mostly of the PTs heart failure, MI, and unstable angina. Given that many of the investigator-reported MIs in the apixaban arm were adjudicated as "no event" and that the disparity with warfarin is 0.6%, the reviewer is not concerned that apixaban may cause serious cardiac disorders.

There were numerically more HLGT neurological disorders nec in subjects treated with apixaban than warfarin (120 (1.3%) vs. 84 (0.9%), respectively). A closer look at the PT associated with this HLGT show that the SAEs were mostly syncope, dizziness, and presyncope **(Table 97)**. The syncope may be worth noting in labeling.

Table 97 ARISTOTLE -Subjects with SAE of Neurological disorders HLGT

MedDRA Preferred Term	On Treatment				Day 3-30 Post Dose			
	Apixaban N=9088		Warfarin N=9052		Apixaban N=9088		Warfarin N=9052	
	n	%	n	%	n	%	n	%
Syncope	77	(0.8)	47	(0.5)	2	(0.0)	3	(0.0)
Dizziness	18	(0.2)	12	(0.1)	0	(0.0)	1	(0.0)
Presyncope	15	(0.2)	12	(0.1)	0	(0.0)	0	(0.0)
Paraesthesia	3	(0.0)	1	(0.0)	0	(0.0)	0	(0.0)
Ataxia	2	(0.0)	2	(0.0)	0	(0.0)	1	(0.0)

Reviewer's analysis \ae\sae\sae, Applicant's data set adae

There were more SAE ischemic stroke PTs in the period post dose (17 on apixaban, 8 on warfarin). See **Sec 6.1.10.2**. for a discussion of these.

7.3.3 Dropouts and/or Discontinuations

For the most part, there were no differences in reasons for treatment discontinuation (by SOC) between treatment arms. However in the trial as a whole, more subjects discontinued for an AE in the warfarin arm (8.4% vs. 7.6%, respectively). This was likely driven in part by a nearly three-fold higher number of warfarin-treated subjects with discontinuations for injury, poisoning, and procedural complications (63 (0.7%) vs. 22 (0.2%) subjects, respectively). The most common reason for treatment discontinuation was in the SOC of nervous system disorders (1.5% of subjects on apixaban vs. 1.7% on warfarin), consisting mostly of stroke/TIA events, followed by gastrointestinal disorders (1.3% of subjects on apixaban vs. 1.2% of subjects on warfarin) (Applicant's analysis, Table S.6.6B1 in ARISTOTLE CSR).

z. Data not shown. Reviewer's analysis \ae\sae\sae, Applicant's dataset adae

Reviewer's comment: This analysis may need to be repeated after the sponsor resubmits the AE data.

7.3.4 Significant Adverse Event

The Applicant discusses hepatic safety, neurological events, and non-traumatic fracture events in this section of their CSR. The first two are reviewed in **Sec 7.3.5.1** and **7.3.5.2**. The incidence of new non-traumatic fractures was 0.8% in both treatment groups during the treatment period. Fractures in the spine, rib, or hip were reported in ≤ 0.2% of subjects in both treatment groups (Applicant's analysis, Table S.6.9.A1 in ARISTOTLE CSR). In the follow-up period, there were 4 fractures in the apixaban group and none in the warfarin group.

Reviewer comment: The reviewer was unable to complete this section of the review according to the safety review template for reasons already described in **Sec 7.1.2***. Analyses outstanding include: severe AE, non-serious adverse events that led to an intervention (significant additional concomitant therapy or temporary drug discontinuation).*

7.3.5 Submission Specific Primary Safety Concerns

The submission specific primary safety concerns are bleeding, drug induced liver injury (DILI) and serious neurologic adverse events (other than stroke) such as Guillain-Barre syndrome (GBS) or amyotrophic lateral sclerosis (ALS).

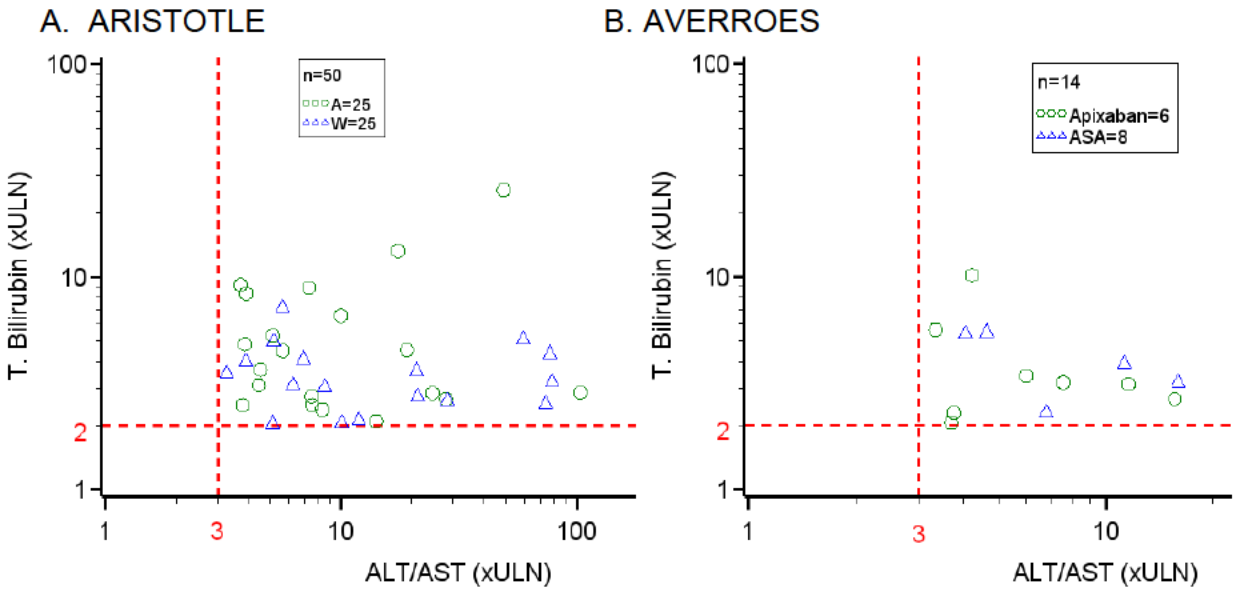
7.3.5.1 Drug induced liver injury (DILI)

Because ximelagatran, an oral direct thrombin inhibitor, was associated with hepatotoxicity, drug induced liver injury (DILI) has been the subject of intense monitoring and interim safety assessments in the development program of subsequent antithrombotics. However, neither dabigatran nor rivaroxaban were associated with DILI in trials with a warfarin comparator. Apixaban does not appear to be associated with DILI either.

Figure 25 of important liver tests that could indicate Hy's Law subjects shows that the number of subjects is the similar in both treatment arms for both ARISTOTLE and AVERROES.[aa] Since the active comparator is not associated with DILI, the distribution of liver tests does not heighten a cause for concern that apixaban may cause DILI.

aa A figure often referred to as an "E-DISH analysis" (maximum ALT and maximum total bilirubin per subject without respect to time) shows 30 subjects on apixaban and 25 subjects on warfarin in the upper right quadrant for ARISTOTLE (**Attachment 8 ARISTOTLE – Maximum ALT and T.Bili per Subject**). The Applicant's analysis agrees.

Figure 25 Potential Hy's Law Subjects
Max ALT/AST > 3x ULN, concurrent T. Bili > 2x ULN & AP < 2x ULN



Reviewer's analysis: alt_tb time, Applicant's dataset: adliver1. "Concurrent" defined as total bilirubin and alkaline phosphatase within 30 days after the ALT or AST. When both ALT and AST were greater than 3xULN with "concurrent" total bilirubin >2xULN and alkaline phosphatase < 2xULN, the ALT with associated total bilirubin were plotted. On treatment defined as ALT/AST within 30 days after drug discontinuation.

Three blinded, independent hepatologists (Drs. (b) (4) (b) (4)) reviewed cases with elevations of ALT>3xULN and total bilirubin>2xULN on the same date and/or preselected hepatic SAEs (jaundice, hepatitis, and hepatic failure).[bb] These hepatologists assessed the likelihood of DILI in 69 cases (**Table 98**). As shown in the table, there were no probable cases. The "possible" cases were often confounded by concomitant drugs known to cause elevations in LFTs (e.g., amiodarone, simvastatin), or disease states (e.g., heart failure, chronic hepatitis). Based on the hepatologists assessments of the Applicant's identified cases the likelihood that apixaban causes DILI is low.

bb. Drs. (b) (4) reviewed one case because Drs. (b) (4) were unblinded on 27Sep2011.

Table 98 ARISTOTLE - Summary of independent, blinded assessment of Applicant's identified hepatic cases

Categories of assessment	Apixaban	Warfarin
Probable	0	0
Possible	6	4
Unlikely	26	30
Inadequate information	1	2

Applicant's data set: adexthep

As mentioned earlier, **Sec 7.2.6**, the Applicant had the potential to miss a case of DILI because of the selection criteria. Of the 50 reviewer identified potential Hy's Law subjects in ARISTOTLE, the hepatologists reviewed 38 of these cases. The causality assessment of the 38 cases indicates that apixaban is unlikely to cause DILI.

Table 99 ARISTOTLE - Summary of independent, blinded assessment of reviewer's identified potential Hy's Law cases

Assessment	Apixaban	Warfarin
Probable	0	0
Possible	3	3
Unlikely	16	15
Inadequate information	1	0

Reviewer's analysis: alt_tb time, hep consult. Applicant's data set: adliver1, adexthep

Of the remaining unassessed cases, 5 were on apixaban and 7 were on warfarin. None had significant ALT elevations (they had significant AST elevations with concurrent (as defined by the reviewer) total bilirubin elevations). The reviewer judged the apixaban cases to be 4 unlikely, and 1 inadequate[cc] information.

There was one fatal liver failure case that occurred during the open label extension phase of study CV185048 (thus, he is not counted in the figures or tables in this review). He was a 92 year old man who was diagnosed with liver failure after 8 months of apixaban. He did not receive apixaban during the blinded phase of the study. In addition to AFib, he also had right-sided heart failure and moderate renal impairment. Three expert hepatopathologists, (Drs. (b) (4)) interpreted and agreed that his post-mortem liver biopsy showed "massive hepatic necrosis consistent with DILI and less consistent with ischemic liver injury".[dd] Approximately 6-7 months prior to hospitalization, the subject started tianeptine 12.5 mg bid, an antidepressant that has been associated with liver injury. He was also taking rilmenidine, a loop diuretic, budesonide, and tiotropium bromide. Drs. (b) (4) agreed that the causality assessment should be "possibly related". Given apixaban's clean track record to date and known liver

cc Subject CV185030-19-1912
dd There was disagreement as to the onset of the event; Dr. (b) (4) dated onset as within one month of death, Dr. (b) (4) stated the onset was possibly two months prior to hospitalization.

liability with tianeptine, they felt that tianeptine was the more likely cause of the subject's liver failure.

Subsequent to their consensus, Dr. (b) (4) discussed the case with two expert hepatologists who have published on tianeptine liver injury, Drs. (b) (4) . These hepatologists indicated that tianeptine is a mitochondrial toxin that can cause idiosyncratic hepatocellular injury (but is quite rare). Additionally, the liver injury from tianeptine is generally a mixed hepatocellular-cholestatic injury. The doctors were also unaware of acute liver failure associated with tianeptine. There was a subsequent teleconference between Drs. (b) (4) where Dr. (b) (4) felt that although the liver biopsy was most consistent with DILI, liver ischemia could not be excluded. However the massive nature of the injury was not typical for ischemia. The paucity of inflammation in the liver favored an acute dose-dependent toxic injury such as is typical for acetaminophen rather than a delayed idiosyncratic type injury that would be expected from a drug exposure for multiple months. There was no history of acetaminophen ingestion, but the family had disposed of all of the subject's medications after his death. The consensus remained that a role for apixaban in the liver injury of this subject remained "possible".

The Applicant's analysis of liver related AEs, SAEs (including deaths) and discontinuations (using liver SMQs) also do not indicate a signal for apixaban DILI. There were five SAEs with a fatal outcome in each arm, no liver transplants, 26 AE leading to discontinuation in the apixaban arm and 35 in the warfarin arm.

Elevations of various categories of liver tests were balanced between arms (reviewer's analysis \hep\alt_tb_time, sponsor's dataset adliver1, data not shown).

In addition to cases in ARISTOTLE, the external hepatologists reviewed cases in APPRAISE-2 (all judged unlikely).

In sum, the data in the application do not suggest that apixaban causes DILI.

7.3.5.2 Neurologic events

Also specific for apixaban, was the potential for serious neurological adverse events. This was an adverse event of interest because of 1 case of amyotrophic lateral sclerosis (ALS) and 1 case of Guillain-Barre Syndrome (GBS) in subjects taking apixaban 10 mg QD and 5 mg QD, respectively, in a Phase 2 study.[ee]

Following the two reports, the Applicant enhanced surveillance for neurological events in all Phase 3 studies by use of a supplemental Clinical Safety Plan (CSP) CRF (See Attachment 9, **p.260**). For ARISTOTLE, this was instituted in Amendment 2 (30 Jul

ee Phase 2 VTE prevention following total knee replacement (study CV185010), subjects CV185010-204-6 and CV185010-131-18

2007) Adverse events of special interest, about 7 months after the first subject was enrolled. All AE reports were to be compared to a list of MedDRA terms that are suggestive of possible neuropathies or other neurological events. If the AE matched any of the terms and lasted for at least 7 days or resulted in a neurology consult, a CSP CRF was to be filled out. This CRF collected additional information to aid in the work up of a potential serious neurologic AE. Neurological consultations were recommended for any SAE that matched the list of MedDRA terms.

Additionally, in the second quarter of 2010 (near study completion of January 31, 2011) the Applicant instituted external, blinded, independent neurologist assessments[ff] of any SAEs with PTs included in the MedDRA high level of terms of "acute polyneuropathies" (acute polyneuropathy, critical illness polyneuropathy, Guillain-Barre Syndrome, Miller Fisher syndrome, and polyneuropathy) and PT amyotrophic lateral. The table below highlights the subjects with diagnosed serious neurologic AEs discussed in the apixaban NDA and SUR. All were blindly assessed except for subject CV185068-279-3923 (critical illness polyneuropathy on apixaban); it is unclear why this subject did not have a blinded assessment. However, after reviewing the CIOMS and neurology consult, the reviewer believes that the event that started 15 days after drug discontinuation is unlikely to be drug related. All blinded consensus assessments deemed the SAE to be unlikely related to the drug. There was not a consensus assessment available for subject CV185010-131-18, one of the subjects that heightened the surveillance for the neurology CSP in the Phase 3 program. The individual assessments were 2 unlikely and 1 possible for this subject.

ff The external neurologists were Drs. (b) (4), who are recognized experts in peripheral nerve disease

Table 100 Subjects with serious neurologic AE in apixaban NDA and SUR

	total	apixaban	Comparator	Blinded
Guillain-Barre Syndrome	6	2 CV185010-204-6 CV185036-43-4645	3 CV185030-1106-20906 CV185030-1534-7633 CV185068-452-6719	1 CV185056-575-890
Amyotrophic Lateral Sclerosis	3	1 CV185010-131-18	1 CV185030-1295-3113	1 CV185057-136-308
Other acute polyneuropathies	7	1 CV185068-279-3923*	5 CV185030-135-195 CV185030-543-2199 CV185030-1301-10540 CV185036-472-5774 CV185068-722-1826	1 CV185057-765-1087

*not assessed by blinded, independent, expert neurologists
AFIB studies: CV185030, CV185048
Non-AFib studies: CV185010 (apixaban 10 mg QD or 5 mg QD), CV185036 (apixaban 2.5 mg BID or enoxaparin 40 mg QD), CV185068 (placebo controlled),
Non-AFib, blinded studies: CV185056 (apixaban 5 mg BID vs. W), CV185057 (apixaban 2.5 mg BID, 5 mg BID or placebo)

In ARISTOTLE, serious neurologic AEs occurred infrequently and were balanced in each group (29 subjects in each arm). Six subjects in the apixaban arm and 8 subjects in the warfarin arm had neurological events that led to treatment discontinuation. The AEs that led to discontinuation in the apixaban arm included paresthesia, burning sensation, peripheral neuropathy, dysarthria, and hypoesthesia. (sponsor's analysis, Table S.6.7.2D1)

Based on the totality of the data, the reviewer believes that apixaban is unlikely to cause serious neurologic AE such as GBS, ALS or acute polyneuropathy. The Applicant appears to have done a good job identifying possible cases.

7.4 Supportive Safety Results

7.4.1 Common Adverse Events

The most common AE was bleeding and it occurred less frequently in the apixaban arm than in the warfarin arm. The annual event rate and relative risk of bleeding for less serious definitions are shown in the next table.

Table 101 ARISTOTLE - Less serious definitions of bleeding

Event	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
	(n)	%/yr	(n)	%/yr	HR	95% CI
Any Bleed	2356	18.1	3060	25.8	0.71	(0.68, 0.75)
Minor Bleed	701	4.72	989	6.98	0.68	(0.62, 0.75)
CRNM	318	2.08	444	3.00	0.70	(0.60, 0.80)
ISTH Major Bleed and CRNM	613	4.07	877	6.01	0.68	(0.61, 0.75)

Reviewer's analysis: erate_HR\erateHR runs bleed, Applicant's dataset adbs2, adefl. HR, n=number of subjects (first event).

Review of the sponsor's summary of common adverse events with onset during the treatment period indicate that there were no SOC or PT adverse events that occurred more frequently in the apixaban arm than in the warfarin arm **Table 102** (Source: Table 8.8 of ARISTOTLE CSR) . A total of 81.5% of subjects on apixaban and 83.1% of subjects on warfarin had an AE during the treatment period. The Applicant's analysis shows that the frequency of subjects with AEs during the treatment period within each SOC and PT was similar to warfarin.

Reviewer's comment: This analysis should be redone after the Applicant cleans up their AE dataset. Based on the Applicant's analyses, the reviewer recommends including only information on bleeding in the common AE section of labeling.

Table 102 ARISTOTLE - Applicant's summary of most common AE with onset during treatment period

System Organ Class (SOC) (%) Preferred Term (PT) (%)	Apixaban N = 9088	Warfarin N = 9052
TOTAL SUBJECTS WITH AN EVENT	7406 (81.5)	7521 (83.1)
INFECTIONS AND INFESTATIONS	3416 (37.6)	3495 (38.6)
NASOPHARYNGITIS	763 (8.4)	779 (8.6)
URINARY TRACT INFECTION	512 (5.6)	532 (5.9)
BRONCHITIS	503 (5.5)	516 (5.7)
UPPER RESPIRATORY TRACT INFECTION	436 (4.8)	456 (5.0)
GASTROINTESTINAL DISORDERS	2471 (27.2)	2641 (29.2)
DIARRHOEA	585 (6.4)	584 (6.5)
RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS	2115 (23.3)	2245 (24.8)
DYSPNOEA	605 (6.7)	649 (7.2)
EPISTAXIS	560 (6.2)	685 (7.6)
COUGH	495 (5.4)	505 (5.6)
CARDIAC DISORDERS	2055 (22.6)	1986 (21.9)
ATRIAL FIBRILLATION	496 (5.5)	473 (5.2)
CARDIAC FAILURE	481 (5.3)	453 (5.0)
MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS	1984 (21.8)	2071 (22.9)
ARTHRALGIA	447 (4.9)	463 (5.1)
BACK PAIN	433 (4.8)	506 (5.6)
NERVOUS SYSTEM DISORDERS	1972 (21.7)	2044 (22.6)
DIZZINESS	663 (7.3)	709 (7.8)
HEADACHE	482 (5.3)	485 (5.4)
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS	1912 (21.0)	1912 (21.1)
OEDEMA PERIPHERAL	611 (6.7)	663 (7.3)
INJURY, POISONING AND PROCEDURAL COMPLICATIONS	1540 (16.9)	1839 (20.3)
CONTUSION	301 (3.3)	482 (5.3)

Source: ARISTOTLE CSR, Table 8.8, page 202-203

7.4.2 Laboratory Findings

The laboratory findings of marked abnormalities were similar between apixaban and warfarin (source: ARISTOTLE CSR, Table 8.9.1). Pooled results for platelet decreases were similar in the apixaban and comparator arms.

Table 103 ARISTOTLE and AVERROES -Decrease in platelets during the treatment period

Decrease in platelets	Apixaban N=10,653	Comparator N=10,538
<100,000/mm^3	118	111
<50,000/mm^3	5	4

[Source: Applicant's Table 3.2, Summary of Clinical Safety], N is number of treated subjects with platelet laboratory results during treatment period

In ARISTOTLE, five subjects on apixaban and seven subjects on warfarin had an SAE of thrombocytopenia during the treatment period.

7.4.3 Vital Signs

Blood pressure and heart rate were assessed at baseline, Month 12, 24 and 36/end of treatment. There was no difference in VS between apixaban and warfarin.

7.4.4 Electrocardiograms (ECGs)

No clinically relevant differences between treatment groups were observed in ECG changes over time in either of the two Phase 3 studies. See Section 7.4.5

7.4.5 Special Safety Studies/Clinical Trials

The FDA QT Inter-Disciplinary Review Team reviewed the Thorough QT study, CV185031 (blinded, placebo-controlled, positive control) and found no effect on the QTc interval at concentrations up to 1000 ng/mL (50 mg QD) (review dated 7/1/2008). The upper bound of the 90% CI for ΔΔ QTc Fridericia was 4.3 ms and 4.6 ms for apixaban 10 mg QD and 50 mg QD, respectively. Apixaban 50 mg QD covers exposures 3-fold higher than 10 mg QD.

7.4.6 Immunogenicity

Not applicable

7.5 Other Safety Explorations

7.5.1 Dose Dependency for Adverse Events

Less than 5% of subjects received apixaban 2.5 mg and a specific "high risk for bleeding population" received that dose, so it is difficult to draw conclusions on dose dependency of adverse events.

7.5.2 Time Dependency for Adverse Events

This was explored for the primary safety concerns (major bleeding, DILI, and serious neurologic events) and is discussed, if relevant, in those respective sections.

7.5.3 Drug-Demographic Interactions

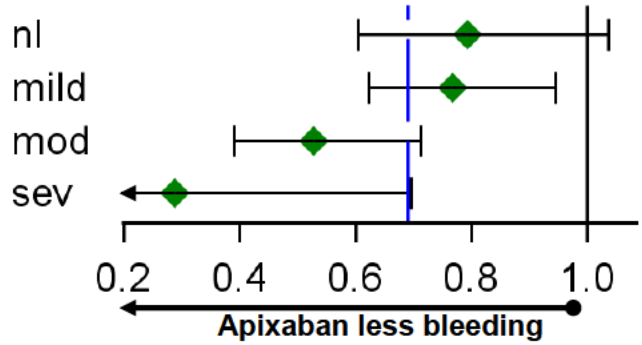
See **Sec 7.3.2.1.5 Subgroup Analysis – Demographics.**

7.5.4 Drug-Disease Interactions

Renal elimination does not play a large role in the excretion of apixaban since its elimination is multimodal. So one would not expect a large effect of renal impairment on PK, and certainly there is not one. However, it is known that subjects with renal impairment are inherently at risk for more adverse events, including bleeds and strokes. As stated earlier in **Sec 7.3.2.1.6**, most subjects with moderate to severe renal

impairment on apixaban received the lower dose. Apixaban had less major bleeds, yet was still effective in all levels of renal impairment (**Figure 26** and **Figure 27**).

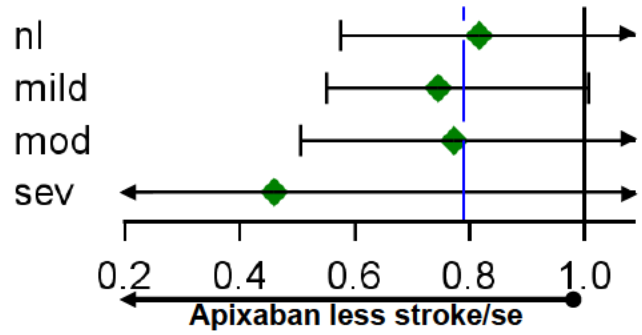
Figure 26 ARISTOTLE – ISTH Major bleed by level of renal impairment



Renal Impairment	Apixaban n/N	Apixaban %/yr	Warfarin n/N	Warfarin %/yr
nl	96 / 3750	1.46	119 / 3746	1.84
mild	157 / 3807	2.45	199 / 3758	3.21
mod	66 / 1357	3.16	123 / 1380	6.01
sev	7 / 136	3.75	19 / 132	11.9

Reviewer's analysis: \bleed\sub\subgroup runs bleed, forestplot_renal bsl rate, Applicant's datasets adbl2. HR=apixaban/warfarin, n=number of events, N=number in each group, CrCL (mL/min): normal > 80, mild 50 to ≤ 80, moderate 30 to ≤ 50, severe ≤ 30.
blue dashed line indicates HR for ISTH major bleed in ARISTOTLE.

Figure 27 ARISTOTLE – Stroke/SE by level of renal impairment



Renal Impairment	Apixaban n/N	Apixaban %/yr	Warfarin n/N	Warfarin %/yr
nl	57 / 3750	0.86	68 / 3746	1.04
mild	75 / 3807	1.16	98 / 3758	1.56
mod	38 / 1357	1.80	49 / 1380	2.34
sev	5 / 136	2.64	9 / 132	5.48

Reviewer's on treatment analysis: \eff\stse renal, forestplot_renal bsl rate stse, Applicant's datasets adbl2. HR=apixaban/warfarin, n=number of events, N=number in each group, CrCL (mL/min): normal > 80, mild 50 to ≤ 80, moderate 30 to ≤ 50, severe ≤ 30.
blue dashed line indicates HR for Stroke/SE in ARISTOTLE.

See **Pharmacokinetics** for more discussion of renal impairment and hepatic impairment.

7.5.5 Drug-Drug Interactions

There is a discussion of specific drug interaction studies in **Sec 4.4.3.**

The section focuses on apixaban use with other drugs that cause bleeding. The next table shows the absolute and relative risk of ISTH major bleed for aspirin or clopidogrel use at baseline in ARISTOTLE. This is, of course, no indication of whether subjects were taking either drug during the trial.

Table 104 ARISTOTLE – ISTH major bleed by baseline ASA or clopidogrel use

	Apixaban N=9088		Warfarin N=9052		Apixaban vs. Warfarin	
Event	(n)	%/yr	(n)	%/yr	HR	95% CI
Aspirin at baseline						
No	198 / 6242	1.87	298 / 6290	2.84	0.66	(0.55, 0.79)
Yes	129 / 2846	2.73	164 / 2762	3.68	0.75	(0.60, 0.95)
Clopidogrel at baseline						
No	316 / 8919	2.10	448 / 8884	3.05	0.69	(0.60, 0.80)
Yes	11 / 169	4.13	14 / 168	5.43	0.78	(0.35, 1.71)

Reviewer's analysis, Cox proportional hazards model stratified by prior VKA status: \bleed\sub\subgroup runs bleed, Applicant's dataset: adbs2.

Thus, the reviewer analyzed the bleeding data from APPRAISE-2 to assess the risk of bleeding in subjects taking apixaban with single or dual antiplatelet therapy. The reader should keep in mind that this was an ACS population and the trial was stopped early because the bleeding risk outweighed the benefit. The relevant concomitant medications in APPRAISE-2 are shown in the next table. Most subjects were on dual antiplatelet drugs, likely aspirin and clopidogrel.

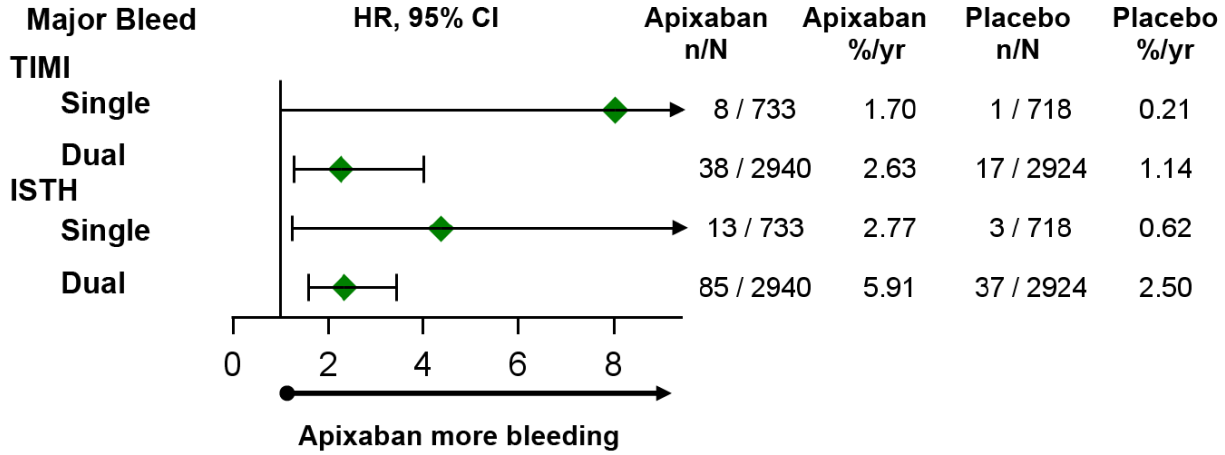
Table 105 APPRAISE2 - Concomitant medications

Medication class (%) Generic Name (%)	Placebo N = 3643	Apixaban N = 3672	Total N = 7315
TOTAL SUBJECTS USING MEDICATION	3643 (100.0)	3671 (100.0)	7314 (100.0)
ORAL ANTICOAGULANTS	36 (1.0)	24 (0.7)	60 (0.8)
ORAL ANTIPLATELETS	3613 (99.2)	3633 (98.9)	7246 (99.1)
ACETYLSALICYLIC ACID	3520 (96.6)	3558 (96.9)	7078 (96.8)
CLOPIDOGREL	2938 (80.6)	2933 (79.9)	5871 (80.3)
CLOPIDOGREL SULFATE	7 (0.2)	12 (0.3)	19 (0.3)
PRASUGREL	91 (2.5)	90 (2.5)	181 (2.5)
TICLOPIDINE	14 (0.4)	19 (0.5)	33 (0.5)
TICLOPIDINE HYDROCHLORIDE	0	1 (<0.1)	1 (<0.1)

Source: APPRAISE-2 CSR, Table 6.3, p 80.

The data from APPRAISE-2 show that apixaban with an antiplatelet increases the risk of bleeding significantly (**Figure 28**). For apixaban with single antiplatelet therapy, the annual rate of TIMI major bleeding is ~1.7 times greater than the rate of TIMI major bleeding in ARISTOTLE (0.96%/year, see **Table 88**), where apixaban was used alone. The TIMI major bleeding rate with dual antiplatelet therapy is 2.6 times greater than in ARISTOTLE. For ISTH major bleeding, the rate of bleeding on apixaban plus single antiplatelet therapy is almost 3 times greater and with dual antiplatelet therapy is 6 times greater than in ARISTOTLE where apixaban was used alone.

Figure 28 APPRAISE-2 - Risk of bleeding on apixaban and antiplatelet



Reviewer's analysis Cox proportional hazards with stratification by type of antiplatelet therapy: \appraise2\runs bleed, forestplot, Applicant's dataset adbs. HR=apixaban/placebo, n=number of events, N=number in each group, TIMI major bleed was the primary safety endpoint.

Reviewer's comment: The patient population in APPRAISE-2 and ARISTOTLE was different. But there certainly may be subjects with AFib that also have ischemic heart disease, and therefore may need both an anticoagulant and an antiplatelet drug. One would have to weigh the risk versus the benefit. Interestingly, the rate of bleeding (TIMI major and ISTH major) on apixaban plus single antiplatelet therapy in APPRAISE-2 was similar to the rate of bleeding on warfarin in ARISTOTLE. So in patients that need an anticoagulant and an antiplatelet it may be reasonable to use apixaban with a single antiplatelet. The rate of bleeding with apixaban and dual antiplatelet drugs is very worrisome; that coupled with the unknown risk:benefit ratio makes the reviewer even less enthusiastic about their combined use in AFib. The reviewer thinks extreme caution should be used when contemplating the combined use of apixaban with dual antiplatelet therapy. Stronger consideration should be given to the use of apixaban with a single antiplatelet treatment (as opposed to dual antiplatelet treatment) because of the higher risk of bleeding. The Applicant recommends that "a careful assessment of the potential benefits against the potential risks be made before combining" apixaban with any antiplatelet (single or dual).

7.6 Additional Safety Evaluations

7.6.1 Human Carcinogenicity

Reviewer attempts to get an accurate assessment of types of cancer were difficult due to the problem with the AE dataset identified in Sec (events on same date were listed as both serious and non-serious. Of the few cases checked, it appears that the higher MedDRA hierarchy terms (HLT and HLGT) were correctly mapped twice, so if one were only interested in cancers that were SAEs then an accurate account might be possible).

A gross look of cancers just by subject shows no imbalance suggestive of apixaban being carcinogenic. Note that the cancer could have been a reoccurrence. The table merely shows the number of subjects that had an AE of cancer in ARISTOTLE.

Table 106 ARISTOTLE -Subjects with cancer in adverse event dataset

	treatment	3-30 d PST	> 30 d PST	> 60 d PST
Apixaban	474	35	12	30
warfarin	609	43	16	16

Reviewer analysis: \cancer\can, sponsor dataset adae

7.6.2 Human Reproduction and Pregnancy Data

Apixaban has not been studied in pregnant or lactating women. Studies in animals suggest that apixaban does not affect fertility or fetal development; an appreciable amount was excreted in the milk.

The summary of safety reports 1 pregnancy in a subject (CV185030-518-4629) who received apixaban 5 mg BID in ARISTOTLE. She was a 32 year old Caucasian with a relevant history of using an intra-uterine contraceptive device, positive pregnancy test on Day 119, apixaban discontinuation on Day 113. Her child was reported as a normal newborn delivered by cesarean section. There were no other reported pregnancies in Phase 2/3 AFib studies.

In short-term studies for VTE prevention or treatment there was one pregnancy reported in a subject (CV185061-1-2) after one day of apixaban. The subject discontinued from the study and the investigator reported that her pregnancy resulted in a live, normal birth at 40 weeks gestation. There was a pregnancy reported in the female partner of a 48 year old male subject receiving apixaban in CV185017. Her pregnancy outcome was not reported.

In ongoing blinded VTE prevention Phase 3 studies, there were a total of 8 reported pregnancies. The outcomes were 2 normal newborns, 3 induced abortions, 2 spontaneous abortions, and 1 unknown.

7.6.3 Pediatrics and Assessment of Effects on Growth

Not done.

7.6.4 Overdose, Drug Abuse Potential, Withdrawal and Rebound

In ARISTOTLE, there were 38 cases of overdose reported in 35 subjects on apixaban; overdose was either from study drug or concomitant nonstudy medications. Four of these subjects had overdoses during the 30 day post-treatment period. Overdose was reported as an SAE in 28 cases. Only 2 events led to treatment discontinuation. In AVERROES, there were 5 cases of overdose reported in 4 subjects in the apixaban group. Four of these cases were SAEs and one led to treatment discontinuation. All events resolved. [Source: Applicant's Summary of Clinical Safety, p.196]

There is no pharmacological treatment to reverse the effects of apixaban. Activated charcoal given up to 3 hours, in dog studies, increased the clearance of apixaban, so activated charcoal may be considered. The Applicant recommends that in the event of hemorrhagic complications, treatment be discontinued, the source of bleeding investigated, and appropriate supportive treatment be initiated (e.g. surgical hemostats or the transfusion of FFP).

There was no evidence suggesting drug abuse/dependence on apixaban.

There is an increase in thrombotic events following cessation of apixaban; strokes/se as well as major bleeds were higher in the apixaban arm than in the warfarin arm. Since most of the events occurred 14 days after apixaban discontinuation, the reviewer thought that this might be the inability to properly transition subjects onto warfarin and achieve therapeutic INR range. However, this hypothesis does not hold given the data in AVERROES. It is unclear why there were more bleeding events after apixaban cessation. Rivaroxaban, another Factor Xa inhibitor was associated with an increase in thrombotic events after drug discontinuation; the label has a black box warning. We are recommending a similar black box warning be included in apixaban's label, if it were to be approved.

7.7 Additional Submissions / Safety Issues

The safety profile for apixaban is consistent between the SUR (Safety Update Report) and the SCS (Summary of Clinical Safety).

7.7.1 Data regarding adjudication of major bleed and MI

The reviewer analyzed the concordance between the investigator's assessment of the adverse event, major bleed and MI, and the final adjudication in attempts to identify any potential bias favoring apixaban. Since the investigator did not assess the type of bleed, the reviewer asked the Applicant to program the type (minor, CRNM, and major) of investigator reported bleed by using the ISTH definition and CRF S4 (Suspected Bleeding Endpoints page), filled out by the investigator. The reviewer compared the events that were "downgraded" and "upgraded" between both arms.

There appears to be reasonable adjudication (concurrence, downgrade, and upgrade) between investigator reported bleeds and adjudicated bleeds in both arms. It does not appear that there were more downgrades in the apixaban arm compared to the warfarin arm or that there were considerably more investigator reported bleeds in the warfarin arm that were adjudicated "down" compared to the apixaban arm.

Table 107 ARISTOTLE - Concurrence between investigator reported bleed and adjudication – apixaban arm

Final adjudication	Inv reported minor	Inv reported CRNM	Inv reported Major	Inv rep Total
Not adjudicated[1,2,3]	2886	1	2	2889
Adj not clinically relevant	181	29	42	252
Adj Minor	676	199	35	910
Adj CRNM	206	191	11	408
Adj Major	115	36	241	392
Adj hemorrhagic stroke	2	0	11	13
Adj Ischemic stroke[4]	0	0	2	2

Reviewer's analysis: analysis\adjud\concurrence, sponsor's data adadj2 (submitted 12/27/2011) Concurrence highlighted in yellow. Inv=investigator, CRNM=clinically relevant non major, rep=reported, Adj=adjudicated, 1.Occured prior to randomization, 2.Investigator reported in error, 3. Did not meet trigger criteria, clinically overt but not clinically relevant, same event as one already adjudicated, 4. Ischemic with hemorrhagic conversion

Table 108 ARISTOTLE - Concurrence between investigator reported bleed and adjudication – warfarin arm

Final adjudication	Inv reported minor	Inv reported CRNM	Inv reported Major	Inv rep Total
Not adjudicated[1,2,3]	3957	2	0	3959
Adj not clinically relevant	221	34	40	295
Adj Minor	942	328	66	1336
Adj CRNM	260	264	24	548
Adj ISTH Major	128	46	290	464
Adj hemorrhagic stroke	0	0	30	30
Adj Ischemic stroke[4]	0	0	2	2

Reviewer's analysis: analysis\adjud\concurrence, sponsor's data adadj2 (12/23/2011)
Concurrence highlighted in yellow. Adj= adjudicated, Inv=investigator, CRNM=clinically relevant non major, rep=reported, Adj=adjudicated, 1.Occured prior to randomization, 2.Investigator reported in error, 3. Did not meet trigger criteria, clinically overt but not clinically relevant, same event as one already adjudicated, 4. Ischemic with hemorrhagic conversion

The adjudication of myocardial infarction tended to favor apixaban; there were more MIs reported in the apixaban arm, however more MIs were adjudicated as "no MI" compared to warfarin. This happened so much so that there were numerically more adjudicated MIs in the warfarin arm.

The adjudication packages appeared to be reasonably cleaned to maintain blindness.

Table 109 ARISTOTLE – Discordance in MI adjudication

MI Event	Apixaban	Warfarin
Investigator reported MI	410	379
Adjudicated as MI	100	118
Adjudicated as no event	323	270
Not adjudicated[1,2]	4	4

Reviewer's analysis: analysis\adjud\concurrence, sponsor's data adadj2 (12/23/2011)
1.occured prior to randomization, 2.Investigator reported in error

8 Postmarketing Experience

As stated in Sec 2.6.1, a lower dose (and shorter duration) of apixaban is available in Europe. Because the recommended dose is only 2.5 mg BID and the recommended longest duration of treatment is short (only 32 to 38 days after hip surgery) relative to that in AF, postmarketing experience was not evaluated other than the information contained in the PSUR.

9 Appendices

9.1 Literature Review/References

See p. 239 for reference list.

9.2 Labeling Recommendations

The clinical reviewers recommend a Complete Response. In the event that this NDA is reviewed without the additional information requested suggested by the reviewers (see Sec.1.1) we recommend that labeling not include numerical c/ information regarding the comparisons of apixaban to warfarin for key safety and efficacy parameters. The reason for this is that given the poor execution of the trial, we think the nominal data provided by the sponsor may be misleadingly optimistic about the benefits of apixaban relative to warfarin. However the data are adequate to determine that apixaban is not worse than warfarin for the primary efficacy endpoint, death, or ISTH major bleeding. .

The following highlights the major discrepancies between our recommendation for labeling and the Applicant's. Substantial new text is underlined.

1. There should be a boxed warning about the risk of stroke after drug discontinuation, similar to the one in rivaroxaban labeling.

2. **Sec 1. Indications and Usage**
ELIQUIS (apixaban) is indicated to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation.

Reviewer's comment: The Applicant (b) (4)

3. **Sec 5. Warnings and Precautions**
5.1. Increased risk of stroke after discontinuation in nonvalvular atrial fibrillation
Discontinuing ELIQUIS in the absence of adequate alternative anticoagulation increases the risk of thrombotic events. An increased rate of stroke was observed during the transition from ELIQUIS (b) (4) to warfarin in clinical trials in atrial fibrillation patients. If ELIQUIS must be discontinued for a reason other than pathological bleeding, consider administering another anticoagulant [see Dosage and Administration (2.1) and Clinical Studies (14.1)]

Reviewer's comment: We are generally in agreement with the Applicant's other proposed text for this section.

4. **Sec 6. Adverse Reactions**

6.1. Clinical Trial Experience

Because clinical trials are conducted under widely varying conditions, adverse reaction rates observed in the clinical trials of a drug cannot be directly compared to rates in the clinical trials of another drug and may not reflect the rates observed in practice.

The safety of ELIQUIS was evaluated in the ARISTOTLE and AVERROES studies, including 11,284 patients exposed to ELIQUIS 5 mg twice daily and 602 patients exposed to ELIQUIS 2.5 mg twice daily [*see Clinical Studies (14)*]. The duration of ELIQUIS exposure was >12 months for 9375 patients and >24 months for 3369 patients in the two studies. In ARISTOTLE, the mean duration of exposure was 89 weeks (b) (4) In AVERROES, the mean duration of exposure was approximately 59 weeks (b) (4) (b) (4) The rate of ISTH major intraocular bleeding was higher with apixaban (0.21%/year) compared to warfarin (0.14%/year).

There was numerically more serious syncope in the apixaban arm (n=77) than in the warfarin arm (n=47).

Reviewer's comment: This section should not include any details of the trial that might imply superiority of apixaban over warfarin for major bleeding. Additionally, the adverse event dataset contains errors in over reporting of events (unique event reported as both serious and non-serious, see **Sec 7.1.2** for details*). This error appeared to be of the type where the event was an SAE, yet it was reported as both. So while the reviewer has more confidence in the SAE data, the reviewer was unable to complete analyses of*

the AE data. Analyses of the submitted data do not indicate an imbalance in AEs, but the reviewer does not have confidence in the accuracy of the data.

5. **Sec 7. Drug Interactions**

7.1. Anticoagulants and thrombolytic agents

Coadministration of ELIQUIS is not recommended with other anticoagulants (e.g., warfarin, heparin, rivaroxaban, dabigatran etexilate mesylate) or thrombolytic agents because of the increased risk of bleeding.

7.2 Antiplatelet Agents

The concomitant use of ELIQUIS with antiplatelet agents increases the risk of bleeding: cyclooxygenase inhibitors (aspirin), adenosine diphosphate (ADP) receptor inhibitors (clopidogrel, prasugrel, ticagrelor, ticlopidine), phosphodiesterase inhibitors (cilostazol), glycoprotein IIB/IIIA inhibitors (abciximab, eptifibatide, tirofiban), adenosine reuptake inhibitors (dipyridamole).

The concomitant use of ELIQUIS with NSAIDs may increase the risk of bleeding.

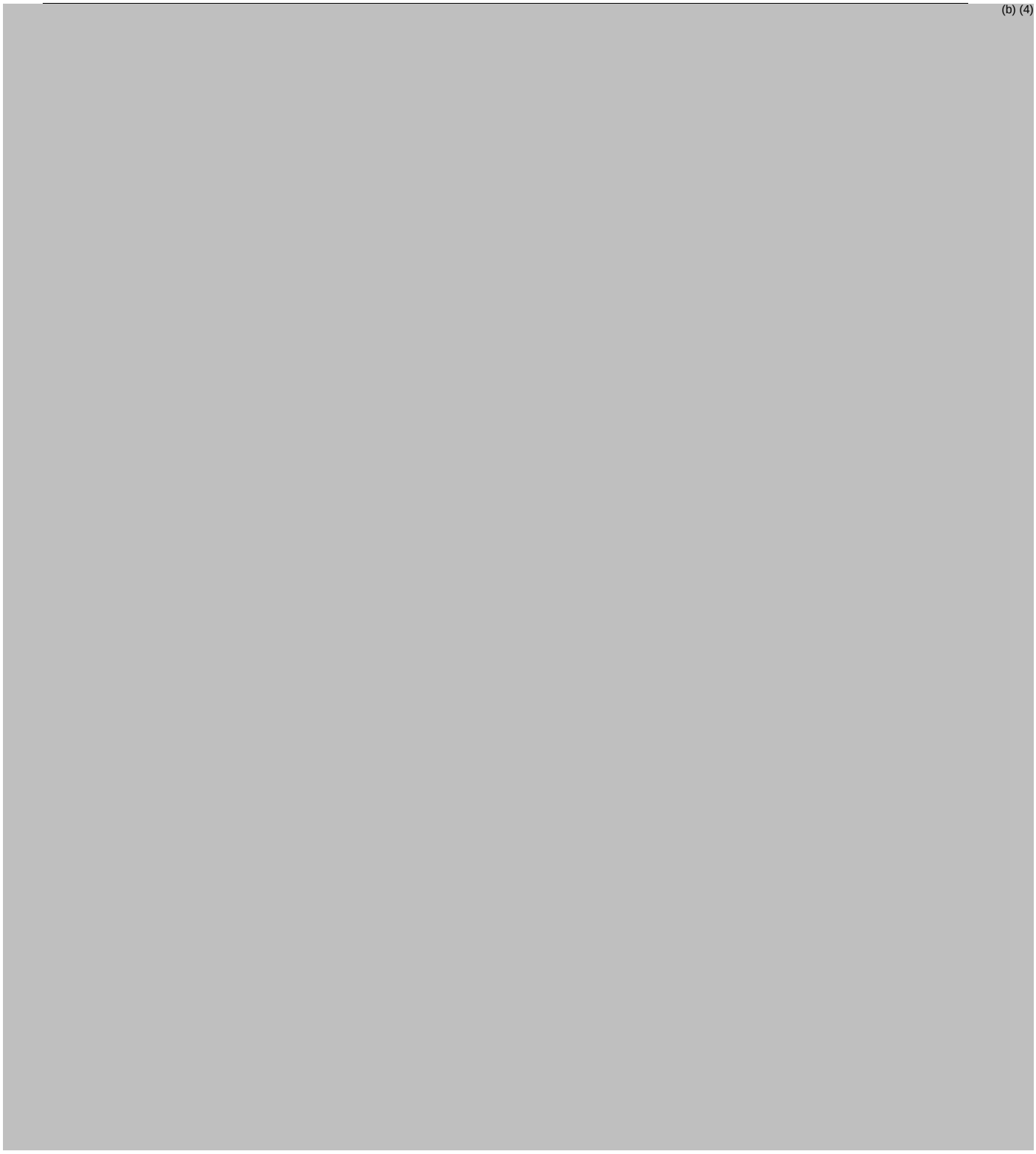
Reviewer's comment: The above preliminary recommendations are from SEALD. The Applicant proposes that (b) (4) [redacted]

The Applicant's proposal is reasonable. It is also reasonable to recommend that in patients that require both an anticoagulant and an antiplatelet, that stronger consideration be given to using only a single antiplatelet because of the lower risk of bleeding compared to dual antiplatelet therapy.

6. **Sec 14. Clinical Studies**

(b) (4) [redacted]

(b) (4)



Reviewer's comment: This section has been rewritten, but the description of the study design is generally in agreement with the Applicant's.

9.3 Advisory Committee Meeting

No advisory meeting has been scheduled.

Reference List

(1) Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007;146:857-867.

(2) Cabral KP, Ansell J, Hylek EM. Future directions of stroke prevention in atrial fibrillation: the potential impact of novel anticoagulants and stroke risk stratification. *J Thromb Haemost* 2011;9:441-449.

(3) Fuster V, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: full text: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 guidelines for the management of patients with atrial fibrillation) developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Europace* 2006;8:651-745.

(4) Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest* 2010;137:263-272.

(5) Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest* 2008;133:160S-198S.

(6) Wallentin L, Yusuf S, Ezekowitz MD et al. Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. *Lancet* 2010;376:975-983.

(7) Connolly SJ, Pogue J, Eikelboom J et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation* 2008;118:2029-2037.

(8) Connolly S, Pogue J, Hart R et al. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. *Lancet* 2006;367:1903-1912.

(9) Camm AJ, Kirchhof P, Lip GY et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010;31:2369-2429.

(10) Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thromb Haemost* 1993;69:236-239.

(11) Connolly SJ, Pogue J, Eikelboom J et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation* 2008;118:2029-2037.

(12) Lind M, Fahlen M, Kosiborod M, Eliasson B, Oden A. Variability of INR and its relationship with mortality, stroke, bleeding and hospitalisations in patients with atrial fibrillation. *Thromb Res* 2012;129:32-35.

(13) Konrad CJ, Schuepfer GK, Gerber H, Rukwied R, Schmelz M, Schley M. Duration of effects of aspirin on platelet function in healthy volunteers: an analysis using the PFA-100. *J Clin Anesth* 2006;18:12-17.

(14) Jimenez AH, Stubbs ME, Tofler GH, Winther K, Williams GH, Muller JE. Rapidity and duration of platelet suppression by enteric-coated aspirin in healthy young men. *Am J Cardiol* 1992;69:258-262.

Attachment 1 ARISTOTLE - Enrollment in Regions and Countries

Region(%) Country(%)	Apixaban N = 9120	Warfarin N = 9081	Total N = 18201
NORTH AMERICA	2249 (24.7)	2225 (24.5)	4474 (24.6)
CANADA	529 (5.8)	528 (5.8)	1057 (5.8)
USA	1720 (18.9)	1697 (18.7)	3417 (18.8)
LATIN AMERICA	1743 (19.1)	1725 (19.0)	3468 (19.1)
ARGENTINA	786 (8.6)	775 (8.5)	1561 (8.6)
BRAZIL	353 (3.9)	347 (3.8)	700 (3.8)
CHILE	128 (1.4)	130 (1.4)	258 (1.4)
COLOMBIA	54 (0.6)	57 (0.6)	111 (0.6)
MEXICO	310 (3.4)	299 (3.3)	609 (3.3)
PERU	103 (1.1)	110 (1.2)	213 (1.2)
PUERTO RICO	9 (<0.1)	7 (<0.1)	16 (<0.1)
EUROPE	3672 (40.3)	3671 (40.4)	7343 (40.3)
AUSTRIA	17 (0.2)	17 (0.2)	34 (0.2)
BELGIUM	96 (1.1)	98 (1.1)	194 (1.1)
CZECH REPUBLIC	83 (0.9)	82 (0.9)	165 (0.9)
DENMARK	169 (1.9)	170 (1.9)	339 (1.9)
FINLAND	15 (0.2)	11 (0.1)	26 (0.1)
FRANCE	17 (0.2)	18 (0.2)	35 (0.2)
GERMANY	431 (4.7)	423 (4.7)	854 (4.7)
HUNGARY	227 (2.5)	228 (2.5)	455 (2.5)
ISRAEL	170 (1.9)	174 (1.9)	344 (1.9)
ITALY	84 (0.9)	94 (1.0)	178 (1.0)
NETHERLANDS	155 (1.7)	154 (1.7)	309 (1.7)
NORWAY	47 (0.5)	43 (0.5)	90 (0.5)
POLAND	157 (1.7)	157 (1.7)	314 (1.7)
ROMANIA	138 (1.5)	136 (1.5)	274 (1.5)
RUSSIA	896 (9.8)	904 (10.0)	1800 (9.9)
SOUTH AFRICA	44 (0.5)	45 (0.5)	89 (0.5)
SPAIN	116 (1.3)	114 (1.3)	230 (1.3)
SWEDEN	111 (1.2)	106 (1.2)	217 (1.2)
SWITZERLAND	0	0	0
TURKEY	3 (<0.1)	3 (<0.1)	6 (<0.1)
UK	216 (2.4)	218 (2.4)	434 (2.4)
UKRAINE	480 (5.3)	476 (5.2)	956 (5.3)
ASIA/PACIFIC	1456 (16.0)	1460 (16.1)	2916 (16.0)
AUSTRALIA	166 (1.8)	156 (1.7)	322 (1.8)
CHINA	422 (4.6)	421 (4.6)	843 (4.6)
HONG KONG	38 (0.4)	38 (0.4)	76 (0.4)
INDIA	302 (3.3)	299 (3.3)	601 (3.3)
JAPAN	161 (1.8)	175 (1.9)	336 (1.8)
MALAYSIA	64 (0.7)	62 (0.7)	126 (0.7)
PHILIPPINES	103 (1.1)	102 (1.1)	205 (1.1)
SINGAPORE	20 (0.2)	20 (0.2)	40 (0.2)
SOUTH KOREA	153 (1.7)	157 (1.7)	310 (1.7)
TAIWAN	27 (0.3)	30 (0.3)	57 (0.3)

“Europe” included all of Russia and all of Turkey, including the Asiatic portions of these countries. It also included Israel and South Africa. Some analyses split Europe into Eastern Europe and Western Europe, defined as follows: Eastern Europe included the Czech Republic, Hungary, Poland, Romania, Russia, and Ukraine. Western Europe included: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, and the UK.

Attachment 2 ARISTOTLE - Warfarin Initiation Algorithm

Source: ARISTOTLE "Guidance for the Use and Dosing of Warfarin for Sites and Investigators"

Age	Day 1: Test	Day 1:	Day 2:	Day 3: Test	Day 3:	Day 4:	Day 5: Test	Day 5: Dose
<80	Baseline	6	6	INR <1.5	8	8	INR <1.5	10
				INR 1.5-1.9	6	4	INR 1.5-1.9	6
				INR 2.0-3.0	2	2	INR 2.0-3.0	4
				INR >3.0	Hold	Hold	INR >3.0	Hold
≥80	Baseline	4	4	INR <1.5	6	6	INR <1.5	8
				INR 1.5-1.9	2	2	INR 1.5-1.9	6
				INR 2.0-3.0	Hold	2	INR 2.0-3.0	2
				INR >3.0	Hold	Hold	INR >3.0	Hold

Figure 3. Suggested algorithm for the initiation of warfarin and testing of INR for atrial fibrillation patients. Dosing decisions on Day 5 must reflect patient's response to previous warfarin doses, in addition to the Day 5 test result.

Attachment 3 Modified Rankin Score

Score Description:

0	No symptoms at all
1	No significant disability despite symptoms: able to carry out all usual duties and activities
2	Slight disability: unable to carry out all previous activities but able to look after own affairs without assistance
3	Moderate disability: requiring some help, but able to walk without assistance
4	Moderately severe disability: unable to walk without assistance and unable to attend to own bodily needs without assistance
5	Severe disability: bedridden, incontinent and requiring constant nursing care and attention
6	Patient death

Attachment 4 ARISTOTLE - Special Dosing Procedures

Elective Procedures (copied from the protocol)

In general, local standards of care for discontinuation of anticoagulation prior to elective procedures/surgery should be employed; these may be informed by current guidelines. These are summarized below based upon the risk of thromboembolism:

Low risk of thromboembolism

- Stop warfarin/warfarin-placebo and apixaban/apixaban-placebo 4 days before the planned procedure.
- Monitor the INR using the encrypted POC device as necessary.
- Once the POC INR has attained a value deemed appropriate for the proposed procedure, this value may be confirmed by locally obtained coagulation studies (e.g. INR, PT, aPTT) conforming to the site's standard of care.
- If the procedure is associated with an increased risk of thrombosis, brief postoperative protection with UFH or LMWH at a prophylactic dose may be considered.
- Restart warfarin/warfarin-placebo (usually the night of the day of surgery) and apixaban/apixaban-placebo postoperatively (usually the day after surgery) when it is deemed safe to do so. If UFH/LMWH is used in the postoperative period, begin apixaban/apixaban-placebo when the INR is therapeutic.

Intermediate risk of thromboembolism

- Stop warfarin/warfarin-placebo 4 days before the planned procedure.
- Monitor the INR using the encrypted POC device as necessary.
- Two days before the planned procedure, stop the apixaban/apixaban-placebo and begin UFH or LMWH. The doses employed should conform to the local standard of care.
- Once the POC INR has attained a value deemed appropriate for the proposed procedure, this value may be confirmed by locally obtained coagulation studies (e.g. INR, PT, aPTT) conforming to the site's standard of care.
- Maintain on UFH/LMWH during the postop period (full dose preferred over prophylactic dose) until INR is therapeutic.
- Restart warfarin/warfarin-placebo (usually the night of the day of surgery) and apixaban/apixaban-placebo postoperatively when the INR is therapeutic and when it is deemed safe to do so. Stop UFH/LMWH.
-

High risk of thromboembolism

- Stop warfarin/warfarin-placebo 4 days before the planned procedure.
- Monitor the INR using the encrypted POC device as necessary.
- Begin full dose UFH or LMWH as the INR falls (approximately 2 days before the planned procedure). The doses employed should conform to the local standard of care. Stop apixaban/apixaban-placebo.

- Once the POC INR has attained a value deemed appropriate for the proposed procedure, this value may be confirmed by locally obtained coagulation studies (e.g. INR, PT, aPTT) conforming to the site's standard of care.
- Maintain on UFH or LMWH in the postoperative period as per the local standard of care (full dose preferred) until INR is therapeutic.
- Restart warfarin/warfarin-placebo (usually the night of the day of surgery) and apixaban/apixaban-placebo postoperatively (when the INR is therapeutic) when it is deemed safe to do so. Stop UFH/LMWH.

Emergency Procedures (Copied from the protocol)

For urgent or emergent invasive procedures, when waiting 4 - 5 days is not an option, management will in part depend on the randomized treatment assignment (warfarin or apixaban) and unblinding may be necessary (see Section 5.4 Blinding/Unblinding). Regardless of treatment, study drugs should be discontinued and standard laboratory coagulation tests (PT/INR, aPTT, platelet count, etc.) performed. The procedure should be carried out and in such a way to minimize the risk of bleeding.

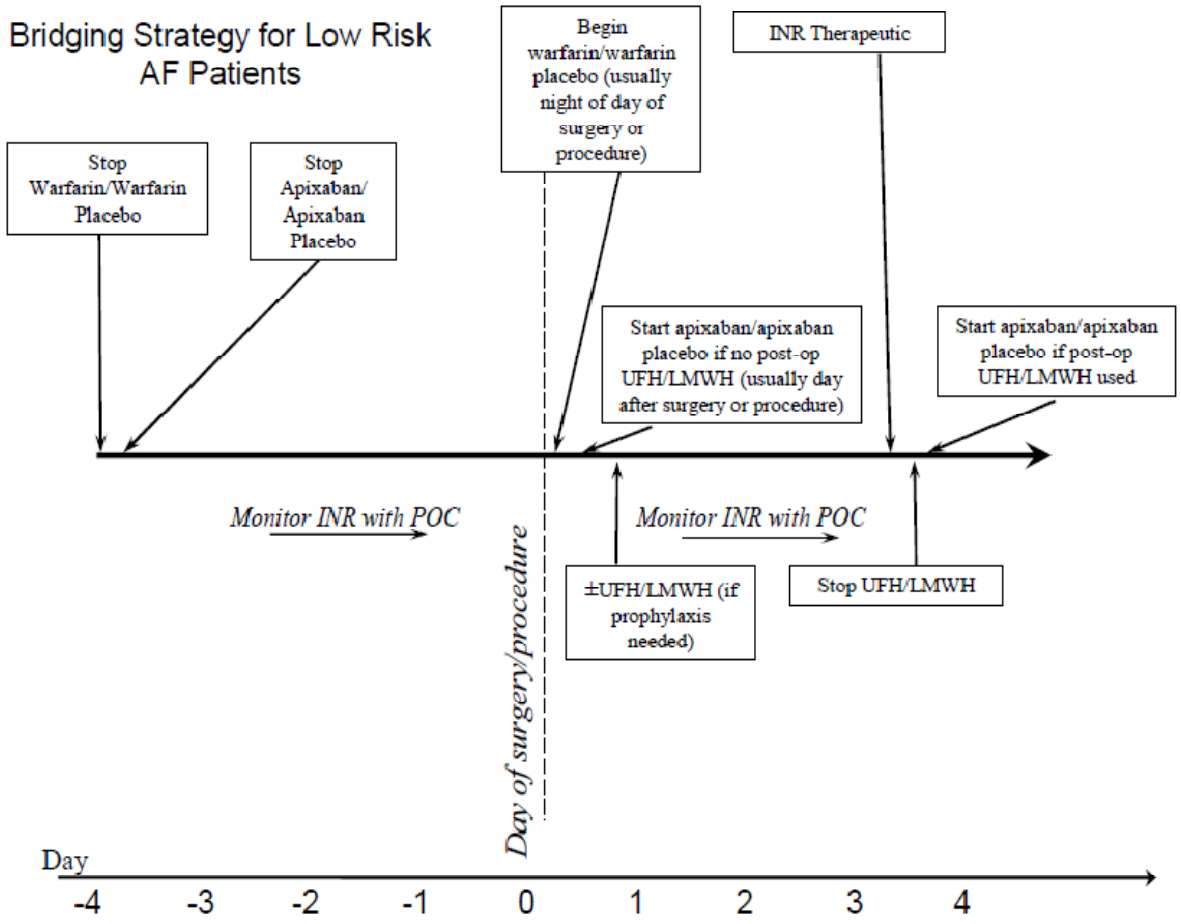
Subjects receiving warfarin should be managed according to the local standard of care. The anticoagulant effects of warfarin will be reflected in the PT and INR and, after discontinuation, will take several days (3 - 5) to return to normal. Warfarin can be reversed more quickly by giving oral or intravenous vitamin K (depending on circumstances and the local standard of care) and/or with fresh frozen plasma (FFP).

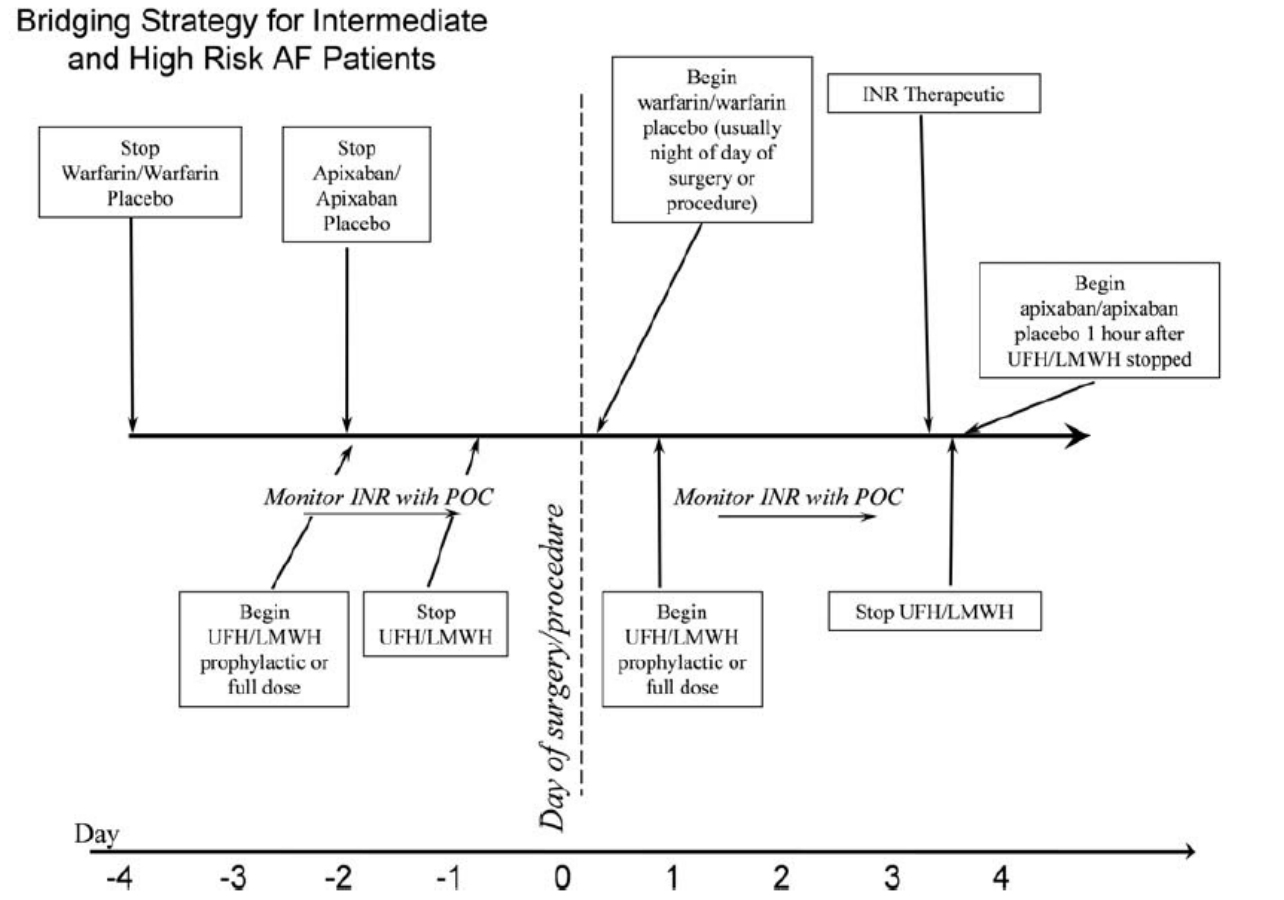
For subjects receiving apixaban, the risk of bleeding with invasive procedures is unknown. At therapeutic doses, the anticoagulant effects of apixaban will not be reflected in standard coagulation tests; there is no reversal agent for apixaban. Vitamin K and protamine sulfate are not expected to affect the anticoagulant effect of apixaban, and may carry some risk. Given its half-life (12 hours), however, the anticoagulant effect of apixaban abates in 24 - 48 hours. Depending on the subject's risk of bleeding with the procedure, subjects receiving apixaban who require an invasive or surgical procedure within 24 hours of their last dose may be treated with prophylactic peri-procedural FFP (2 units IV every 6 hours) at the discretion of the local physician and investigator.

If treatment with an alternative open label anticoagulant/antithrombotic is indicated for the procedure, it should be used at the lowest therapeutic dose (if at all) in the 12 hours following last dose of apixaban. Interactions between apixaban and other antithrombotics (with the exception of aspirin and clopidogrel) have not been evaluated.

Several figures depicting bridging strategies for patients undergoing invasive procedures were provided:

Figure 5.4.1.3A: Bridging Strategy for Low Risk AF Subjects





Management of Cardioversion (copied from the protocol)

> Cardioversion, both spontaneous and as the result of medical intervention, is an important clinical issue for subjects with atrial fibrillation. Many of the subjects randomized to the study will spontaneously convert to and from atrial fibrillation during the trial, often on more than one occasion. It is important to assess the effectiveness of apixaban as compared to warfarin in preventing stroke in these subjects, and in subjects in whom cardioversion is induced either electrically or by the use of an antiarrhythmic drug. In general, subjects entered in the trial should receive blinded oral anticoagulation with therapeutic INRs for at least 3 weeks prior to undertaking elective cardioversion (either electrical or chemical) as is recommended in current guidelines.
>
> In certain subjects at higher risk for left atrial or left atrial appendage thrombus, transesophageal echocardiography may be a useful adjuvant in guiding clinical decision making.

Attachment 5 ARISTOTLE - Adjudication Trigger Document

(copy of original text)

ARISTOTLE CEC Suspected Event Queries
(i.e. "Triggers")

This document will serve as an addendum to the *Clinical Events Classification (CEC) Process Guideline Document for the ARISTOTLE trial.* It describes the process of how patients with suspected endpoint events, including Stroke, Systemic Embolism, Bleeding, Myocardial Infarction and Death will be identified for CEC review in the ARISTOTLE Trial.

Suspected events for review will be identified ("triggered") by an electronic and manual review of the clinical data captured on the CRF. The initial set of "triggers" described in this document, are based on exclusive review of the trial protocol, CRF, and general CEC experience in prior trials. However, the development of clinical trial "triggers" is best viewed as an iterative process. If potential changes in the triggers are identified after events have been reviewed, the triggers may be revised during the course of the trial.

Notes:

- **All Suspected Events will be triggered at randomization and until the trial ends.**
- **Events that occur prior to randomization should not trigger for adjudication.**
- **Triggers will be programmed in a hierarchical order as described by the conditions below.**

Stroke

If either of the following Conditions 1, 2 or 3 is satisfied a stroke will be triggered for CEC adjudication

Condition 1

Suspected Stroke Endpoint Details Page 700 or 701 is provided

Condition 2

Within the Clinical Event Assessment pages 27, 36, 40, 44, 48, 52, 56, 66, 71, 81, 86, 91, 102, 107, 112, 122, 127, 132, 144, 149, 154, 164, 169, 174, 185, 190, 195, 205, 210, 215, 227, 232, 237, 247, 252, 257, 268, 273, 278, 288, 293, 298, 310, 315, 320, 330, 360, 908

either of ischemic stroke, hemorrhagic stroke, ischemic stroke with hemorrhagic conversion or TIA is checked yes between current visit and previous visit

*Strokes and TIA's that occur on the same day are considered one event

Condition 3

Within the suspected bleeding endpoint page 704 primary location of bleeding is checked as Intracranial

*Intracranial bleeds and strokes that occur within 24 hours of one another are considered one event

Systemic Embolism

If either of the following Conditions 1 or 2 is satisfied a systemic embolism will be triggered for CEC adjudication

Condition 1

Systemic Embolism Endpoint Details Page 702 is provided

Condition 2

Within the Clinical Event Assessment pages 27, 36, 40, 44, 48, 52, 56, 66, 71, 81, 86, 91, 102, 107, 112, 122, 127, 132, 144, 149, 154, 164, 169, 174, 185, 190, 195, 205, 210, 215, 227, 232, 237, 247, 252, 257, 268, 273, 278, 288, 293, 298, 310, 315, 320, 330, 360, 908

systemic embolism is checked yes between current visit and previous visit

Death

If either of the following Conditions 1, 2, or 3 is satisfied a death will be triggered for CEC adjudication

Condition 1

Death Endpoint Details Page 706 is provided

Condition 2

SAE with Outcome of Death

Condition 3

Within the Clinical Event Assessment pages 27, 36, 40, 44, 48, 52, 56, 66, 71, 81, 86, 91, 102, 107, 112, 122, 127, 132, 144, 149, 154, 164, 169, 174, 185, 190, 195, 205, 210, 215, 227, 232, 237, 247, 252, 257, 268, 273, 278, 288, 293, 298, 310, 315, 320, 330, 360, 908

death is checked yes between current visit and previous visit

Myocardial Infarction (MI)

If either of the following Conditions 1, 2, 3 or 4 is satisfied an MI will be triggered for CEC adjudication

Condition 1

Suspected MI Details Page 703 is provided

Condition 2

Within the Clinical Event Assessment pages 27, 36, 40, 44, 48, 52, 56, 66, 71, 81, 86, 91, 102, 107, 112, 122, 127, 132, 144, 149, 154, 164, 169, 174, 185, 190, 195, 205, 210, 215, 227, 232, 237, 247, 252, 257, 268, 273, 278, 288, 293, 298, 310, 315, 320, 330, 360, 908

myocardial infarction or unstable angina is checked yes between current visit and previous visit

* myocardial infarction or unstable angina that occur on the same day are considered one event

Condition 3

Within the Unscheduled Cardiac marker page 921, either of the following are satisfied

- CK-MB or Troponin is above the ULN
- CK values $\geq$ 2 X ULN
- Troponin I or T positive

* Enzyme triggers should only trigger once the first time it meets criteria and should not look at any more records within that visit period

Condition 4

Within the Unscheduled ECG page 902 or Scheduled ECG pages 23, 139, 222, 305, 337

Pathological Q wave is checked yes and pathological Q wave is checked no on previous ECG page between current visit and previous visit

Bleeding

If either of the following Conditions 1, 2, 3 or 4 is satisfied when either Part A or Part B is also satisfied then a bleeding event will be triggered for CEC adjudication

PART A

Within the Clinical Event Assessment pages 27, 36, 40, 44, 48, 52, 56, 66, 71, 81, 86, 91, 102, 107, 112, 122, 127, 132, 144, 149, 154, 164, 169, 174, 185, 190, 195, 205, 210, 215, 227, 232, 237, 247, 252, 257, 268, 273, 278, 288, 293, 298, 310, 315, 320, 330, 360, 908

bleeding is checked yes between current visit and previous visit

OR

PART B

Bleeding Endpoint Details Pages 704 or 705 is provided

AND

Condition 1

Within the suspected bleeding endpoint details page 704, any of the following occurs

- Did the bleeding event lead to death is answered “YES”
- Primary location of bleeding is marked as either of intraspinal, intraocular, pericardial, intra-articular, intramuscular with compartment syndrome, or retroperitoneal

Condition 2
Within the Transfusion page 922, the subject received at least 1 transfusion of either whole blood or packed cells

* If transfusion is within 14 days of a reported bleeding event then it should trigger once.

Condition 3
If the following lab changes occur:

- Chronological evaluation of local labs demonstrated a drop in hemoglobin of $\geq$ 2 g/dL from the most recent prior core lab value

Condition 4
Within the Suspected bleeding endpoint details page 705 any of the following questions are answered "YES"

- Subject received medical/surgical consultation for evaluation of the bleed
- Subject required a medical or surgical intervention to stop the bleed
- The bleed was associated with hemodynamic compromise
- Was there a fall in hemoglobin $\geq$ 2 g/dL
- Was there a change in antithrombotic therapy
- Did the bleeding lead to hospitalization
- Did the bleeding lead to a transfusion

Attachment 6 ARISTOTLE - Demographic and Disposition Data at US Sites

Demographic Data

Table 5.3.1-USA: Demographic Characteristics Summary at Baseline - Subjects Randomized in USA (CV185030)

	Apixaban N = 1720	Warfarin N = 1697	Total N = 3417
AGE (YRS)			
N	1720	1697	3417
MEAN	70.9	70.9	70.9
MEDIAN	72.0	72.0	72.0
MIN , MAX	35 , 93	40 , 95	35 , 95
Q1 , Q3	65.0 , 78.0	64.0 , 78.0	65.0 , 78.0
STANDARD DEVIATION	9.48	9.74	9.61
AGE CATEGORY (%)			
<65	425 (24.7)	427 (25.2)	852 (24.9)
65-<75	620 (36.0)	581 (34.2)	1201 (35.1)
>=75	675 (39.2)	689 (40.6)	1364 (39.9)
NOT REPORTED	0	0	0
GENDER (%)			
MALE	1176 (68.4)	1153 (67.9)	2329 (68.2)
FEMALE	544 (31.6)	544 (32.1)	1088 (31.8)
NOT REPORTED	0	0	0
FEMALE AGE CATEGORY (%)			
<=50	7 (0.4)	9 (0.5)	16 (0.5)
>50	537 (31.2)	535 (31.5)	1072 (31.4)
NOT APPLICABLE(MALE)	1176 (68.4)	1153 (67.9)	2329 (68.2)
NOT REPORTED	0	0	0
RACE (%)			
WHITE	1635 (95.1)	1629 (96.0)	3264 (95.5)
EUROPEAN	1061 (61.7)	1029 (60.6)	2090 (61.2)
MIDDLE EASTERN OR NORTH AFRICAN	2 (0.1)	13 (0.8)	15 (0.4)
OTHER WHITE	572 (33.3)	587 (34.6)	1159 (33.9)
NOT REPORTED	0	0	0
BLACK/AFRICAN AMERICAN	71 (4.1)	57 (3.4)	128 (3.7)
ASIAN	7 (0.4)	4 (0.2)	11 (0.3)
ASIAN INDIAN	0	1 (<0.1)	1 (<0.1)
CHINESE	2 (0.1)	1 (<0.1)	3 (<0.1)
JAPANESE	1 (<0.1)	0	1 (<0.1)
OTHER ASIAN	4 (0.2)	2 (0.1)	6 (0.2)
NOT REPORTED	0	0	0
AMERICAN INDIAN/ALASKA NATIVE	2 (0.1)	2 (0.1)	4 (0.1)
NATIVE HAWAIIAN/OTHER PACIFIC ISLANDER	2 (0.1)	1 (<0.1)	3 (<0.1)
OTHER	3 (0.2)	4 (0.2)	7 (0.2)
NOT REPORTED	0	0	0
ETHNICITY (%)			
HISPANIC/LATINO	29 (1.7)	41 (2.4)	70 (2.0)
NOT HISPANIC/LATINO	1691 (98.3)	1655 (97.5)	3346 (97.9)
NOT REPORTED	0	1 (<0.1)	1 (<0.1)

Level of Renal Impairment (CrCL)	Apixaban N = 1720	Warfarin N = 1697	Total N = 3417
SEVERE (<=30 ML/MIN) (%)	24 (1.4)	21 (1.2)	45 (1.3)
MODERATE (>30 - <=60 ML/MIN) (%)	249 (14.5)	242 (14.3)	491 (14.4)
MILD (>50 - <=80 ML/MIN) (%)	647 (37.6)	627 (36.9)	1274 (37.3)
NORMAL (> 80 ML/MIN) (%)	792 (46.0)	803 (47.3)	1595 (46.7)
NOT REPORTED (%)	8 (0.5)	4 (0.2)	12 (0.4)

TYPE OF RISK FACTOR AT ENROLLMENT (%)			
AGE >= 75 YEARS	675 (39.2)	689 (40.6)	1364 (39.9)
PRIOR STROKE, TIA, OR SYSTEMIC EMBOLISM	289 (16.8)	251 (14.8)	540 (15.8)
PRIOR STROKE	145 (8.4)	118 (7.0)	263 (7.7)
PRIOR TIA	145 (8.4)	122 (7.2)	267 (7.8)
PRIOR SYSTEMIC EMBOLISM	22 (1.3)	24 (1.4)	46 (1.3)
SYMPTOMATIC CHF WITHIN 3 MONTHS OR LVEF <= 40%	410 (23.8)	438 (25.8)	848 (24.8)
SYMPTOMATIC CHF	274 (15.9)	317 (18.7)	591 (17.3)
LVEF <=40%	262 (15.2)	246 (14.5)	508 (14.9)
DIABETES MELLITUS	562 (32.7)	537 (31.6)	1099 (32.2)
HYPERTENSION WITH PHARMACOLOGICAL TREATMENT	1517 (88.2)	1494 (88.0)	3011 (88.1)

	Apixaban N = 1720	Warfarin N = 1697	Total N = 3417
NUMBER OF RISK FACTORS AT ENROLLMENT (%)			
NOT REPORTED	0	0	0
0	0	2 (0.1)	2 (<0.1)
1	562 (32.7)	546 (32.2)	1108 (32.4)
2	697 (40.5)	700 (41.2)	1397 (40.9)
3	360 (20.9)	349 (20.6)	709 (20.7)
4	88 (5.1)	84 (4.9)	172 (5.0)
5	13 (0.8)	16 (0.9)	29 (0.8)
<= 1	562 (32.7)	548 (32.3)	1110 (32.5)
>= 2	1158 (67.3)	1149 (67.7)	2307 (67.5)
CHADS-2 SCORE AT ENROLLMENT (%)			
NOT REPORTED	0	0	0
0	9 (0.5)	14 (0.8)	23 (0.7)
1	580 (33.7)	570 (33.6)	1150 (33.7)
2	643 (37.4)	637 (37.5)	1280 (37.5)
3	291 (16.9)	296 (17.4)	587 (17.2)
4	141 (8.2)	125 (7.4)	266 (7.8)
5	47 (2.7)	43 (2.5)	90 (2.6)
6	9 (0.5)	12 (0.7)	21 (0.6)
<= 1	589 (34.2)	584 (34.4)	1173 (34.3)
>= 3	488 (28.4)	476 (28.0)	964 (28.2)
MEAN (SD)	2.1 (1.09)	2.1 (1.08)	2.1 (1.08)
MEDIAN	2.0	2.0	2.0
MIN, MAX	(0.0, 6.0)	(0.0, 6.0)	(0.0, 6.0)

	Apixaban N = 1720	Warfarin N = 1697	Total N = 3417
ATRIAL FIBRILLATION HISTORY (%)			
ONSET OF ATRIAL FIBRILLATION PRIOR TO RANDOMIZATION			
< 3 MONTHS	326 (19.0)	339 (20.0)	665 (19.5)
3 - 12 MONTHS	259 (15.1)	250 (14.7)	509 (14.9)
> 12 MONTHS	1126 (65.5)	1097 (64.6)	2223 (65.1)
NOT REPORTED	9 (0.5)	11 (0.6)	20 (0.6)
RHYTHM AT ENROLLMENT			
ATRIAL FIBRILLATION	1344 (78.1)	1342 (79.1)	2686 (78.6)
ATRIAL FLUTTER	88 (5.1)	92 (5.4)	180 (5.3)
SINUS RHYTHM	241 (14.0)	231 (13.6)	472 (13.8)
PACED RHYTHM	230 (13.4)	220 (13.0)	450 (13.2)
ATRIAL	28 (1.6)	47 (2.8)	75 (2.2)
VENTRICULAR	154 (9.0)	134 (7.9)	288 (8.4)
BOTH	40 (2.0)	39 (2.3)	87 (2.5)
OTHER	77 (4.5)	82 (4.8)	159 (4.7)
DURATION OF ATRIAL FIBRILLATION OR FLUTTER			
<14 DAYS	545 (31.7)	505 (29.8)	1050 (30.7)
>=14 DAYS	1167 (67.8)	1186 (69.9)	2353 (68.9)
NOT REPORTED	8 (0.5)	6 (0.4)	14 (0.4)
TREATMENT STRATEGY			
RATE CONTROL	1413 (82.2)	1430 (84.3)	2843 (83.2)
RHYTHM CONTROL	513 (29.8)	520 (30.6)	1033 (30.2)
CARDIOVASCULAR DISEASE HISTORY (%)			
CURRENT NYHA CLASS			
I	1284 (74.7)	1259 (74.2)	2543 (74.4)
II	369 (21.5)	364 (21.4)	733 (21.5)
III	65 (3.8)	70 (4.1)	135 (4.0)
IV	2 (0.1)	3 (0.2)	5 (0.1)
DOCUMENTED CORONARY ARTERY DISEASE	691 (40.2)	694 (40.9)	1385 (40.5)
UNSTABLE ANGINA	191 (11.1)	183 (10.8)	374 (10.9)
MYOCARDIAL INFARCTION	305 (17.7)	274 (16.1)	579 (16.9)
CONGESTIVE HEART FAILURE	317 (18.4)	340 (20.0)	657 (19.2)
AT LEAST MODERATE VALVULAR HEART DISEASE	488 (28.4)	491 (28.9)	979 (28.7)
MITRAL REGURGITATION	432 (25.1)	426 (25.1)	858 (25.1)
MITRAL STENOSIS	20 (1.2)	16 (0.9)	36 (1.1)
AORTIC REGURGITATION	115 (6.7)	111 (6.5)	226 (6.6)
AORTIC STENOSIS	61 (3.5)	66 (3.9)	127 (3.7)
PERICARDITIS	17 (1.0)	20 (1.2)	37 (1.1)
HYPERTROPHIC CARDIOMYOPATHY	63 (3.7)	63 (3.7)	126 (3.7)
CONGENITAL HEART DISEASE	44 (2.6)	29 (1.7)	73 (2.1)
SUSTAINED VENTRICULAR TACHYCARDIA	41 (2.4)	36 (2.1)	77 (2.3)
VF/VT CARDIAC ARREST	27 (1.6)	21 (1.2)	48 (1.4)
SYNCOPE IN THE LAST 5 YEARS	131 (7.6)	137 (8.1)	268 (7.8)
HYPERTENSION REQUIRING PHARMACOLOGICAL RX	1499 (87.2)	1484 (87.4)	2983 (87.3)
FIRST DIAGNOSIS			
< 3 MONTHS	46 (2.7)	51 (3.0)	97 (2.8)
3 - 12 MONTHS	55 (3.2)	56 (3.3)	111 (3.2)
> 12 MONTHS	1359 (79.0)	1345 (79.3)	2704 (79.1)
NOT REPORTED	39 (2.3)	32 (1.9)	71 (2.1)
PERIPHERAL ARTERY DISEASE	126 (7.3)	139 (8.2)	265 (7.8)
AORTIC ANEURYSM	53 (3.1)	63 (3.7)	116 (3.4)
THORACIC	9 (0.5)	9 (0.5)	18 (0.5)
ABDOMINAL	44 (2.6)	54 (3.2)	98 (2.9)

Disposition Data

Discontinued Treatment

NUMBER OF RANDOMIZED SUBJECTS, N	1720	1697
NUMBER OF SUBJECTS DISCONTINUED, n (%)	511 (29.7)	511 (30.1)
REASON FOR DISCONTINUATION, n (%)		
DEATH	41 (2.4)	48 (2.8)
ADVERSE EVENT	154 (9.0)	139 (8.2)
STROKE	18 (1.0)	16 (0.9)
SYSTEMIC EMBOLISM	6 (0.3)	1 (<0.1)
MYOCARDIAL INFARCTION	6 (0.3)	3 (0.2)
BLEEDING	35 (2.0)	37 (2.2)
OTHER	92 (5.3)	85 (5.0)
SUBJECT REQUEST TO DISCONTINUE STUDY TREATMENT	202 (11.7)	193 (11.4)
INCONVENIENCE	67 (3.9)	63 (3.7)
INABILITY TO COMPLY WITH INR MONITORING	19 (1.1)	21 (1.2)
PERCEIVED SIDE EFFECTS	40 (2.3)	44 (2.6)
OTHER	101 (5.9)	83 (4.9)
NOT REPORTED	2 (0.1)	1 (<0.1)
LOST TO FOLLOW-UP	9 (0.5)	4 (0.2)
POOR/NON-COMPLIANCE	15 (0.9)	16 (0.9)
PREGNANCY	0	0
SUBJECT NO LONGER MEETS STUDY CRITERIA	22 (1.3)	35 (2.1)
ADMINISTRATIVE REASON BY SPONSOR	0	2 (0.1)
OTHER - PHYSICIAN REFUSED TO CONTINUE TREATMENT	31 (1.8)	28 (1.6)
PERCEIVED RISK	10 (0.6)	8 (0.5)

DESIRE TO HAVE SUBJECT ON OPEN-LABEL WARFARIN	16 (0.9)	16 (0.9)
OTHER	4 (0.2)	6 (0.4)
NOT REPORTED	1 (<0.1)	0
OTHER	26 (1.5)	37 (2.2)
NOT REPORTED	11 (0.6)	9 (0.5)

Discontinued Follow-up

NUMBER OF RANDOMIZED SUBJECTS, n	1720	1697
COMPLETED FOLLOW-UP, n (%)	1477 (85.9)	1452 (85.6)
DID NOT COMPLETE FOLLOW-UP, n (%)	243 (14.1)	245 (14.4)
ENTERED FOLLOW-UP BUT DID NOT COMPLETE FOLLOW-UP, n (%)	80 (4.7)	86 (5.1)
DEATH	38 (2.2)	38 (2.2)
SUBJECT WITHDREW CONSENT	25 (1.5)	35 (2.1)
LOST TO FOLLOW-UP	9 (0.5)	7 (0.4)
OTHER	5 (0.3)	3 (0.2)
NOT REPORTED	3 (0.2)	3 (0.2)
DID NOT ENTER FOLLOW-UP, n (%)	163 (9.5)	159 (9.4)
DEATH	63 (3.7)	76 (4.5)
SUBJECT WITHDREW CONSENT	67 (3.9)	49 (2.9)
LOST TO FOLLOW-UP	9 (0.5)	5 (0.3)
OTHER	15 (0.9)	17 (1.0)
NOT REPORTED	9 (0.5)	12 (0.7)

Vital Status at end of Study

VITAL STATUS AT THE END OF STUDY		
ALIVE	1539 (89.5)	1514 (89.2)
DEATH	112 (6.5)	127 (7.5)
UNKNOWN	69 (4.0)	56 (3.3)
SUBJECT WITHDREW CONSENT	36 (2.1)	24 (1.4)
LOST TO FOLLOW-UP	6 (0.3)	4 (0.2)
OTHER	27 (1.6)	28 (1.6)

Assigned Dose

	Apixaban N = 1720	Warfarin N = 1697	Total N = 3417
APIXABAN OR MATCHING PLACEBO DOSE AT RANDOMIZATION (%)			
2.5 MG BID	105 (6.1)	101 (6.0)	206 (6.0)
5.0 MG BID	1615 (93.9)	1596 (94.0)	3211 (94.0)
CALCULATED DOSE REDUCTION CRITERIA AT RANDOMIZATION (%)			
AGE >= 80 YEARS	335 (19.5)	336 (19.8)	671 (19.6)
WEIGHT <= 60 KG	90 (5.2)	76 (4.5)	166 (4.9)
SERUM CREATININE >= 1.5 MG/DL	181 (10.5)	153 (9.0)	334 (9.8)
0 CRITERION	1221 (71.0)	1229 (72.4)	2450 (71.7)
1 CRITERION	395 (23.0)	373 (22.0)	768 (22.5)
2 CRITERIA	101 (5.9)	93 (5.5)	194 (5.7)
3 CRITERIA	3 (0.2)	2 (0.1)	5 (0.1)

Attachment 7 AVERROES - Subgroup Analyses of Primary Endpoint

Figure 7.1.2: Forest Plot for Adjudicated Stroke or Systemic Embolism During the Intended Treatment Period - Randomized Subjects

Subgroup	Apixaban n (%/yr)	ASA n (%/yr)	Hazard Ratio (95% CI)
VKA Unsuitable			
Demonstrated	17 (1.37)	52 (4.20)	0.33(0.19,0.56)
Expected	34 (1.79)	61 (3.25)	0.55(0.36,0.84)
Reason VKA unsuitable			
Subj. refused trt w/VKA	6 (1.25)	17 (3.84)	0.33(0.13,0.83)
CHADS2 Scr=1/Phy not rec	4 (1.25)	2 (0.57)	2.18(0.40,11.91)
All other reasons	41 (1.75)	94 (4.06)	0.43(0.30,0.62)
Apixaban Dose			
Apix/Plc 2.5 mg BID	3 (1.63)	12 (6.24)	0.26(0.07,0.93)
Apix/Plc 5 mg BID	48 (1.62)	101 (3.46)	0.47(0.33,0.66)
ASA Dose			
ASA/Plc 81 mg QD	39 (1.92)	85 (4.30)	0.45(0.31,0.65)
ASA/Plc 162 mg QD	11 (1.39)	20 (2.41)	0.57(0.27,1.20)
ASA/Plc 243 mg QD	1 (1.31)	2 (2.88)	0.41(0.04,4.48)
ASA/Plc 324 mg QD	0	5 (2.18)	<0.01(NE)
Geographic Region			
North America	5 (0.94)	18 (3.43)	0.27(0.10,0.74)
Latin America	8 (1.28)	31 (5.06)	0.25(0.12,0.55)
Europe	23 (1.57)	46 (3.20)	0.49(0.30,0.81)
Asia/Pacific	15 (2.87)	18 (3.35)	0.86(0.43,1.71)
Age			
<65 yrs	7 (0.73)	19 (1.93)	0.38(0.16,0.89)
>=65 and <75 yrs	24 (2.02)	29 (2.78)	0.73(0.43,1.25)
>=75 yrs	20 (2.00)	65 (6.00)	0.34(0.20,0.56)
Gender			
Male	26 (1.40)	49 (2.72)	0.52(0.32,0.83)
Female	25 (1.93)	64 (4.89)	0.40(0.25,0.63)
Race			
White	37 (1.44)	93 (3.71)	0.39(0.26,0.57)
Black/African American	0	0	NE
Asian	14 (2.71)	18 (3.43)	0.79(0.39,1.59)
Other	0	2 (3.90)	<0.01(NE)

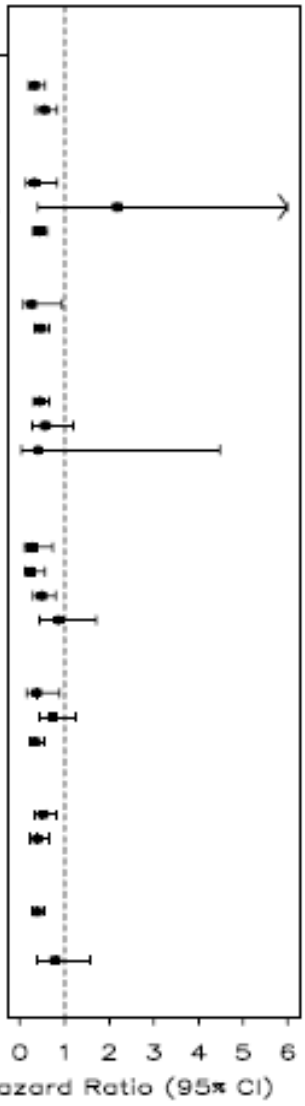
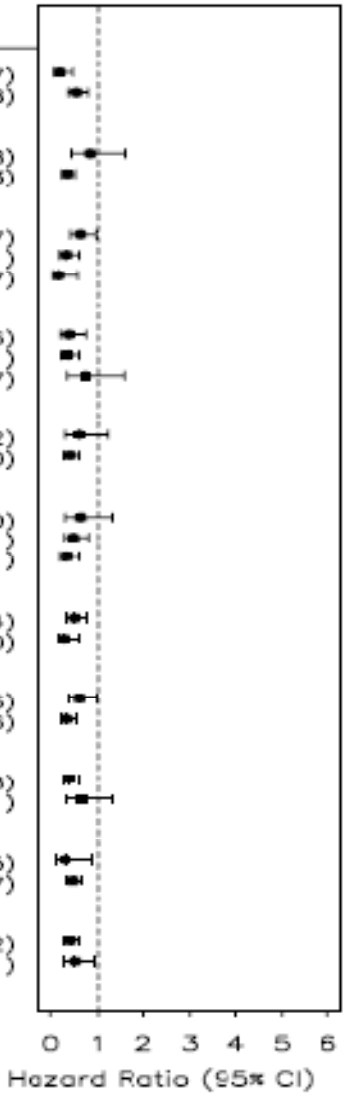


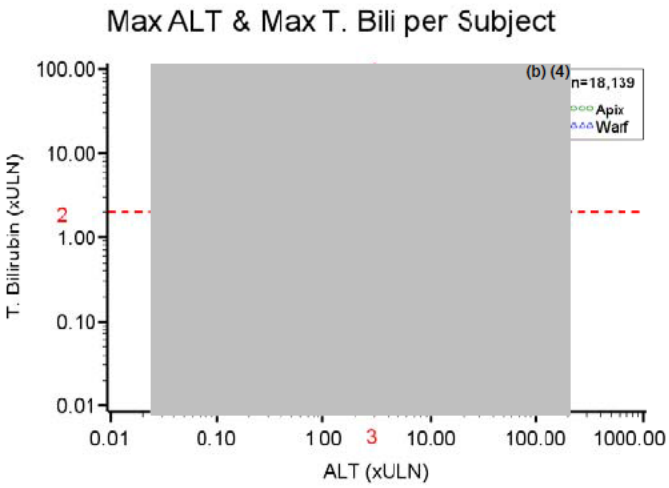
Figure 7.1.2: Forest Plot for Adjudicated Stroke or Systemic Embolism During the Intended Treatment Period - Randomized Subjects

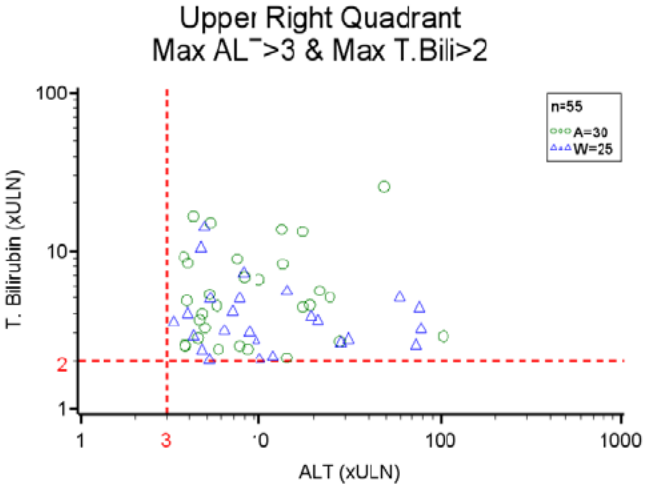
Subgroup	Apixaban n (%/yr)	ASA n (%/yr)	Hazard Ratio (95% CI)
Ethnicity			
Hispanic/latino	6 (1.62)	30 (5.27)	0.20(0.08,0.47)
Not hispanic/latino	45 (1.78)	82 (3.29)	0.54(0.38,0.78)
Weight			
Weight <= 60 kg	18 (3.88)	20 (4.67)	0.84(0.44,1.58)
Weight > 60 kg	33 (1.23)	93 (3.48)	0.36(0.24,0.53)
BMI			
BMI <= 28 kg/m2	34 (2.20)	56 (3.49)	0.63(0.41,0.97)
BMI > 28 to 33 kg/m2	14 (1.42)	40 (4.30)	0.33(0.18,0.61)
BMI > 33 kg/m2	3 (0.50)	17 (2.99)	0.17(0.05,0.57)
Level of Renal Impairment			
Severe/Moderate	13 (2.25)	32 (5.67)	0.40(0.21,0.76)
Mild	22 (1.83)	58 (4.95)	0.37(0.23,0.61)
Normal	12 (1.09)	16 (1.48)	0.74(0.35,1.57)
Number of risk factors			
<=1	13 (1.07)	21 (1.75)	0.61(0.31,1.22)
>=2	38 (1.97)	92 (4.87)	0.41(0.28,0.60)
CHADS2 Score			
<=1	12 (1.00)	19 (1.58)	0.63(0.31,1.30)
2	23 (1.95)	48 (4.04)	0.49(0.29,0.81)
>=3	16 (2.09)	51 (6.04)	0.35(0.20,0.61)
Prior Stroke or TIA			
No	41 (1.50)	80 (2.95)	0.51(0.35,0.74)
Yes	10 (2.45)	33 (8.29)	0.29(0.14,0.60)
Age >=75 yrs			
No	31 (1.44)	48 (2.37)	0.61(0.39,0.96)
Yes	20 (2.00)	55 (6.00)	0.34(0.20,0.56)
Diabetes Mellitus			
No	37 (1.44)	91 (3.66)	0.40(0.27,0.58)
Yes	14 (2.40)	22 (3.54)	0.67(0.34,1.31)
Hypertension req. Rx			
No	5 (1.12)	15 (3.69)	0.31(0.11,0.85)
Yes	46 (1.70)	98 (3.62)	0.47(0.33,0.67)
Heart Failure			
No	32 (1.53)	78 (3.72)	0.41(0.27,0.62)
Yes	19 (1.79)	35 (3.45)	0.52(0.30,0.91)



n = number of randomized subjects with an event
Horizontal bars represent 95% CIs for hazard ratios (Apixaban/ASA)
Program Source: S:\RHO\BMS\CV185048\Figures\rg-ef-ftstrkse.sas
02FEB2011:14:20:32

Attachment 8 ARISTOTLE – Maximum ALT and T.Bili per Subject





Reviewer's analysis: hep\mxALT_TB, applicant's data set adliver1; Lab values within 30 days after drug discontinuation are included in the above analyses. Analysis without respect to time.

Attachment 9 ARISTOTLE - Clinical Safety Plan (CSP)

CSP Events of Interest	Applicable CV185 Protocols***	Algorithm for Identification of CSP Qualified Event
Decrease in Thrombocyte Count*	017, 023, 030, 048	<100,000 cells/mm^3 , 2 consecutive occasions. The CRF will be sent on the lowest platelet level.
Neurology	017, 023, 030, 034, 035, 036, 047, 048, 056, 057	1. Identify events using selected MedDRA term filters. 2. Send CRF to SAEs/NSAEs with duration ≥ 7days. 2. Query remaining SAEs/NSAEs: "Based on event, was a neurological consult obtained?" 4. If query response is "yes", send CSP CRF. If "no", no additional follow-up required.
Elevated Creatine Kinase**	017, 023, 030, 036, 048, 056, 057	≥ 5X ULN
Hepatobiliary	023, 030, 034, 035, 036, 047, 048, 056, 057	**AEs/SAEs:** 1. Identify qualifying events using selected MedDRA term filters. 2. Send CSP CRFs on all events identified using the MedDRA term filters (with or without associated lab abnormals). **Lab abnormalities:** 1. Identify lab abnormals (local/central) meeting any of the following specified algorithms: ▪ AT ≥ 3X ULN (+) T-Bili ≥ 2X ULN: drawn same day ▪ AT ≥ 5X ULN: any single occurrence ▪ AT ≥ 3X ULN : 2 consecutive* occurrences ▪ T-Bili ≥ 2X ULN: any single occurrence Where AT = AST or ALT *consecutive = blood draw of, at least, next calendar day (no maximum). 2. CSP CRF follow-up will be on qualified lab abnormals with no reported associated AEs/SAEs. 3. For multiple ongoing qualified lab abnormalities, one CSP CRF will be sent on the first reported lab values. 4. Additional CSP CRFs will be sent if lab abnormals meet any of the lab criteria again at a later date.

* CSP not utilized in protocols mandating adjudication of these EOIs.
** CSP not utilized in post-surgical studies.
*** CSP will not be applied to following protocols: CV185027, CV185068

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

/s/

MARTIN ROSE
05/22/2012

BACH N BEASLEY
05/22/2012

CLINICAL FILING CHECKLIST FOR NDA/BLA or Supplement

NDA/BLA Number: 202155 **Applicant: BMS** **Stamp Date: Sept. 30, 2011**

Drug Name: Apixaban **NDA/BLA Type: 1**

On initial overview of the NDA/BLA application for filing:

	Content Parameter	Yes	No	NA	Comment
FORMAT/ORGANIZATION/LEGIBILITY					
1.	Identify the general format that has been used for this application, e.g. electronic CTD.			X	ECTD
2.	On its face, is the clinical section organized in a manner to allow substantive review to begin?	X			
3.	Is the clinical section indexed (using a table of contents) and paginated in a manner to allow substantive review to begin?	X			
4.	For an electronic submission, is it possible to navigate the application in order to allow a substantive review to begin (*e.g.*, are the bookmarks adequate)?	X			
5.	Are all documents submitted in English or are English translations provided when necessary?	X			
6.	Is the clinical section legible so that substantive review can begin?	X			
LABELING					
7.	Has the applicant submitted the design of the development package and draft labeling in electronic format consistent with current regulation, divisional, and Center policies?	X			
SUMMARIES					
8.	Has the applicant submitted all the required discipline summaries (*i.e.*, Module 2 summaries)?	X			
9.	Has the applicant submitted the integrated summary of safety (ISS)?	X			
10.	Has the applicant submitted the integrated summary of efficacy (ISE)?	X			
11.	Has the applicant submitted a benefit-risk analysis for the product?	X			
12.	Indicate if the Application is a 505(b)(1) or a 505(b)(2). If Application is a 505(b)(2) and if appropriate, what is the reference drug?			X	505(b)(1)
DOSE					
13.	If needed, has the applicant made an appropriate attempt to determine the correct dosage and schedule for this product (*i.e.*, appropriately designed dose-ranging studies)? Study Number: CV 185010 Study Title: A PHASE 2 RANDOMIZED, DOUBLE-BLIND (BMS-562247 AND ENOXAPARIN), ACTIVE-CONTROLLED (ENOXAPARIN AND WARFARIN), PARALLEL-ARM, DOSE-RESPONSE STUDY OF THE ORAL FACTOR XA INHIBITOR BMS-562247 IN SUBJECTS UNDERGOING ELECTIVE TOTAL KNEE REPLACEMENT SURGERY Sample Size: 1217 / Arms: 8: 6 apixaban + 2 control arms Location in submission: mod. 5.3.5.4 Study Number: CV 185017 Study Title: A PHASE 2 RANDOMIZED, PARALLEL-ARM STUDY OF ORAL DIRECT FACTOR Xa-INHIBITOR	X			FDA agreed that dose ranging data from VTE prevention and treatment studies could be extrapolated to AF studies. Study CV 185010 appears to be the primary support for the 5 mg bid dose.

CLINICAL FILING CHECKLIST FOR NDA/BLA or Supplement

	Content Parameter	Yes	No	NA	Comment
	APIXABAN AND LOW MOLECULAR WEIGHT HEPARIN, OR FONDAPARINUX WITH A VITAMIN K ANTAGONIST IN SUBJECTS WITH ACUTE SYMPTOMATIC DEEP VEIN THROMBOSIS Sample Size: 520 / Arms: 4: 3 apixaban + 1 control arm Location in submission: mod. 5.3.5.4				
EFFICACY					
14.	Do there appear to be the requisite number of adequate and well-controlled studies in the application? Pivotal Study #1 CV 185030 - ARISTOTLE (apixaban vs warfarin). Indication: Reduction of risk of stroke and systemic embolism in patients with non-valvular AF	X			Study CV 185048 (AVERROES, apixaban vs. aspirin) is supportive of the efficacy of apixaban in the target population.
15.	Do all pivotal efficacy studies appear to be adequate and well-controlled within current divisional policies (or to the extent agreed to previously with the applicant by the Division) for approvability of this product based on proposed draft labeling?	X			
16.	Do the endpoints in the pivotal studies conform to previous Agency commitments/agreements? Indicate if there were not previous Agency agreements regarding primary/secondary endpoints.	X			
17.	Has the application submitted a rationale for assuming the applicability of foreign data to U.S. population/practice of medicine in the submission?	X			
SAFETY					
18.	Has the applicant presented the safety data in a manner consistent with Center guidelines and/or in a manner previously requested by the Division?	x			
19.	Has the applicant submitted adequate information to assess the arythmogenic potential of the product (*e.g.*, QT interval studies, if needed)?	x			CSR CV185031, TQT, blinded, PC, pos control
20.	Has the applicant presented a safety assessment based on all current worldwide knowledge regarding this product?	x			
21.	For chronically administered drugs, have an adequate number of patients (based on ICH guidelines for exposure[1]) been exposed at the dose (or dose range) believed to be efficacious?	x			
22.	For drugs not chronically administered (intermittent or short course), have the requisite number of patients been exposed as requested by the Division?			x	
23.	Has the applicant submitted the coding dictionary[2] used for mapping investigator verbatim terms to preferred terms?	x			For hepatic, bleeding, and neurologic events
24.	Has the applicant adequately evaluated the safety issues that				This seems like a

[1] For chronically administered drugs, the ICH guidelines recommend 1500 patients overall, 300-600 patients for six months, and 100 patients for one year. These exposures MUST occur at the dose or dose range believed to be efficacious.

[2] The "coding dictionary" consists of a list of all investigator verbatim terms and the preferred terms to which they were mapped. It is most helpful if this comes in as a SAS transport file so that it can be sorted as needed; however, if it is submitted as a PDF document, it should be submitted in both directions (verbatim -> preferred and preferred -> verbatim).

CLINICAL FILING CHECKLIST FOR NDA/BLA or Supplement

	Content Parameter	Yes	No	NA	Comment
	are known to occur with the drugs in the class to which the new drug belongs?				review issue.
25.	Have narrative summaries been submitted for all deaths and adverse dropouts (and serious adverse events if requested by the Division)?	x			
OTHER STUDIES					
26.	Has the applicant submitted all special studies/data requested by the Division during pre-submission discussions?	X			
27.	For Rx-to-OTC switch and direct-to-OTC applications, are the necessary consumer behavioral studies included (*e.g.*, label comprehension, self selection and/or actual use)?			X	
PEDIATRIC USE					
28.	Has the applicant submitted the pediatric assessment, or provided documentation for a waiver and/or deferral?	X			
ABUSE LIABILITY					
29.	If relevant, has the applicant submitted information to assess the abuse liability of the product?			X	
FOREIGN STUDIES					
30.	Has the applicant submitted a rationale for assuming the applicability of foreign data in the submission to the U.S. population?	X			
DATASETS					
31.	Has the applicant submitted datasets in a format to allow reasonable review of the patient data?	X			
32.	Has the applicant submitted datasets in the format agreed to previously by the Division?	X			
33.	Are all datasets for pivotal efficacy studies available and complete for all indications requested?	X			
34.	Are all datasets to support the critical safety analyses available and complete?	X			APPRAISE-2 (CAD) data expected in 8 weeks
35.	For the major derived or composite endpoints, are all of the raw data needed to derive these endpoints included?	X			
CASE REPORT FORMS					
36.	Has the applicant submitted all required Case Report Forms in a legible format (deaths, serious adverse events, and adverse dropouts)?	X			
37.	Has the applicant submitted all additional Case Report Forms (beyond deaths, serious adverse events, and adverse drop-outs) as previously requested by the Division?	X			
FINANCIAL DISCLOSURE					
38.	Has the applicant submitted the required Financial Disclosure information?	X			
GOOD CLINICAL PRACTICE					
39.	Is there a statement of Good Clinical Practice; that all clinical studies were conducted under the supervision of an IRB and with adequate informed consent procedures?	X			

CLINICAL FILING CHECKLIST FOR NDA/BLA or Supplement

IS THE CLINICAL SECTION OF THE APPLICATION FILEABLE? _YES

If the Application is not fileable from the clinical 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.

Martin Rose and Nhi Beasley	11/2/11
Reviewing Medical Officer	Date
Clinical Team Leader	Date

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

/s/

MARTIN ROSE
11/02/2011

BACH N BEASLEY
11/02/2011