CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

202439Orig1s000

ROCKET AF Reanalysis Reviews

Clinical Review
Statistical Review
Clinical Pharmacology Review

Division of Cardiovascular and Renal Products
CDTL Review

By: Martin Rose, MD, JD
Date: September 26, 2016
Re: NDA 202439 – Rivaroxaban (Xarelto®, Janssen Pharmaceuticals, Inc.) to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation: Impact of use of the HemoSense/Alere INRatio PT/INR monitoring system in the confirmatory ROCKET AF trial on the interpretation of the trial results[1]
Also affects: IND 75,238

1 Summary and Recommendations

1.1 Summary

Rivaroxaban was approved for use to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation on November 4, 2011. Evidence for the efficacy of rivaroxaban came from the warfarin-controlled ROCKET AF (ROCKET) trial, which also contributed much of the safety data in the application (see Sec. **2.1** and Sec. **2.2** for summaries of the relevant design features and results of ROCKET, respectively). In September 2015 we received from Janssen Pharmaceuticals, Inc. (Janssen), the NDA holder, information that had the potential to affect our interpretation of the data from ROCKET. Janssen informed us that the hand-held, point-of-care (POC) INRatio® device that was used at all study sites to monitor INR and adjust warfarin dose in ROCKET had been subject to a Class 1 recall event involving a "voluntary correction" of its labeling in December 2014, about 3 years after the approval of rivaroxaban by DCRP. The recall was based on post-marketing information indicating that INR values reported by the device were lower than near-contemporaneous readings from a laboratory-based device in certain patient groups identified by the current manufacturer of this device, Alere, Inc. Some patients with substantially discordant INR readings were hospitalized for bleeding episodes.

The rationale for the recall raised the possibility that patients in the warfarin arm of ROCKET were over-anticoagulated as a result of use of the INRatio device, potentially distorting the results of the study by increasing the rate of bleeding in the warfarin arm. This might include an increased rate of hemorrhagic stroke, a component of the primary efficacy endpoint. These distortions would tend to bias the results in favor of rivaroxaban in comparisons of bleeding rates vs. warfarin, and possibly also bias the primary endpoint in favor of rivaroxaban. Over-anticoagulation in the warfarin arm might have reduced the rate of ischemic stroke, also a component of the primary endpoint. This might have distorted the study efficacy results in the opposite direction, i.e., against rivaroxaban.

After learning of the recall, Janssen and FDA independently performed a variety of analyses intended to characterize the impact of use of the INRatio device on the safety and efficacy results of ROCKET. Most of these analyses involved data from a large, embedded PK/PD sub-study in ROCKET that involved sparse INR sampling at Weeks 12 and 24 of treatment in the

[1] This review draws heavily on the excellent and innovative modeling analyses performed by Drs. Florian and McDowell (described in the Clinical Pharmacology review) and Lawrence and Hung (described in the Biostatistics review). In this review, Dr. McDowell also contributed substantially analyses of the clinical data from ROCKET and the discussion of the modeling performed by Janssen.

study. About 87% of the 7125 warfarin arm patients who were treated in the study had at least one pair of same-day INR samples that were run using the INRatio device at each patient's study site (POC INR) and a laboratory based device at Duke University (LAB INR). We consider the LAB INR as the standard in the comparison of LAB INR to POC INR. The matched pairs of INR values enabled comparative analyses by FDA and Janssen of the performance of the two devices, and also analyses of the effects of differences between POC INR and LAB INR on rates of important outcomes in ROCKET. FDA and Janssen also built mathematical models of INR vs. outcomes that generated expected bleeding and ischemic stroke rates in the two study arms and rivaroxaban vs. warfarin comparisons for these endpoints. The models estimated thrombotic and bleeding outcomes in ROCKET in a hypothetical situation in which the true INR was the same as the reported LAB INR, so that warfarin arm patients were less intensely anticoagulated they were in ROCKET. These rates and summary statistics were compared to the reported results of ROCKET.

The various analyses referred to above indicated:

- Overall, observed POC INR averaged about 13% less than LAB INR, with modest variability in this relationship except when LAB INR was ≤ 2 (**Figure 3**).
- In 6225 warfarin arm patients with a POC INR/LAB INR pair at Week 24 (or Week 12, if there was not a pair at Week 24), 52% of patients had a POC INR within the target range of 2.0 to 3.0. Of this subgroup, 35% had a same-day LAB INR above the target range, compared to 3% that had a LAB INR below the target range (**Table 10**).
- In Janssen's categorical analysis of POC INR vs. LAB INR, patients in the warfarin arm with a LAB INR category that was 1 or more categories higher than their same-day POC INR category had a somewhat higher rate of major bleeding (3.17 events per 100 patient-years) than those whose LAB INR and POC INR were in the same category (2.74 events per 100 patient years, **Table 7**). We performed our own categorical analysis, with directionally consistent results (**Table 11**).
- FDA used two mathematical modeling approaches to estimate the clinical outcomes results that might have occurred in ROCKET if a more accurate INR assay had been used to guide warfarin dosing, i.e., one that that reported results similar to the laboratory-based assay at Duke. Four such models were created by Janssen. Of these six models, three were based on the INR results obtained closest to the time of endpoint events (our preferred methodology) in ROCKET or other studies of direct-acting oral anticoagulants (DOACs) vs. warfarin. Each of these analyses predicted a small decrease in the expected rate of major bleeding in the warfarin arm compared to the observed rate of 3.45 events per 100 patient years in ROCKET (reductions in the three models ranged from 7% to 10% of the observed rate). In each analysis, the hazard ratio or rate ratio for major bleeding for rivaroxaban vs. warfarin was reciprocally increased. Estimated rates of other types of bleeding in the warfarin were also reduced to a small extent. Overall, these estimated reductions in the rates of bleeding events in the warfarin arm were small enough so that the benefits of rivaroxaban would still outweigh its risks if efficacy were not affected. Notably, two of these three analyses also estimated the rate of ischemic stroke, and one estimated the rate of the primary efficacy endpoint of total stroke + systemic embolism. Both these models predicted that the rate of ischemic stroke would be increased in the warfarin arm, resulting in an expected improvement of the efficacy of rivaroxaban relative to warfarin. Also, the one analysis that estimated the rate of the primary endpoint found an increase in the rate of this parameter in the warfarin arm. *A fortiori*, compared to warfarin, the benefits of rivaroxaban would still outweigh its risks (Sec. **5**, **Table 12** and **Table 15**).

- The EINSTEIN matched pair of randomized trials in patients with venous thromboembolism and pulmonary embolism supports rivaroxaban's current indications for acute use in these two conditions. In these studies, rivaroxaban was transiently given at a higher dose than in ROCKET and then reduced to a dose identical to the one used in ROCKET, and control arm patients received enoxaparin transiently and then were switched to warfarin, which was dosed as in ROCKET. In the pooled trial results described in labeling, the overall bleeding profile for rivaroxaban was not worse than for the control, and major bleeding, including intracranial hemorrhage, was less frequent with rivaroxaban (**Table 16**). The INRatio device was not used in these trials. The trials' results are reassuring with regard to the risk of bleeding with rivaroxaban compared to dose-adjusted warfarin.

The information summarized above indicates that it is quite likely that patients in the warfarin arm of ROCKET unintentionally received higher doses of warfarin than they would have received if the INRatio device had provided results similar to those provided by the laboratory-based device at Duke. However, the effects of this increased intensity of anticoagulation on clinical outcomes were likely to have been quite modest. It seems very unlikely that if the device had performed similarly to the INR assessment device at Duke, the benefit/risk profile of rivaroxaban compared to warfarin would have been notably different from the profile based on the observed results of ROCKET. Accordingly, the conclusion we made in 2011 that the benefits of rivaroxaban in patients with non-valvular atrial fibrillation outweigh its risks should not be changed.

1.2 Recommendations:

1) No changes in rivaroxaban labeling to reflect the impact of use of the INRatio device in ROCKET are warranted. No other major regulatory action should be taken with respect to rivaroxaban.
2) Our conclusions regarding the issues addressed by this review should be communicated to the public in a suitable manner, but not through any changes in labeling.

Contents

2 Background

Rivaroxaban, an orally available inhibitor of the activated form of coagulation factor X (FXa), was approved for use to reduce the risk of stroke and systemic embolism in patients with non-valvular atrial fibrillation (AF) on November 4, 2011. This indication is sometimes colloquially referred to as "SPAF," an acronym for stroke prevention in (nonvalvular) atrial fibrillation. Our finding of efficacy to support this use was based entirely on the results of the large ROCKET AF trial (ROCKET), which also supplied most of the clinical safety information for this use. Rivaroxaban was the second DOAC approved for SPAF; dabigatran, an inhibitor of Factor IIa (thrombin), had been approved for this use in 2010. Since the approval of rivaroxaban for SPAF, two other DOACs, apixaban and edoxaban, have been approved for this use. Like rivaroxaban, these latter two drugs are inhibitors of FXa.

2.1 Key Design Features of ROCKET

ROCKET was a randomized, double-blind, event-driven, confirmatory trial of rivaroxaban vs. dose-adjusted warfarin that was performed at 1187 study sites on 6 continents. The target was 405 primary endpoint events. It was expected that about 14,000 patients would be enrolled and treated for up to 32 months to meet the event target. Attainment of the target would mark the end of the study.

The primary objective of ROCKET was to demonstrate that the efficacy of rivaroxaban for the intended indication, reduction in the rate of stroke and systemic embolism in patients with nonvalvular AF, was non-inferior that of dose-adjusted warfarin.[2] The primary safety objective was to demonstrate that rivaroxaban is superior to dose-adjusted warfarin as assessed by the composite of major and non-major clinically relevant bleeding events. The proposed primary efficacy endpoint was time to the composite of stroke (any type) or systemic embolism. Efficacy and safety outcomes (i.e., bleeding events) were blindly adjudicated by a centralized adjudication committee. ROCKET was planned as the only study performed to support the efficacy of rivaroxaban for approval of the AF indication in the US and all other regions except for Japan, where a study with a slightly different design was conducted.

After being screened, patients were randomized using a telephonic IVRS system. Randomization was 1:1 to treatment with rivaroxaban or warfarin. The dose of rivaroxaban was 20 mg once daily for most subjects. It was 15 mg daily for those with creatinine clearance (CrCL) 30 to 50 mL/min; those with CrCL<30 mL/min were excluded. Warfarin was dose-adjusted with an INR target of 2.0 to 3.0 for all subjects. The study included both patients who were or were not taking warfarin at enrollment. INRs were to be performed at least every 4 weeks during the study, and considerably more often at the start of the study when warfarin was initiated in some patients.

A double-dummy dosing scheme was used to maintain the blind. In addition, to avoid unblinding based on knowledge of INR, INR was to be measured at the study sites for all enrolled subjects using a hand-held point-of-care (POC) device provided to the site by Janssen. The device was the INRatio PT Monitoring System, a marketed device in the US. It was modified specially for use in the trial by its manufacturer, HemoSense Inc. The modification involved only the software relating to reporting of the INR readout. The device and associated procedures were designed to minimize the likelihood of unblinding based on INR data. After

[2] Janssen specified a non-inferiority margin of 1.46, but we used a margin of 1.38 in our analyses.

analyzing a blood sample, this device displayed a code number instead of the actual INR value. This code number was entered into the telephonic IVRS by site staff along with the subject's study identification number. The IVRS decoded the INR code number and then issued a standardized report which contained either:

- the INR value if the subject was assigned to warfarin or
- a sham INR value if the subject was assigned to rivaroxaban.

The site was notified of the sham or true INR during the phone call; a fax of the result was also generated by the IVRS and sent to site. The INR was not entered into the CRF, but was kept separately at the site. There was a data transfer from the IVRS to the study database of the INR information, including the coded ("encrypted") INR, the "decoded" (true) INR and the "randomly generated" (sham) INR. The database contains all versions of the INR for each measurement, but only the true INR was reported to the site by the IVRS for warfarin arm patients and only the sham INR was reported to the site for rivaroxaban arm patients.

Decoded (true) INR values were reported to the site for warfarin arm patients as follows:

- INR values less than 1 were reported as "less than 1.0", but the true value is in the study database
- INR values >6.0 were all reported as "greater than 6.0" and entered into the database as "6.1".

Janssen selected the HemoSense INRatio system for use at each site world-wide. The system involved a testing cuvette that could be stored at room temperature (unlike alternative systems, which required analogous materials to be refrigerated). A drop of blood was obtained from the patient with a capillary tube from a skin puncture site and placed on the test strip. The strip was then inserted into a slot on the hand-held monitor and results were reported as described above. Janssen states that between the start of the study in 2006 and 2009, over 3000 INRatio devices were purchased for use in ROCKET by Janssen from HemoSense and Inverness Medical Innovations, the company that acquired HemoSense in August 2007.

Importantly, sites were instructed to use only the INRatio device to adjust warfarin dose in ROCKET. However, this was not the only source of INR information from ROCKET. Sporadically, patients had open-label INR determinations in hospitals or emergency rooms, usually in connection with bleeding or invasive procedures. Most importantly, over 12,000 of the 14,236 treated patients in ROCKET took part in a sparse PK/PD study, with sampling at Weeks 12 and 24 of treatment during routine study visits. Blood samples were drawn at the study site and frozen for analysis at the Duke Hemostasis and Thrombosis Center after all samples had been gathered to examine rivaroxaban blood levels and their relationship to coagulation parameters, including prothrombin time (PT). Duke also reported INR with the PT results. Because the study was blinded, patients in both treatment arms were included in this PK/PD study, even though all of the PK/PD study objectives concerned only rivaroxaban. Data from this study were not provided to the study sites at any time and could not have been used to guide warfarin dosing in the trial.

Because the PK/PD blood samples were taken at routine study visits, in nearly all cases there was a study INR performed using the INRatio device on the same day as the PK/PD blood draws. The NDA submission in 2011 included the massive dataset containing INR values analyzed with the INRatio device and a much smaller dataset that included the INR values analyzed by a laboratory-based device at Duke. Both datasets were analyzed by Janssen, and

results related to rivaroxaban PK/PD relationships based on the PK/PD dataset were included in the NDA. However, Janssen states that they did not assess the degree of concordance between INRs obtained the same day in the two datasets until 2015, after learning of recall of the INRatio device because of reporting of low INR values compared to results from laboratory-based devices (see below). Likewise, DCRP did not assess INR concordance between the two datasets until after learning of the recalls in 2015. These assessments are discussed below in Sections **3** and **4** of this review, respectively.

2.2 Overview of Results of ROCKET

Relevant results of efficacy and safety outcomes in ROCKET are described in rivaroxaban labeling and the NDA clinical review. Key efficacy endpoint results are summarized below (**Table 1**). The ITT population and on-treatment analyses of the primary endpoint (time to first stroke or systemic embolism) each favored rivaroxaban numerically and satisfied the prespecified and also the FDA-preferred non-inferiority criteria for rivaroxaban vs. warfarin, but each missed demonstrating superiority of rivaroxaban by a small margin. It is notable that the entire observed advantage of rivaroxaban for stroke was based on results for hemorrhagic stroke. Only the ITT results are shown below. Results for all-cause mortality numerically favored rivaroxaban over warfarin.

The only notable safety risk of rivaroxaban was bleeding. Rates of ISTH major bleeding on-treatment (i.e., up to the last dose of study drug + 2 days) are shown in **Table 2.** Overall, major bleeding occurred at similar rates in the treatment arms. As in the other studies of the DOACs, intracranial hemorrhage (ICH), including hemorrhagic stroke and other forms of ICH, was more frequent in the warfarin arm. Fatal bleeding was also more frequent with warfarin. However, major GI bleeding was more frequent with rivaroxaban than with warfarin.

Table 1 ROCKET AF: Efficacy Results
All randomized patients followed to site notification[1]

	Rivaroxaban		Warfarin		Rivaroxaban vs. Warfarin
Event	**N=7081 n (%)**	**Event Rate (%/year)**	**N=7090 n (%)**	**Event Rate (%/year)**	**Hazard Ratio (95% CI)**
Primary Endpoint[2]	**269 (3.8)**	**2.1**	**306 (4.3)**	**2.4**	**0.88 (0.74, 1.03)**
Stroke	253 (3.6)	2.0	281 (4.0)	2.2	
Hemorrhagic Stroke	33 (0.5)	0.3	57 (0.8)	0.4	
Ischemic Stroke	206 (2.9)	1.6	208 (2.9)	1.6	
Unknown Stroke Type	19 (0.3)	0.2	18 (0.3)	0.1	
Non-CNS Systemic Embolism	20 (0.3)	0.2	27 (0.4)	0.2	
All-Cause Mortality†	**619 (8.8)**		**667 (9.4)**		
Vascular Death††	397 (5.6)		421 (5.9)		
Non-vascular Death	160 (2.2)		167 (2.3)		
Death of Unknown Cause	62 (0.9)		79 (1.1)		

1 Data are shown for randomized patients followed to the date of notification of sites that the study would end.
2 The primary endpoint was the time to first occurrence of stroke (any type) or non-CNS systemic embolism The upper limit of the 95% CI is less than 1.38, thus satisfying FDA's preferred test of non-inferiority. Data from one site with significant GCP issues were excluded from all efficacy analyses.
† Adjudicated deaths regardless of treatment or timing
†† Includes fatal bleeding and sudden or unwitnessed death
Note: In describing event rates, "events per 100 patient-years" and '%/year" yield identical results and are used interchangeably.
Note: Results for the primary endpoint include both fatal and non-fatal events.
Source: US labeling.

Table 2 ROCKET AF: Bleeding Results
(Treated patients followed to last dose of study drug + 2 days)

Parameter	Rivaroxaban N = 7111 n (%/year)	Warfarin N = 7125 n (%/year)	Rivaroxaban vs. Warfarin HR (95% CI)
Major Bleeding†	**395 (3.6)**	**386 (3.5)**	**1.04 (0.90, 1.20)**
Intracranial Hemorrhage (ICH)‡	55 (0.5)	84 (0.7)	0.67 (0.47, 0.93)
Hemorrhagic Stroke§	36 (0.3)	58 (0.5)	0.63 (0.42, 0.96)
Other ICH	19 (0.2)	26 (0.2)	0.74 (0.41, 1.34)
Gastrointestinal (GI)	221 (2.0)	140 (1.2)	1.61 (1.30, 1.99)
Fatal Bleeding#	27 (0.2)	55 (0.5)	0.50 (0.31, 0.79)
ICH	24 (0.2)	42 (0.4)	0.58 (0.35, 0.96)
Non-intracranial	3 (0.0)	13 (0.1)	0.23 (0.07, 0.82)

Abbreviations: HR = Hazard Ratio, CI = Confidence interval
* Major bleeding events within each subcategory were counted once per patient, but patients may have contributed events to multiple subcategories.
† Defined as clinically overt bleeding associated with a decrease in hemoglobin of ≥2 g/dL, a transfusion of ≥2 units of packed red blood cells or whole blood, bleeding at a critical site, or with a fatal outcome (ISTH major bleeding).

Source: US labeling
‡ Intracranial bleeding events included intraparenchymal, intraventricular, subdural, subarachnoid and/or epidural hematoma.
§ Hemorrhagic stroke in this table specifically refers to non-traumatic intraparenchymal and/or intraventricular hematoma.
# Fatal bleeding is adjudicated death, with a determination that bleeding was the primary cause of death

One aspect of the conduct of ROCKET AF deserves discussion here because it might be relevant to the impact on study outcomes of use of a device that reported falsely low INR values. The quality of anticoagulation control in warfarin arm of ROCKET was lower than in other trials of DOACs. Control of anticoagulation was assessed by the metric of "time in therapeutic range" (TTR) using the linear interpolation technique of Rosendaal. This technique has been used as a quality metric to assess anticoagulation in many clinical trials as well as in clinical practice. In essence, the TTR is an approximation of the percentage of time during treatment with warfarin when the parameter is within the target range. Mean TTR of individuals in the warfarin arms of industry-sponsored of DOACs in patients with AF is shown in **Table 3**. Data that were available to FDA during the review of the rivaroxaban NDA are highlighted in yellow. Italicized and bolded entries denote trials of drugs that have been approved in the US for use to prevent stroke and systolic embolism in patients with AF. Note that in all the studies included in the table, the target range of INR was 2.0 to 3.0, as it was in ROCKET.

Table 3 Mean TTR in Warfarin Arms of Industry-Sponsored Trials of DOACs in Patients with AF

Study Name (Experimental Drug)	Mean TTR (%)	Study Name (Experimental Drug)	Mean TTR (%)
ROCKET (rivaroxaban)*	***55***	***RE-LY (dabigatran)****	***64***
SPORTIF III (ximelagatran)	66	***ARISTOTLE (apixaban)****	***62***
SPORTIF V (ximelagatran)	68	***ENGAGE AF (edoxaban)****	***65***

* TTR was calculated individually for each patient in the study's warfarin arm and then averaged. For studies without an asterisk, it is not known how the mean TTR was calculated.
Italicized and bolded entries denote studies of products indicated in the US to reduce the rate of stroke and systemic embolism in patients with AF.
Data for ARISTOTLE and ENGAGE were not yet available at the time of the approval of rivaroxaban.
All studies had an INR target range of 2.0 to 3.0.

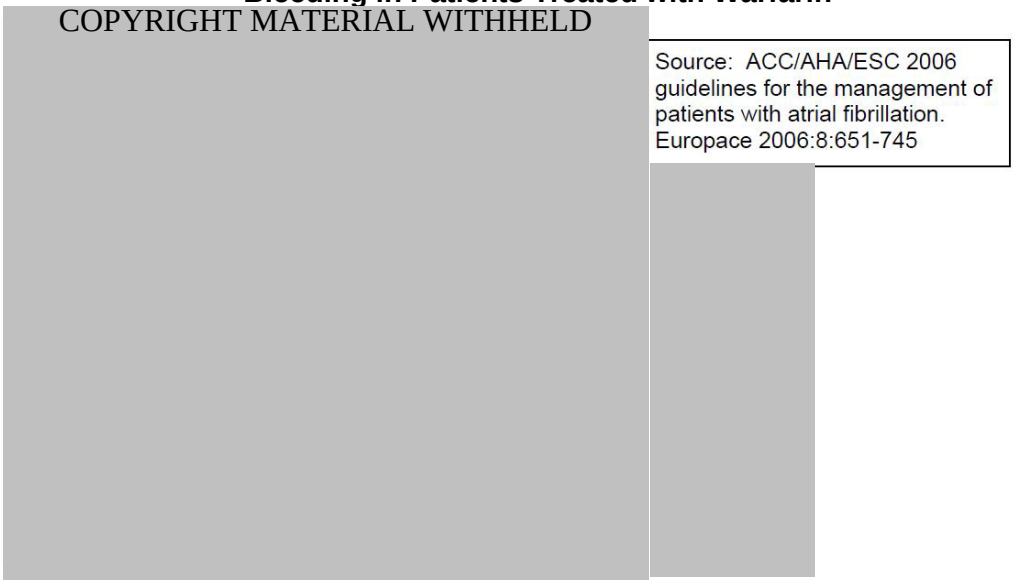
The mean TTR in the ROCKET warfarin arm, 55%, is the lowest value listed in the table. All other studies listed in the table had a mean TTR in their warfarin arms between 62% and 68%. The three warfarin-controlled studies for the other approved DOACs, namely, RE-LY (dabigatran), ARISTOTLE (apixaban) and ENGAGE AF (edoxaban), had warfarin arm mean TTRs of 64%, 62% and 65%, respectively.

The warfarin arm TTR calculations for ROCKET and the other approved DOACs were performed using the relevant study INR database. For ROCKET, as previously noted, the INR database was populated with data download from the IVRS system that was based on readings made by the INRatio devices at each study site.

The mean TTR of 55% in ROCKET means that about 45% of the time, warfarin arm subjects were not in the therapeutic range of INR. Janssen's data from the POC INR database indicate that over the course of the study, INR was below the therapeutic range (i.e., <2.0) 29% of the

time and above the therapeutic range (>3.0) 16% of the time. This is of concern with respect to antithrombotic efficacy (i.e., ischemic stroke prevention) because the ischemic stroke risk of patients with AF increases sharply as INR falls below 2.0, as shown in **Figure 1**, reproduced from the 2006 consensus guidelines for the management of AF. The risk of other types of bleeding in patients taking warfarin has the same basic relationship to INR as intracranial bleeding – as INR increases over the range of values usually obtained in clinical practice, the rate of bleeding increases.

Figure 1 Relationship between INR and Risks of Ischemic Stroke and Intracranial Bleeding in Patients Treated with Warfarin



If the true INR were higher than the POC INR, then the true time below the therapeutic range would likely be less than what was observed (29%). The true time in range and/or the time above range would tend to be increased over the observed values, and the warfarin arm, as a whole, would be somewhat more anticoagulated, and perhaps more patients were over-anticoagulated, than the POC INR results suggest. If this were the case, there would be a lower rate of ischemic stroke in the warfarin arm than what one might expect with the observed POC INR results. Also, there would be a higher rate of all types of bleeding than what one would expect in the warfarin arm with a 29% rate of time below therapeutic INR range. These perturbations in warfarin arm event rates from what one would expect with accurate INR readings would tend to affect the rivaroxaban vs. warfarin comparisons by making rivaroxaban look relatively less effective in preventing ischemic stroke compared to warfarin but also relatively less harmful compared to warfarin with respect to bleeding risk.

2.3 Commercial and Regulatory History of the INRatio Family of Devices

As noted above, when Janssen elected to use the INRatio device in ROCKET, the device was marketed by HemoSense, Inc. In August 2007, HemoSense was acquired by Inverness Medical Innovations, Inc. In 2010, Inverness changed its corporate name to Alere Inc. Various iterations of the INRatio device have been marketed by these companies from 2002 to the present day (**Table 4**). Recently, Alere announced it plans to remove INRatio products from

the US market (see **below**).

Table 4 Marketing History of the INRatio Family of Devices

Device Name	Applicant	510(K) Number	Decision Date
INRatio2 PT/INR Monitoring System;	Alere San Diego, Inc. (Formerly Biosite, Inc.)	K110212	05/01/2012
INRatio/INRatio2 Test Strips	Biosite Incorporated	K092987	06/11/2010
INRatio 2 Pt Monitoring System	HemoSense, Inc.	K072727	10/26/2007
INRatio Self-Test	HemoSense, Inc.	K021923	10/24/2002
INRatio PT Monitoring System	HemoSense, Inc.	K020679	05/06/2002

Source: CDRH Website
https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?start_search=1&Center=&Panel=&ProductCode=GJS&KNumber=&Model=&Applicant=HEMOSENSE%2C%20INC%2E&DeviceName=&Type=&ThirdPartyReviewed=&ClinicalTrials=&ExpeditedReview=&Decision=&DecisionDateFrom=&DecisionDateTo=07%2F22%2F2016&DeNovo=&IVDProducts=&CombinationProducts=&ZNumber=&PAGENUM=10&sortcolumn=DecisionDateDESC

The specific device used in ROCKET was the "INRatio PT Monitoring System®," cleared on 05/06/2002 under 510(k) K020679. The device was determined by CDRH to be substantially equivalent to the Roche CoaguChek S® system for PT and INR in whole blood.

Warning Letters sent to HemoSense, 2005 and 2006

Prior to Janssen's decision to use the INRatio device in ROCKET, HemoSense received two letters from FDA regarding deficiencies in its reporting of post-marketing adverse experiences associated with use of this device. We learned of these letters late in 2015.

Letter of 4 October 2005: Following an inspection from 16 May 2005 to 1 June 2005, the San Francisco District sent a warning letter to HemoSense with the following findings relevant to the subject of this review:

- Failure to submit an MDR (medical device report) to FDA regarding several cases where the INRatio device read substantially higher or lower than a subsequent INR result from a laboratory-based device.

Letter of 29 November 2006: Following an inspection from 15 May 2006 to 13 July 2006, the San Francisco District sent a warning letter to HemoSense with the following findings relevant to the subject of this review:

- Failure to report, evaluate, and/or investigate complaints of discrepant INR results between the INRatio device and a laboratory based device or retests with the INRatio device. These included cases where the INRatio results were lower than a lab-based device and also one case where the INRatio result was higher than a lab-based device.

> Two cases where the INRatio device read low were associated with hospitalization of the patient.

We view these two letters as being concerned with deficiencies in the reporting and investigational practices of HemoSense after receiving complaints regarding the INRatio device. While the letters describe cases where the INRatio device provided INR readings that were discrepant from near-contemporaneous readings from a laboratory-based device, the letters indicate that these cases should have been reported to CDRH and/or investigated by HemoSense. The letters did not state or imply that the observed discrepancies necessarily signaled that there was a deficiency in the performance of the device that required remediation.

Device Recalls, 2014

The next notable regulatory events relevant to the INRatio device – two recalls affecting the INRatio and INRatio2 systems – occurred in 2014. Janssen states that they became aware of the latter of these recalls in September 2015 and notified us about it later that month.

Recall of April 16, 2014
(http://www.fda.gov/MedicalDevices/Safety/ListofRecalls/ucm397509.htm)

Alere implemented a voluntary Class 1 recall in the US of the INRatio2 PT/INR Professional Test Strips. The recall resulted from "complaints of patients who had a therapeutic or near therapeutic INR with the Alere INRatio2 PT/INR Professional Test Strip but a significantly higher INR (outside of therapeutic range) when performed by a central laboratory." The root cause of this issue was not known, and Alere "could not determine the circumstances that might contribute to the discrepancy in results." The CDRH recall web page indicates that,

> "Alere San Diego received 9 reports of malfunctions; 6 injuries and 3 three deaths caused by bleeding. The firm is recalling the Alere INRatio2 PT/INR Professional Test Strips due to complaints of patients who had a therapeutic or near-therapeutic INR result with the Alere INRatio2 PT/INR Professional Test Strip but a significantly higher INR result (outside of therapeutic range) when re-testing was performed by a central laboratory because of deterioration in the patient's clinical condition.
>
> Use of this affected product may cause serious adverse health consequences, including death."

Alere notified customers to stop using the identified test strips immediately and use alternative methods to perform PT/INR testing, including use of an "alternative Alere product." Alere asked customers to return unused strips. Alere stated that they would "transition customers from the current Alere INRatio2 PT/INR Professional Test Strip to the Alere INRatio PT/INR Test Strip (PN 100139)," which was not affected by this recall.

Janssen stated that they did not notify us about the May 2014 recall because the recalled test strips were not used in ROCKET.

"Voluntary correction" implemented on December 5, 2014 (also a Class 1 recall)
(http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm396324.htm)

This action was aimed at US users of the Alere INRatio and INRatio2 PT/INR Monitor system (INRatio Monitor or INRatio2 Monitor and INRatio Test Strips). The monitors and test strips described in this correction were the only POC PT/INR monitors then marketed by Alere. Notably, INRatio Monitor and Test Strip versions that were used in ROCKET were included in this correction.

The correction indicated that,

> "In certain cases an INRatio and INRatio2 PT/INR Monitor system may provide an INR result that is clinically significantly lower than a result obtained using a reference INR system (laboratory method). This issue can arise if the patient has certain medical conditions or can occur if the instructions in the labelling for performing the test are not followed.
>
> "The INRatio and INRatio2 PT/INR Monitor system should not be used on patients with any of the following conditions:
>
> - Anemia of any type with hematocrit less than 30%
> - Any conditions associated with elevated fibrinogen levels including:
>   - Acute inflammatory conditions (examples may include acute viral or bacterial infections such as pneumonia or influenza)
>   - Chronic inflammatory conditions (examples may include rheumatoid arthritis, Crohn's disease, ulcerative colitis, infectious liver diseases such as hepatitis, or inflammatory kidney diseases such as diabetic nephropathy and glomerulonephritis)
>   - Severe infection (e.g., sepsis)
>   - Chronically elevated fibrinogen for any reason
>   - Hospitalized or advanced stage cancer or end stage renal disease patients requiring hemodialysis
> - Any bleeding or unusual bruising, clinically observed or reported by the patient
>
> Patients with any of the conditions listed above should immediately be transitioned to a laboratory INR method for monitoring their INR and warfarin therapy."

Under the heading "Reason for Recall," the notice indicates that "Alere received a total of 18,924 complaints of all types for the INRatio Test Strip (Optimize) from 2013-2014." A linked notice indicates that this was a Class 1 recall and implicated "Manufacturing and Distribution Dates: April 1, 2008 to December 4, 2014."

We understand that there was skepticism in 2014 within the group in CDRH with responsibility for devices that measure coagulation parameters, the Division of Immunology and Hematology Devices, regarding the validity of limiting the scope of this correction to the medical conditions named above (the "listed conditions"). It is possible that the listed conditions may simply represent those known to be present in the patients who had discrepant results between the INRatio device and another test of INR. It is notable that if this recall had had the analogous scope as the April 2014 recall (i.e., if it affected all use of the named products), then the INRatio product line would have been removed from the market as a practical matter. It is also notable that data from ROCKET indicate that there was no difference in the degree of discordance between INR assessed by the INRatio device and INR assessed by a laboratory-based device when warfarin arm patients with one of the conditions described in in the December 5, 2014 recall notice ("listed conditions") were compared to patients without one of these conditions.

Accordingly, we believe that Alere's claim that malfunctions of the INRatio ratio device are limited to occasions when it is used in patients with one of the listed conditions is spurious. There is more information regarding this issue in Section **3.**

Planned Device Withdrawal Announced July 11, 2016

In an announcement dated July 11, 2016, Alere stated that it would voluntarily withdraw all remaining INRatio products from the market in the US. A date for this withdrawal was not specified in the announcement. Alere indicated that the rationale for the planned withdrawal was FDA's request for such action, based on the Agency's position that Alere's submission in 2015 of information regarding "software enhancements" for the INRatio device does "…not adequately demonstrate the effectiveness of the software modification." The software enhancements were intended by Alere to ameliorate the INR discrepancies described above in connection with the December 2014 "voluntary correction."

3 Review of the Submissions by Janssen

3.1 Timing of Submissions

Janssen notified us of the December 5, 2014 INRatio recall in writing on September 29, 2015, after receiving a letter from Alere on September 24, 2015 indicating that some INRatio products used in ROCKET were affected by the recall. Janssen had telephoned the rivaroxaban NDA project manager several days prior to writing to us to tell us to expect a letter about this issue. This was more than 10 months after the recall. Janssen indicated to us that Alere did not notify Janssen about the December 2014 recall when it occurred because Alere limited notification to a list of customers with orders dating back to some unspecified date in 2009. Janssen stated that they had made no orders during the covered period. Janssen informed us that the covered period was crafted so that all Alere customers who could have unexpired INRatio Test Strips (which have 15 month expiration dating) would be notified about the recall.

Janssen stated that they learned of the possible applicability of the recall to ROCKET when they were advised of the recall by an unnamed "third party" and then contacted Alere on their own initiative for information regarding the INRatio products used the trial.

Since the first contact from Janssen in September 2015 regarding the recall they have made numerous submissions of information relative to the recall and the impact of the INRatio device malfunction on the results of ROCKET. Most of these submissions were made on their own initiative. They also made several submissions in response to information requests from us and EMA. Submissions to us were all made to IND 75,238.

3.2 Submitted Data

3.2.1 Relationship of POC INR to LAB INR

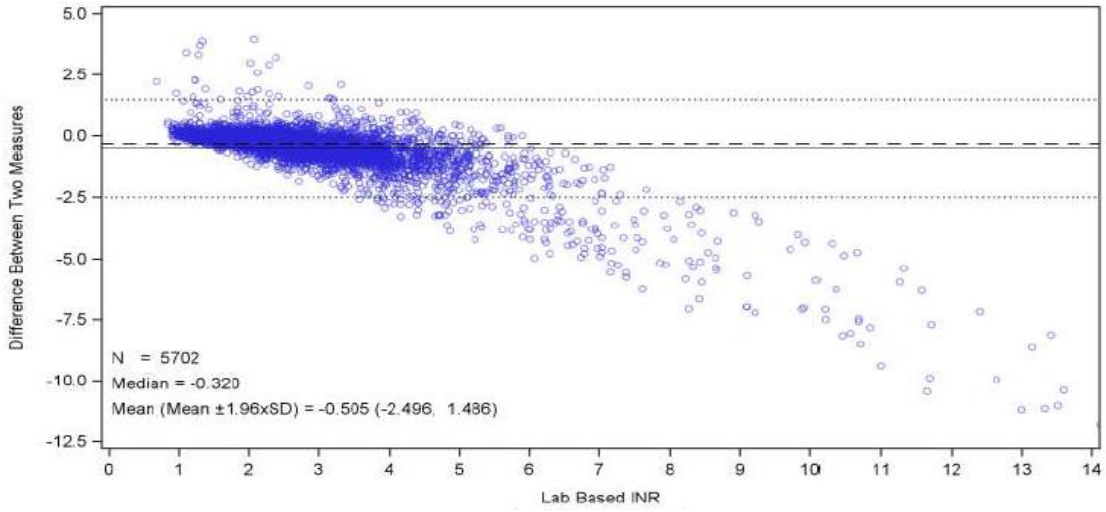
Data from ROCKET comparing POC INR values to same-day LAB INR values from the PK/PD dataset indicate that POC INR, on the average was about 13% lower than same-day LAB INR. This pattern is consistent with the information that supported the two recalls of the INRatio device in 2014.

As noted earlier, the ROCKET PK/PD study involved over 6000 patients in the warfarin arm. About 87% of patients in the warfarin arm had at least one same-day pair of INRs read with the INRatio device at the site (point-of-care or POC INR) and also read near the end of the study with a laboratory-based device at Duke University (LAB INR). Most of these patients had two such pairs of INRs.

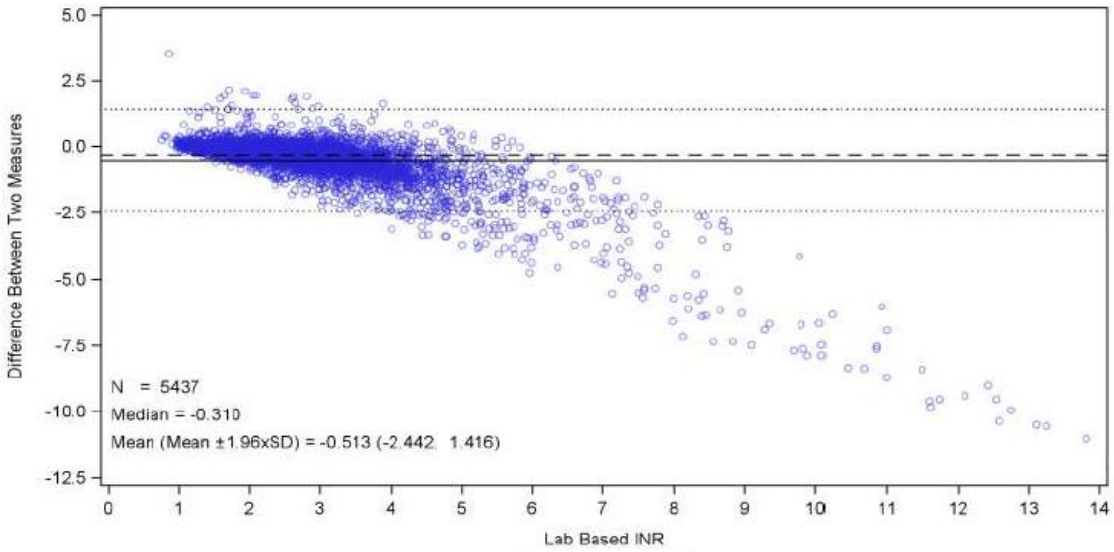
Figure 2 is a scatter plot that depicts the relationship between POC INR minus LAB INR (y axis) and LAB INR (x axis) for warfarin arm patients in ROCKET at Week 12 (A) and Week 24 (B). Note that patients with POC INR > 6.0 are excluded because such LAB INR values were truncated at 6.1, making the difference between POC INR and LAB INR difficult to interpret. At Week 12 (A), the mean ± 1.96 x SD for this parameter was -0.505 (-2.496, 1.486), and the median was -0.32. Mean and median values, respectively, were similar to these at Week 24 (B).

Figure 2 ROCKET AF: POC INR Minus LAB INR vs. LAB INR
Treated Patients, Warfarin Arm

A. Week 12



B. Week 24

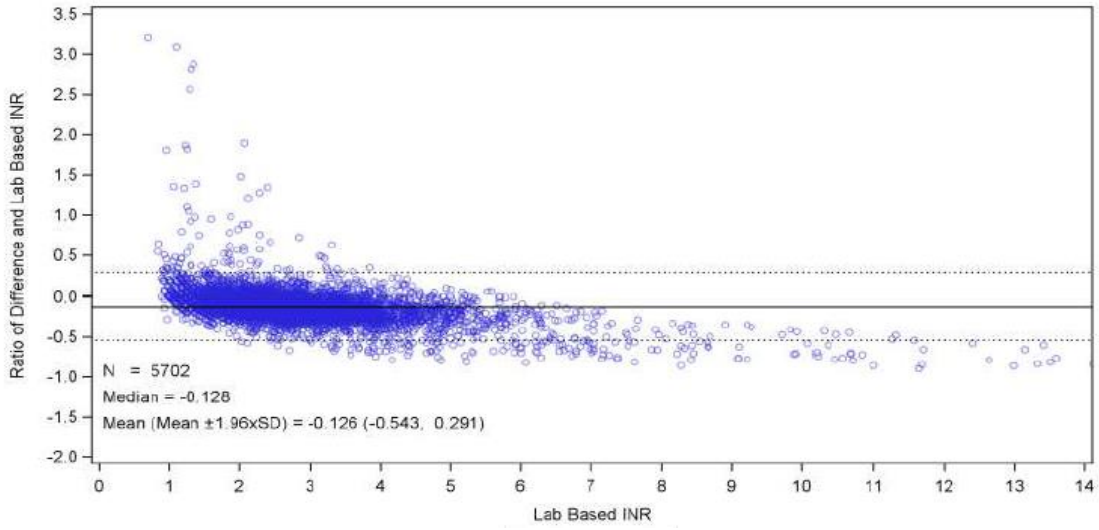


Note: Device based INR at least 6.1 is excluded.

The solid horizontal line is mean difference. The dashed horizontal line is median difference.
The upper and lower dotted horizontal lines are mean difference + and - 1.96 x Standard Deviation.
Source: Submission of December 14, 2015

The ratio of the difference of POC INR minus LAB INR to LAB INR at Week 12 is shown in **Figure 3**. The mean value of this parameter was -0.13 (i.e., POC INR was a mean of 13% less than LAB INR) and had a substantially less steep slope over the range of reported LAB INR than POC INR minus LAB INR. Analogous data for Week 24 were quite similar to the Week 12 data (data not shown).

Figure 3 ROCKET AF: Ratio of POC INR Minus LAB INR to LAB INR at Week 12
Treated Patients, Warfarin Arm



Note: Device based INR at least 6.1 is excluded.
The solid horizontal line is mean difference. The dashed horizontal line is median difference. (They are overlaid).
The upper and lower dotted horizontal lines are mean difference + and - 1.96 x Standard Deviation.
Source: Submission of December 14, 2015

Janssen also analyzed the paired POC and LAB INR values by categories at Week 12 (**Table 5**) and Week 24 (**Table 6**). The two tables show similar patterns with respect to the relationship between POC INR (one category per row) and LAB INR (one category per column). Ignoring the total rows and column, for each POC INR category row, the cell with the largest value is the one with a concordant LAB INR value. For example, in either table, in the row for POC INR 2 – 3, the column with the largest value is LAB INR 2 – 3. However, when the POC INR and LAB INR values are not concordant, LAB INR is likely to be higher than POC INR except when POC INR is the in the highest category (>4). This is consistent with the data for POC INR minus LAB INR in previous tables and figures. Also, the total percent of patients with POC INR in range (2 – 3) was higher than the total with LAB INR in range in each table. As one might expect from the data previously shown, at Week 12, those who were out of range were more likely have to have LAB INR > 3 (35% of all subjects) than POC INR > 3 (18%). A similar pattern was observed at Week 24 (38% vs. 18%, respectively).

Table 5 ROCKET AF: Janssen's Categorical Analysis of POC INR vs. LAB INR at Week 12

Treated Patients, Warfarin Arm

	Lab INR				
	<2 n (%)	**2 - 3 n (%)**	**>3 - 4 n (%)**	**>4 n (%)**	**Row Total n (%)**
Column Total → ↓POC INR	**1356 (23.5)**	**2405 (41.7)**	**1238 (21.5)**	**767 (13.3)**	**5766 (100)**
<2	1239 (21.5)	604 (10.5)	54 (0.9)	47 (0.8)	**1944 (33.7)**
2 – 3	96 (1.7)	1740 (30.2)	793 (13.8)	172 (3.0)	**2801 (48.6)**
3 – 4	8 (0.1)	50 (0.9)	358 (6.2)	291 (5.0)	**707 (12.3)**
>4	13 (0.2)	11 (0.2)	33 (0.6)	257 (4.5)	**314 (5.4)**

Note: Percentages calculated with the total number of subjects as denominator.
Cells with concordant results for LAB INR and POC INR are highlighted in yellow
Source: Submission of December 14, 2015

Table 6 ROCKET AF: Janssen's Categorical Analysis of POC INR vs. LAB INR at Week 24

Treated Patients, Warfarin Arm

	Lab INR				
	<2 n (%)	**2 - 3 n (%)**	**>3 - 4 n (%)**	**>4 n (%)**	**Row Total n (%)**
Column Total → ↓POC INR	**1103 (20.0)**	**2394 (43.5)**	**1260 (22.9)**	**750 (13.6)**	**5507 (100)**
<2	983 (17.9)	548 (10.0)	61 (1.1)	48 (0.9)	**1640 (29.8)**
2 – 3	101 (1.8)	1790 (32.5)	805 (14.6)	200 (3.6)	**2896 (52.6)**
3 – 4	9 (0.2)	37 (0.7)	369 (6.7)	276 (5.0)	**691 (12.5)**
>4	10 (0.2)	19 (0.3)	25 (0.5)	226 (4.1)	**280 (5.1)**

Note: Percentages calculated with the total number of subjects as denominator.
Cells with concordant results for LAB INR and POC INR are highlighted in yellow
Source: Submission of December 14, 2015

FDA also performed categorical analyses of INR, but we used 6 categories rather than 4 to better appreciate the rates of gross discordance of POC and LAB INR (**Table 10**). There is additional information regarding the degree of INR discordance in the discussion of that table.

3.2.2 Relationship between POC INR vs. LAB INR Differences and Bleeding

Janssen's analyses of bleeding rates in patients with and without discordance between POC INR and LAB INR indicated that patients whose LAB INR was in a higher INR group than their POC INR group had higher rates of major bleeding and non-major bleeding than patients whose LAB INR and POC INR groups were concordant, although the difference in rates was not large.

Janssen analyzed the effects of discordancy between POC INR group and LAB INR group on bleeding rates using the same groupings as in **Table 6.** Results at Week 12 for patients with no

discordance, LAB INR at least one group higher than POC INR, and LAB INR at least 2 groups higher than POC INR, are shown in **Table 7**. Rates for all treated patients in the rivaroxaban and warfarin arms are shown for contrast. In general, warfarin arm bleeding rates are increased in patients with discordance vs. no discordance. This is most evident in the most common grade of bleeding, major + clinically relevant non-major bleeding. The trend is less evident for the less common types of bleeding shown in the table, including major GI and intracranial bleeding. Bleeding rates in patients with same-day POC and LAB INR pairs at Week 24 showed the same pattern as the Week 12 data (data not shown). Of note, patients without a pair of same-day POC and LAB INR values at Week 12 (N=1359) or Week 24 (N=1618) had higher rates of major + clinically relevant non-major bleeding than any of the groups included in Table 7, 23.44 and 24.98 events per 100 patient-years, respectively.[3] They also had higher rates of the primary endpoint than patients who had matched INR pairs at these time points (data not shown). These findings are probably due in some part to the fact that patients with study endpoints before Week 12 or Week 24 might discontinue follow-up prior to the blood draws for INR.

Table 7 ROCKET AF: Warfarin Arm Bleeding Rates in Patients with Varying Degrees of Discordancy between POC and LAB INR at Week 12

(Treated patients followed to last dose of study drug + 2 days)

	Rivaroxaban	Warfarin	Warfarin - Degree of Discordancy[1]		
Bleeding Type[2]	**All treated patients N=7111**	**All treated patients N=7125**	**No Discordancy N=3594**	**LAB INR Group ≥1 Higher than POC INR Group N=1961**	**LAB INR Group ≥2 Higher than POC INR Group N=273**
Major	3.60	3.45	2.74	3.17	2.92
Intracranial	0.49	0.74	0.62	0.61	0.83
GI	2.00	1.2	1.10	1.63	0.42
Major + clinically relevant non-major	14.91	14.52	12.88	14.48	16.59

[1] Assessed as LAB INR Group minus POC INR Group (Groups: 1, <2; 2, 2-3; 3, >3-4; 4, >4)
[2] Bleeding rates are expressed as events per 100 patient-years.
Source: Submission of December 23, 2015

3.2.3 Analyses of POC INR vs. LAB INR in Patients with or without the Medical Conditions listed in the December 2014 Recall of the INRatio Device.

Note that the above data are not informative with respect to the validity of Alere's claim in the recall notice of December 4, 2014 regarding the medical history-specific nature of the difference between INRs measured by the INRatio device and a laboratory based device. However, Janssen performed analyses of the ROCKET data that compared differences between POC INR and LAB INR in patients with vs. without one of the conditions listed in the recall notice. These analyses showed that there was no difference between the two cohorts of patients in the performance of the INRatio device relative a laboratory based device, thus suggesting that

[3] Source: EMA>CHMP report, February 5, 2016. No other bleeding types were included in this subgroup analysis.

Alere's claim that malfunctions of the INRatio device are limited to patients with one of the listed conditions is false.

In the most simple analyses, Janssen compared the mean and median of the differences between POC INR and LAB INR, as well as the ratio of the difference to LAB INR, in warfarin arm patients with or without "recall-related time" (RRT, defined in the next paragraph) that affected an individual INR determination. As discussed below, these various metrics did not vary notably between patients with or without INRs affected by RRT.

Whether an individual INR determination at Week 12 and/or 24 was affected by RRT was determined in a blinded fashion by two physicians based on each subject's trial-related medical history, laboratory data, and AE information. This information was used to determine whether the subject had one of the conditions associated with malfunction of the device that were listed in the Alere INRatio recall notice of December 5, 2014. Note that some of these of conditions are chronic and some are time-limited. Our understanding is that if the subject had a qualifying condition that was chronic, all INRs drawn after the onset date of the condition were considered to occur during recall-related time. If the qualifying condition was acute, the INR was considered to be drawn in recall-related time if (1) the start date of the condition was prior the INR draw date and (2) the condition had not resolved by the date of the last INR draw previous to the one at issue.[4]

Data for the various metrics described above at Week 12 and Week 24 are shown in **Table 8**, which contrasts results for INR determinations that were and were not affected by RRT. If Alere's claim that the INRatio device reported low INRs only in patients with one of the conditions listed in the December 5, 2014 recall notice, then results in the YES column would show appreciably greater differences between POC and Lab INR than the NO column. However, the data in the two columns are very similar for the mean difference between POC and LAB INR, the median difference, and ratio of the difference to LAB INR. All cells are consistent in showing that POC INR tended to be lower than LAB INR. These data do not support the validity of Alere's claim.

[4] To illustrate how this worked: Assume that a patient had POC and LAB INR draws at both Week 12 and 24, with regular POC draws at Weeks 16 and 20. Also assume that the patient had one of the transient conditions listed in the announcement of the December 5, 2014 recall (such as pneumonia) that lasted from week 15 to 17. Then, neither the Week 12 or Week 24 INR data would be considered to be affected by RRT because the condition started after the Week 12 draw and its end date was prior to the last INR drawn before the Week 24 draw, i.e., the Week 20 draw. However, if the pneumonia lasted from Week 19 to Week 21, the Week 24 INR would be considered to be affected by RRT because the Week 20 draw occurred before the end date of the pneumonia. The Week 12 INR would not be affected by RRT.

Table 8 ROCKET AF: Janssen's Analyses of INR Parameters in Warfarin Arm Patients With or Without "Recall-Related Time"
Treated Patients, Warfarin Arm

	INR Affected by Recall-Related Time?	
Parameter	**YES (N=769 WK 12) (N=735 WK 24)**	**NO (N=4933 WK 12) (N=4702 WK 24)**
Week 12 POC INR-LAB INR Median	-0.32	-0.32
Week 12 POC INR-LAB INR Mean ± 1.96 x SD	-0.53 (-2.7, 1.7)	-0.50 (2.5, 1.5)
Week 12(POC INR-LAB INR)/LAB INR Median	-0.13	-0.13
Week 12(POC INR-LAB INR)/LAB INR Mean ± 1.96 x SD	-0.13 (0.56, 0.31)	-0.13 (-0.54, 0.29)
Week 24 POC INR-LAB INR - Median	-0.36	-0.31
Week 24 POC INR-LAB INR) Mean ± 1.96 x SD	-0.50 (-2.2, 1.2)	-0.51 (-2.5, 1.5)
Week 24(POC INR-LAB INR)/LAB INR Median	-0.14	-0.12
Week 24(POC INR-LAB INR)/LAB INR Mean ± 1.96 x SD	-0.14 (-0.49, 0.23)	-0.13 (-0.51, 0.25)

Source: Submission of December 14, 2015

Also, Janssen used a conservative modification of ISO 2007 standard for assessment of INR measuring devices as the basis of determining whether there was a "discrepancy" between the POC INR assessments at Weeks 12 and 24 and same-day assessment of INR that were analyzed at Duke for the PK/PD study in ROCKET. The percentage of patients with an INR discrepancy was compared in patients with our without RRT. A discrepancy was defined as follows:

- For LAB INR < 2, POC INR more than 0.5 units higher or lower than LAB INR
- For LAB INR ≥ 2, POC INR more than 30% higher or lower than LAB INR

Results are shown in **Table 9**. One would expect subjects with recall-related time to have a higher percentage of discrepant results (Discrepancy = YES) than those without recall-related time, but the discrepancy rate is quite similar in the two cohorts, and is actually slightly lower in the former than the latter subjects. This is true both for Week 12 and Week 24. Thus, like the results of Janssen's other analyses described above, these results do not support the validity of Alere's claim that only subjects with one of the conditions listed in the recall notice of Dec. 5, 2014, are prone to discrepant results with use of the INRatio device. However, it does not tell us anything about how the discrepant INR results in some patients affected the results of ROCKET.

Table 9 ROCKET AF: Impact of "Recall-Related Time" on Discrepancy of INR Results
Treated Patients, Warfarin Arm

Subject Group	Week 12 (N=5766) n (%)	Week 24 (N=5507) n (%)
With Recall-related Time	**775 (100)**	**745 (100)**
Discrepancy -		
YES	97 (12.5)	90 (12.1)
NO	672 (86.7)	645 (86.6)
Without Recall-related Time	**4991 (100)**	**4762 (100)**
Discrepancy -		
YES	653 (13.1)	630 (13.2)
NO	4280 (85.8)	4072 (85.5)

Note: Percentages calculated with the N of subjects per arm in each recall-related time category as denominator.
Note: Discrepancy is defined as POC INR is outside of LAB INR +/- 0.5 if LAB INR < 2, and POC INR is outside of LAB INR +/- 30% of LAB based INR if LAB INR ≥ 2
Subjects with POC INR > 6 are not included
Source: Submission of Nov. 16, 2015

Janssen and their consultants, the Duke Clinical Research Institute, have made several analyses of clinical outcomes in ROCKET that contrast results in patients with vs. without recall-related time. As one might expect from the data presented above, these analyses showed that there was no effect of recall-related time on the safety and efficacy of rivaroxaban relative to warfarin. The recall-related time analyses of outcomes are not discussed further in this review.

At our request, Janssen also modeled the effects of inaccurate readings by the INRatio device on clinical outcomes in ROCKET AF. These analyses are discussed below along with FDA's models in Sec. **5.**

4 Analyses of Janssen's Data Conducted by FDA

We performed our own analyses of the relationship between POC INR and a same-day LAB INR, as well as the effect of differences between POC and LAB INR on the rate of bleeding. Although our analytic methodology was different from that of Janssen, our results are consistent with those of Janssen regarding both of these issues.

4.1 Relationship of POC INR to LAB INR

Table 10 is a display of POC vs. LAB INR categories for 6225 warfarin patients who had a Week 12 or Week 24 matched pair of INR values.[5] For patients with pairs at both visits, only the Week 24 value was used in the analysis, so no patient is represented more than once.

[5] A handful of INR pairs that included a reported POC INR value >6.1 were excluded from this analysis because such values should have reported as 6.1 in the database according to information regarding data standards for POC INR information in the ROCKET AF study report. We suspected that such values might have reported by some device other than the Alere INRatio device provided by Janssen to the study sites.

Categories were INR <1.5, 1.5 to <2, 2 to 3, >3 to 4, >4 to 6, and >6. This analysis differs from the analysis performed by Janssen in that there is one additional INR category at the low end of INR and one additional category at the high end of INR. These additional categories are associated with patients at high risk of thromboembolic events and bleeding, respectively. Such patients are often candidates for immediate warfarin dose adjustment (rather than waiting for a confirmatory out-of-range INR value at a subsequent draw before adjusting dose), or if INR is high, temporary interruption of warfarin therapy until INR is in the therapeutic range.

In this table, each row represents a POC INR category and each column represents a LAB INR category. The percentage value to the right of each n was calculated using a denominator that is the total number of patients in the analysis, 6225.

As before, in row the cell representing the largest number of subjects is the one with a concordant category for POC and LAB INR. However, that is not true in all the columns. In the column for LAB INR category 4 (INR >3 to 4), the modal value is in the row for POC INR category 3 (INR 2 to 3, the target range). If we assume that the LAB INR is closer to the truth than the POC INR, which is a reasonable assumption, then a person with an INR between 3 and 4, i.e., that is above range and probably at increased risk of bleeding and who probably would benefit from a warfarin dose reduction, would be more likely to have a POC reading that indicates that he or she is in range. The investigator then would have no reason to adjust this patient's dose or bring the patient back in a week or two for a confirmatory blood draw.

In the column corresponding to LAB INR category 5 (INR >4 to 6), bleeding risk is more than slightly elevated. Here the modal value for POC INR is category 4 (INR of >3 to 4). This value is above the target range, but an investigator looking at the INR readout might not think that a dose reduction is urgent. Also, of 628 patients with a LAB INR category 5 reading, 172 (27%) had a POC INR in category 3, in the target range. The patients with this level of LAB INR are at moderately elevated risk of bleeding, and the proper course of action would be to lower the warfarin dose immediately or at a minimum, bring the patient back early for a repeat INR. However, the study physician for these patients received a POC INR value that indicated that the INR was in the target range. There probably would be no reason to believe otherwise unless the patient had active bleeding or bruising at the time of the blood draw.

The data from this table also indicate that on the day of the reading used in the analysis, about 52% of patients were in the target range of INR on the basis of the POC INR reading, but only 43% were in range on the basis of the LAB INR reading. Most of the patients who were not in the therapeutic range on the basis of LAB INR had LAB INR values above the therapeutic range.

These data suggest that warfarin arm patients in ROCKET may have been at increased risk of bleeding due to higher than recognized INR values. While the bleeding in the rivaroxaban arm would not have been affected by increased INR in the warfarin arm, the comparison of bleeding rates would have been tilted against warfarin, thus making rivaroxaban look comparatively better than it might have looked in a trial that used an accurate device to assess INR. In addition, depending on the true shape of the INR response curve for ischemic stroke and the extent to which patients were deliberately or inadvertently maintained with a POC INR values less than 2, patients in the warfarin arm might have benefited from a higher degree of anticoagulation and had fewer ischemic strokes than they would have had if INR had been reported accurately at the study sites. This would have tilted the comparison in favor of warfarin

for ischemic stroke, and possibly made rivaroxaban look relatively worse than it really is in preventing ischemic stroke.

For a more visually oriented approach to the analysis of INR categories, see Appendix 1, **below**.

Table 10 ROCKET AF: Categorical Analysis of POC vs. LAB INR at Week 12 or 24
Treated Patients, Warfarin Arm

	LAB INR Group													
POC INR Group↓	1		2		3		4		5		6		Total	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
1	376	6.0	137	2.2	58	0.9	17	0.3	12	0.2	7	0.1	**607**	**9.8**
2	68	1.1	580	9.3	569	9.1	51	0.8	34	0.5	10	0.2	**1312**	**21.1**
3	5	0.1	104	1.7	1990	32.0	908	14.6	172	2.8	56	0.9	**3235**	**52.0**
4	0	0.0	9	0.1	37	0.6	400	6.4	253	4.1	33	0.5	**732**	**11.8**
5	2	0.0	0	0.0	8	0.1	30	0.5	150	2.4	70	1.1	**260**	**4.2**
6	4	0.1	6	0.1	12	0.2	5	0.3	7	0.1	45	0.7	**79**	**1.3**
Total	**455**	**7.3**	**836**	**13.4**	**2674**	**43.0**	**1411**	**22.7**	**628**	**10.1**	**221**	**3.6**	**6225**	**100**

Notes: INR Groups: 1, <1.5; 2, 1.5 to <2; 3, 2 to 3; 4, >3 to 4; 5, >4 to 6; 6, > 6
Includes patients with a same-day pair of POC and LAB INR values at the Week 12 or Week 24 study visits. For patients with pairs at both visits, the Week 24 value was used.
For cells in the "Total" row or column, the denominator for calculating percentages is 6225. For other cells, the denominator is the Total n in the cell at the bottom of the relevant column.
Cells highlighted in yellow represent patients with concordant values for POC and Lab INR groups.

4.2 Relationship between POC INR vs. LAB INR Discordance and Bleeding

To determine whether malfunction of use of INRatio device might have affected outcomes in ROCKET, as a preliminary analysis we examined outcome rates in subgroups of patients with same-day pairs of POC INR and LAB INR at Week 12 and/or Week 24. These subgroups include those with LAB INR >3 and POC INR <3; LAB INR >4 and POC INR <3; and LAB INR group – POC INR group ≥ 2, using the same INR groups as in **Table 10.** Endpoints analyzed included key study safety and efficacy parameters (see **Table 11**). Our expectation was that warfarin arm subjects with these characteristics probably would have been over-anticoagulated and thus would have a higher rate of bleeding and a possibly a lower rate of ischemic stroke than the study population as a whole or the entire subgroup with at least one same-day pair of POC and LAB INR values, whose data are also included in the table for each endpoint of interest. Because INR from any source was not used to adjust rivaroxaban dose, we did not expect patients in the rivaroxaban arm with LAB INR>POC INR to be different from those in entire subgroup of patients with at least one same-day POC and LAB INR pair.

Table 11 ROCKET AF: Selected Outcomes in Patients with Same-Day POC INR/LAB INR Pairs at Week 12 and/or Week 24

Safety Population

	Rivaroxaban		Warfarin		R vs. W	
ENDPOINT / Population or Subgroup	**n/N**	**ER**	**n/N**	**ER**	**HR**	**95% CI**
MAJOR BLEEDING						
All treated patients	395 / 7111	3.60	386 / 7125	3.45	**1.04**	(0.90, 1.20)
With ≥1 POC/LAB INR same –day pair[1]	337 / 6133	3.23	314 / 6225	2.95	**1.09**	(0.94, 1.28)
With LAB INR >3 & POC INR ≤3[2]	35 / 520	4.05	94 / 1698	3.16		
With LAB INR >4 & POC INR ≤3[2]	14 / 233	3.68	15 / 381	2.25		
With LAB INR Group – POC INR Group ≥2[3]	17 / 284	3.60	29 / 464	3.68		
FATAL BLEEDING						
All treated patients	38 / 7111	0.34	63 / 7125	0.56	**0.61**	(0.41, 0.92)
With ≥1 POC/LAB INR same –day pair[1]	34 / 6133	0.32	45 / 6225	0.42	**0.77**	(0.49, 1.20)
With LAB INR >3 & POC INR ≤3[2]	5 / 520	0.57	13 / 1698	0.43		
With LAB INR >4 & POC INR ≤3[2]	2 / 233	0.52	4 / 381	0.59		
With LAB INR Group – POC INR Group ≥2[3]	2 / 284	0.42	5 / 464	0.62		
PRIMARY ENDPOINT						
All treated patients	191 / 7111	1.71	244 / 7125	2.15	**0.79**	(0.66, 0.96)
With ≥1 POC/LAB INR same –day pair[1]	160 / 6133	1.51	189 / 6225	1.75	**0.86**	(0.70, 1.06)
With LAB INR >3 & POC INR ≤3[2]	15 / 520	1.70	51 / 1698	1.69		
With LAB INR >4 & POC INR ≤3[2]	7 / 233	1.81	16 / 381	2.38		
With LAB INR Group – POC INR Group ≥2[3]	8 / 284	1.67	21 / 464	2.63		
ISCHEMIC STROKE						
All treated patients	151 / 7111	1.35	162 / 7125	1.43	**0.95**	(0.76, 1.18)
With ≥1 POC/LAB INR same –day pair[1]	127 / 6133	1.20	125 / 6225	1.16	**1.03**	(0.81, 1.32)
With LAB INR >3 & POC INR ≤3[2]	10 / 520	1.14	30 / 1698	0.99		
With LAB INR >4 & POC INR ≤3[2]	5 / 233	1.30	8 / 381	1.19		
With LAB INR Group – POC INR Group ≥2[3]	5 / 284	1.04	11 / 464	1.38		
HEMORRHAGIC STROKE						
All treated patients	29 / 7111	0.26	50 / 7125	0.44	**0.59**	(0.37, 0.93)
With ≥1 POC/LAB INR same –day pair[1]	23 / 6133	0.22	37 / 6225	0.34	**0.63**	(0.38, 1.07)
With LAB INR >3 & POC INR ≤3[2]	3 / 520	0.34	11 / 1698	0.36		
With LAB INR >4 & POC INR ≤3[2]	1 / 233	0.26	4 / 381	0.59		
With LAB INR Group – POC INR Group ≥2[3]	1 / 284	0.21	7 / 464	0.87		

(Continued on next page)

	Rivaroxaban		Warfarin		R vs. W	
	n/N	ER	n/N	ER	HR	95% CI
ALL-CAUSE DEATH						
All treated patients	210 / 7111	1.88	252 / 7125	2.22	**0.85**	(0.70, 1.02)
With ≥1 POC/LAB INR same –day pair[1]	152 / 6133	1.43	193 / 6225	1.79	**0.80**	(0.65, 0.99)
With LAB INR >3 & POC INR ≤3[2]	15 / 520	1.70	53 / 1698	1.75		
With LAB INR >4 & POC INR ≤3[2]	8 / 233	2.07	15 / 381	2.23		
With LAB INR Group – POC INR Group ≥2[3]	11 / 284	2.29	17 / 464	2.12		

Notes: All endpoints are analyzed as time to first event
ER=event rate (events per 100 patient-years); HR= hazard ratio; CI=confidence interval
Primary Endpoint is first event of stroke (any type) or systemic embolism
[1] Patients with same-day of POC INR and LAB INR values at Week 12 and/or Week 24
[2] The specified criteria were met at Week 12, Week 24 or both.
[3] Lab INR group minus POC INR group ≥ 2; criteria met at Week 12, Week 24 or both; INR groups are defined as in **Table 10** (INR Group 1, <1.5; 2, 1.5 to <2; 3, 2 to 3; 4, >3 to 4; 5, >4 to 6; 6, > 6)

Major bleeding rates in both treatment arms were lower in the large subgroup of patients with at least one same-day POC INR/LAB INR pair than in all treated patients. This might be due to differences in the patients that took part in the PK/PD study and those that did not, or it might simply be due to the fact that patients who bled prior to Week 12 or 24 may have dropped out of the study prior to their blood draws, but these conjectures have not been explored. As we expected, warfarin arm patients with LAB INR>3 and (on the same day) POC INR <3 had a higher rate of major bleeding than the large warfarin arm subgroup with same-day pairs. Bleeding events were sparse in the considerably smaller subgroups that were limited to those with larger differences between LAB and POC INR (those with LAB INR >4 and POC INR <3, and those with LAB INR Group – POC INR Group ≥2), and the major bleeding data cannot be interpreted.

There were relatively few fatal bleeding events and hemorrhagic strokes in patients with the analyzed INR disparities, and rates of these endpoints did not show the expected pattern.

As with major bleeding, the rate in the warfarin arm of primary endpoint events was lower in each treatment arm in the population with at least one same-day POC INR/LAB INR pair of values. In the warfarin arm, the subgroup with LAB INR >3 and POC INR <3 had a lower event rate than the entire subgroup of patients with at least one matched pair of POC/LAB INR values. Events were sparse in groups limited to those with patients with greater degrees of difference between LAB INR and POC INR. Rates of ischemic stroke, the endpoint that comprised the majority of primary endpoint events, had the same pattern as primary endpoint events in the warfarin arm, as one might expect if time below therapeutic range was low or if there is a persistent negative slope of the INR/ischemic stroke response curve as INR increases. Rates of all-cause death in the warfarin arm had a similar pattern as the primary endpoint and ischemic stroke in cells with at least 50 deaths.

These data are not dispositive of the question of whether use of the INRatio device in ROCKET affected the outcome of the study. However, the pattern of the event rates for major bleeding, and perhaps the rates of the primary endpoint and ischemic stroke, suggest that these outcomes might have been affected in the manner we expected in the patients with LAB INR >

POC INR. Section **5** of this review explores modeling of the data by FDA and Janssen to explore this issue further.

4.3 Effects of Storage Time on Ratio between LAB INR and POC INR

ROCKET, like other studies of the DOACs in patients with atrial fibrillation, was a large and lengthy study. Because LAB INR samples were frozen for various periods and analyzed in batches near the end of the study, we were concerned that the time between the LAB INR blood draw and analysis of the sample at Duke might have affected the reported LAB INR and thus perhaps accounted for some of the differences between POC INR and LAB INR. However, our analyses indicate that it is very unlikely that there was a meaningful effect of time in storage on the observed disparities between POC INR and same-day LAB INR in ROCKET.

Dr. John Lawrence of the Office of Biostatistics 1 investigated the potential impact of storage time on the reported results of the LAB INR samples analyzed at Duke. As noted above, these samples were frozen and stored regionally and then shipped to Duke for analysis near the end of the study. Accordingly, the overall timeline of the ROCKET trial is relevant:

- First patient randomized: 18 December 2006
- Last patient randomized: 17 June 2009
- Site notification that the event target had been reached and sites were to schedule end-of-study visits and perform other close-out procedures: 17 June 2010 [6]
- Last patient contact: 15 September 2010

Blood was drawn for the LAB INR measurements in the PK/PD study at Week 12 and Week 24 of study treatment. Janssen describes sample handling as follows:

> "…samples were collected in tubes containing 3.2% buffered sodium citrate additive. Samples were to be centrifuged within 1 hour of collection and plasma separated from cells immediately after centrifugation. The plasma sample tubes were to be covered with dry ice (i.e., shock frozen at -78º C), if available, and stored at < -18º C until shipment to the [regional] (b) (4) t. Once at the (b) (4), samples were stored at -70º C. Samples were sent from the (b) (4) to the reference laboratory (Duke Hemostasis & Thrombosis Center Core Laboratory) on dry ice."

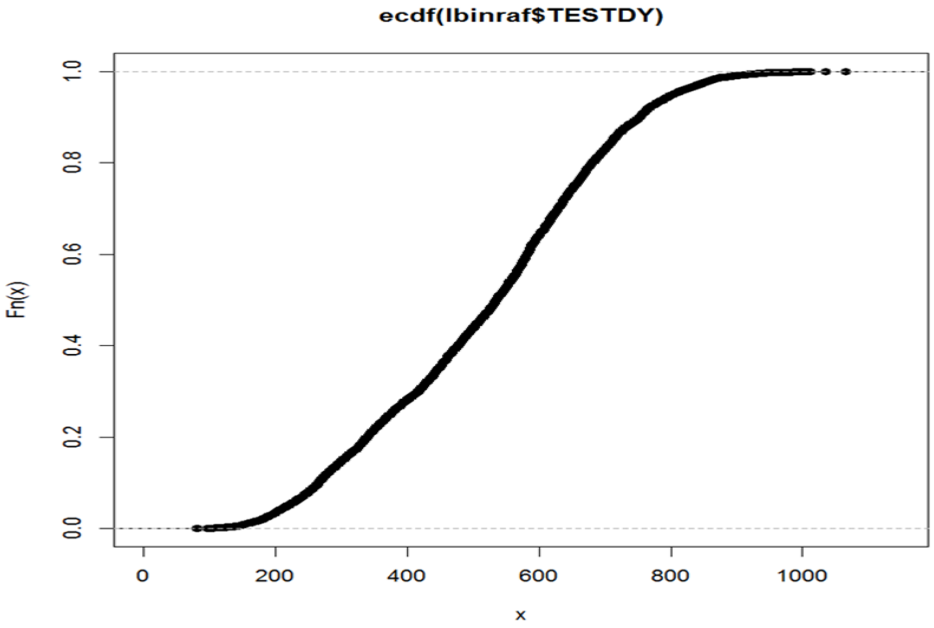
It is not clear how long the samples were kept at the study sites before shipment to (b) (4). The LAB INR samples were run at Duke in batches on 93 dates over the period from 22 December 2009 to 8 July 2010.

Janssen provided us with the dates that individual LAB INR samples were drawn as well as the dates that each sample was run at Duke for 5766 Week 12 samples and 5507 Week 24 samples (a total of 11,273 samples), all with matching same-day POC INR values. The following statistics summarize the number of days between the date of the sample blood draw and the date the sample was analyzed at Duke:

[6] This date applied to all countries except South Africa, where site notification occurred on 01 April 2010. Study sites in South Africa were closed early to avoid potential disruptions in study procedures caused by the FIFA (soccer) World Cup, which involved matches at multiple venues in South Africa in June and July of 2010.

- mean, 520;
- minimum, 80;
- 1st quartile, 375;
- median, 520;
- 3rd quartile 654; and
- maximum, 1066 (see **Figure 4**).

Figure 4 Days between Blood Draw and Analysis for LAB INR Samples
Cumulative Function Plot (N=11,273)

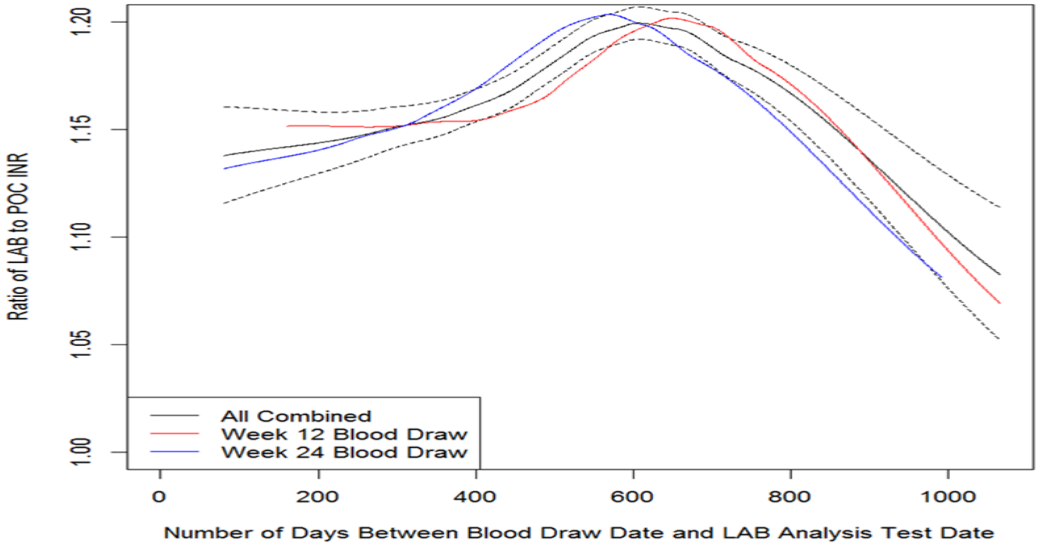


Source: FDA analysis of data provided by Janssen.

Figure 5 is a plot of the relationship between the ratio of LAB INR to POC INR vs. the N of days between the LAB INR blood draw and its run date for samples drawn at Week 12, Week 24, and the pooled samples. The curves are oddly shaped, with a peak at about day 600, but variability is not large between peak (a ratio of 1.20) and nadir (between 1:05 and 1.10). The dotted line is the 95% CI for the pooled set of samples. Note that about 5% of samples were stored for more than 800 days prior to their run date. If these samples are ignored, the variability over time becomes quite small.

It is difficult to imagine a storage time-related process that would produce a curve with this shape, i.e. an upward inflection point between 400 and 500 days, and a downward inflection point at around 600 days.

Figure 5 Ratio of LAB INR to POC INR vs. Days between Blood Draw and Analysis for LAB INR Samples



Finally, Figure 6 is a plot of the relationship between the ratio of LAB INR to POC INR vs. the calendar date of the LAB INR analysis, without regard to when the sample was drawn. This curve is similar in shape to the curve in Figure 5, which suggests the possibility that differences in the ratio of LAB INR to POC INR may be might be related to changes over time in the assay system than changes related to storage time.

The analytical report for the PK/PD study data in the NDA submission has the following information for the lots of thromboplastin used to analyze LAB INR at Duke:

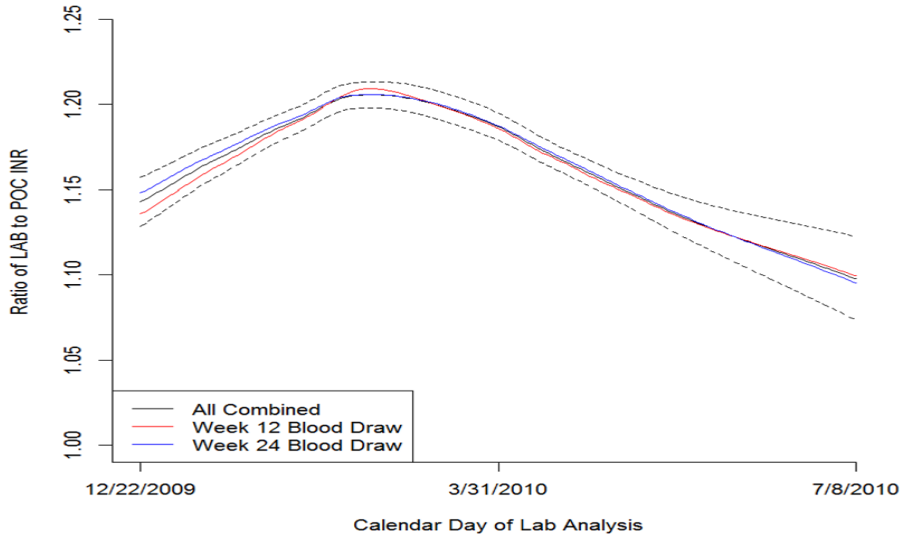
Analysis Date[7]	Lot No.	ISI
12/22/09	104730	1.21
01/12/10	102387	1.24
04/30/10	104079	1.25

The ISI is the "International sensitivity index," which is used to calculate INR from the prothrombin time of the patient and a reference prothrombin time. The data indicate that different lots of thromboplastin were used for analysis runs over the course of the analysis period, which might have affected the results slightly, despite that fact that adjusting prothrombin time for ISI is intended to yield a stable INR. Nonetheless, the overall ratio of LAB INR to POC INR is about 1.15, and does not vary much over the course of the study except for a small fraction of samples with storage time greater than about 800 days. For such samples, the ratio was closer to 1.0 than for other samples. These data indicate that the observed

[7] It is not clear what these dates signify. It could be that the first lot was used until 1/12/10, the date given for the second lot. Then the second lot was used until 4/30/10, when the third lot began to be used. However, it seems unnecessary to understand the cause of variability over time in the ratio of LAB INR to POC INR because the variability was not clinically meaningful.

differences between LAB INR and same-day POC INR over the course of the study probably were at most minimally affected by the duration of storage.

Figure 6 Ratio of LAB INR to POC INR vs. Calendar Date of Analysis



5 Modeling of the Impact of the INRatio Device on Clinical Outcomes in ROCKET

5.1 Modeling by FDA

Modeling was performed by FDA to estimate how the use of the INRatio device affected the major clinical outcomes of ROCKET. Specifically, two distinctly different modeling approaches were used to estimate modified warfarin arm major bleeding rates that would be expected if the INRatio device had produced readings similar to those of the laboratory based device at Duke. Both analyses indicated that the modified warfarin arm major bleeding rate would be expected to be slightly lower than the observed rate, with a minimal increase in the rivaroxaban vs. warfarin hazard ratio for this endpoint (i.e., a small change favoring warfarin). One of the analyses also predicted a modest increase in the rate of ischemic stroke in the warfarin arm from the observed rate, which would lead to a reduction in the hazard ratio for this endpoint (i.e., a modest change favoring rivaroxaban). The other model did not estimate this parameter. Neither model suggested that our previous assessment that the benefits of rivaroxaban outweighed its risks relative to warfarin would have been altered if ROCKET had yielded the modeled results.

Modeling by Drs. Florian and McDowell

The first FDA analysis was created by Dr. Jeffry Florian of the Division of Biometrics in the Office of Clinical Pharmacology and Dr. Tzu-Yun McDowell in the Division of Cardiovascular and Renal Products. The analysis was based on data from 22,063 patients in the warfarin arms

of the trials that supported approval of the three DOACs other than rivaroxaban approved for the SPAF indication in the US since 2010: RE-LY (dabigatran), ARISTOTLE (apixaban) and ENGAGE AF (edoxaban). In each of these trials, warfarin dose was adjusted at least once a month on the basis of INR. As in the ROCKET trial, the INR target was 2.0 to 3.0. The median INR value in patients from these trials was 2.3 with an interquartile range of 1.9 to 2.8. As of this date, we have received no information to suggest that there was a systematic concern with the accuracy of the INR assessments in any of the three trials used to create the modeling base.[8]

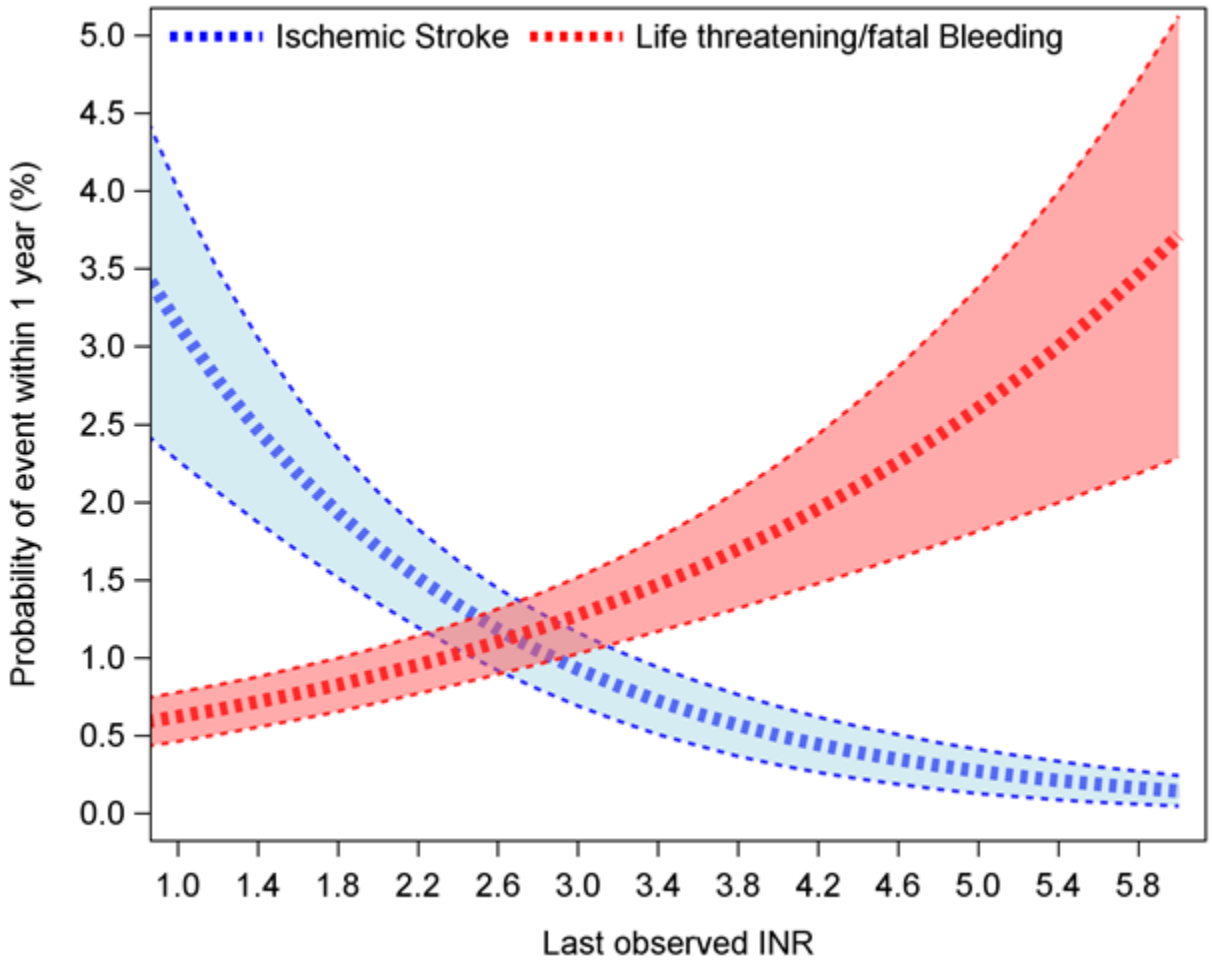
The review of the modelers states,

> "For each outcome of interest, a multivariate Cox proportional hazard (PH) model was developed to examine the time to the first occurrence of an on-treatment event as a function of INR and other patient covariates…. We used the last observed INR, which was defined as the last measured INR value prior to or on the date of the first outcome event of interest (INR value closest to the censored date if no event) to explore the INR-outcome event relationship. This INR value was selected as a best representation of an individual patient's INR reading proximal to the time of event or censoring. In addition, INR values >6 were truncated to 6. A set of common baseline covariates collected in these studies, which could be potentially associated with the outcome of interest was obtained and tested in Cox PH model. These covariates included age, sex, race (white/non-white), baseline body weight, baseline aspirin use, baseline antiplatelet use, baseline CHADS2 score, history of stroke or transient ischemic attack (TIA), diabetes, baseline creatinine clearance (categorical as normal, mild, moderate, and severe based on Cockcroft-Gault equation), smoking history and alcohol use. Covariates in the Cox PH model were selected using stepwise forward addition followed by backward elimination based on Bayesian information criteria (BIC) and the stepAIC function from the 'MASS' package. Considering that INR management varies geographically, sensitivity analysis was conducted using North American patients alone. All the analyses and plots were conducted and generated in R (version 3.1.2) and/or SAS 9.3."

The model-based risk of ischemic stroke and life-threatening or fatal bleeding across a broad range of INR values is shown in **Figure 7**. Note that unlike the plot depicted in **Figure 1,** the curve for ischemic stroke is not flat as INR increases above the target range; instead, the higher levels of INR are associated with lower rates of ischemic stroke across the range depicted in the plot.

[8] In the double-blind ARISTOTLE and ENGAGE AF trials, sites were provided with hand-held ITC Protime devices to assess INR. These modifications were analogous to the modifications of the INRatio device used in ROCKET to preserve the study blind. In RE-LY, which was an open-label trial, the decision of how to assess INR was left to the investigators' discretion, although TTR was monitored.

Figure 7 Relationships of the Rates of Ischemic Stroke and Life-Threatening or Fatal Bleeding vs. INR in Pooled Data from the Warfarin Arms of Three Confirmatory Trials of DOACs



Note: The trials were RE-LY (dabigatran), ARISTOTLE (apixaban) and ENGAGE AF (edoxaban). All were warfarin-controlled, and had a pooled N in their warfarin arms of 22,063 patients.
Note: "Last observed INR" refers the last measured INR value prior to or on the date of the first outcome event of interest, or in the case of patients with no event of interest, the INR closest to the censor date.
Source: Analysis by FDA (Drs. Tzu-Yun McDowell and Jeffry Florian).

The INR data from ROCKET obtained with use of the INRatio device were adjusted ("modified") as described above and run through the model. The resulting modeled outcome data are shown in **Table 12**. The reported ROCKET data are provided for contrast next to the modified data.

The modified warfarin arm event rates and the modified risk ratios varied from the observed results in the directions that we expected. The modified major bleeding rate was reduced by 10% from the observed rate, thereby increasing the rivaroxaban vs. warfarin risk ratio (RR) for this event by 11%. The RRs for life-threatening or fatal bleeding and also hemorrhagic stroke increased by roughly the same proportion (i.e., rivaroxaban appeared to cause slightly more bleeding relative to warfarin than reported in labeling or in the publication of the trial results).

However, the estimated results for ischemic stroke moved in the opposite direction. The observed rate of this endpoint with warfarin was 1.42 %/year, while the modified was 1.70 %/year. The observed RR was 0.94, while the modified RR is 0.79, representing a 16% reduction in risk for rivaroxaban relative to warfarin. Thus, the model predicts that if INR had been reported accurately, one would expect that the results for ischemic stroke would be somewhat more favorable for rivaroxaban than the reported results. The effect of use of the INRatio device on the primary study endpoint (total stroke + systemic embolism) was not modeled. However, the data indicate that if the primary endpoint results had been modeled, the results would move in the same direction as the results for ischemic stroke, because the modeled absolute increase in the rate of ischemic stroke in the warfarin arm is 0.28 events per 100 patient-years, and the modeled absolute reduction in the rate of hemorrhagic stroke is only 0.04 events per 100 patient-years, yielding a net increase in the stroke event rate of 0.24 events per 100 patient-years.[9] By increasing the rate in the denominator, the hazard ratio for rivaroxaban vs. warfarin would be reduced, making the results more favorable for rivaroxaban.

Table 12 ROCKET AF: Reported and Model-Generated Outcomes Data for Selected Outcomes

	ROCKET AF Trial Results†				Modified Results	
Outcome Event	**Rivaroxaban ER**	**Warfarin ER**	**HR (95% CI)**	**RR**	**Modified warfarin ER**	**Modified RR**
Major Bleeding	3.61	3.45	1.05 (0.91, 1.20)	1.04	3.11	1.16
Life Threatening or Fatal Bleeding	1.64	1.93	0.85 (0.70, 1.04)	0.85	1.74	0.94
Hemorrhagic Stroke	0.26	0.44	0.59 (0.37, 0.93)	0.59	0.40	0.65
Ischemic Stroke	1.34	1.42	0.94 (0.74,1.17)	0.94	1.70	0.79

ER=Event rate (events per 100 patient-years); HR=Hazard Ratio; RR=Rate Ratio;
† ROCKET trial results was on treatment (last dose plus 2 days) analysis in the safety population
Note: The modified RR was calculated using the observed event rate for rivaroxaban and the modified ER for warfarin.

Modeling by Dr. Lawrence

Another analysis was created by Dr. John Lawrence of the Division of Biostatistics 1. His model was based on the relationship between "true" (adjusted) INR and bleeding, using transformed values of POC INR to establish the base model. His goal was to assess the impact of use of

[9] One would reach the same conclusion if one considered the effect of the INRatio device on the rate of systemic embolism because (1) the rate of this event is small in comparison to either ischemic stroke or hemorrhagic stroke in patients with atrial fibrillation taking warfarin and (2) one would expect it to move in the same direction as ischemic stroke if it were to be modeled.

the INRatio on major bleeding in ROCKET. He did not assess its impact on other clinical outcomes. He describes his model as follows in his review:

> "I first used the matched pairs of POC INR and LAB INR that were taken on the same day within the same patient. These allowed me to model the relationship between true INR (assumed equal to the LAB INR) and the POC INR. For a given observed POC INR, I imputed a true INR from the distribution observed in the model. I used LASSO (least absolute shrinkage and selection operator) to select the best covariates in the Cox regression model to predict Major Bleeding from the rivaroxaban arm alone [Tibshirani, Robert. 1997. "The lasso Method for Variable Selection in the Cox Model". Statistics in Medicine, Vol. 16, 385—395 (1997)]. Then, I added a time varying covariate for INR to produce a prediction model for Major Bleeding in the rivaroxaban arm. I then used multiple imputation to impute Major Bleeding events in the warfarin arm given that these patients' warfarin dose would have been titrated to achieve a true INR equal to their observed POC INR. Only subjects who had a Major Bleeding event could have had an imputed event. For each of those patients, either they would still have an event at theta time, or they would be censored at that time. The probability of having an event is equal to the ratio of the hazard rate given the observed POC INR compared to their true imputed INR. I found 100 such imputed data sets. I then combined the results from those 100 imputed datasets using the formulas from [Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys. New York: Wiley & Sons].

Using this methodology, he estimated that if a device with the performance characteristics of the device at Duke that was used to assess PK/PD study INR samples and warfarin was titrated to the same apparent INR as was observed in the study, the number of major bleeding events in the warfarin would be reduced by about 7%. This would result in an increase in the rivaroxaban vs. warfarin HR for major bleeding, to 1.12 (95% CI 0.97, 1.30). This is similar to the results of the Florian-McDowell model (see **Table 12**).

Thus, the two FDA analyses are fairly consistent in their results with respect to the potential impact that the use of the INRatio device on major bleeding. They estimated that if a more accurate device had been used, the rate of major bleeding in the warfarin arm might have been 7% to 10% less than observed, with correspondingly small changes in the rivaroxaban vs. warfarin hazard ratio that would disfavor rivaroxaban. In addition, the Florian/McDowell model estimated that the hazard ratios for life-threatening or fatal bleeding and hemorrhagic stroke, which both favor rivaroxaban numerically, would be minimally increased, but both would still favor rivaroxaban. The hazard ratio for ischemic stroke would be moderately changed in favor of rivaroxaban. If ROCKET had produced results similar to the results of either of these models, we would have approved rivaroxaban for the SPAF indication.

5.2 Modeling by Janssen

Janssen has developed analyses intended to explore the questions that FDA's analyses explored. The first such analysis was submitted to EMA at that agency's request, and then sent to us. EMA's request was,

> "Calculate "population attributable risk" in the warfarin treatment arm of the study population based on 34% expected to have discordant INR results and the highest increased risk as estimated in "Analysis 3" of safety outcomes (Both week 12 & 24 increase from 12.21 to 14.63 / 100 PY).

> "Apply the result to explore to what extent the event rates for the warfarin group in the table for safety analyses results as summarized in the SmPC (Table 2) would be reduced.
>
> "Calculate corresponding risk ratios comparable to the available hazard ratios (HR). In recalculating HRs a worst case assumption should be made and consequently remove the earliest events."

Note that this request applies only safety endpoints.

Janssen's analysis was constructed as follows:

> "For each of the safety outcomes summarized in the SmPC (Table 4: Safety results from phase III ROCKET AF) a modified warfarin event rate (events per 100 pt-yr) was derived for the total Warfarin cohort of the ROCKET AF trial. This modified warfarin event rate is contrasted to the reported rate for the total Rivaroxaban cohort as given in the SmPC (Table 4) via a rate ratio and the resulting rate ratio point estimate would then be contrasted to the SmPC reported hazard ratio as a descriptive measure.
>
> "The modified warfarin event rate is based on the Device and Lab based INR inconsistency criterion of the Device based INR value lower than the Lab based INR by at least one category using the categories of <2, 2-3, 3-4, >4 and is derived as follows:
>
> "The warfarin event rate from the two subgroups of subjects WITH (here denoted R1) and WITHOUT (here denoted R0) Device INR Lower Than Lab INR by at Least One Category at Both Week 12 and Week 24 …. was taken to derive a total warfarin rate (here denoted R_{TOT}) as a weighted average of these two rates assuming that 34% (taking the percentage as requested by EMA) of the weight is for the subgroups of subjects WITH Device INR Lower Than Lab INR by at Least One Category at Both Week 12 and Week 24 (i.e.=0.34*R1+0.66*R0).
>
> "A "population attributable (risk) fraction" (here denoted as PARF) is calculated as the ratio of the difference of the total warfarin rate minus the rate for subjects WITHOUT Device INR Lower Than Lab INR by at Least One Category at Both Week 12 and Week 24 to the total warfarin rate (i.e. PARF=(R_{TOT}-R0/R_{TOT})
>
> "The warfarin rate reported in the SmPC for the total warfarin cohort (which includes the two warfarin subgroups above who had paired INRs as well as a third subgroup of warfarin subjects who did not have paired INRs) was then "deflated" by a factor of (1-PARF) to get a modified warfarin event rate (i.e. multiply the annualized warfarin event rate reported in the SmPC by the factor (1-PARF)) which was then used to calculate the rate ratio.
>
> "Please note that the rate "deflation factor" (1-PARF) is also applied to warfarin subjects who do not have paired INR measurements at Both Week 12 and Week 24, which is one of the limitations of this approach."

Basically, this approach focusses on the event rates for bleeding endpoints in patients whose Week 12 and Week 24 POC INR both were lower by at least one category (using the categories defined in the quote above) than the LAB INR category (R1) and in patients who do not meet this requirement (R0). A weighted average of R1 and R0 rate is used to create a "deflator" that is used to adjust the rate of bleeding in the warfarin arm. This average weights R1 at 34%,

even though patients who met the criteria for R1 totaled only 782 of 5048 subjects with Week 12 and Week 24 INR pairs (15.5%).

Results of this analysis are shown in **Table 13**. Compared to the observed data for the warfarin arm, the adjusted data showed at most a small decrease in the rate of various types of bleeding that were analyzed. The risk ratio for rivaroxaban vs. warfarin was either essentially unchanged or at most modestly elevated from the observed values.

Table 13 ROCKET AF: Attributable Risk Analysis
Effects on Warfarin Arm Bleeding Rates

	Rivaroxaban n (%/year)	Warfarin n (%/year)	HR (95% CI)	Rate ratio	Warfarin. modified rate (%/year)	Rate ratio R vs W modified
Major & non-major CR bleeding	1,475 (14.91)	1,449 (14.52)	1.03 (0.96, 1.11)	**1.03**	13.60	**1.10**
Major bleeding	395 (3.60)	386 (3.45)	1.04 (0.90, 1.20)	**1.04**	3.33	**1.08**
Death due to bleeding	27 (0.24)	55 (0.48)	0.50 (0.31, 0.79)	**0.50**	0.44	**0.55**
Intracranial hemorrhage	55 (0.49)	84 (0.74)	0.67 (0.47, 0.93)	**0.66**	0.74	**0.66**

Safety population, on treatment
See text for explanation of modified rates
Source: Janssen submission dated 16 January 2016; EMA Assessment Report dated 10 February 2016 (http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000944/WC500201726.pdf)

We were concerned that the validity of this analysis might be affected by the small number of subjects included in the high-risk cohort of the analysis, as well as the use of Week 12 and Week 24 data to predict the course of events over the entire study. Accordingly, FDA requested Janssen to create one or more models with the same aim as the attributable risk analysis, but with different methodologies and modeling techniques. We suggested that they model ischemic stroke, the primary endpoint, hemorrhagic stroke, major bleeding, and life-threatening/fatal bleeding in warfarin arm subjects. We requested that they include any covariates that they thought were appropriate and to account for the observed variability in the POC to Lab INR ratio within and between patients. On the basis of the developed models, modified warfarin arm event rates would be derived and compared to the observed rivaroxaban arm event rates for the specified endpoints.

Janssen then responded on June 23, 2016 by submitting 4 analyses. The methodology of one of these analyses (Analysis 4) corresponded closely to our analyses, but the other 3 did not. Features of the 4 analyses are described in **Table 14**. All analyses had three main steps: (1) Using the Cox proportional hazard model, they examined the relationship between INR (each analysis used a different INR value) and the outcome event of interest while controlling for other covariates, which are listed in **Table 14**. (2) On the basis of the developed Cox model for each outcome event, they calculated the predicted changes in warfarin event rate based on the assumption that "true INR" was 0.3 higher than POC INR (median difference between the POC and laboratory INR values). [10] (3) They calculated the adjusted warfarin event rates and adjusted rate ratios between rivaroxaban and warfarin using the observed rivaroxaban event rates and the adjusted warfarin event rates.

Table 14 INR Value and Covariates Used to Examine the Relationship between INR and the Outcome Event of Interest

	Choice of INR value	Covariates	Outcome event
Analysis 1	Mean Lab INR at Week 12 and Week24	For analysis of bleeding endpoints: age, sex, diastolic blood pressure, baseline aspirin use, history of chronic obstructive pulmonary disease, history of gastrointestinal bleed, and history of anemia For analysis of efficacy endpoints: baseline CHADS2 score	-Major bleeding -Life threatening and fatal bleeding -Hemorrhagic stroke -Ischemic stroke -Primary efficacy endpoint
Analysis 2	Mean POC INR		
Analysis 3	Proportions of POC INR values in different ranges: <1, 1to <1.5, 1.5 to <2, 2 to <2.5, 2.5 to 3, >3 to <3.5, 3.5 to <4, ≥4		
Analysis 4	Last INR before or on the day of event or censoring		

Among the four analyses, Janssen commented that Analysis 4 using the last INR value before or on the day of event or censoring is most likely to be informative regarding rates of the modeled events (ischemic stroke or hemorrhagic events) due to the close temporal relationship of observed events with the INR values used in the model.

We agree with Janssen that the approach used in Analysis 4 is probably the most valid way of building a model using the ROCKET results to estimate the effects of INR on outcomes in

[10] A median INR difference of 0.3 between POC and lab values was used for Analysis 1, Analysis 2 and Analysis 4. For Analysis 3, Janssen fitted a regression model to the paired (i.e. POC and laboratory-based) INR samples at visit Week 12 and 24 and used this model to transform the POC INR values and assess potential changes of warfarin INR event rate. Janssen did not provide additional details about this particular analysis.

patients with AF taking warfarin. Analysis 1 used mean Lab INR to create the model. Analysis 2 used mean POC INR for this purpose. Both mean LAB INR and mean POC INR may not reflect the peaks and valleys of INR over the course of the study that we believe contribute to bleeding and ischemic events, respectively, in some patients. In addition, mean LAB INR is derived from values at Weeks 12 and 24, which may not predict INR and the risk of events after a year or more on treatment. Analysis 3 may attempt to account for the variability of INR over time, but it yielded a modified warfarin arm major bleeding rate that was higher than the observed rate, along with a large percentage reduction in the rate of hemorrhagic stroke. The increase in major bleeding is illogical because with a lower INR, one would expect less major bleeding, not more. The hemorrhagic stroke rate moved in the expected direction, but the percentage change was quite large and not consistent with a modest reduction in INR. Accordingly we think that Analysis 4 is the most reasonable approach, and the discussion below focusses on the results obtained with this analysis. For results with the other models, see Appendix 2 **below**.

Results for Analysis 4 are shown in **Table 15**. The adjusted warfarin event rates and modified hazard ratios are only modestly different from those reported in the trial, and all move in the direction consistent with what one would have predicted when modeling to a lower INR. These findings are consistent with the FDA analyses suggesting that the impact of erroneous POC device on the ROCKET trial results was small.

Table 15 ROCKET AF: Modeling with Last POC INR as Covariate - Modified Hazard Ratios Using the Adjusted Warfarin Arm Rates and the Observed Rivaroxaban Arm Rates
(Janssen's Analysis 4; Safety Population, On-Treatment)

Endpoint	Riva Rate per 100 pt yr	Warfarin Rate per 100 pt yr	Hazard Ratio	95% CI	Adjusted Warfarin Rate per 100 pt yr*	Modified Hazard Ratio
Major bleeding	3.60	3.45	**1.04**	0.90, 1.20	3.22	**1.12**
Fatal bleeding	0.24	0.48	**0.50**	0.31, 0.79	0.42	**0.57**
Hemorrhagic stroke	0.26	0.44	**0.59**	0.37, 0.93	0.38	**0.68**
Ischemic stroke	1.34	1.42	**0.94**	0.75,1.17	1.58	**0.85**
Primary efficacy endpoint	1.7	2.15	**0.79**	0.65, 0.95	2.23	**0.76**

*The adjustment reflected the median difference of 0.3 units for LAB INR minus POC INR that was observed in ROCKET.
Source: Submission of June 23, 2016

6 Data from Other Studies of Rivaroxaban

Three other studies of rivaroxaban have been performed that used a dose similar to or higher than the dose used in ROCKET. These studies had comparator arms or cohorts that included warfarin dosed in the same manner as in ROCKET or another DOAC, dabigatran. None of these studies used the Alere device to assess INR, and none suggested that rivaroxaban had reduced efficacy or a markedly worse bleeding profile than the control arm or cohort. By

implication, they support the conclusion that the effect of the Alere device on the outcome of ROCKET was not substantial and should not affect our interpretation of the study results.

6.1 EINSTEIN DVT/PE

The EINSTEIN DVT and EINSTEIN PE (pulmonary embolism) studies were two active controlled, open-label RCTs in patients with acute deep vein thrombosis and pulmonary embolism, respectively. These studies supported the approval of rivaroxaban for its DVT and PE treatment indications. The goal of each study was to demonstrate the non-inferiority of rivaroxaban to an active control with regard to the rate of recurrent thrombotic events (either DVT or PE in each of the studies). Many features of the two studies were similar. Patients were randomized to open-label treatment (with a duration of 3, 6, or 12 months at discretion of investigator) with either:

- Rivaroxaban 15 mg bid (30 mg daily) with food X 3 weeks, then 20 mg daily with food to end of treatment, or
- Enoxaparin 1 mg/kg SC bid for >= 5 d, + a vitamin K antagonist started no later than day 2 and maintained to the end of treatment. Enoxaparin was discontinued when INR >=2.0 x 2 days, but no earlier than day 5. The INR target was 2.0-3.0, similar to ROCKET.

Notably, thrombotic and bleeding events were adjudicated centrally in a blinded fashion. This would reduce, but not eliminate, the potential impact of investigator bias in assessing study outcomes on the study results in these open-label studies.

In each of the two studies, the treatment arms had very similar distributions with respect to the percentage of patient who were selected for the 3 protocol specified treatment durations:

Assigned duration of treatment:	3 mo	6 mo	12 mo
Einstein DVT			
Rivaroxaban (% of patients)	12.0	62.6	25.4
Enoxaparin/VKA (% of patients)	11.8	63.0	25.1
Einstein PE			
Rivaroxaban (% of patients)	5.3	57.3	37.4
Enoxaparin/VKA (% of patients)	5.1	57.5	37.5

TTR in each of the EINSTEIN studies during VKA treatment was better than in ROCKET. Each study met its non-inferiority goal for prevention of recurrent DVT or PE. Bleeding data in the pooled studies favored the rivaroxaban regimen over the enoxaparin/VKA regimen with respect to ISTH major bleeding, but the treatment arms had similar rates of less serious bleeding and overall bleeding (**Table 16**). Note that the rivaroxaban dose used in this study, 15 mg bid for 3 weeks, followed by 20 mg once daily for the duration of treatment, was more intense than the regimen used in ROCKET, which was 20 mg once daily throughout the study. This study is reassuring with respect to the bleeding risk of rivaroxaban 20 mg daily.

Table 16 EINSTEIN DVT/PE: Bleeding Results

Bleeding Type	Rivaroxaban N=4130 n (%)	Enoxaparin/VKA N=4116 n (%)
ISTH major bleeding	40 (1.0)	72 (1.7)
Fatal bleeding	**3 (<0.1)**	**8 (0.2)**
Intracranial	**2 (<0.1)**	**4 (<0.1)**
Non-fatal critical organ bleeding	**10 (0.2)**	**29 (0.7)**
Intracranial	**3 (<0.1)**	**10 (0.2)**
Retroperitoneal	**1 (<0.1)**	**8 (0.2)**
Intraocular	**3 (<0.1)**	**2 (<0.1)**
Intra-articular	**0**	**4 (<0.1)**
Clinically relevant non-major bleeding	357 (8.6)	357 (8.7)
Any bleeding	1169 (28.3)	1153 (28.0)

Source: Rivaroxaban labeling

6.1 Mini-Sentinel Observational Study: Rivaroxaban vs. Warfarin

Analysis of this study is still underway. This is a retrospective study of FDA's Sentinel database of insurance claims information from commercial insurers. New users of any dose of rivaroxaban or warfarin with a diagnosis of atrial fibrillation were identified and propensity-matched. Outcomes of interest were GI bleeding, intracranial hemorrhage (including hemorrhagic stroke) and ischemic stroke, based on inpatient primary discharge codes and were measured using a sequential design. During the study period from November 2011 to April 2015, 41,648 new users of rivaroxaban and 89,080 new users of warfarin who met the study inclusion criteria were identified. After propensity matching based on demographic and disease-based factors, the two cohorts were reduced to 36,129 and 88,982 closely matched subjects, respectively. The average follow-up time on treatment was 0.23 years for the rivaroxaban cohort and 0.19 years for the warfarin cohort.

In the final analysis period, the HR for GI bleeding and intracranial hemorrhage were directionally similar to those observed in ROCKET. In contrast to ROCKET's results, a reduction in risk of ischemic stroke for rivaroxaban compared to warfarin was observed , a finding that is consistent with the modeling exercises described above (rivaroxaban might have more bleeding risk and more antithrombotic efficacy relative to warfarin than was observed in ROCKET). The beneficial effect on intracranial hemorrhage observed with rivaroxaban in ROCKET appeared slightly attenuated in the Sentinel analysis (**Table 17**). The HR for GI bleeding in the Sentinel study (1.30) is lower than that observed in ROCKET (1.61) suggesting that rivaroxaban is associated with less GI bleeding compared to warfarin than ROCKET would predict. The GI outcome was not modeled in the previous analyses. Although these analyses are still being finalized, they do not raise safety signals regarding bleeding risks for rivaroxaban. It is

important to remember that the Sentinel analysis is not a randomized study and may be confounded by differences between the cohorts that are not understood.

Table 17 Mini-Sentinel Observational Study of Rivaroxaban vs. Warfarin: Bleeding and Ischemic Stroke Outcomes

Study Outcome	IR Difference per 1000 pt-years (Rivaroxaban - Warfarin)	Risk Ratio (95% CI) (Reference is warfarin)	ROCKET HR (95% CI) (Reference is warfarin)
GI bleeding	16.59	1.30 (1.18, 1.43)	1.61 (1.30, 1.99)
Intracranial hemorrhage	-3.92	0.73 (0.57, 0.94)	0.67 (0.47, 0.93)
Ischemic stroke	-13.59	0.81 (0.72, 0.92)	0.94 (0.75, 1.17)

IR=Incidence rate
Results from ROCKET shown for comparison

6.2 Medicare Database Observational Study: Rivaroxaban vs. Dabigatran

To be complete in describing FDA's pharmacovigilance efforts involving rivaroxaban, this study of rivaroxaban vs. warfarin in Medicare is presented. However, without inclusion of a control/warfarin arm, it is unclear how relevant these results are to the investigation of INR device issues in the ROCKET trial. This study was performed using the CMS Medicare claims database and is pending publication. Subjects included persons age 65 or older with atrial fibrillation who were new users of rivaroxaban 20 mg (~66,000) or dabigatran 150 mg bid (~52,000) during the study period of November 2011 to June 2014. Outcomes of interest were the same as in the previously discussed study: GI bleeding, intracranial hemorrhage and ischemic stroke.

After propensity matching, the two treatment cohorts were similar in terms of demography and relevant medical history. Only10% of subjects were followed for more than 240 days. Results of this study indicate that there was a greater rate of observed GI bleeding and intracranial hemorrhage and a reduced rate of ischemic stroke with rivaroxaban compared to dabigatran (**Table 18**). Like for the Mini-Sentinel study, this study was not randomized and the results should be interpreted cautiously.

Table 18 Medicare Database Observation Study of Rivaroxaban vs. Dabigatran: Bleeding and Ischemic Stroke Outcomes

	IR Difference per 1000 years (Dabigatran - Rivaroxaban	HR (95% CI) (Reference is Rivaroxaban)
GI bleeding	-9.4	0.71 (0.63, 0.81)
Intracranial hemorrhage	-2.3	0.61 (0.44, 0.83)
Ischemic stroke	1.8	1.24 (0.99, 1.55)

IR=Incidence rate

7 Discussion

The available evidence from ROCKET, based on comparisons between same-day pairs of POC INR determined by the INRatio device and LAB INR, indicates that POC INR, which the investigators used to guide warfarin management, was a mean of 13% less than the INR analyzed at Duke with a laboratory-based assay at the end of the study. FDA's analysis indicates that storage of these samples did not seem to affect INR. Janssen's analyses indicate that patients with or without one of the conditions listed in Alere's December 2014 recall notice and described as being associated with low INR readings with the INRatio device had a similar rate of low INRs and also similar rates of bleeding events. This suggests that Alere's claim that inaccuracy of the device occurs only in patients with the specific conditions described in the recall notice is spurious. Instead, the data suggest that all warfarin arm patients in ROCKET appeared to be affected by the inaccuracy of the device.

The data thus imply that patients in the warfarin arm of ROCKET may have been unintentionally over-anticoagulated to some degree. This would lead to an increased risk of bleeding compared to what would be expected if the INRatio device produced results similar to those produced by the laboratory-based device at Duke. Conversely, it is possible that some of these patients benefitted from being over-anticoagulated by having a reduced rate of ischemic stroke compared to what would be expected with a device that did not report falsely low INR values.

Both FDA and Janssen performed mathematical modeling of the ROCKET data to estimate how the rates of major bleeding, hemorrhagic stroke, ischemic stroke and several other important outcomes were affected by use of the INRatio device. Several approaches to modeling were used. Three of the models used INR values closest to a bleeding or efficacy event to establish INR vs. response relationships for the relevant event. We believe that this approach most accurately estimates the risk of bleeding or thrombotic events. These three models estimated that use of the INRatio had modest effects on study outcomes in warfarin arm patients in the directions described above. In particular, the models estimated that use of an INR monitoring device that did not read low would have resulted in rates of major bleeding that were 7% to 10% lower than the observed rate of major bleeding in the ROCKET warfarin arm, along with reciprocal increases in the hazard or risk ratios for rivaroxaban vs. warfarin, i.e., changes disfavoring rivaroxaban.

Of the three models that based on INRs closest to endpoint events, the two models that included results for hemorrhagic stroke estimated that the rate of this event would be negligibly changed. However, the estimated increase in bleeding risk would be too small to change our prior assessment that the benefits of rivaroxaban, compared to warfarin, outweighed its risks.

Notably, the two models that analyzed ischemic stroke estimated that rate of this event in the warfarin arm rate would have been modestly increased with use a device that did not produce falsely low INR results. This would result in a reciprocal decrease in the hazard or risk ratio for ischemic stroke, thus favoring rivaroxaban. The net effect on the primary endpoint would be driven by the ischemic stroke results, and would also favor rivaroxaban. Thus, *a fortiori*, we would conclude that based on these hypothetical results, the benefits of rivaroxaban would still outweigh its risks relative to warfarin. Data from other studies of rivaroxaban support this conclusion.

Accordingly, the reviewers see no need for regulatory action at this time. We think that a labeling change to describe the modeling results would very difficult to write in a concise manner and might be more likely to confuse than to edify, and is not warranted. FDA might make a brief announcement regarding our conclusions and make this review or a summary of it available to the public online. A publication of our analyses in a journal might be desirable. If others think it is important to change labeling, we might add a simple statement that FDA has reviewed the INR and outcomes information in ROCKET and determined that the effect of use of the INRatio device on outcomes in ROCKET was too small to affect our prior conclusion that the benefits of rivaroxaban outweigh its risks relative to warfarin.

Appendix 1 Proportional Color Block Display of POC INR vs. LAB INR

The two plots in **Figure 8** are visual representations of the data in **Table 10** (reproduced below for convenience), which is a cross-tabulation of categorized values of POC INR vs. same-day LAB INR. In plot A, each of the 6 numbered vertical columns represents one of the 6 POC INR categories (groups) shown to the right of the plot. The width of each column is proportional to the ratio of the total N of patients represented in the column to 6225, the total number of patients represented in the plot. The height of each color block in the column is proportional to the ratio of the number of patients in the relevant group of LAB INR to the total number of patients represented in the column. Each group of LAB INR is represented by a color defined in the color block key at the top right of the plot, labeled "Lab Group B." The color blocks are arrayed within each column with Group 1 (LAB INR <1.5) at the bottom of each column, progressing up to Group 6 (LAB INR >6) at the top. There is one cross hatched color block in each column that represents the block where an individual's POC INR and LAB INR groups are concordant.

Plot B is analogous to Plot A, but the axes are reversed: each column represents a group of LAB INR and each color block represents patients within the column whose POC INR group corresponds to the color block key. INR groups and the color block key are the same in the two plots, as are rules for the width of columns and height of each color block.

It is easy to appreciate in Plot A that when POC INR is the target range or near it (POC INR groups 2, 3 and 4) the number of patients with a LAB INR group number greater than the value concordant with the POC INR group (represented by the area in the column above the hatched block) is much greater than the number of patients whose LAB INR is in one of the groups with a number less than that of the concordant block (corresponding to the area in the column below the hatched block).

The inverse is in true for Plot B, where the columns represent LAB INR groups. Also, it is obvious that for patients in LAB INR group 4 (INR >3 to 4), the number of patients with a concordant POC INR is less than the number patients with POC INR in group 3. The situation is similar for those with LAB INR in group 5, where the number of patients with a concordant POC INR is smaller than then number with a POC INR in group 4, and there is a substantial fraction of patients with LAB INR in group 5 and POC INR in group 3 (i.e., the nominal INR is in the target range but the LAB INR is in the range of 4 to 6). These results suggest that the risk of bleeding related to elevated INR might be increased by use of the INRatio device.

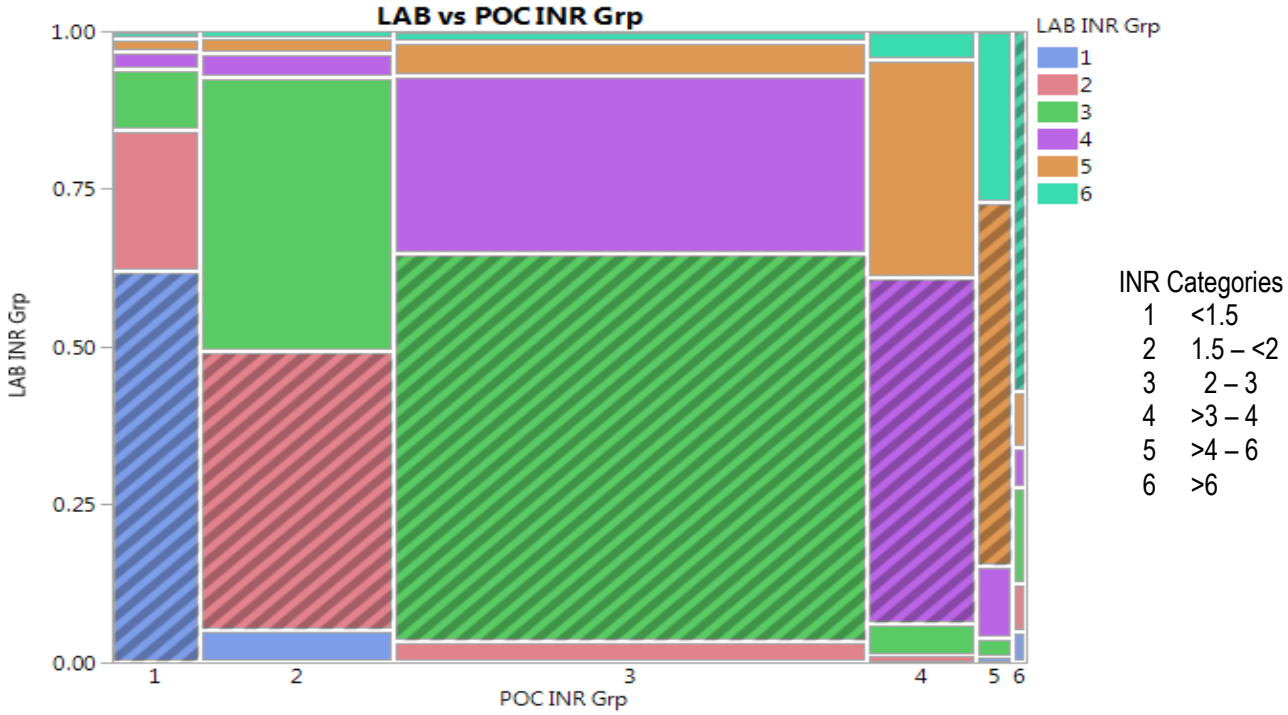
Table 10 ROCKET AF: Categorical Analysis of POC vs. LAB INR at Week 12 or 24; Treated Patients, Warfarin Arm

	LAB INR Group													
POC INR	**1**		**2**		**3**		**4**		**5**		**6**		**Total**	
Group↓	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
1	376	6.0	137	2.2	58	0.9	17	0.3	12	0.2	7	0.1	**607**	**9.8**
2	68	1.1	580	9.3	569	9.1	51	0.8	34	0.5	10	0.2	**1312**	**21.1**
3	5	0.1	104	1.7	1990	32.0	908	14.6	172	2.8	56	0.9	**3235**	**52.0**
4	0	0.0	9	0.1	37	0.6	400	6.4	253	4.1	33	0.5	**732**	**11.8**
5	2	0.0	0	0.0	8	0.1	30	0.5	150	2.4	70	1.1	**260**	**4.2**
6	4	0.1	6	0.1	12	0.2	5	0.3	7	0.1	45	0.7	**79**	**1.3**
Total	**455**	**7.3**	**836**	**13.4**	**2674**	**43.0**	**1411**	**22.7**	**628**	**10.1**	**221**	**3.6**	**6225**	**100**

Notes: INR Groups: 1, <1.5; 2, 1.5 to <2; 3, 2 to 3; 4, >3 to 4; 5, >4 to 6; 6, > 6

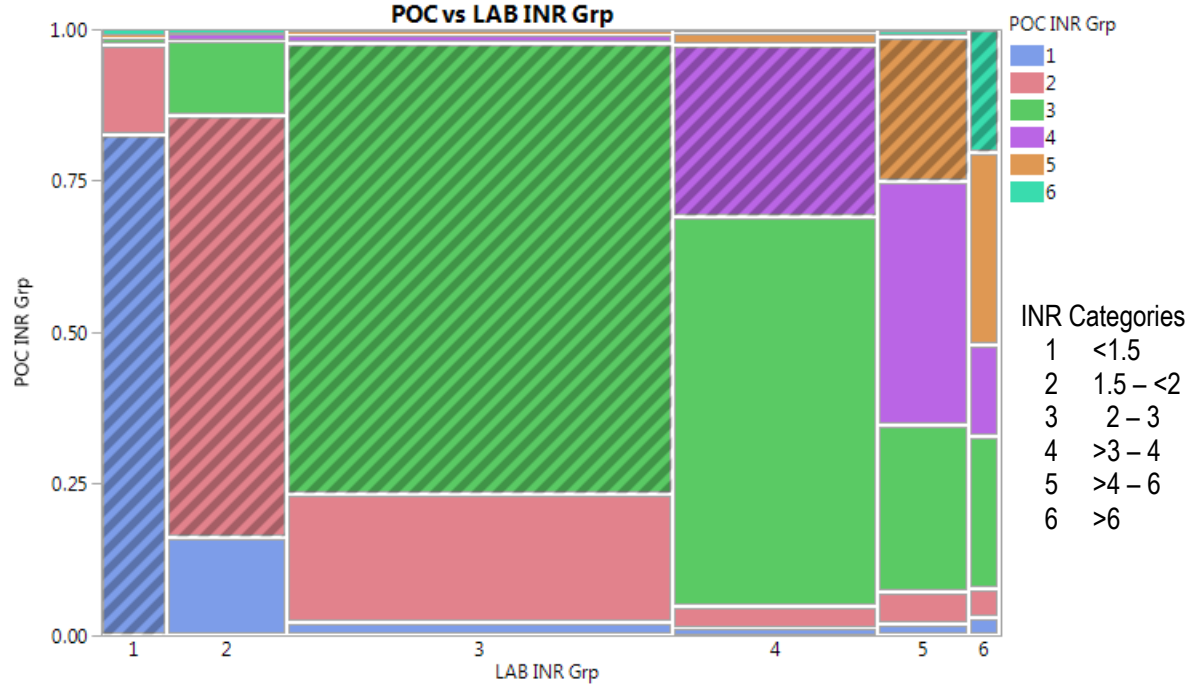
Figure 8 ROCKET AF: Proportional Color Block Plot of Lab INR vs. POC INR
Treated Patients, Warfarin Arm (N=6225)

Plot A



POC INR Group

Plot B



LAB INR Group

Appendix 2 Modeling by Janssen

The information below is from a submission of modeling data dated June 23, 2016.

Analysis 1: Mean Lab INR as Covariate - modified hazard ratios using the adjusted warfarin event rates and the observed rivaroxaban event rates (Adjustment based on 0.3 median difference in INR)

Endpoint	Riva. /100 pt yr	Warfarin /100 pt yr	HR	95% CI	Adjusted Warfarin /100 pt yr	Modified HR
Major bleed	3.6	3.45	1.04	0.90, 0.20	3.39	1.06
Fatal bleed	0.24	0.48	0.5	0.31, 0.79	0.46	0.52
Hem. Stroke	0.26	0.44	0.59	0.37, 0.93	0.41	0.64
Ischemic stroke	1.34	1.42	0.94	0.75, 1.17	1.39	0.96
Primary efficacy endpoint	1.7	2.15	0.79	0.65, 0.95	2.09	0.81

Analysis 2: Mean POC INR as Covariate - modified hazard ratios using the adjusted warfarin arm rates and the observed rivaroxaban arm rates

Endpoint	Riva. /100 pt yr	Warfarin /100 pt yr	HR	95% CI	Adjusted Warfarin /100 pt yr	Modified HR
Major bleed	3.6	3.45	1.04	0.90, 0.20	2.96	1.22
Fatal bleed	0.24	0.48	0.5	0.31, 0.79	0.31	0.77
Hem. Stroke	0.26	0.44	0.59	0.37, 0.93	0.31	0.85
Ischemic stroke	1.34	1.42	0.94	0.75, 1.17	1.85	0.73
Primary efficacy endpoint	1.7	2.15	0.79	0.65, 0.95	2.52	0.68

Analysis 3: Proportions of Imputed POC INR - modified hazard ratios using the adjusted warfarin event rates and the observed rivaroxaban event rates

Endpoint	Riva. /100 pt yr	Warfarin /100 pt yr	HR	95% CI	Adjusted Warfarin /100 pt yr	Modified HR
Major bleed	3.6	3.45	1.04	0.90, 0.20	4.16	0.87
Fatal bleed	0.24	0.48	0.5	0.31, 0.79	0.35	0.69
Hem. Stroke	0.26	0.44	0.59	0.37, 0.93	0.26	1.00
Ischemic stroke	1.34	1.42	0.94	0.75, 1.17	2.02	0.66
Primary efficacy endpoint	1.7	2.15	0.79	0.65, 0.95	3.04	0.56

Analysis 4: Last POC INR as Covariate - modified hazard ratios using the adjusted warfarin arm rates and the observed rivaroxaban arm rates

Endpoint	Riva. /100 pt yr	Warfarin /100 pt yr	HR	95% CI	Adjusted Warfarin /100 pt yr	Modified HR
Major bleed	3.6	3.45	1.04	0.90, 0.20	3.22	1.12
Fatal bleed	0.24	0.48	0.5	0.31, 0.79	0.42	0.57
Hem. Stroke	0.26	0.44	0.59	0.37, 0.93	0.38	0.68
Ischemic stroke	1.34	1.42	0.94	0.75, 1.17	1.58	0.85
Primary efficacy endpoint	1.7	2.15	0.79	0.65, 0.95	2.23	0.76

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
09/26/2016

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 202-439
Supplement #:	
Drug Name:	Xarelto (rivaroxaban)
Indication(s):	Delay time to stroke or systemic embolic event in patients with Atrial Fibrillation
Applicant:	Janssen Pharmaceuticals
Date(s):	9/10,16
Review Priority:	Standard
Biometrics Division:	DBI
Statistical Reviewer:	John Lawrence, Ph D
Concurring Reviewers:	Jim Hung
Medical Division:	Cardiorenal.
Clinical Team:	Tzu-Yun McDowell PhD, Martin Rose MD
Project Manager:	Bridget Kane
Keywords:	active control/non-inferiority, Cox regression, l

Table of Contents

LIST OF TABLES

No table of figures entries found.

LIST OF FIGURES

EXECUTIVE SUMMARY

Xarelto (rivaroxaban, NDA 202439) was approved for patients with atrial fibrillation based on the results of the ROCKET-AF study. ROCKET-AF was a double-blind non-inferiority study using the active control warafarin. Warfarin dose needs to be adjusted in patients based on the observed INR (International Normalized Ratio), a measure of the clotting tendency of the patient's blood. Many different factors can affect the INR; monitoring and warfarin dose adjustment is frequent. If the INR is too high, the blood cannot clot and bleeding events will occur more frequently, sometimes fatal. Conversely, if the INR is too low, then blood will clot too quickly and ischemic strokes and SEE (systemic embolic events) will occur more frequently. The goal of warfarin therapy is to keep the INR between 2.0 and 3.0. In the ROCKET-AF study, a POC (Point-Of-Care) device was used to measure the patient's INR quickly in the clinic. A dummy INR was generated and given for the patients randomized to the test arm of the study. In addition to these POC generated INR values, there were a small number of blood samples (approximately 2 per subject) sent to a central laboratory where the INR was measured in a more reliable way. After the study was completed and after Xarelto was approved, it was found that the POC INR was consistently lower than these more reliable laboratory INR values. It is reasonable to assume that the warfarin dose was, in many patients, higher than it would have been if the POC device had been operating correctly. If that is true, then there were more bleeding events in the warfarin arm than there should have been. Hence, the Xarelto arm appeared safer in comparison to this warfarin arm than it would have appeared if the patients had been dosed correctly with warfarin. In the trial, the hazard ratio for Major Bleeding, Life Threatening Bleeding, and other ways of defining bleeding were all less than 1 (in favor of Xarelto). I re-examined these results in light of the knowledge about the POC device. I imputed "true" INR values and also imputed bleeding events assuming lower hazard rate in the warfarin arm. For example, if the observed POC INR for a patient at a particular time was 2.5, then a true INR might be 3.5. If the patient actually had a bleeding event 3 days later, would the patient have still had a bleeding event if the doctor had known their true INR was 3.5 and had been able to adjust the warfarin dose to make the INR 2.5? How much could these results change under various assumptions. I used multiple imputation for the true INR values and for the bleeding event responses. INR was used as a time varying covariate in a Cox regression model. The FDA has also looked at other sources of data available such as post-marketing data about the rate of reported bleeding for Xarelto and bleeding rates in the warfarin arm from other atrial fibrillation studies. Given all of these analyses, it appears that the faulty device had a minor impact on the bleeding rate in the warfarin arm in the ROCKET –AF trial. Rivaroxaban is still judged to be safe and effective based on the results from that trial.

INTRODUCTION

1.1 Overview

In the ROCKET-AF trial, 14264 patients were randomized to rivaroxaban or warfarin and followed for a median of 1.6 years. This was an active control trial with a non-inferiority margin of 1.38 for the primary endpoint of time to stroke or systemic embolism (SEE). Rivaroxaban was deomstrated non-inferior to warfarin [HR 0.88 with 95% CI of (0.74, 1.03)]. During the trial, a point of care device was used to guide warfarin dosing. In addition, blood samples were taken from most patients at Week 12 and Week 24 and frozen for future PK analysis. Approximately 6000 patients from the warfarin arm had at least one blood sample. These samples were later analyzed and the POC INR was found to be biased relative to the laboratory INR. This review describes the re-analysis of the results by imputing true INR in the trial to examine what might have happened had the POC device been operating properly.

1.2 Data Sources

Electronic datasets and Study Reports:

\\cdsesub1\evsprod\NDA202439\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\afib\5351-stud-rep-contr\39039039afl3001

\\cdsesub1\evsprod\NDA202439\0052\m5\datasets\39039039afl3001\analysis\datasets

\\cdsesub1\evsprod\NDA202439\0000\m5\datasets\39039039afl3001\analysis\datasets

\\CDSESUB1\evsprod\NDA202439\0079\m5\datasets\39039039afl3001\analysis\datasets

STATISTICAL EVALUATION

1.3 Data and Analysis Quality

NA

1.4 Evaluation of Efficacy

The issue with the POC INR does not affect the efficacy assessment from the original NDA review. Patients in the warfarin arm were receiving an effective dose of warfarin, possibly too high.

1.4.1 Study Design and Endpoints

The endpoint used in this review is the time to first Major Bleeding event. Major bleeding was defined as clinically overt bleeding associated with a decrease in hemoglobin of ≥ 2 g/dL, transfusion of ≥ 2 units of packed red blood cells or whole blood, bleeding at a critical site, or with a fatal outcome.

1.4.2 Statistical Methodologies

I first used the matched pairs of POC INR and LAB INR that were taken on the same day within the same patient. These allowed me to model the relationship between true INR (assumed equal to the LAB INR) and the POC INR. For a given observed POC INR, I imputed a true INR from the distribution observed in the model. I used LASSO (least absolute shrinkage and selection operator) to select the best covariates in the Cox regression model to predict Major Bleeding from the rivaroxaban arm alone [Tibshirani, Robert. 1997. "The lasso Method for Variable Selection in the Cox Model". Statistics in Medicine, Vol. 16, 385—395 (1997)]. Then, I added a time varying covariate for INR to produce a prediction model for Major Bleeding in the rivaroxaban arm. I then used multiple imputation to impute Major Bleeding events in the warfarin arm given that these patients's warfarin dose would have been titrated to achieve a true INR equal to their observed POC INR. Only subjects who had a Major Bleeding event could have had an imputed event. For each of those patients, either they would still have an event at theta time, or they would be censored at that time. The probability of having an event is equal to the ratio of the hazard rate given the observed POC INR compared to their true imputed INR. I found 100 such imputed data sets. I then combined the results from those 100 imputed datasets using the formulas from [Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys. New York: Wiley & Sons].

1.4.3 Patient Disposition, Demographic and Baseline Characteristics

NA.

1.4.4 Results and Conclusions

The scatterplot in Figure 1 shows the relationship between the POC INR and the LAB INR from the sample taken on the same day within the same patient. The POC INR were truncated at about 6. Figure 2 shows the same scatterplot but the LAB INR have been truncated and a diagonal line is included to show where the LAB INR would equal the POC INR.

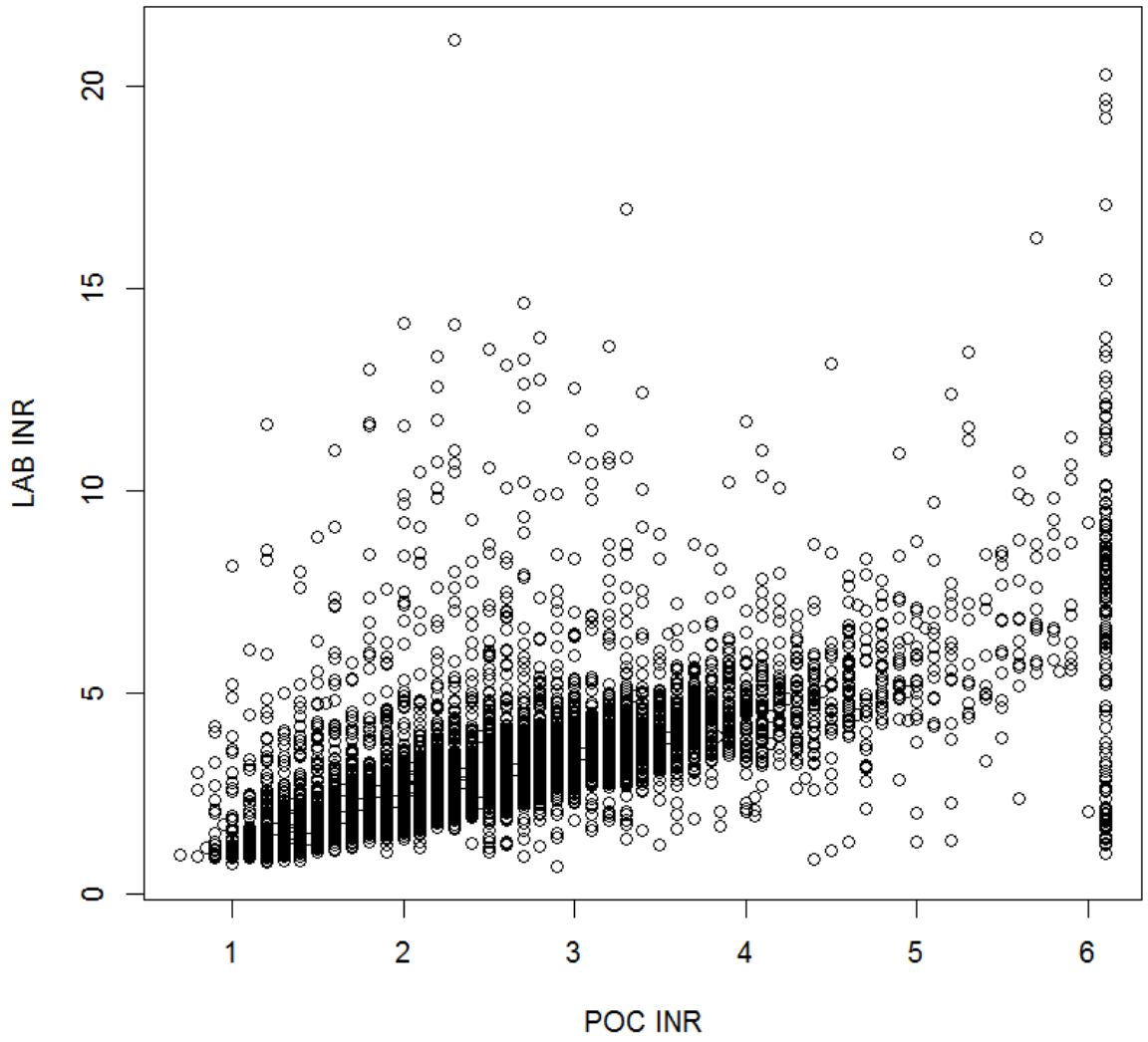


Figure 1 Scatterplot of POC INR versus LAB INR at matched time points for patients in the warfarin arm.

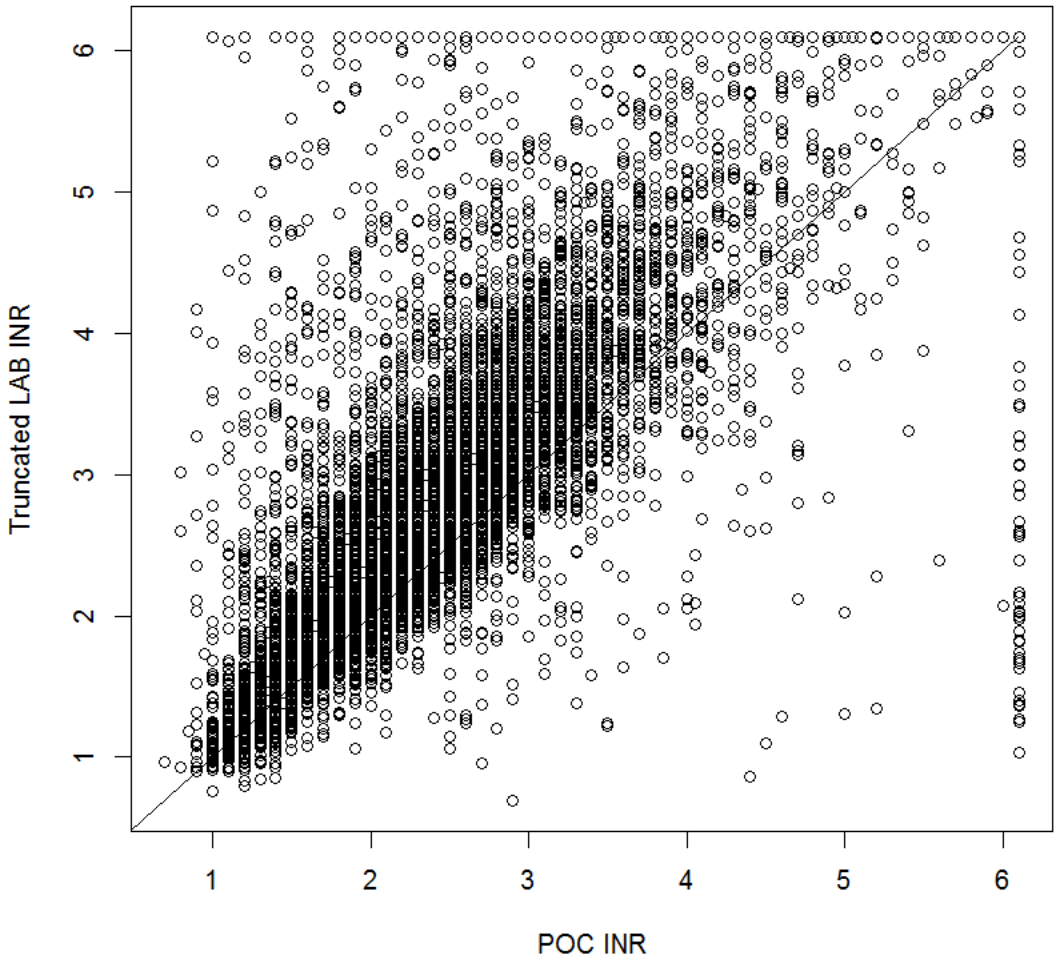


Figure 2 Scatterplot of POC INR versus LAB INR at matched time points for patients in the warfarin arm.

Figure 3 shows effect of freezing time on the ratio of LAB INR to POC INR using locally weighted regression. There seems to be a trend toward a higher ratio around 600 days freezing time. Figure 4 shows the same curves with the x-axis changed to the Calendar date of the LAB INR analysis. Because the samples from Week 12 and Week 24 seem to be the same when using calendar date on the x-axis, I believe the effect of time on the ratio may not be an effect of freezing time, but an effect due to conditions in the laboratory. The effect is minor and the ratio is always greater than 1. The mean and median time in the freezer were just over 500 days and 95% of the samples had less than 800 days of freezing time.

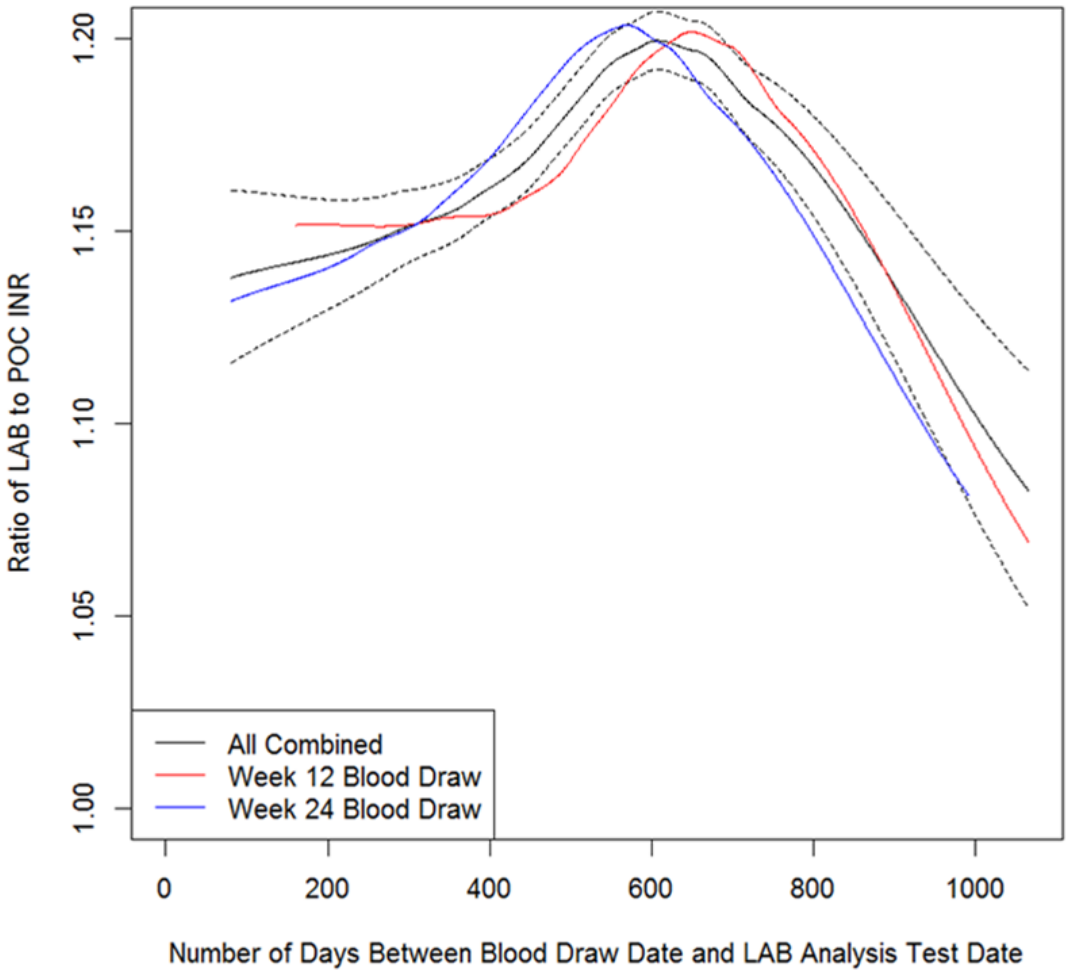


Figure 3 Loess regression model showing relationship between Days in Freezer and ratio of LAB INR to POC INR.

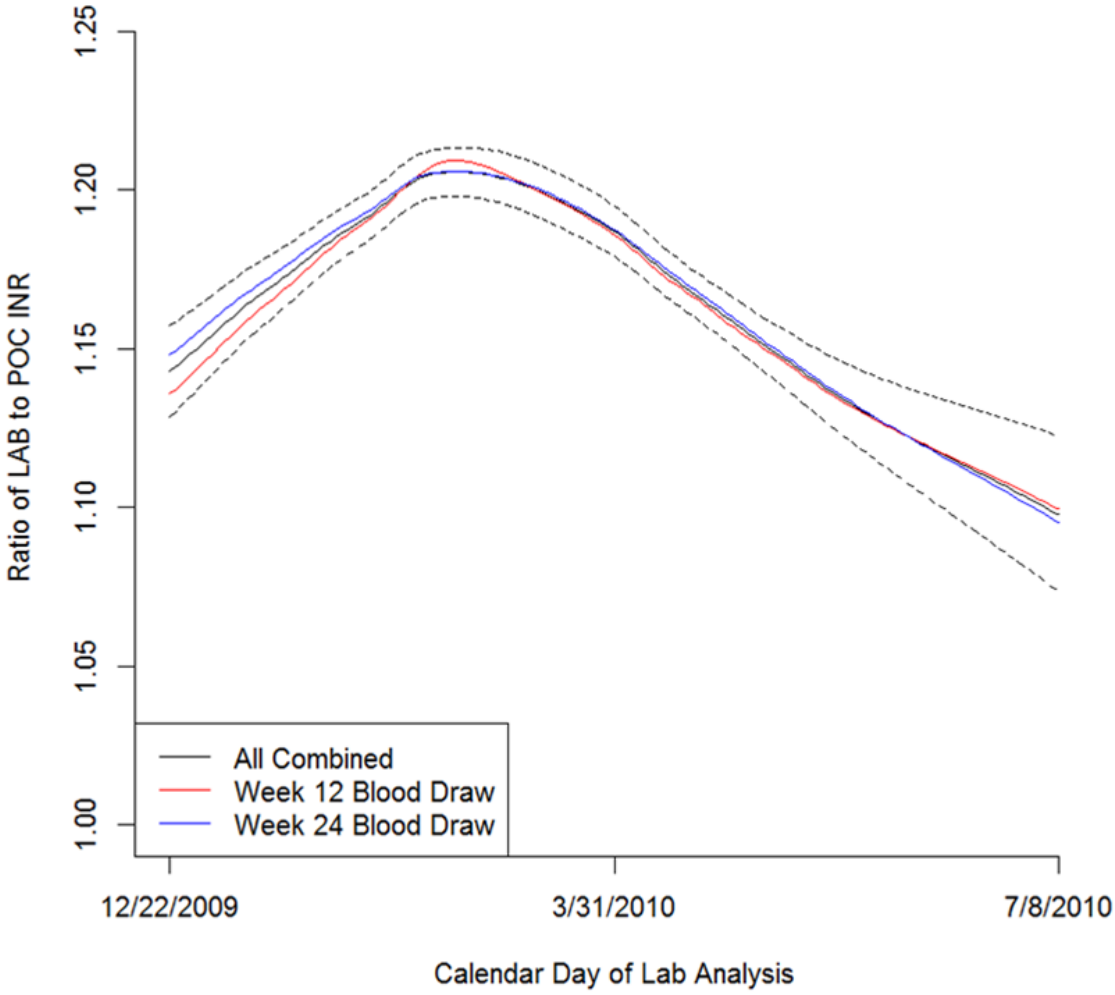


Figure 4 Loess regression model showing relationship between Calendar Date of LAB INR analysis and ratio of LAB INR to POC INR.

On average, about 360 events imputed in warfarin arm (compared to 386 observed). The estimate of Hazard Ratio for Major Bleeding from the model and the imputed datasets is 1.12 with 95% CI (0.97, 1.30).

1.5 Evaluation of Safety

NA.

1.6 Benefit-Risk Assessment (Optional)

Given the effectiveness of Xarelto observed in the ROCKET-AF trial, the risk benefit of Xarelto remains favorable and worthy of approval. Xarelto has not been demonstrated to be superior to warfarin at reducing strokes (this is stated in the label and remains true). Xarelto has also not been shown to be superior to warfarin in terms of safety (major bleeding).

FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

1.7 Gender, Race, Age, and Geographic Region

NA.

1.8 Other Special/Subgroup Populations

NA.

SUMMARY AND CONCLUSIONS

1.9 Statistical Issues

Multiple imputation and modeling cannot replace data from a well-controlled clinical trial. In this case, there was some concern about the warfarin arm when Xarelto was approved. The time in therapeutic range for the warfarin arm was approximately 55%, where we would expect it to be- and have seen it in other trials- close to 70%. Xarelto was approved despite that. If there had been any new issues about efficacy, then this type of analysis used in the review could not save the trial. In this case, we can still rely on the efficacy data from the trial. Given the effectiveness and the relative importance of stroke and SEE relative to Major Bleeding, there is a moderately large margin for error on safety. Xarelto could increase the rate of major bleeding by as much as 50% compared to warfarin and it would still be considered an approvable therapy.

1.10 Collective Evidence

NA.

1.11 Conclusions and Recommendations

No actions are recommended regarding Xarelto.

1.12 Labeling Recommendations (as applicable)

The label already states that "There is insufficient experience to determine how XARELTO and warfarin compare when warfarin therapy is well-controlled." That statement is sufficient. No further labeling changes are recommended.

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

/s/

JOHN P LAWRENCE
09/10/2016
Impact of Alere INR device on interpretation of ROCKET-AF trial

JEONGSOOK L KIM
09/11/2016

Clinical Pharmacology Review

NDA	202439
Submission type	Technical Report – Post-Marketing Safety
Brand name	Xarelto
Generic name	Rivaroxaban
Sponsor	Jansen Pharmaceuticals Inc
Reviewer	Tzu-Yun McDowell, PhD
	Jeffry Florian, PhD
OND Division	Division of Cardiovascular and Renal Products

Background

Rivaroxaban (Xarelto®), a direct oral anticoagulant (DOAC), was approved for the prevention of stroke and systemic embolism in patients with non-valvular atrial fibrillation (AF) in 2011 based on the confirmatory trial: "Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCEKT AF)"[1]. The Food and Drug Administration (FDA) was informed in September 2015 by Janssesn Research & Development LLC (JRD), the sponsor of rivaroxaban, that the point-of-care (POC) warfarin monitoring device (i.e. Alere INRatio Monitor system) that was used to measure the international normalized ratio (INR) and guide the warfarin dosing in the ROCKET AF trial was subject a class 1 recall on December 5, 2014[2]. The recall correction notice was issued based on post-marketing information indicating that the POC device may provide INR reading lower than a plasma-based laboratory INR in patients with certain medical conditions (e.g. Conditions associated with raised fibrinogen levels). Analysis comparing central-laboratory INR values (Lab INR) on blood samples collected on the same day that the POC device were performed at 12 weeks and 24 weeks (paired INR samples) in the ROCKET AF trial also showed that POC INR was on average 13% lower compared to the Lab INR; this difference was observed in patients regardless the medical conditions listed in the recall [3]. As a result, warfarin-treated patients in the ROCKET AF may have been over-anticoagulated due to the use of the inaccurate POC monitoring device, potentially increasing the risk of bleeding –related events , including hemorrhagic stroke, but decreasing risk of ischemic strokes compared to a properly functioning POC device. These distortions might have affected interpretation of both safety and efficacy results in the ROCKET AF trial. Furthermore, the use of this POC device may have

potentially altered the benefit-risk of rivaroxaban relative to warfarin as assessed within ROCKET AF.

Goals

To evaluate how use of the Alere POC device in ROCKET-AF may have impacted trial results based on the relationship between INR and outcomes developed using the warfarin data from other approved NOAC trials

Methods

FDA has approved four DOAC products to reduce the risk of stroke and systemic embolism in patients with non-valvular atrial fibrillation (AF) during 2010-2015. In order to predict outcomes in ROCKET-AF assuming a different INR POC device was used, subject-level data from the other three warfarin-controlled and randomized DOAC clinical trials were combined in the present analyses: (1) Randomized Evaluation of Long-term Anticoagulant Therapy (RE-LY), (2) Apixaban for Reduction in Stroke and Other Thromboembolic Events (ARISTOTLE) and (3) The Effective Anticoagulation with Factor Xa Next Generation in AF (ENGAGE-AF). The justification to pool the warfarin arm data from these trials was based on the comparable nature of the trial design, recent start and completion of all studies relative to each other, and trial construct that follows rigorous regulatory standards. In addition, there is no information to date to suggest that there was a systemic concern with the accuracy of the INR assessment in any of these trials.

A total of 22,063 warfarin-treated subjects from RE-LY, ARISTOTLE and ENGAGE-AF trials were pooled to assess the relationship between INR and the clinical outcome events of interest including major bleeding, as defined by the International Society of Thrombosis and Haemostasis (ISTH)[4], life threatening and fatal bleeding using the definition for Global Use of Strategies to Open Occluded Arteries (GUSTO)[5] sever major bleeding along with fatal bleeding, hemorrhagic stroke, and ischemic stroke. In these trials, INR was measured at least monthly for a majority of warfarin-treated subjects and warfarin was dose-adjusted to achieve an INR between 2.0 to 3.0. Median INR value in patients from these trials based on last INR proximal to the time of event or censoring was 2.3 with an interquartile range of 1.9 to 2.8.

For each outcome of interest, a multivariate Cox proportional hazard (PH) model was developed to examine the time to the first occurrence of an on-treatment event as a function of INR and other patient covariates. On-treatment was defined as the period between administration of the first dose of the study drug and the pre-specified days (between 2-5 days, it slightly varied across trials based on half-life of each study drug) after the receipt of the last dose in each trial. We used the last observed INR, which was defined as the last measured INR value prior to or on the date of the first outcome event of interest (INR value closest to the censored date if no event) to explore the INR-outcome event relationship. This INR value was selected as a best representation of an individual patient's INR reading proximal to the time of event or censoring. In addition, INR values >6 were truncated to 6. A set of common baseline covariates collected

in these studies, which could be potentially associated with the outcome of interest was obtained and tested in Cox PH model. These covariates included age, sex, race (white/non-white), baseline body weight, baseline aspirin use, baseline antiplatelet use, baseline CHADS2 score, history of stroke or transient ischemic attack (TIA), diabetes, baseline creatinine clearance (categorical as normal, mild, moderate, and severe based on Cockcroft-Gault equation), smoking history and alcohol use. Covariates in the Cox PH model were selected using stepwise forward addition followed by backward elimination based on Bayesian information criteria (BIC) and the *stepAIC* function from the 'MASS' package. Considering that INR management varies geographically, sensitivity analysis was conducted using Nnorth American patients alone. All the analyses and plots were conducted and generated in R (version 3.1.2) and/or SAS 9.3.

Results

Table 1 shows demographic and clinical characteristics of warfarin-treated patients in the four NOAC trials. For the most part, characteristics of patients were similar across the trials except that ROCKET-AF trial included a higher proportion of patients with CHARDS2 score > 3 and with prior history of stroke/TIA. Table 2-Table 5 show the parameter estimates for various Cox PH models based on warfarin data from ENGAGE, ARISTOTLE, and RE-LY. As expected, last observed INR was a significant predictor for each outcome of interest. Figure 1 shows the event rate for ischemic stroke and life threatening/fatal bleeding by INR for a typical patient (i.e. 70 year old white man with a prior history of stroke/TIA and CHADS score >2 and without use of aspirin) on the basis of our Cox models.

Inferences

Based on these established INR-outcome event relationships, we then estimated the potential impact of the Alere POC device on the ROCKET trial results, assuming POC INRs measured in the trial were on average 13% lower than the "true" INRs (the "true" INRs would be on average 15% higher than the POC INRs). Our models estimate that the bleed risk for warfarin arm in the ROCKET trial would be reduced by ~10%, at most, and the risk for ischemic stroke would be increased by ~20% if the POC device had not underreported the INR. "Modified" warfarin event rates and rate ratios (RR) compared to the observed event rates for rivaroxaban were calculated and presented along with the observed HRs/ RRs for each outcome of interest to assess the impact of POC device (Table 6). Our analyses show that the modified RR for bleeding-related endpoints was only slightly higher compared to the observed results in the ROCKET trail, suggesting that the impact of Alere POC device on the ROCKET bleeding results was modest. Sensitivity analysis on North American patients alone revealed that INR values were not a significant covariate except on major bleeding and ischemic stroke (Table 7-Table 10). This observation may have been due to a lack of power in the subset analysis to identify significant covariates. For these two endpoints, a 22% increase in ischemic stroke and 9% decrease in major bleeding are predicted based on the assumption that the INR values in the trial would average 13% lower than the "true" INRs. These predictions are in agreement with those from the full population.

Limitations

While our analyses demonstrated a modest impact of erroneous POC device on the ROCKET trial results, there are some limitations one should consider in interpreting the results. First, it is important to note that all our models rely on the last observed INR prior to the event or the censored date, assuming this INR value would be best related to the outcome of interest. However, it is possible, in some cases, that the last observed INR value may differ from the INR at the time of the event if there has been a recent change in dosing. Secondly, we know that INR values may vary over the course of the study. Rather than including warfarin as a time-varying covariate, we used only the INR value closest to the time of the event. Thirdly, we assessed the impact of POC INR using a fixed adjustment in each INR value (i.e. true INR was 15% higher compared to the POC INR in ROCKET). This approach assumes a constant positive bias from POC device for every subject in ROCKET, which is a straightforward way to assess the impact of the POC but it did not take into account individual variations and could potentially underestimate the impact for some individuals. With the limited data provided, a reliable estimate of variability seems unlikely.

Table 1 Demographic and Clinical Characteristics of Warfarin Patients in the Four NOAC Trials

Variables	ARISTOTLE N=9052	ENGAGE-AF N=7012	RE-LY N=5999	ROCKET-AF N=7082
Age, mean (SD)	69 (9.7)	71 (9.4)	72 (8.6)	71 (9.4)
Sex, n (%) Male	5879 (65%)	4383 (63%)	3796 (63%)	4283 (61%)
Race, n (%) White	7469 (83%)	5679 (81%)	4158 (69%)	5909 (83%)
CHADS2 score, n (%) ≤ 1 2-3 4-6	 3076 (34%) 4834 (53%) 1142 (13%)	 5 (0.07%) 5422 (77%) 1585 (23%)	 1860 (31%) 3418 (57%) 721 (12%)	 2 (0.03%) 4062 (57%) 3018 (43%)
Prior Stroke/TIA, n (%) Yes	1735 (19%)	1983 (28%)	1191 (20%)	3692 (52%)
CRCL* (mL/min), n (%) <30 30-50 >50-<80 ≥80	 132 (1%) 1380 (15%) 3757 (42%) 3747 (42%)	 51 (1%) 1307 (19%) 3045 (43%) 2609 (37%)	 29 (1%) 1047 (18%) 2796 (49%) 1872 (33%)	 4 (0.1%) 1581 (22%) 3164 (45%) 2324 (33%)
Prior Warfarin use, n (%) Yes	5180 (57%)	4124 (59%)	4035 (67%)	4437 (63%)
Prior Aspirin use, n (%) Yes	2762 (31%)	2083 (30%)	2431 (41%)	2606 (37%)

* Estimated using Cockcroft-Gault

Table 2 Final Cox proportional hazards model and hazard ratios for major bleeds using all warfarin data from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	0.321	1.378	(1.234, 1.538)	<0.001
Age (years)	0.047	1.048	(1.041,1.055)	<0.001
Race (White)	-0.3	0.741	(0.653, 0.84)	<0.001
Diabetes (Yes)	0.225	1.252	(1.115, 1.407)	<0.001
Last observed INR	0.297	1.345	(1.262, 1.434)	<0.001

Table 3 Final Cox proportional hazards model and hazard ratios for life-threatening or fatal bleed using all warfarin data from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	0.489	1.631	(1.362, 1.953)	<0.001
Age (years)	0.042	1.043	(1.031, 1.054)	<0.001
Race (White)	-0.498	0.608	(0.499, 0.741)	<0.001
Last observed INR	0.366	1.442	(1.305, 1.594)	<0.001
CHADS2 $\leq$2	-0.364	0.695	(0.579, 0.833)	<0.001

Table 4 Final Cox proportional hazards model and hazard ratios for ischemic stroke using all warfarin data from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Age (years)	0.041	1.042	(1.029, 1.055)	<0.001
Race (White)	-0.567	0.567	(0.452, 0.711)	<0.001
Prior stroke/TIA (yes)	0.913	2.492	(2.013, 3.085)	<0.001
Last observed INR	-0.605	0.546	(0.462, 0.645)	<0.001

Table 5 Final Cox proportional hazards model and hazard ratios for hemorrhagic stroke using all warfarin data from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	0.486	1.626	(1.218, 2.17)	<0.001
Age (years)	0.039	1.039	(1.022, 1.057)	<0.001
Race (White)	-0.888	0.412	(0.306, 0.554)	<0.001
Prior stroke/TIA (yes)	0.699	2.011	(1.496, 2.704)	<0.001
Last observed INR	0.441	1.554	(1.324, 1.824)	<0.001

Figure 1 Probability of ischemic stroke and life threatening/fatal bleeding within one year as a function of the last observed INR using all warfarin data from RE-LY, ARISTOLE, and ENGAGE for a typical patient (i.e. 70 year old white male with a prior history of stroke/TIA and CHADS score >2 and without use of aspirin). The shaded region represents the 95% confidence interval

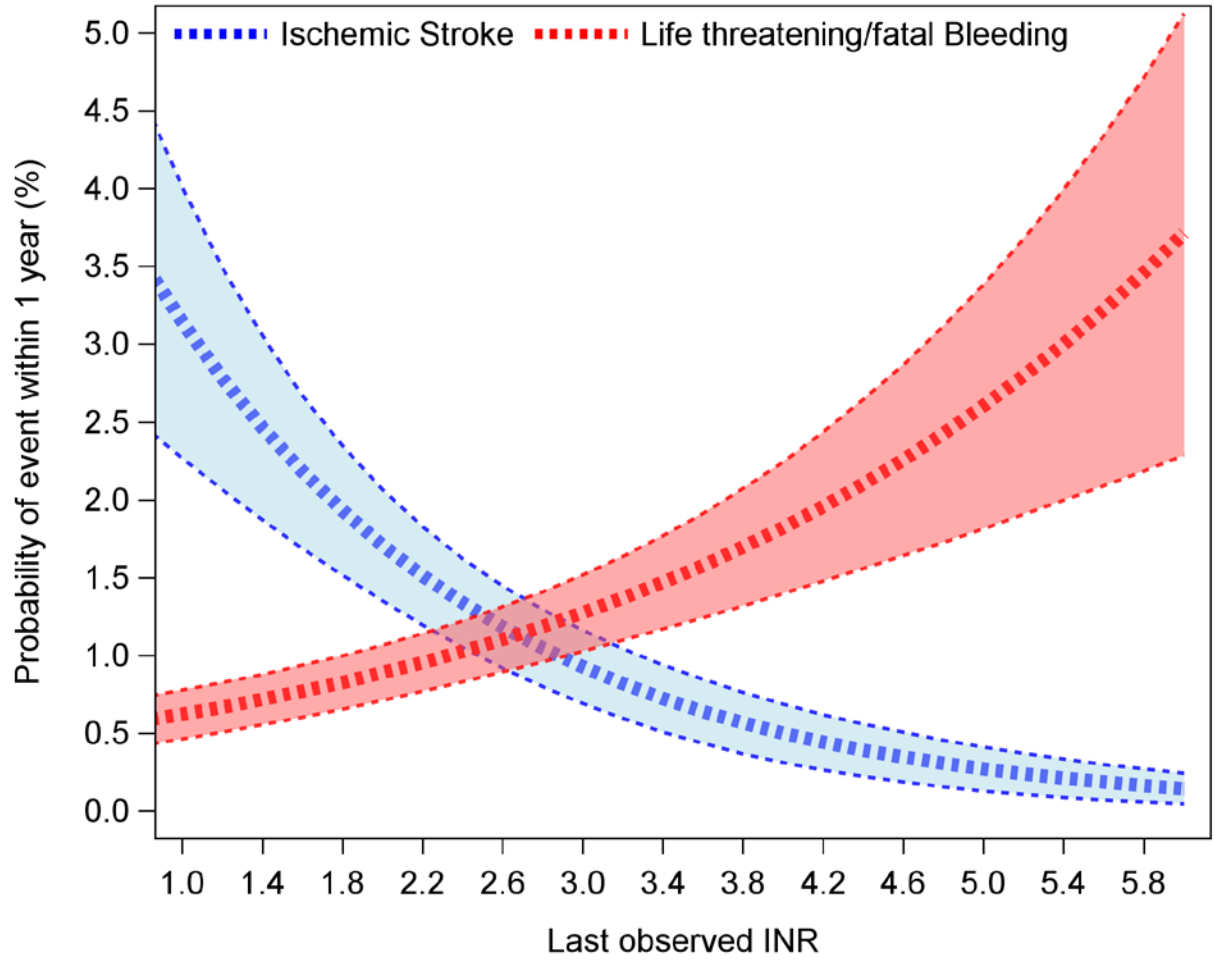


Table 6 Comparison of ROCKET AF Trial Results and Results based on a modified warfarin event rate (ER)

	ROCKET AF Trial Results†				Modified Results	
Outcome Event	**Rivaroxaban ER(%pt-yr)**	**Warfarin ER (%pt-yr)**	**HR (95% CI)**	**RR**	**Modified warfarin ER(%pt-yr)**	**Modified RR**
Major Bleeding	3.61	3.45	1.05 (0.91, 1.20)	1.04	3.11	1.16
Life Threatening/Fatal Bleeding	1.64	1.93	0.85 (0.70, 1.04)	0.85	1.74	0.94
Hemorrhagic Stroke	0.26	0.44	0.59 (0.37, 0.93)	0.59	0.40	0.65
Ischemic Stroke	1.34	1.42	0.94 (0.74,1.17)	0.94	1.70	0.79

† ROCKET trial results was on treatment (last dose plus 2 days) analysis in the safety population

Table 7 Final Cox proportional hazards model (exponential distribution) and hazard ratios for major bleeds using warfarin data from patients enrolled in **North America** from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	0.327	1.386	(1.153, 1.667)	<0.001
Age (years)	0.039	1.039	(1.028,1.051)	<0.001
Chronic Heart Failure (Yes)	0.308	1.361	(1.118, 1.656)	0.002
Last observed INR	0.288	1.334	(1.191, 1.495)	<0.001

Table 8 Final Cox proportional hazards model (exponential distribution) and hazard ratios for life-threatening or fatal bleed using warfarin data from patients enrolled in **North America** from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	0.699	2.011	(1.438, 2.812)	<0.001
Age (years)	0.044	1.045	(1.024, 1.067)	<0.001

Table 9 Final Cox proportional hazards model (exponential distribution) and hazard ratios for ischemic stroke using warfarin data from patients enrolled in **North America** from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Age (years)	0.041	1.042	(1.02, 1.07)	0.002
Prior stroke/TIA (yes)	1.043	2.838	(1.81, 4.45)	<0.001
Last observed INR	-0.624	0.536	(0.37, 0.77)	<0.001

Table 10 Final Cox proportional hazards model (exponential distribution) and hazard ratios for hemorrhagic stroke using warfarin data from patients enrolled in **North America** from RE-LY, ARISTOLE, and ENGAGE.

Coefficient	Estimate	HR	HR 95% CI	p-value
Baseline aspirin use (Yes)	1.056	2.874	(1.505, 5.489)	0.001
Body weight (kg?)	-0.028	0.973	(0.956, 0.989)	0.001

References

1. Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 2011;365:883-91.
2. Food and Drug Administration. Alere INRatio2 PT/INR Professional Test Strips: recall — higher INR when performed by central laboratory. 2014:http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm396324.htm.
3. European Medicines Agency assessment report: Xarelto. 2016:http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000944/WC500201726.pdf.
4. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. J Thromb Haemost 2005;3:692-4
5. GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. N Engl J Med.1993;329:673–682.

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

/s/

JEFFRY FLORIAN
09/12/2016

TZU-YUN C MCDOWELL
09/12/2016