Apixaban for the prevention of stroke and systemic embolism in people with non-valvular atrial fibrillation
STA REPORT

This report was commissioned by the NIHR HTA Programme as project number 11/49

BMJ Technology Assessment Group

Title: Apixaban for the prevention of stroke and systemic embolism in people with non-valvular atrial fibrillation

Produced by: BMJ Technology Assessment Group (BMJ-TAG)

Authors: Steve Edwards, Head of BMJ-TAG, London

Victoria Hamilton, Health Technology Assessment Analyst, BMJ-TAG, London

Nicola Trevor, Health Economics Lead, BMJ-TAG, London

Leo Nherera, Health Economist, BMJ-TAG, London

Charlotta Karner, Health Technology Assessment Analysis Lead, BMJ-TAG, London

Elizabeth Thurgar, Health Economist, BMJ-TAG, London

Correspondence to: Steve Edwards, Head of BMJ-TAG, BMJ-TAG, BMJ Evidence Centre, BMJ Group, BMA House, Tavistock Square, London, WC1H 9JP.

Date completed: 23/10/2012

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 11/49

Declared competing interests of the authors:
None.

Acknowledgements:
The ERG would like to thank Dr Derick Todd (Consultant Cardiologist) for providing clinical advice throughout the project. Thanks also to Dr James Nicholson (Consultant Haematologist) for providing feedback on the clinical sections of the report.

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows: Edwards SJ, Hamilton V, Trevor N, Nherera L, Karner C, Thurgar E. Apixaban for the prevention of stroke and systemic embolism in people with non-valvular atrial fibrillation: A Single Technology Appraisal. BMJ-TAG, London, 2012.

Contributions of authors:
Please refer to the International Committee of Medical Journal Editors (ICMJE) **Uniform Requirements for Manuscripts Submitted to Biomedical Journals see** http://www.icmje.org/

Steve Edwards	Project lead: supervised the production of the final report; critical appraisal of the manufacturer's submission; critical appraisal of the clinical evidence; and critical appraisal of the economic evidence
Victoria Hamilton	Critical appraisal of the manufacturer's submission; critical appraisal of the clinical evidence; cross checking of manufacturer's search strategies; and drafted the background and clinical effectiveness sections
Charlotta Karner	Critical appraisal of the manufacturer's submission; and critical appraisal of the clinical evidence
Leo Nherera	Critical appraisal of the manufacturer's submission; critical appraisal of the economic model; critical appraisal of the economic evidence; and drafted the economic analysis sections
Elizabeth Thurgar	Critical appraisal of the manufacturer's submission; and critical appraisal of the economic evidence
Nicola Trevor	Critical appraisal of the manufacturer's submission; critical appraisal of the economic model; critical appraisal of the economic evidence; and drafting and finalisation of the economic analysis sections

All authors read and commented on draft versions of the ERG report

Table of contents

Abbreviations

ACE	Angiotensin-converting enzyme
AF	Atrial fibrillation
ARB	Angiotensin-receptor blocker
ARISTOTLE	Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation
ARISTOTLE-J	Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation – Japan
AVERROES	Apixaban Versus Acetylsalicylic Acid [ASA] to Prevent Stroke in Atrial Fibrillation Patients Who Have Failed or Are Unsuitable for Vitamin K Antagonist Treatment
BD	Twice daily
$CHADS_2$	Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score
CHA2DS2-VASc	Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled], Vascular Disease, Age 65–74, and Sex category [female]) risk score
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CNS	Central nervous system
CRNM	Clinically relevant non-major
CSR	Clinical study report
cTTR	Study centre time in therapeutic range
CV	Cardiovascular
DMC	Drugs and safety monitoring committee
DVT	Deep vein thrombosis
ECG	Electrocardiogram
EMA	European Medicines Agency
ERG	Evidence review group
ESC	European Society of Cardiology
FDA	Food and Drugs Administration
GI	Gastrointestinal
GP	General practitioner
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value
ICH	Intracranial haemorrhage
INR	International normalised ratio
ISTH	International Society on Thrombosis and Haemostasis
ITT	Intention to treat
LVEF	Left ventricular ejection fraction
MI	Myocardial infarction
MS	Manufacturer's submission
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NMA	Network meta-analysis
NOAC	Novel oral anticoagulant
NSAID	Non-steroidal anti-inflammatory drug
NVAF	Non-valvular atrial fibrillation
NYHA	New York Heart Association

OD	Once daily
PE	Pulmonary embolism
POC	Point of care
RCT	Randomised controlled trial
RE-LY	Randomized Evaluation of Long-Term Anticoagulation Therapy
ROCKET-AF	Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation
SD	Standard deviation
SE	Systemic embolism
SG	Standard gamble
SMR	Standardised mortality ratio
SR	Systematic review
STA	Single technology appraisal
TA	Technology appraisal
TIA	Transient ischaemic attack
TTO	Time trade off
TTR	Time in therapeutic range
UK	United Kingdom
VKA	Vitamin K antagonist
VTE	Venous thromboembolic events
QOF	Quality and outcomes framework

1 SUMMARY

1.1 Critique of the decision problem in the manufacturer's submission

The manufacturer of apixaban (Eliquis®, Bristol-Myers Squibb and Pfizer) submitted to the National Institute for Health and Clinical Excellence (NICE) clinical and economic evidence in support of the effectiveness of apixaban for the prevention of stroke and systemic embolism (SE) in people with atrial fibrillation (AF). At the time of writing apixaban did not have marketing authorisation for the indication that is the focus of this single technology appraisal (STA). However marketing authorisation is expected in December 2012 and anticipated to be for "the prevention of stroke and systemic embolism (SE) in adult patients with non-valvular atrial fibrillation (AF) with one or more risk factors". In addition, on the 20th September 2012 the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion of the use of Eliquis® (apixaban) in this indication.

The final scope issued by NICE for this STA requested a population of "adults with non-valvular AF who are at risk of stroke or systemic embolism". The evidence review group (ERG) notes that the manufacturer presented disaggregated evidence for the population specified in the NICE scope. The manufacturer presented clinical effectiveness and cost-effectiveness evidence in patients who are suitable for and patients who are unsuitable for vitamin K antagonist (VKA) therapy; both patient groups are at risk of stroke or SE. However, the ERG notes that no evidence was presented for low risk patients (i.e. $CHADS_2$ score of 0) as this population is outside of the licensed indication anticipated for apixaban. Furthermore, in addition to the comparators listed in the final scope, the manufacturer included clinical and economic comparisons with aspirin. The manufacturer's rationale for including aspirin as a comparator was that it is currently recommended and widely used to treat VKA unsuitable patients in England and Wales. Therefore, with respect to population and comparators, the ERG considers that the manufacturer has both adhered to and gone beyond the scope of the decision problem.

Regarding outcomes, the manufacturer excluded transient ischaemic attack (TIA) from the clinical and cost-effectiveness evidence submitted. However, this is because data on TIA were not collected in either of the key trials that informed the manufacturer's submission (MS); ARISTOTLE and AVERROES. Furthermore, health-related quality of life data (HRQoL) were not collected in ARISTOTLE or AVERROES. Therefore, the manufacturer submitted HRQoL data identified from a systematic review of the literature.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The manufacturer submitted clinical effectiveness evidence for the comparison of apixaban versus warfarin based on a trial named ARISTOTLE. In addition, the manufacturer presented data for apixaban versus aspirin in people for whom warfarin (a VKA) is unsuitable based on a trial named AVERROES. The ERG notes that the comparator (aspirin) in AVERROES was not requested in the final scope issued by NICE. Reporting of AVERROES is thus restricted to Appendix 9.3. A total of 20,098 patients were enrolled in ARISTOTLE and 18,201 were randomised to treatment. The intention to treat (ITT) population for the apixaban treatment group consisted of 9,120 patients and the warfarin group 9,081. In the MS the manufacturer reported that 9,088 patients were treated in the apixaban group and 9,052 in the warfarin group; i.e. they received at least one dose of study drug and thus were included in the safety analyses.

Time in therapeutic range (TTR) is a measure of warfarin control and not necessarily the same as treatment compliance but nonetheless important in assessing warfarin treatment efficacy. The target international normalised ratio (INR) for ARISTOTLE was an INR in the range of 2.0–3.0. The mean TTR for patients in the warfarin arm of ARISTOTLE was 62.2% and the median TTR was 66.0%.

Regarding discontinuations, significantly fewer (defined as $p<0.5$) patients permanently discontinued treatment in the apixaban group compared with the warfarin group (25.3% vs 27.5% respectively; $p=0.001$).

The primary objective of ARISTOTLE was to prove the non-inferiority of apixaban versus warfarin in the prevention of stroke and SE. Non-inferiority was proven based on the definition of the upper bound of the 2-sided 95% confidence interval (CI) for the hazard ratio (HR) <1.38 [redacted] and ITT populations (HR 0.79; 95% CI 0.66–0.95). In addition, superiority was proven as the HR for the primary efficacy end point of stroke and SE was 0.79 and the upper bound of the 95% CI was <1 (95% CI 0.66–0.95). This suggests that apixaban was associated with significantly fewer stroke and systemic emboli when compared with warfarin ($p=0.01$).

Apixaban was associated with statistically fewer haemorrhagic strokes compared with warfarin (HR 0.51; 95% CI 0.35–0.75; $p<0.001$) although there was no statistically significant difference between apixaban and warfarin for the individual outcomes of ischaemic stroke (HR 0.92; 95% CI 0.74–1.13; $p=0.42$) and SE (HR 0.87; 95% CI 0.44–1.75; $p=0.70$). In addition, apixaban was associated with significantly fewer all-cause deaths than warfarin (HR 0.89; 95% CI 0.80–0.99; $p=0.047$). However, the differences observed in myocardial infarction (MI) and pulmonary embolism (PE) or deep vein thrombosis (DVT) were not statistically significant ($p=0.37$ and $p=0.63$, respectively).

The primary safety outcome of ARISTOTLE was International Society on Thrombosis and Haemostasis (ISTH) major bleeding and apixaban was proven to be superior to warfarin in reducing these bleeding events (HR 0.69; 95% CI 0.60–0.80, p<0.01). The ERG notes that apixaban was associated with a non-significant (p>0.05) reduction in gastrointestinal (GI) bleeding events when compared with warfarin (HR 0.89; 95% CI 0.70–1.15, p=0.37).

The overall adverse event and safety profile of apixaban was comparable or better when compared with warfarin for the outcomes reported in the MS.

The results of ARISTOTLE for the post-hoc subgroup analysis of Western Europe, demonstrated [redacted]. However, ARISTOTLE was not powered to detect differences in efficacy and safety in the different geographical subgroups.

A subgroup analysis based on study centre INR control was conducted using the median of each study centre's individual patients' TTR (cTTR). Four ranges of cTTR were considered for analysis in accordance with the quartiles of cTTR observed in ARISTOTLE. The results for the cTTR subgroup analyses reported in ARISTOTLE suggested that, regardless of INR control, apixaban was associated with fewer stroke or SE and major bleeding events than warfarin. However, this difference was not statistically significant.

As a result of the absence of head-to-head trials, the manufacturer carried out two network meta-analyses (NMAs). The aim of these NMAs was to compare apixaban with dabigatran and rivaroxaban, as specified in the final scope issued by NICE. NMA 1 was comprised of patients suitable for treatment with warfarin and compared apixaban, warfarin, dabigatran and rivaroxaban. NMA 2 was reported to be in a population of patients unsuitable for VKA therapy and compared apixaban, dabigatran, rivaroxaban and aspirin. However, the ERG notes that the majority (3 out of 4) of the trials in NMA 2, included patients who were suitable for treatment with warfarin (ARISTOTLE, RE-LY and ROCKET-AF). The ERG considers it inappropriate to combine results from trials in VKA suitable patient populations with trial results for VKA intolerant patient populations as similar efficacy for novel oral anticoagulants (NOACs) has yet to be demonstrated. Thus the focus of this ERG report is NMA 1.

NMA 1 was comprised of the following three randomised controlled trials (RCTs):

- ARISTOTLE: apixaban 5 mg twice daily (BD) vs warfarin dosed to achieve a target INR 2.0–3.0;
- RE-LY: dabigatran 110 mg BD vs dabigatran 150 mg BD vs warfarin dosed to achieve a target INR 2.0–3.0;
- ROCKET-AF: rivaroxaban 20 mg once daily (OD) vs warfarin dosed to achieve a target INR 2.0–3.0.

The base case results of NMA 1 suggested that apixaban was associated with a significantly lower incidence of MI compared with dabigatran 150 mg or 110 mg. [redacted]

There were no further statistically significant differences for the efficacy outcomes between apixaban and its comparators in NMA 1.

For the bleeding safety outcomes, apixaban was associated with significantly fewer events for the following outcomes:

- intracranial haemorrhage (ICH) compared with rivaroxaban [redacted];
- major bleeding compared with rivaroxaban, with dabigatran 150 mg [redacted];
- GI bleeding compared with rivaroxaban and with dabigatran 150 mg;
- other major bleeding compared with rivaroxaban, with dabigatran 150 mg [redacted];
- clinically relevant non-major (CRNM) bleeding compared with rivaroxaban [redacted];
- any bleeding compared with rivaroxaban, with dabigatran 150 mg, with dabigatran 110 mg [redacted].

In addition, apixaban was associated with significantly fewer discontinuations compared with dabigatran 150 mg, dabigatran 110 mg, rivaroxaban [redacted]

The manufacturer carried out two sensitivity analyses. Sensitivity analysis 1 (SA 1) utilised post-hoc data from RE-LY (published as an updated analysis in 2010) rather than the original data published in 2009. The results of SA 1 were generally consistent with those of the base case NMA 1. Sensitivity analysis 2 (SA 2) utilised the safety on treatment dataset from ROCKET-AF instead of the ITT data which are used in the base case NMA 1. [redacted]

[redacted]In addition, subgroup analyses were conducted by the manufacturer "to explore the consistency of treatment effects across stroke risk severity ($CHADS_2$) and cTTR patient subgroups in accordance with the NICE scope". These subgroup analyses were limited to the outcomes of stroke or SE and major bleeds. The subgroups of the cTTR analyses were defined differently in each of the included trials. Therefore, the ERG does not consider the subgroup analyses to be directly comparable.[redacted]

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

For this STA, two RCTs (ARISTOTLE and AVERROES) were included in the clinical effectiveness section of the MS to provide clinical data on apixaban. Based on the final scope issued by NICE, the ERG considers that of the two trials only ARISTOTLE met the inclusion criteria for this STA. The ERG considers that apixaban vs aspirin was not a comparison of interest specified in the NICE final scope and thus do not consider AVERROES to meet the inclusion criteria for this STA. However, based on clinical advice, the ERG acknowledges that aspirin is utilised in clinical practice for some patients in the United Kingdom (UK).

The ERG considers that the inclusion and exclusion criteria for ARISTOTLE were acceptable to address the trial's objectives. In addition, the ERG notes that the baseline characteristics of the randomised populations of ARISTOTLE appeared to be well balanced between trial arms. Also, based on the advice of clinicians regarding the TTR expected in a UK patient population, the ERG considers that the mean TTR for ARISTOTLE (62.2%) was acceptable.

The ERG acknowledges that the outcome data reported from ARISTOTLE appeared to be consistent with the data collected in the trial. However, the ERG notes that TIA and HRQoL data requested in the NICE final scope were not collected in the trial. In terms of duration of follow-up and the statistical data analysis plan, the ERG considers that both were acceptable for the outcomes assessed.

The results of ARISTOTLE demonstrated that apixaban was superior to warfarin in the reduction of stroke or SE. The ERG considers that the overall adverse event and safety profile of apixaban was comparable or better when compared with warfarin for the outcomes reported in the MS. Subgroup analyses suggested that European patients may have derived slightly less benefit from apixaban for both the efficacy and safety outcomes when compared with the whole trial population. However, the

results of the subgroup analyses by cTTR suggested that the safety and efficacy of apixaban compared with warfarin were independent of the level of warfarin control i.e. %TTR. With respect to subgroup analyses by $CHADS_2$ score categories, [redacted]. However, the lack of detailed individual $CHADS_2$ score data, particularly for the higher $CHADS_2$ scores (i.e. 3, 4, 5 and 6) limits the ability of the ERG to comment on any potential variation in apixaban treatment effect for these subgroups.

Regarding the NMAs, the ERG considers it important to highlight the following potential sources of clinical heterogeneity between the three trials included in NMA 1:

- ARISTOTLE and ROCKET-AF were double-blind, double dummy trials whereas treatment allocation to dabigatran or warfarin was not concealed in RE-LY;
- ROCKET-AF was comprised of a higher stroke/SE risk population based on baseline $CHADS_2$ score compared with ARISTOTLE and RE-LY (mean baseline $CHADS_2$ 3.6, 2.1 and 2.1, respectively);
- mean %TTR was lower in ROCKET-AF compared with ARISTOTLE and RE-LY (55%, 62%, and 64%, respectively);
- a statistically significant difference ($p<0.05$) in baseline MI between treatment groups in ROCKET-AF (16.6% in the rivaroxaban arm versus 18.0% in the warfarin arm);
- the "on treatment" population of ROCKET-AF were used for some outcomes analysed in NMA 1 due to the absence of published ITT data.

The manufacturer stated that due to the presence of only one study for each comparator, no studies were excluded from any of the analyses. This is because exclusion of any single study would have resulted in the exclusion of one of the treatments from the analysis.

Regarding the subgroup analyses for NMA 1, the ERG is concerned that the number of patients within each $CHADS_2$ score category for each of the trials may be disproportionately different. Therefore, aggregation of multiple scores may provide misleading results. In addition, for the cTTR subgroup analyses, the cTTR subgroups were defined differently in each of the included trials. Therefore, the ERG does not consider the cTTR subgroups to be directly comparable. The ERG considers that the results of the manufacturer's aggregation of the data for the $CHADS_2$ and cTTR subgroup analyses should be interpreted with caution.

1.4 Summary of cost-effectiveness evidence submitted by the manufacturer

Within the published literature, the manufacturer identified six economic evaluations of apixaban or relevant comparators in an AF patient population. In addition, the manufacturer summarised the economic evidence from two previous technology appraisals (TA249 and TA256). However, none of the economic evaluations identified by the manufacturer served to answer the decision problem regarding the cost-effectiveness of apixaban in a UK AF population. Therefore, the manufacturer constructed a *de novo* economic evaluation to assess the incremental cost-effectiveness of:

- apixaban, dabigatran blend (150 mg switching to 110 mg at the age of 80 years), dabigatran 110 mg, rivaroxaban and warfarin in a VKA suitable population;
- apixaban, dabigatran blend (150 mg switching to 110 mg at the age of 80 years), dabigatran 110 mg, rivaroxaban and aspirin in a VKA unsuitable population.

The manufacturer's economic evaluation was carried out within a Markov cohort model that consisted of 18 health states, including the absorbing state of death. Patients transitioned between health states in cycles of 6 weeks and only one clinical event permitted per cycle. The model captured all relevant thromboembolic (except TIA) and adverse events. Modelled events were categorised as temporary or permanent in accordance with the long-term impact associated with the event. A variety of treatment switches were permitted within the model. However, only a switch from first line anticoagulation therapy to second-line therapy with aspirin altered patients' risk profile. A switch from first-line to second-line therapy was either event-related (event-related discontinuation) or a result of other causes (other-cause discontinuation).

VKA suitable and VKA unsuitable patient populations were considered separately within the manufacturer's model. However, the baseline characteristics of both populations were assumed to be equivalent to the characteristics reported in the UK General Practice Research Database (GPRD). Data from the manufacturer's NMA 1 (based on data from RE-LY, ROCKET-AF and ARISTOTLE) and NMA 2 (based on data from RE-LY, ROCKET-AF, ARISTOTLE and AVERROES) were used to inform the clinical effectiveness of treatments in the VKA suitable and VKA unsuitable populations, respectively. HRQoL data were identified from a comprehensive systematic review of health state utility value (HSUV) literature and wherever possible EQ-5D quality of life data were used. In the absence of EQ-5D data, data elicited with time trade off (TTO) or standard gamble (SG) methods were used. Where available, resource use and cost data were derived from data reported in National Health Service (NHS) Reference Costs 2010/11. However, published literature was used as a complementary source for some included costs. Costs and benefits were discounted at 3.5% and a lifetime time horizon was adopted.

The base case incremental cost-effectiveness results in the VKA suitable patient population (with warfarin as the reference treatment) indicated that:

- dabigatran 110 mg was strictly dominated by (i.e. is less costly and less effective than) dabigatran blend;
- rivaroxaban and dabigatran blend were extendedly dominated (i.e. resulted in a lower incremental cost-effectiveness ratio (ICER) versus warfarin despite having higher total costs) by apixaban;
- apixaban had an ICER versus warfarin of £11,008 per quality adjusted life year (QALY).

In the VKA unsuitable patient population the manufacturer's base case incremental cost-effectiveness results (with aspirin as the reference treatment) were as follows:

- dabigatran 110 mg was strictly dominated by (i.e. is less costly and less effective than) dabigatran blend;
- dabigatran blend had an ICER of £1,111 per QALY versus aspirin;
- rivaroxaban had an ICER of £2,326 per QALY versus aspirin and an incremental ICER of £23,027 per QALY versus dabigatran blend;
- apixaban had an ICER of £2,903 per QALY versus aspirin and an incremental ICER of £8,401 per QALY versus rivaroxaban.

1.5 Summary of the ERG's critique of cost-effectiveness evidence submitted

The ERG notes that the modelling approach taken by the manufacturer was reasonable and consistent with other published economic evaluations of prophylactic interventions used in AF. Furthermore, the ERG notes that the model was well constructed, transparent, accurate and easy to navigate. However, the ERG considers that with the exception of the comparison of apixaban with aspirin, the economic evidence submitted in the VKA unsuitable patient population was potentially flawed. The ERG notes that this was a result of the paucity of efficacy data for dabigatran or rivaroxaban in a VKA unsuitable patient population. However, as aspirin was not included in the scope of the decision problem, the focus of the ERG report is the results of the manufacturer's economic evaluation in a VKA suitable population.

Generally, the ERG considers that the economic evidence submitted by the manufacturer in the VKA was robust and conservative; i.e. generally, bias potentially associated with model assumptions was likely to be against apixaban. However, the ERG identified some areas where the face validity and accuracy of the manufacturer's economic model could have been improved; although, the ERG acknowledges that the impact of these on the cost-effectiveness results was limited.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

1.6.1 Strengths

The ERG considers ARISTOTLE, the key trial addressing the final scope issued by NICE for this STA, to be a large, well conducted, international, double-blind, RCT. The primary efficacy outcome of ARISTOTLE was reduction in stroke and SE, for which apixaban was proved to be superior to warfarin. The ERG notes that the primary outcome of ARISTOTLE was also the key outcome of interest specified in the final scope issued by NICE for this STA.

The manufacturer conducted extensive subgroup analyses to demonstrate consistency of the treatment effect of apixaban. In addition, the manufacturer submitted additional data for a comparator, aspirin, which was not specified in the final scope issued by NICE.

NMAs were conducted by the manufacturer and enabled an indirect comparison of apixaban with rivaroxaban, dabigatran 150 mg, dabigatran 110 mg and warfarin.

The manufacturer submitted a well constructed, transparent, accurate and easy to navigate economic model. In addition, the economic evidence submitted by the manufacturer in the VKA was robust and conservative.

1.6.2 Weaknesses and areas of uncertainty

TIA was listed as an outcome of interest in the final scope issued by NICE, although no TIA data were presented in the MS. However, the ERG acknowledges that no data on TIA were collected in ARISTOTLE or AVERROES. The effectiveness of apixaban in reducing TIAs is thus an area of uncertainty.

In addition, the HRQoL data presented within the MS was limited to generic AF HRQoL data identified from a systematic review of the literature presented in the cost-effectiveness section of the MS. No treatment specific HRQoL data for apixaban were collected in ARISTOTLE or AVERROES and thus the effect of apixaban on HRQoL is a further area of uncertainty.

The ERG considers that as a result of differences between the trials included in NMA 1 there is potential clinical heterogeneity. The ERG acknowledges that there was only one study for each comparator in the network and that the network is ‘star-shaped’. The ERG thus considers that the conclusions drawn from NMA 1 would benefit from the addition of further head-to-head trials but acknowledges that no further data are currently available.

With respect to NMA 2, the ERG notes that the population of the included rivaroxaban and dabigatran trials consisted of patients suitable for treatment with warfarin. Therefore, the ERG considers that

NMA 2 was potentially flawed. The ERG considers that NMA 2 does not address the question of the effectiveness of apixaban compared with rivaroxaban or dabigatran in people for whom warfarin is unsuitable. However, the ERG also acknowledges that there are currently no published RCTs suitable for inclusion to address this question.

Finally, the ERG considers that there are some areas where the face validity and accuracy of the manufacturer's economic model could have been improved. However, the ERG acknowledges that the impact of these on the cost-effectiveness results was limited.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG (VKA suitable population only)

The ERG carried out numerous sensitivity analyses, in response to points of critique identified. However, the ERG notes that the incremental cost-effectiveness results were unaffected by any of the sensitivity analyses carried out. That is, dabigatran 110 mg continued to be strictly dominated by dabigatran blend and rivaroxaban and dabigatran blend remained extendedly dominated by apixaban. However, some of the ERG's sensitivity analyses were combined to form a revised ERG base case for the comparison of apixaban and warfarin. The ERG's revised base case included the following amendments:

- other-cause mortality was assumed to be independent of treatment received;
- utility was adjusted for age;
- stroke severity and bleed type distribution were assumed to be independent of treatment received;
- SE and MI patients were assumed to be at risk of recurrent stroke;
- the acute cost of SE was assumed to be equal to the cost used in the rivaroxaban HTA submission;
- a 26-year time horizon was used.

The ERG's revised base case resulted in an ICER of £12,757 for apixaban compared with warfarin; i.e. just under £2,000 greater than the manufacturer's base case ICER.

In addition, the ERG notes that there is uncertainty surrounding the choice of second-line treatment. Therefore, the ERG considered the following incremental analyses:

- apixaban, dabigatran blend, dabigatran 110 mg and rivaroxaban: using warfarin as second-line treatment;
- apixaban, dabigatran blend, dabigatran 110 mg and warfarin: using rivaroxaban as second-line treatment;
- apixaban, rivaroxaban and warfarin using dabigatran blend as second-line treatment.

The results of these analyses were highly variable, with incremental ICER for apixaban varying between £287 (versus warfarin when dabigatran 110 mg was chosen as second-line treatment) and

£60,366 (versus dabigatran blend, when rivaroxaban was chosen as second-line treatment). However, the ERG notes that within the manufacturer's model, patients on second-line treatment were exposed to a constant risk of events. Therefore, the results of these analyses should be interpreted with caution as the main driver of the ICERs was discontinuation rates associated with first-line therapy. That is, patients who discontinued treatment fared far better than in the base case. Therefore, treatments with higher discontinuation rates (e.g. dabigatran) appeared more effective than in the manufacturer's base case.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problems

In the Context section of the manufacturer's submission (MS; Section 2), the manufacturer provides an overview of the key aspects of atrial fibrillation (hereafter referred to as AF) relevant to the decision problem. These include pathophysiology, prevalence and stroke risk of AF. In addition, the manufacturer outlines the implications of stroke in AF patients, from a clinical and financial perspective. The evidence review group (ERG) notes that the focus of this single technology appraisal (STA) is on patients with non-valvular atrial fibrillation (NVAF) who represent a subgroup of patients with AF. The ERG also notes that most of the published background data available and reported by the manufacturer are for the whole AF population rather than the NVAF subgroup. The ERG considers this to be acceptable as NVAF represents the majority of cases of AF.[(1)]

Boxes 1 to 3 present summaries of pathophysiology (Box 1), prevalence (Box 2) and stroke risk of AF (Box 3), and Box 4 discusses the impact of stroke in AF.

All information presented in boxes is taken directly from the MS unless otherwise stated and the references have been renumbered.

Box 1. Pathophysiology of AF

AF is the most common cardiac arrhythmia[(2)] and is characterised by an irregularly irregular heartbeat. AF leads to deterioration in the mechanical function of the atria preventing complete expulsion of blood from the heart. This lack of movement of blood can lead to the formation of a thrombus (blood clot), which can become mobile (embolus), potentially resulting in stroke or SE.
Abbreviations used in box: AF, atrial fibrillation; SE, systemic embolism.

Box 2. Prevalence of AF

The prevalence of AF is 1.4% in England[(3)] and 1.7% in Wales[(4)] according to data collected as part of the NHS QOF for 2009/2010. Prevalence of AF increases exponentially with age;[(5)] according to UK

epidemiological studies AF is uncommon in people aged under 50 years, it then increases to ~1% in individuals 55–64 years, and to 7–13% in individuals 85 years and above.[5-9] While AF is known to increase the risk of overall mortality by as much as 60% [SMR 1.6 (95% CI: 1.4-1.8) compared to the general population],[10] its most serious manifestation is through the increased risk of stroke.

Abbreviations used in box: AF, atrial fibrillation; NHS, National Health Service; QOF, Quality and Outcomes Framework; SMR, Standardised Mortality Ratio; 95% CI, 95% confidence interval.

Box 3. Stroke risk of AF

AF increases the risk of stroke by approximately 5-fold,[11] and more than 20% of all strokes are attributed to this arrhythmia.[12] Strokes can cause a wide spectrum of clinical sequelae ranging from asymptomatic, minor events to life-changing disabilities, or even death. AF is also associated with an increased risk of SE. Although rare, SE can be devastating, causing severe complications including ischaemic bowel, renal infarction and lower limb ischaemia, which itself may lead to amputation.[13]
Abbreviations used in box: AF, atrial fibrillation; SE, systemic embolism.

Box 4. Clinical and financial consequences of stroke in AF patients

Stroke is the most important consequence of AF, with the greatest impact on morbidity and mortality. The risk of stroke is dependent on a number of factors and ranges from an annual risk of 1% in patients aged over 65 with no risk factors, to over 12% per year in patients with multiple risk factors.[2] Such risk factors include age, hypertension, diabetes, heart failure, or history of prior stroke. Furthermore, the risk of recurrent stroke within 5 years of the first stroke is up to 43%.[2] Strokes associated with AF are generally more severe than strokes in patients who do not have AF.[14;15] The risk of symptomatic intracranial haemorrhage is significantly greater in stroke patients with chronic AF compared to stroke patients without AF (16% vs. 5%; OR 2.95; 95% CI: 1.12-9.30;[16]). Strokes caused by AF are often fatal,[14] with Marini et al (2005) showing the 1-year mortality rate of AF-related strokes to be approximately 50%.[17] Similarly, the Framingham study suggests a 30-day mortality rate of 30% with AF-associated stroke.[18] Those patients who survive suffer increased levels of disability and longer hospital stays compared with stroke patients without AF.[2;12] For example, AF increases the risk of death, disability and handicap by approximately 50% at 3 months, independently of any other risk factors.[14] Surviving stroke is associated with significant levels of psychological distress on the part of both patients and their caregivers.[19] The financial implications to the NHS of treating and managing stroke are substantial. Luengo-Fernandez et al (2006) showed that, in a study of 2004 patients with stroke, the cost to the UK economy was £8 billion (including healthcare productivity and informal care costs) of which £4.6 billion was incurred by the NHS.[20] It is also worth noting that the cost of acute stroke in patients with a history of AF is 66% higher than in patients with no history of AF.[20] Thus, managing AF-associated strokes is more costly than managing strokes in patients without AF, showing that reducing the incidence of strokes in patients with AF will have wide clinical, economic and societal implications.
Abbreviations used in box: AF, atrial fibrillation; SE, systemic embolism.

The ERG consider it important to highlight that the reference cited by the manufacturer in support of the statement "Prevalence of AF increases exponentially with age" (Box 2) does not mention an exponential relationship between the increase in AF prevalence with increasing age.[5] However, the ERG acknowledges that the reference[5] does support the statement that AF increases with increasing age, i.e. there is a positive correlation between age and AF.

Based on expert clinical advice, the ERG considers the manufacturer's overview of the underlying health problem to be accurate.

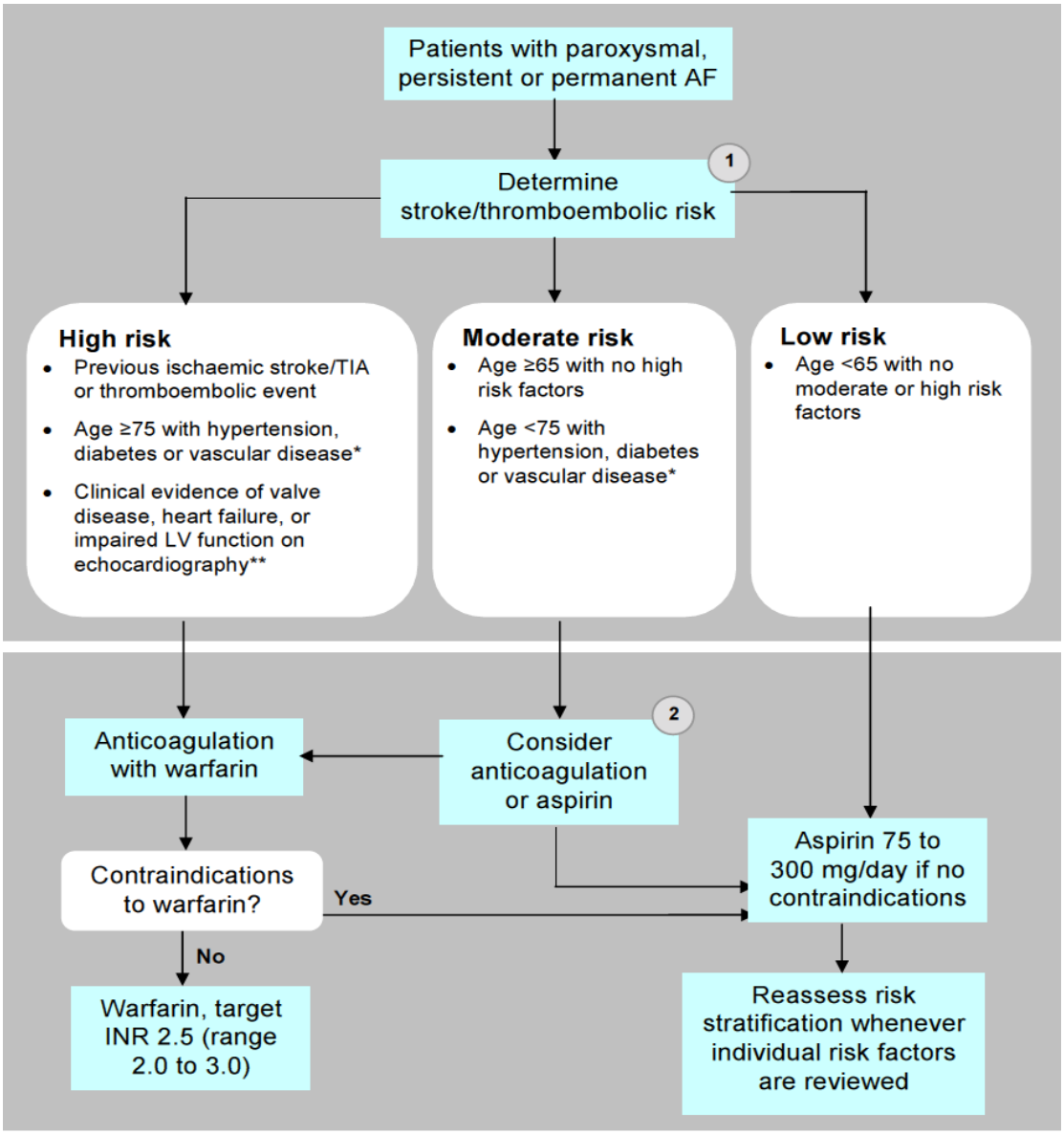
2.2 Critique of manufacturer's overview of current service provision

The manufacturer's overview of current service provision included an overview of the current National Institute for Health and Clinical Excellence (NICE) guideline for the management of AF (CG36)[2] (Box 5 and Figure 1), along with a summary of the relevant NICE technology appraisals[21;22] (Table 1). In addition, the manufacturer outlined the proposed position of apixaban in the current treatment pathway for NVAF (Box 6), and estimated the number of patients in England and Wales who would be eligible for treatment with apixaban (Table 2).

Box 5. Manufacturers overview of CG36[2]

In 2006 NICE published a clinical guideline on the diagnosis and management of AF.[2] Within the full guideline,[23] the anti-thrombotic therapy section included reviews of the evidence for warfarin and aspirin for stroke prevention in AF, which concluded that stroke risk in people with AF can be reduced with anti-thrombotic treatment. The guideline also reviewed the evidence for stroke risk, and Appendix B of the guideline provided a summary of the published stroke risk stratification algorithms. Based on the review of the stroke risk evidence, the NICE guideline adopted an algorithm based on a modified scheme specifically adapted for use in the UK. The stroke risk stratification algorithm presented in Figure 1 below is taken from NICE CG36 which currently recommends that people with AF at high risk of stroke should receive anticoagulation with warfarin.[2] In patients with AF at low risk of stroke – such as those under the age of 65 years with no risk factors – or in those patients who are unsuitable for warfarin therapy, treatment with aspirin is recommended.[2]
Abbreviations used in box: NICE, National Institute for Health and Clinical Excellence; AF, atrial fibrillation.

Figure 1. NICE CG36[2] stroke risk stratification algorithm



1. Note that risk factors are not mutually exclusive, and are additive to each other in producing a composite risk. Since the incidence of stroke and thromboembolic events in patients with thyrotoxicosis appears similar to that in patients with other aetiologies of AF, antithrombotic treatments should be chosen based on the presence of validated stroke risk factors.
2. Owing to lack of sufficient clear-cut evidence, treatment may be decided on an individual basis, and the physician must balance the risk and benefits of warfarin versus aspirin. As stroke risk factors are cumulative, warfarin may, for example, be used in the presence of two or more moderate stroke risk factors. Referral and echocardiography may help in cases of uncertainty.

*Coronary artery disease or peripheral artery disease.

** An echocardiogram is not needed for routine assessment, but refines clinical risk stratification in the case of moderate or severe LV dysfunction and valve disease.

The manufacturer highlighted that there is currently a lack of consensus in the United Kingdom (UK) AF community on the most appropriate stroke risk stratification scheme. In particular, the manufacturer commented on the recently-published European Society of Cardiology (ESC) guidelines for the management of AF[12] that recommend the use of the new CHA_2DS_2-VASc (Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled], Vascular Disease, Age 65–74,

and Sex category [female]) risk score. This risk score differs from the high, moderate and low risk classification system used in CG36.[2] The CHA_2DS_2-VASc score is instead based on a point system in which 2 points are assigned for a history of stroke or transient ischaemic attack (TIA), or age ≥75 years; and 1 point each is assigned for age 65–74 years, a history of hypertension, diabetes, recent cardiac failure, vascular disease (e.g. myocardial infarction or peripheral arterial disease), and female gender. The total score possible is 9 and the minimum, a score of 0. In addition, the ERG notes that the ESC guideline[12] along with older guidelines such as the National Health Service (NHS) Improvement Programme "GRASP-AF" tool[24] mention the older $CHADS_2$ [Cardiac failure, Hypertension, Age, Diabetes, Stroke (doubled)] risk classification system. The $CHADS_2$ score is also used to assess stroke risk in the key trials reported in the MS.[25-28] The $CHADS_2$ risk index is based on a 6-point system in which 2 points are assigned for a history of stroke or TIA and 1 point each is assigned for age >75 years, a history of hypertension, diabetes, or recent cardiac failure. The 2010 ESC guidelines[12] highlighted that a key difference between the $CHADS_2$ score and CHA_2DS_2-VASc is that the CHA_2DS_2-VASc system takes into account a larger array of the stroke risk factors that may influence the decision to anti-coagulate a patient. Therefore, the CHA_2DS_2-VASc provides a more comprehensive stroke risk assessment when compared with the $CHADS_2$ score.

The ERG and manufacturer both note that CG36[2] is currently in the process of being updated by NICE and could result in changes to the AF antithrombotic treatment pathway in the UK. In particular, the ERG consider that the presence of the new CHA_2DS_2-VASc stroke risk scoring system may influence the stroke risk stratification algorithm used in the updated version of CG36.[2]

Table 1: NICE technology appraisals evaluating treatments for NVAF

Technology Appraisal Number	Date issued	Title	NICE Recommendation
TA 249[21]	March 2012	Dabigatran etexilate for the prevention of stroke and SE in AF.	Dabigatran etexilate is recommended as an option for the prevention of stroke and SE within its licensed indication, that is, in people with NVAF with one or more of the following risk factors: - previous stroke, TIA or SE, - left ventricular ejection fraction below 40%, - symptomatic heart failure of NYHA class 2 or above, - age 75 years or older, - age 65 years or older with one of the following: diabetes mellitus, coronary artery disease or hypertension.
TA 256[22]	May 2012	Rivaroxaban for the prevention of stroke in AF.	Rivaroxaban is recommended as an option for the prevention of stroke and SE within its licensed indication, that is, in people with NVAF with one or more risk factors such as: - congestive heart failure, - hypertension, - age 75 years or older, - diabetes mellitus, - prior stroke or TIA.
Abbreviations used in table: TA, technology appraisal; SE, systemic embolism; AF, atrial fibrillation; NVAF, non-valvular atrial fibrillation; TIA, transient ischaemic attack; NYHA, New York Heart Association.			

As a consequence of the NICE technology appraisals (TAs) TA 249[21] and TA 256,[22] dabigatran etexilate and rivaroxaban are included as comparators within this technology appraisal. In addition, the ERG acknowledges the presence of the following NICE interventional guidance:

- Interventional Procedures Guidance No. 349, June 2010, "Percutaneous occlusion of the left atrial appendage in for the prevention of thromboembolism".
- Interventional Procedures Guidance No. 400, June 2011 "Thoracoscopic exclusion of the left atrial appendage (with or without surgical ablation) for non-valvular atrial fibrillation for the prevention of thromboembolism".

Box 6. Manufacturer's proposed position of apixaban in the NVAF treatment pathway

Apixaban is expected to be licensed for patients with non-valvular AF and one or more risk factors for stroke. Patients at moderate or high risk of stroke would be eligible for apixaban, representing an alternative option to warfarin, dabigatran and rivaroxaban. Apixaban is also an option for those patients at moderate risk of stroke who are unsuitable for warfarin.
Abbreviations used in box: AF, atrial fibrillation.

The ERG agrees with the manufacturer's description of the likely positioning of apixaban in the current treatment pathway of stroke prevention in UK NVAF patients (Box 6). In addition, the ERG agrees with the manufacturer that apixaban may provide an alternative treatment option for VKA unsuitable patients at moderate risk of stroke. However, based on expert clinical opinion, the ERG

notes that not all patients considered unsuitable for VKA therapy would be suitable for treatment with apixaban.

As part of the assessment of the impact of apixaban on current service provision the manufacturer described the NHS resource use likely to be associated with apixaban (MS; pg 28-29). The manufacturer anticipated that initially apixaban will be initiated in secondary care with follow-up in primary care and stated that apixaban will not require any additional infrastructure to be put in place. This is because, by contrast with warfarin therapy, which is associated with ongoing international normalised ratio (INR) monitoring costs, apixaban does not require INR monitoring. Therefore, apixaban will not require the NHS resource associated with warfarin INR monitoring and testing. However, the ERG considers it important to highlight that the use of apixaban instead of warfarin is unlikely to result in the redeployment of resources that are currently used to support warfarin monitoring. This is because warfarin is used for additional clinical indications (e.g. anticoagulation in heart valve replacement patients) to those for which apixaban is currently licensed or expected to be used in.

Table 2. Manufacturer's estimate of the number of patients in England and Wales who would potentially be eligible for treatment with apixaban (adapted from MS; Table 4; pg 20)

	Rate	2013	2014	2015	2016	2017
Total population of England and Wales (aged 18+)	-	44,694,105	45,049,027	45,405,281	45,738,826	46,054,429
AF prevalence	1.45%	-	-	-	-	-
AF mortality	2.7%	-	-	-	-	-
AF incidence	0.05%	-	-	-	-	-
Net AF patients		646,892	651,951	657,050	662,180	667,328
Patients with NVAF	80%	517,514	521,560	525,640	529,744	533,862
Patients with NVAF and $CHADS_2$ ≥1 (i.e. eligible for apixaban)	87.4%	452,462	456,000	459,567	463,155	466,756
Abbreviations used in table: AF, atrial fibrillation; NVAF, non-valvular atrial fibrillation; $CHADS_2$, cardiac failure, hypertension, age, diabetes, stroke (doubled).						

Regarding the manufacturer's estimate of the number of patients in England and Wales who would potentially be eligible for treatment with apixaban, the ERG and its clinical advisors consider that the manufacturer's estimate of the rate of NVAF being 80% is likely to be an underestimate. In addition, the ERG notes that the reference the manufacturer cites to support their estimate of 80% of the UK population having NVAF is based on an international survey which reports that valvular AF was observed in 26.7% of patients.[29] Assuming that the remainder of the survey have what would be classified as NVAF, the ERG calculates that NVAF would represent 73.3% of the survey population and not 80%. However, based on clinical advice, the ERG considers that over 90% of the UK population with AF are likely to have NVAF. In addition, the ERG identified a 2012 conference

abstract that presents the results of a national audit of AF in GP practices in Scotland and reports that 93.3% of AF patients had NVAF.[1] Using the estimate of 93.3% for the number of patients with NVAF and applying the manufacturer's other assumptions, the ERG has calculated the number of patients potentially eligible for treatment with apixaban between 2013 and 2017. The ERG's estimates are presented in Table 3.

Table 3. ERG estimate of the number of patients in England and Wales who would potentially be eligible for treatment with apixaban

	Rate	2013	2014	2015	2016	2017
Net AF patients	–	646,892	651,951	657,050	662,180	667,328
Patients with NVAF	93.3%	603,550	608,270	613,028	617,814	622,617
Patients with NVAF and $CHADS_2$ ≥1 (i.e. eligible for apixaban)	87.4%	527,503	531,628	535,786	539,969	544,167
Abbreviations used in table: AF, atrial fibrillation; NVAF, non-valvular atrial fibrillation; $CHADS_2$, Cardiac failure, Hypertension, Age, Diabetes, Stroke (doubled).						

In addition, the ERG considers it important to highlight that not all patients with NVAF and $CHADS_2$ ≥1 would be eligible for treatment with apixaban as some will have contraindications to anticoagulation therapy with apixaban. The figures presented in Table 3 are thus likely to be over estimates, although the ERG notes that the number of patients who are likely to be contraindicated to apixaban is likely to be small.

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

The manufacturer provided a summary of the final decision problem issued by the National Institute for Health and Clinical Excellence[30] (NICE; MS, pg 32), together with the rationale for any deviation from the decision problem (Table 4).

Table 4. Summary of decision problem as outlined in the manufacturer's submission (reproduced from MS; Section 5; pg 32)

Key parameter	Final scope issued by NICE[30]	Decision problem addressed in the submission	Rationale if different from the scope
Population	Adults with NVAF who are at risk of stroke or systemic embolism	As per the final scope	
Intervention	Apixaban	As per the final scope	
Comparator(s)	Warfarin (in people for whom warfarin is suitable) Dabigatran etexilate Rivaroxaban	As per the final scope plus aspirin for people for whom warfarin is suitable	As outlined in Sections 2.3 and 2.5 above, aspirin is currently recommended for patients unsuitable for warfarin or those at low risk of strokes, and is also still widely used in clinical practice in England and Wales. Aspirin remains therefore, a relevant comparator in this submission.
Outcomes	Stroke non-CNS systemic embolism Myocardial infarction Mortality Transient ischaemic attacks Adverse effects of treatment including haemorrhage Health-related quality of life	As per the final scope with the exception of TIAs	TIAs were not recorded in the ARISTOTLE trial[28]
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per the final scope	

Subgroups to be considered	If evidence allows, consideration will be given to subgroups defined by: INR time in TTR on warfarin Patients with different level of stroke/ thromboembolic risks. Guidance will only be issued in accordance with the marketing authorisation.	As per final scope	
Special considerations, including issues related to equity or equality	None	As per final scope	
Abbreviations used in table: NICE, National Institute for Health and Clinical Excellence; NVAF, non-valvular atrial fibrillation; CNS, central nervous system; TIA, transient ischaemic attack; ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation; NHS, National Health Service; INR, international normalised ratio; TTR, time in therapeutic range.			

3.1 Population

The key trial presented by the manufacturer to address the question in the NICE scope[30] of apixaban versus warfarin, is the Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) trial.[28] The population of ARISTOTLE consisted of patients who were ≥18 years of age with AF and at least one additional risk factor for stroke. The additional risk factors for stroke were as follows:

- age ≥75 years;
- prior stroke, TIA or systemic embolism (SE);
- either symptomatic congestive heart failure within 3 months or left ventricular dysfunction with a left ventricular ejection fraction ≤40% by echocardiography, radionuclide study or contrast angiography;
- diabetes mellitus;
- heart failure (NYHA class 2 or higher);
- hypertension requiring pharmacological treatment.

Approximately 20% of the randomised population of ARISTOTLE had a history of prior stroke, TIA or SE, and over 85% had hypertension requiring treatment. With respect to the distribution of the trial population by $CHADS_2$ score: 34% had a $CHADS_2$ score ≤1; 36% had a $CHADS_2$ score = 2, and the remaining 30% of patients had a $CHADS_2$ score of ≥3.

The final scope issued by NICE for this STA[30] requested a population of "adults with non-valvular AF who are at risk of stroke or systemic embolism". The ERG notes that current NICE guidance for the management of AF (CG36) suggests that all patients with paroxysmal, permanent or persistent AF are at risk of stroke. CG36 recommends that patients at moderate or high risk of stroke (defined as age

≥65 years or those with hypertension, diabetes, vascular disease, previous stroke/TIA/thromboembolic event or clinical evidence of valve disease, heart failure or impaired LV function on echocardiography) are considered for anticoagulation treatment. Whereas patients at low risk of stroke (defined as those aged <65 years with no moderate or high stroke risk factors) are recommended for treatment with aspirin. The ERG considers that this population at low risk of stroke is equivalent to a patient population with a $CHADS_2$ score of 0; patients with a CHADS2 score of 0 would not be eligible for inclusion in ARISTOTLE. Furthermore, the ERG notes that the population of ARISTOTLE are ≥18 years with AF and at moderate to high risk of stroke. The manufacturer's anticipated licence for the use of apixaban in AF is 'in adult patients with NVAF with one or more risk factors'. Therefore, the ERG considers that the population of ARISTOTLE accurately reflects the manufacturer's anticipated licence for the use of apixaban in AF.

In addition to the population stated in the NICE final scope[30] (adults with non-valvular AF who are at risk of stroke or systemic embolism), the manufacturer has included data for a further patient population based on a trial named AVERROES (Apixaban Versus Acetylsalicylic Acid [ASA] to Prevent Stroke in Atrial Fibrillation Patients Who Have Failed or Are Unsuitable for Vitamin K Antagonist Treatment). The AVERROES trial consisted of a patient population with AF and at least one additional risk factor for stroke who had also failed vitamin K antagonist (VKA) therapy, or had been deemed as unsuitable for VKA therapy (i.e. warfarin unsuitable). The ERG notes that this population was not included in the final scope issued by NICE. In addition, the ERG considers it important to highlight that the comparator in AVERROES (aspirin) was also not included in the NICE final scope for this STA. However, this will be discussed in more detail in Section 3.3. Furthermore, the ERG notes that approximately 40% of the randomised population of AVERROES is considered unsuitable for VKA therapy based on "patients' refusal to take VKA". The ERG does not consider that all patients refusing to take a VKA would be contraindicated to VKA therapy. In addition, the ERG acknowledges that approximately 50% of the patients in AVERROES had multiple reasons for VKA unsuitability. Therefore, the proportion of patients in AVERROES who are contraindicated to warfarin may be higher than 60%. Clinical advisors to the ERG commented that the reasons provided for patients' VKA unsuitability in AVERROES reflect the reasons seen in UK clinical practice. Clinical advisors also commented on the lack of consensus for a standard definition of "VKA unsuitable". The ERG is thus unable to comment further on the applicability of AVERROES to a UK VKA unsuitable population.

3.2 Intervention

The intervention that is the subject of this STA is apixaban; an oral anticoagulant that inhibits factor Xa in the clotting cascade, through a direct and highly selective mode of action. In addition, apixaban indirectly inhibits thrombin-induced platelet aggregation and thus prevents thrombus development.

The prevention of thrombus development in turn results in the prevention of emboli and thus apixaban reduces the risk of stroke and systemic embolism.

At present, apixaban does not have regulatory approval outside of the UK for use in AF; although, the manufacturer reports that a request for approval has been submitted to the FDA. In the MS, the manufacturer stated that UK marketing authorisation is not currently held for the use of apixaban for stroke prevention in AF; although, approval is expected to be granted in December 2012. The licensed indication anticipated for apixaban in the UK is: 'the prevention of stroke and systemic embolism (SE) in adult patients with NVAF with one or more risk factors'. However, the ERG notes that at the time of submission the 2.5 mg dose of apixaban was licensed for use in the UK, for the prevention of venous thromboembolic events (VTE) in adult patients who have undergone elective hip or knee replacement surgery. Furthermore, on the 20th September 2012 the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion of apixaban (Eliquis®);[31] recommending a variation to the terms of the marketing authorisation. The CHMP recommended a new indication for the existing 2.5 mg apixaban strength and recommended the use of the new 5 mg strength as follows:

"Prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation (NVAF), with one or more risk factors, such as prior stroke or transient ischaemic attack (TIA); age ≥75 years; hypertension; diabetes mellitus; symptomatic heart failure (NYHA Class ≥II)".[31]

The ERG notes that the indication anticipated by the manufacturer is inline with the indication recommended by the CHMP;[31] however, marketing authorisation by the European Commission for use in this indication is still awaited. The ERG notes that 5 mg (administered orally twice daily) is the recommended dose of apixaban for the prevention of stroke and SE in patients with NVAF. In addition, a dose reduction to 2.5 mg (administered orally twice daily) is recommended in patients with NVAF and at least two of the following:

- age ≥80years;
- body weight ≤60 kg;
- serum creatinine ≥1.5 mg/dl (133 mcmol/l).

To summarise, the ERG considers the intervention in the MS to be consistent with the anticipated licence and the NICE final scope for this STA.[30]

3.3 Comparators

The comparators listed in the final scope issued by NICE[30] for this STA were as follows:

- warfarin (in people for whom warfarin is suitable);
- dabigatran etexilate (hereafter referred to as dabigatran);
- rivaroxaban.

In the MS, the manufacturer presented data from the ARISTOTLE trial; a two-armed randomised controlled trial comparing apixaban with warfarin in people with AF. To be eligible for inclusion in ARISTOTLE, patients were required to be suitable for treatment with warfarin. Therefore, the ERG considers that ARISTOTLE addresses the comparison of apixaban with warfarin in people for whom warfarin is suitable.

However, there is an absence of head-to-head trials comparing apixaban with either dabigatran or rivaroxaban. Consequently, to assess the relative effect of apixaban versus dabigatran and versus rivaroxaban, the manufacturer carried out network meta-analyses (NMAs); a systematic review of the literature was used to identify trials suitable for inclusion in the manufacturer's NMAs. Based on the absence of head-to-head trial data, the ERG considers the manufacturer's decision to synthesise relative treatment effects using NMAs to be appropriate.

In addition to the above comparators listed in the final scope issued by NICE,[30] the manufacturer has included evidence for apixaban versus aspirin in people for whom warfarin is unsuitable. The justification by the manufacturer for the inclusion of this comparison within the submission is that "aspirin is currently recommended for patients unsuitable for warfarin or those at low risk of strokes, and is also still widely used in clinical practice in England and Wales". The key trial presented by the manufacturer within the MS to provide the evidence for this comparison is the AVERROES trial comparing apixaban with aspirin. In addition the ERG notes that the manufacturer presents data from an NMA (NMA 2) for the comparisons of apixaban versus rivaroxaban, and apixaban versus dabigatran in people for whom warfarin is unsuitable. The ERG considers this NMA to be potentially flawed as the population of the trials of rivaroxaban and dabigatran informing the network consist of patients suitable for treatment with warfarin. The ERG considers that NMA 2 breaks the fundamental assumption of transitivity; transitivity is vital for indirect comparisons.[32] The assumption of transitivity is that the "indirect comparison validly estimates the unobserved head-to-head comparison".[32] Transitivity cannot be tested statistically; instead its plausibility is evaluated conceptually and epidemiologically. The ARISTOTLE trial of apixaban versus warfarin is also included in NMA 2 and comprises patients suitable for treatment with warfarin. The NMA thus mainly consists of patients suitable for treatment with warfarin. The AVERROES trial is the only trial

in the NMA that includes patients unsuitable for warfarin and thus the ERG does not consider that the NMA addresses the question of apixaban versus dabigatran or apixaban versus rivaroxaban in patients unsuitable for warfarin.

In summary, the ERG considers that the comparators specified in the final scope issued by NICE[30] have been addressed within the MS in patients for whom warfarin is suitable. In addition the manufacturer has submitted evidence for apixaban in people for whom warfarin is unsuitable, although this was not a requirement of the final scope issued by NICE.[30] The ERG consider the network meta-analysis submitted by the manufacturer, for apixaban in people for whom warfarin is unsuitable, not to address the question of apixaban versus dabigatran or apixaban versus rivaroxaban in patients unsuitable for warfarin.

3.4 Outcomes

The outcomes listed in the final scope issued by NICE for this STA[30] were as follows:

- stroke;
- non-CNS systemic embolism;
- myocardial infarction;
- mortality;
- transient ischaemic attacks (TIAs);
- adverse effects of treatment including haemorrhage;
- health-related quality of life (HRQoL).

However, as no data on TIAs were collected in ARISTOTLE or AVERROES, no data on TIAs were presented in the MS. In addition, the ERG notes that treatment-specific health-related quality of life (HRQoL) data were not collected in either ARISTOTLE or AVERROES. Consequently, HRQoL data presented within the MS were limited to generic AF HRQoL data identified from a systematic review of the literature presented in the cost-effectiveness section of the MS (MS Section 7.4) and in Section 5.3.10 of this report. Furthermore, the ERG also notes that there are no publically available HRQoL data for either dabigatran or rivaroxaban in patients with NVAF. Therefore, the ERG is unable to comment on the potential impact of treatment on HRQoL.

Within the MS, adverse effects data from ARISTOTLE and AVERROES were comprised mainly of data for bleeding outcomes; although, aggregate data were presented for total adverse events and serious adverse events. In addition, the adverse effects data presented for AVERROES included a breakdown of serious adverse events. Furthermore, the ERG notes that the adverse effects data presented from the manufacturer's network meta-analyses (NMAs) are limited to bleeding outcomes.

The pre-specified primary efficacy outcome of both the ARISTOTLE and AVERROES trials was the composite of stroke and SE. The ERG notes that the individual components of the composite outcome were presented within the MS and the primary publications for both trials.[25;28] However, the individual components of the composite end point do not appear to have been pre-specified secondary outcomes.

To summarise, given the available clinical data for apixaban, the ERG considers that the outcome data presented by the manufacturer are appropriate. However, the ERG notes that no clinical data for the outcome of TIA were presented within the MS.

3.5 Timeframe

In ARISTOTLE, the mean duration of exposure to double-blind study drug was approximately 1.8 years in each treatment group (apixaban and warfarin). In addition, the manufacturer stated that the mean duration of exposure was similar when treatment groups were compared based on prior warfarin/VKA status. However, no further details were provided within the MS.

Following a recommendation from the data and safety monitoring committee (DMC) the AVERROES trial (apixaban versus aspirin) was terminated early as a result of the superior efficacy of apixaban. Consequently, the mean duration of follow-up for patients included in the primary analyses of AVERROES was 1.1 years.

The ERG considers the duration of follow-up in both ARISTOTLE and AVERROES to be suitable for assessing the short-term safety and efficacy outcomes of treatment with apixaban.

3.6 Other relevant factors

The ERG notes that the final scope issued by NICE[30] specified that evidence permitting, consideration should be given to the following subgroups:

- INR time in therapeutic range (TTR) on warfarin;
- patients with different levels of stroke/ thromboembolic risks.

The ERG notes that the manufacturer presents subgroup data from ARISTOTLE for INR TTR based on centre TTR for the primary efficacy and safety outcomes only. In addition, the manufacturer presents subgroup data from both ARISTOTLE and AVERROES for the primary efficacy and safety outcomes for patients with different levels of stroke/thromboembolic risks based on baseline $CHADS_2$ scores. The results of these data will be discussed in further detail in Section 4.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods used by the manufacturer to systematically review clinical effectiveness evidence

The manufacturer conducted two systematic reviews to identify published reports of trials relevant to the decision problem that is the focus of this STA. Respectively, the aim of the two systematic reviews (SRs) was as follows:

- to identify RCT evidence on the efficacy and safety of apixaban and relevant comparators for stroke prevention in patients with AF at moderate to high risk for stroke (hereafter referred to as RCT evidence SR);
- to identify non-RCT evidence on the efficacy and safety of apixaban for stroke prevention in patients with AF at moderate to high risk for stroke (hereafter referred to as non-RCT evidence SR).

4.1.1 Description and critique of manufacturer's search strategy

For each SR, the manufacturer carried out electronic database searches, accompanied by further searches that included hand searching of selected conference proceedings. A summary of the sources searched for each review is presented in Table 5.

Table 5. Sources searched for each systematic review carried out by the manufacturer.

Source type	RCT evidence SR		non-RCT evidence SR	
	Sources searched	**Dates searched**	**Sources searched**	**Dates searched**
Electronic databases:	Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)	1948 to 20th April 2011	Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)	1948 to 13th December 2011
	EMBASE (Ovid)	1980 to Week 15, 2011	EMBASE (Ovid)	1980 to Week 49, 2011
	The Cochrane Library	Inception to 20th April 2011	The Cochrane Library	1968 to 13th December 2011
	CINAHL	Inception to 5th May 2011		
	BIOSIS	1969 to Week 01, 2011		
Conference proceedings:	European Congress of Cardiology and meetings of the Joint Working Groups of the European Society of Cardiology (published in European Heart Journal)	2006–2010	ESC – European Congress of Cardiology	2009–2011, inclusive
	Scientific sessions of the American Heart Association (published in Circulation)	2006–2010	American Heart Association (AHA)	2009–2011, inclusive

	Annual meeting of the American College of Cardiology (published in The Journal of the American College of Cardiology)	2006–2010	American College of Cardiology (ACC)	2009–2011, inclusive
			International Society for Pharmacoeconomics and Outcomes Research (ISPOR)	2009–2011, inclusive
			International Health Economics Association (iHEA)	2009–2011, inclusive
			Heart Rhythm Society	2009–2011, inclusive
Other sources:	Clinicaltrials.gov	Not reported in MS	Relevant systematic reviews	Not applicable
	NCI clinical trial database	Not reported in MS		
	ISRCTN Register	Not reported in MS		
	UKCCR Register of Cancer Trials	Not reported in MS		
	EORTC	Not reported in MS		
	UK Clinical Trials Gateway	Not reported in MS		
	MetaRegister (mRCT) of Controlled Trials	Not reported in MS		
	Reference lists of retrieved articles	Not applicable		

The ERG notes that for both SRs, searches of the electronic databases were updated on 28th February 2012, using a date restriction of 2011 to present.

Within the MS, the manufacturer provided details of the terms used to search each electronic database. The search strategies included terms for AF, relevant pharmacological interventions and study design. The ERG notes that search terms for the identification of both RCTs and SRs were included in the RCT evidence SR. In addition, the ERG notes that the RCT evidence SR included search terms for clopidogrel, vitamin K antagonists in addition to warfarin, as well as edoxaban and betrixaban (two new direct factor Xa inhibitors). The ERG notes that edoxaban and betrixaban do not currently have UK marketing authorisation in the indication that is the focus of this STA. Consequently, edoxaban and betrixaban were not included in the final scope for this STA issued by NICE.[30] However, the ERG notes that no studies that focused on the interventions of clopidogrel, edoxaban or betrixaban were included in the MS.

The ERG also notes that for both of the manufacturer's SRs, search terms for atrial flutter were included in the electronic database search strategies. The ERG considers this to be acceptable as there is known to be an association between atrial flutter[33] and AF; therefore, studies indexed as atrial flutter may also report on AF. Furthermore, only those studies reporting on AF were included at the study selection stage; the ERG considers this to be appropriate based on the final scope issued by NICE for this STA.[30]

Because of time constraints for the completion of this report, the ERG has been unable to fully validate the manufacturer's searches and confirm the results. However, the ERG considers that the manufacturer's searches were comprehensive and the search strategies used for each SR were appropriate. In addition, the ERG is not aware of any relevant studies that have been missed by the manufacturer's search.

4.1.2 Inclusion/exclusion criteria used in study selection

The manufacturer provided details of the inclusion and exclusion criteria applied to each SR (Tables 6 and 7). In addition, the manufacturer presented justifications for any deviations from the PICO (Population, Intervention, Comparators, Outcomes) specified by NICE in the final scope[30] (Tables 6 and 7).

Table 6. Eligibility criteria applied to the search results of the RCT evidence SR (reproduced from MS; Table 5; pg 36)

	Description	Justification
Inclusion criteria		
Population	Adults with NVAF who are at risk of stroke or systemic embolism	Consistent with final scope[30]
Interventions	VKA including adjusted-dose warfarin Aspirin [acetylsalicylic acid (ASA)] (in VKA unsuitable patients only) Rivaroxaban Dabigatran Apixaban	Consistent with final scope[30] *Although not in the final scope, aspirin is still widely used in clinical practice in England and Wales and therefore is a relevant comparator in this submission
Outcomes	Stroke Systemic embolism Myocardial infarction (fatal and non-fatal) Composite outcomes (e.g. all strokes, myocardial infarction or vascular death) Major/minor bleeding Intracranial bleeding Gastrointestinal bleeding Mortality Re-admission rates	Consistent with final scope[30] with the exception that studies were not filtered for TIA as this was not in the original draft scope, and health-related quality of life (which was captured in the economic systematic review)
Study design	Prospective randomised controlled trials	Non-RCT studies were identified through a separate search
Language restrictions	No restriction	
Exclusion criteria		
Population	Subjects <18 years of age, patients with valvular/rheumatic AF	
Interventions	Studies not investigating apixaban or relevant comparator	
Study design	Non-RCT	Non-RCT studies were identified through a separate search
Language restrictions	No restriction	
Abbreviations used in table: MS, manufacturer's submission; NVAF, non-valvular atrial fibrillation; RCT, randomised controlled trial; SR, systematic review; VKA, Vitamin K antagonist.		

Table 7: Eligibility criteria used in the search strategy for the non-RCT evidence SR (reproduced from MS; Table 6; pg 36)

	Description	Justification
Inclusion criteria		
Population	Adults with NVAF who are at risk of stroke or systemic embolism	Consistent with final scope[30]
Interventions	Apixaban No restriction on comparator	Consistent with final scope[30]
Outcomes	Stroke Systemic embolism Myocardial infarction (fatal and non-fatal) Composite outcomes (e.g. all strokes, myocardial infarction or vascular death) Major/minor bleeding Intracranial bleeding Gastrointestinal bleeding Mortality Re-admission rates	Consistent with final scope[30] with the exception that studies were not filtered for TIA as this was not in the original draft scope, and health-related quality of life (which was captured in the economic systematic review)
Study design	Non-RCTs including: Prospective cohorts Case-control/case-referent studies Retrospective cohorts Database studies Cross-sectional studies	RCTs were identified through a separate search
Exclusion criteria		
Population	Subjects <18 years of age, patients with acute AF	
Interventions	Studies not investigating apixaban	
Study design	RCTs	RCTs were identified through a separate search
Language restrictions	Non-English publications	
Abbreviations used in table: AF, atrial fibrillation; MS, manufacturer's submission; NVAF, non-valvular atrial fibrillation; RCT, randomised controlled trial; SR, systematic review;		

As discussed in Section 3.3, the ERG notes that the manufacturer has deviated from the final scope issued by NICE for this STA;[30] the manufacturer has included aspirin as an additional comparator in patients unsuitable for VKA therapy. In addition, studies for either SR were not filtered for the outcomes of TIA and HRQoL. The impact of the omission of these outcomes from the clinical effectiveness review is discussed in Section 3.4.

The ERG also notes that for the non-RCT evidence SR a language restriction was imposed. This limited the results of the review to studies published in English. The manufacturer reported that based on the language restriction, only one study was excluded from the non-RCT evidence SR (MS; Appendix 10.6.7, pg 209). However, details of the excluded study were not provided in the MS, therefore, the ERG is unable to comment on the potential impact of the excluded study.

4.1.3 Details of studies included in the review of clinical effectiveness

The manufacturer presented PRISMA diagrams in the appendices of the MS to depict the inclusion/exclusion of studies for each SR; the findings are summarised below and the PRISMA diagrams can be found in Appendix 9.1. The ERG notes that in line with the decision problem, studies that focused on interventions other than apixaban were excluded from the main clinical effectiveness review. However, studies that considered comparators of interest were included in the clinical effectiveness review intended to inform the manufacturer's network meta-analyses. The network meta-analyses are discussed further in Section 4.4.

To summarise, the manufacturer reported that two relevant RCTs were identified by the RCT evidence SR (Table 8), whereas no studies suitable for inclusion were identified by the non-RCT evidence SR.

Table 8: List of relevant RCTs (reproduced from MS; Table 7; pg 38)

Trial	Phase	Intervention	Comparator	Population	Primary study ref.
ARISTOTLE	III	Apixaban 5 mg BD (2.5 mg BD in selected patients)	Warfarin INR target range 2.0–3.0	Subjects with AF and at least one additional risk factor for stroke	Granger et al, 2011[28] CSR[34]
AVERROES	III	Apixaban 5 mg BD (2.5 mg BD in selected patients)	Aspirin 81–324 mg OD	Subjects with AF and at least one additional risk factor for stroke who have failed or are unsuitable for VKA therapy	Connolly et al, 2011[25] CSR[35]
Abbreviations used in table: AF, atrial fibrillation; BD, twice daily; INR, international normalised ratio, OD, once daily; VKA, vitamin K antagonist					

The ERG considers ARISTOTLE to be the key study addressing the decision problem specified in the final scope[30] for this STA. However, as discussed in Section 3, the ERG considers that based on the final scope issued by NICE,[30] AVERROES does not meet the inclusion criteria for this STA.

The ERG also notes that a phase II study of 222 Japanese patients with NVAF, randomised to treatment with apixaban or warfarin (ARISTOTLE-J) is mentioned in the MS (MS Section 1.5). However, the manufacturer stated that ARISTOTLE-J is not relevant to this STA. The manufacturer reported that the study was small (222 subjects), of short duration (12 weeks), and primarily a safety investigation. The ERG notes that in ARISTOTLE-J, the target INR ranges for warfarin treatment are 2.0–3.0 and 2.0–2.6 for people aged <70 years and ≥70 years, respectively. Furthermore, the ERG notes that the mean age of patients in the warfarin arm of ARISTOTLE-J was 71.7 years. Therefore, over 50% of the patients randomised to warfarin would have the lower target INR range of 2.0–2.6. Based on clinical advice, the ERG notes that the target INR range for patients in the UK with NVAF

is 2.0–3.0 regardless of patient age. Therefore, the ERG agrees with the manufacturer's decision to omit ARISTOTLE-J from the review of clinical effectiveness for this STA.

The ERG is not aware of any additional studies potentially relevant to this STA that have been omitted by the manufacturer.

4.1.4 Quality assessment

The manufacturer assessed the ARISTOTLE and AVERROES trials against criteria adapted from guidance for undertaking reviews in health care issued by the Centre for Reviews and Dissemination,[36] as provided in the NICE template for manufacturer/sponsor submission of evidence to the STA process.[37] The ERG independently validated ARISTOTLE and AVERROES, and agrees with the manufacturer's assessments (Appendix 9.2). The ERG considers both ARISTOTLE and AVERROES to be well-designed RCTs. The ERG acknowledges that there was an imbalance in treatment discontinuations between the treatment groups in each of the RCTs (fewer discontinuations with apixaban in both RCTs) and notes that the manufacturer has used an ITT analysis to report the results of each study. The ERG considers the manufacturer's approach to data analysis to be appropriate. Treatment discontinuations will be discussed in more detail in Sections 4.3.1 and 9.3.3.

4.2 Summary and critique of submitted clinical effectiveness evidence

The manufacturer presents data for two RCTs, ARISTOTLE and AVERROES, in the clinical effectiveness section of the MS. The ERG considers that only ARISTOTLE met the inclusion criteria for this STA based on the final scope issued by NICE;[30] only ARISTOTLE will be discussed in further detail below. Details and results of AVERROES are presented in Appendix 9.3.

4.2.1 Description of ARISTOTLE trial

ARISTOTLE was an international, multicentre, randomised double-blind phase III trial comparing the clinical efficacy and safety of apixaban with warfarin (vitamin K antagonist, VKA). The patient population of ARISTOTLE had AF and at least one additional risk factor for stroke. The primary objective of ARISTOTLE was to determine whether apixaban was non-inferior to warfarin (INR target range 2.0–3.0) for the combined end point of stroke and SE.

ARISTOTLE population

ARISTOTLE included 18,201 randomised patients across 39 countries, of which 41 sites were within the UK. Subjects were randomised 1:1 to apixaban or warfarin via an interactive voice response system (IVRS), with randomisation stratified by clinical site and prior warfarin status (naive and experienced).

The inclusion and exclusion criteria for ARISTOTLE are detailed in Table 9 and the baseline characteristics for the randomised population are presented in Table 10. The ERG agrees with the

manufacturer's assessment that the two treatment groups appear well balanced with respect to their baseline characteristics.

Table 9: Inclusion/exclusion criteria of ARISTOTLE (adapted from MS; Table 9; pg 40)

Trial	Inclusion criteria	Exclusion criteria
ARISTOTLE	Males or females ≥18 year of age, with AF or atrial flutter not due to a reversible cause documented by ECG at time of enrolment, or AF/flutter documented on 2 separate occasions ≥2 weeks apart in the 12 months prior to enrolment, and presenting with ≥1 additional risk factor for stroke. Risk factors for stroke: Age ≥75 years Prior stroke, TIA or SE Either symptomatic congestive heart failure within 3 months or left ventricular dysfunction with a left ventricular ejection fraction ≤40% by echocardiography, radionuclide study or contrast angiography Diabetes mellitus Heart failure (NYHA class 2 or higher at time of enrolment) Hypertension requiring pharmacological treatment	AF due to reversible causes Moderate or severe mitral stenosis Conditions other than AF that required anticoagulation Stroke within the previous 7 days A need for aspirin at a dose of >165 mg/day or for both aspirin and clopidogrel Severe renal insufficiency (serum creatinine level of >2.5 mg/dL or calculated creatinine clearance of <25 mL/min
Abbreviations used in table: AF, atrial fibrillation; ECG, electrocardiogram; MI, myocardial infarction; NYHA, New York Heart Association; SE, systemic embolism; TIA, transient ischaemic attack; ULN, upper limit of normal; VKA, vitamin K antagonist		

Table 10: Characteristics of participants in ARISTOTLE across randomised groups (reproduced from MS; Table 10; pg 42)

	Apixaban (N = 9120)	**Warfarin (N = 9081)**
Age (years)		
Mean±SD	69.1±9.61	69.0±9.74
Gender, n (%)		
Male	5886 (64.5)	5899 (65.0)
Region, n (%)		
North America	2249 (24.7)	2225 (24.5)
Latin America	1743 (19.1)	1725 (19.0)
Europe	3672 (40.3)	3671 (40.4)
Asian Pacific	1456 (16.0)	1460 (16.1)
Median systolic blood pressure (mm Hg)	130	130
Median weight (kg)	82	82
Prior myocardial infarction, n (%)	1319 (14.5)	1266 (13.9)
Prior clinically relevant or spontaneous bleeding, n (%)	1525 (16.7)	1515 (16.7)
Type of AF, n (%)		
Paroxysmal	1374 (15.1)	1412 (15.5)
Persistent/ permanent	7744 (84.9)	7668 (84.4)
Prior use of VKA for >30 consecutive days, n (%)	5208 (57.1)	5193 (57.2)
Qualifying risk factors, n (%)		
Age ≥75 years	2850 (31.2)	2828 (31.1)
Prior stroke, TIA, or systemic embolism	1748 (19.2)	1790 (19.7)
Heart failure or reduced LVEF	3235 (35.5)	3216 (35.4)
Diabetes	2284 (25.0)	2263 (24.9)
Hypertension requiring treatment	7962 (87.3)	7954 (87.6)
$CHADS_2$ score[†] at enrolment, n (%)		
≤1	3100 (34.0)	3083 (34.0)
2	3262 (35.8)	3254 (35.8)
≥3	2758 (30.2)	2744 (30.2)
Mean±SD	2.1±1.1	2.1±1.1
Medications at time of randomisation, n (%)		
ACE inhibitor/ARB	6464 (70.9)	6368 (70.1)
Amiodarone	1009 (11.1)	1042 (11.5)
Beta-blocker	5797(63.6)	5685 (62.6)
Aspirin	2859 (31.3)	2773 (30.5)
Clopidogrel	170 (1.9)	168 (1.9)
Digoxin	2916 (32.0)	2912 (32.1)
Calcium blocker	2744 (30.1)	2823 (31.1)
Statin	4104 (45.0)	4095 (45.1)
NSAID	752 (8.2)	768 (8.5)
Gastric antacid	1683 (18.5)	1667 (18.4)

Renal function, creatinine clearance, n (%)		
Normal (>80 mL/min)	3761 (41.2)	3757 (41.4)
Mild impairment (>50 to 80 mL/min)	3817 (41.9)	3770 (41.5)
Moderate impairment (>30 to 50 mL/min)	1365 (15.0)	1382 (15.2)
Severe impairment (≤30 mL/min)	137 (1.5)	133 (1.5)
Not reported	40 (0.4)	39 (0.4)
Study doses of 2.5 mg BD apixaban (or placebo)	428 (4.7)	403 (4.4)

Abbreviations used in table: ACE, angiotensin-converting enzyme; AF, atrial fibrillation; ARB, angiotensin-receptor blocker; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; TIA, transient ischaemic attack; VKA, vitamin K antagonist

[†]$CHADS_2$ score is an index of the risk of stroke in patients with atrial fibrillation. Scores range from 1 to 6, with higher scores indicating a greater risk of stroke. Congestive heart failure, hypertension, age ≥75 years, and diabetes are each assigned 1 point, and previous stroke or transient ischaemic attack is assigned 2 points. The score is calculated by summing all the points for a given patient

The ERG notes that of the ARISTOTLE trial population 65% were male and the mean age was 69 years. Around 4% of the study population in each treatment arm received the 2.5 mg dose of apixaban, suggesting that 4% of patients had ≥2 of the following criteria: aged 80 years or older; a body weight of ≤60 kg; or a serum creatinine level of ≥1.5 mg/dL. The mean $CHADS_2$ score at baseline was 2.1 in both treatment arms and 65% of patients had a $CHADS_2$ score ≥2. During the clarification stage, the ERG requested a more detailed breakdown of patients' baseline $CHADS_2$ distribution (Table 11). The ERG notes that the majority of patients in ARISTOTLE had a $CHADS_2$ score of [redacted]. Therefore, the ERG considers that the $CHADS_2$ score distribution of ARISTOTLE is comparable with the UK population for whom apixaban treatment would be expected to be considered.

Table 11: Baseline $CHADS_2$ scores for ARISTOTLE (adapted from manufacturer's response to clarification questions; Table 11; pg 13)

$CHADS_2$ score	**ARISTOTLE**	
	Apixaban (N=9,120)	**Warfarin (N=9,081)**
	N (%)	**N (%)**
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]

Abbreviations used in table: $CHADS_2$, Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score.

ARISTOTLE intervention and comparator

The intervention under investigation in ARISTOTLE was apixaban and the comparator warfarin. The study medications were administered in a double dummy design using placebo tablets to match the active treatments. The apixaban treatment group received 5 mg of apixaban twice daily and warfarin placebo tablets. Patients with ≥2 of the following criteria: aged 80 years or older; a body weight of ≤60 kg, or a serum creatinine level of ≥1.5 mg/dL received a lower 2.5 mg dose of apixaban twice daily. The warfarin group received warfarin 2 mg tablets in one daily dose of up to 6 mg and apixaban placebo tablets twice daily. The warfarin dose was adjusted to give an INR within the range of 2.0–3.0. Patients meeting the criteria for the 2.5 mg apixaban dose received a matching 2.5 mg apixaban placebo instead of the 5 mg apixaban placebo.

Patients who were receiving warfarin or another VKA prior to randomisation were required to discontinue this three days before randomisation. The randomised study drugs were commenced once patients' INR level was <2.0.

The study dosing for warfarin and warfarin-placebo tablets was based on INR monitoring using a blinded, encrypted, point-of-care INR device. A treatment algorithm was provided to the study personnel to guide the adjustment of the warfarin dose according to the patient's INR level. The ERG notes that the treatment algorithm was only a guide, and the final decision for INR dosing was up to the study investigator.

Treatment with certain drugs was prohibited while taking the study drug and so if treatment with these drugs was required during the study then the patient's study drug was stopped temporarily. The prohibited drugs were: potent inhibitors of CYP3A4, aspirin at a dose >165 mg/day, other antithrombotic agents and glycoprotein IIb/IIIa inhibitors.

ARISTOTLE outcomes

The primary efficacy end point in ARISTOTLE was the composite of time to first occurrence of confirmed stroke (ischaemic, haemorrhagic or of uncertain type) or SE during the treatment period. The secondary efficacy end points pre-specified in the protocol for ARISTOTLE[38] were time to first occurrence of confirmed:

- ischaemic or of unspecified type stroke;
- haemorrhagic stroke;
- SE;
- all-cause death;
- composite of stroke, SE, major bleeding;
- composite of stroke, SE, all-cause death;

- composite of stroke, SE, all-cause death, major bleeding;
- composite of stroke, SE, MI, all-cause death;
- composite of stroke, SE and major bleeding in warfarin-naive subjects.

The primary safety end point was time from first dose of study drug to first occurrence of confirmed International Society on Thrombosis and Haemostasis (ISTH) major bleeding. The secondary safety end point defined in the protocol was days from first dose of study drug to first occurrence of the composite of confirmed ISTH major bleeding and confirmed clinically relevant non-major (CRNM) bleeding. In addition, it was specified in the protocol that other safety measures such as minor bleeds, would also be assessed.

The definitions for the key outcomes in ARISTOTLE were as follows:

- Stroke: a focal neurologic deficit, from a non-traumatic cause, lasting at least 24 hours and was categorised as ischaemic (with or without haemorrhagic transformation), haemorrhagic, or of uncertain type (in patients who did not undergo brain imaging or in whom an autopsy was not performed);
- Systemic embolism: clinical history consistent with an acute loss of blood flow to a peripheral artery (or arteries), which was supported by evidence of embolism from surgical specimens, autopsy, angiography, vascular imaging, or other objective testing;
- Bleeding: defined according to ISTH guidelines as follows:
    - major bleeding:
        - clinically overt bleeding accompanied by a decrease in haemoglobin of ≥2 g/dL and/or transfusion of ≥2 units of packed red blood cells;
        - bleeding that occurred in a critical site;
        - bleeding that was fatal;
    - CRNM bleeding: clinically overt bleeding that did not satisfy the criteria for major bleeding and that led to either:
        - hospital admission;
        - physician-guided medical or surgical treatment;
        - a change in antithrombotic therapy;
    - minor bleeding: all acute clinically overt bleeding events not meeting criteria for major bleeding or CRNM bleeding.

Generally, the ERG considers the clinical outcome definitions to be consistent with those used in UK clinical practice. However, the ERG notes that haemorrhagic stroke may be classified as an efficacy outcome and an adverse event. Bleeding events are important adverse effects associated with anticoagulation treatment were thus captured in the safety outcomes of ARISTOTLE. Therefore,

given the nature of the event, haemorrhagic stroke may have been captured in both the efficacy and safety outcomes of ARISTOTLE. However, the ERG notes that all-cause stroke (including both ischaemic and haemorrhagic stroke) is a frequently reported efficacy outcome in clinical trials in the disease area of AF (e.g. RE-LY and ROCKET-AF). Therefore, the ERG considers the inclusion of haemorrhagic strokes in the efficacy outcomes of ARISTOTLE to be appropriate.

The ERG notes that the primary and secondary efficacy and safety outcomes in ARISTOTLE were adjudicated by a clinical events committee. The clinical events committee were blinded to patients' study-group assignments and classified outcomes on the basis of pre-specified criteria. The ERG considers that this has reduced the risk of investigator bias affecting the results in terms of outcome assessment.

The ERG acknowledges that data on the outcomes of TIA and HRQoL, specified in the final scope issued by NICE,[30] were not collected in ARISTOTLE and thus could not be presented in the MS. The ERG considers that all other outcomes specified in the final scope issued by NICE[30] were captured in ARISTOTLE and reported appropriately in the MS for this STA.

ARISTOTLE subgroup analyses

A large number of subgroup analyses were pre-specified in ARISTOTLE and in addition several post-hoc subgroup analyses were conducted (Table 12).

Table 12: Subgroup analyses conducted in ARISTOTLE (adapted from MS; Table 15; pg 49)

Characteristic	Subpopulations in ARISTOTLE
Prior warfarin/VKA status	Experienced; Naive
Apixaban dose	2.5 mg BD or matching placebo; 5 mg BD or matching placebo
Geographic region	North America; Latin America; Europe; Asia/Pacific; US†, Eastern EU†,‡, Western EU†,‡
Age	<65 years; ≥65 to <75 years; ≥75 years
Gender	Male; Female
Female age group	≤50 years; >50 years
Race	White; Black or African American; Asian; Other
Ethnicity	Hispanic/Latino; Not Hispanic/Latino
Weight	≤60 kg; >60 kg
Body mass index	≤28 kg/m^2; >28 to 33 kg/m^2; >33 kg/m^2
Level of renal impairment	Severe or moderate: ≤50 mL/min; Mild >50 to 80 mL/min; Normal >80 mL/min
Number of risk factors	≤1; ≥2
$CHADS_2$ score	≤1; 2; ≥3
Prior stroke or TIA	Yes; No
Age ≥75 years	Yes; No
Diabetes mellitus	Yes; No
Hypertension requiring pharmacological treatment	Yes; No
Heart failure	Yes; No
Aspirin at randomisation	Yes; No
Clopidogrel at randomisation†	Yes; No
Type of AF†	Permanent or persistent; paroxysmal
Abbreviations used in table: AF, atrial fibrillation; BD, twice daily; TIA, transient ischaemic attack; VKA, vitamin K antagonist †Post-hoc analysis; ‡Eastern Europe: Czech Republic, Hungary, Poland, Romania, Russia, Ukraine; Western EU: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, UK.	

In addition to the subgroup analyses described in Table 12, a subgroup analysis based on study centre INR control was conducted. TTR (time in therapeutic range) was used as a measure of INR control and a study centre's INR control was approximated using the median of the centre's individual patients' TTR (cTTR). Four ranges of cTTR were considered for analysis in accordance with the quartiles of cTTR observed in ARISTOTLE.

The ERG notes that the subgroup analyses were limited to the primary efficacy and safety outcomes. In addition, the ERG notes that patients were only stratified at randomisation by prior warfarin/VKA status and that ARISTOTLE was not statistically powered to draw conclusions for any of the subgroup analyses reported.

ARISTOTLE follow-up

The treatment period in ARISTOTLE lasted until the attainment of approximately 448 primary efficacy events and the median duration of follow-up was 1.8 years. The ERG notes that a total of 477 primary efficacy events occurred in ARISTOTLE but the ERG is unable to quantify how many of these occurred in the 30-day follow-up period following the study termination.

The ERG notes that in the protocol for ARISTOTLE the manufacturer stated that study visits would occur monthly for INR monitoring. The INR monitoring visits were intended to monitor the patients INR, assess for outcomes, assess for AEs, and assess study medication compliance. Furthermore, the protocol specified that there would be quarterly visits during the treatment period (Months 3, 6, 9, 15, 18, 21, 27, 30, 33, 39, 42, 45, 51, 54 and 57) for laboratory assessments and the assessment of changes in concomitant medications and vital signs. In addition, yearly visits during the treatment period (Months 12, 24, 36 and 48) were specified to take place for physical measurements, and 12 lead electrocardiograms (ECGs). All patients were to be followed up for the development of stroke (ischaemic, haemorrhagic, or of unspecified type), SE, myocardial infarction, death, bleeding, hospitalisation or treatment discontinuation until the end of the study.

The ERG notes that follow-up of subjects who prematurely discontinued on the study drug was attempted. Follow-up attempts were made quarterly by telephone and where possible a final follow-up visit was made in person. Final visits took place within approximately 30 days after the attainment of 448 primary efficacy events in the study. Patients who completed double-blind treatment were followed up for study outcomes by telephone approximately 30 days after the last dose of double-blind study drug.

The ERG notes that there was a treatment algorithm in place for the dosing of warfarin although the final decision was at the discretion of the investigator. The protocol stated that INR monitoring was to commence by the fourth day following initiation of warfarin or warfarin placebo. INR monitoring was to be performed twice a week for 2 weeks, then once a week for 2 weeks and monthly following the attainment of a stable INR. An investigator was able to increase the frequency of INR monitoring if this was considered to be clinically indicated. However, titration of warfarin or warfarin-placebo was based on central monitoring of INR measurements utilising encrypted point of care (POC) devices and centralised dosing recommendations. In the UK INR monitoring is carried out around every 4 weeks for patients with a stable INR. Therefore, the ERG considers the INR monitoring in ARISTOTLE to be consistent with that which would routinely occur in the UK.

4.2.2 Description and critique of statistical approaches used in ARISTOTLE

The primary objective of ARISTOTLE was to determine whether apixaban was non-inferior to warfarin (INR target range 2.0–3.0) for the combined end point of stroke and SE; in subjects with AF and at least one additional risk factor for stroke. The non-inferiority margin used was 1.38; the upper bound of the two-sided 95% CI for the HR <1.38. The ERG notes that a non-inferiority margin of 1.38 is conservative in relation to other novel oral anticoagulant (NOAC) RCTs (RE-LY & ROCKET-AF AF), where the non-inferiority margin was 1.46. However, the ERG notes that ARISTOTLE was appropriately powered to have at least 90% power to meet the pre-specified 1.38 definition of non-inferiority for the HR. The ERG thus considers the manufacturer's choice of non-inferiority margin to be appropriate in ARISTOTLE.

The ERG notes that it has been reported that the choice of non-inferiority margin and population used in the statistical analysis of non-inferiority trials can result in the introduction of bias in the results.[39] In addition, it has been reported previously that non-inferiority trials are often only considered positive if non-inferiority is demonstrated in both the ITT and per protocol populations.[40] The non-inferiority results for ARISTOTLE were not presented in the MS and it was not specified which population(s) non-inferiority was assessed in. As part of the clarification process the ERG requested that the manufacturer provide the results of ARISTOTLE for the primary efficacy outcome using the per protocol population. The results of this will be discussed in Section 4.3.1.

Once non-inferiority was confirmed a hierarchical testing strategy was followed to control the type 1 error in the study to ≤5%. The strategy specified sequential testing for superiority (i.e. moving to the next analysis if and only if superiority had been proven in the previous analysis) in the following order:

- the primary efficacy end point at the one-sided α=0.025;
- ISTH major bleeding at the one-sided α=0.025;
- all-cause death at the one-sided α=0.025.

The manufacturer reported that the "primary and key secondary analyses were performed with the use of the Cox proportional hazards model, with previous warfarin status and geographic region (North America, South America, Europe or Asian Pacific) used as strata in the model" (MS Section 6.3.5).

The ERG notes that the manufacturer conducted the primary and secondary efficacy analyses on all patients who underwent randomisation (ITT population). Therefore, all events from the time of randomisation until the pre-defined cut-off date for efficacy outcomes were included. However, the safety analyses of bleeding events were limited to patients who received at least one dose of a study

drug; all events from the time the first dose of a study drug was received until 2 days after the last dose was received were included.

In general, the ERG considers the manufacturer's approach to the statistical analysis of the data in ARISTOTLE to be appropriate.

4.2.3 Summary statement

For this STA, two RCTs (ARISTOTLE and AVERROES) were included in the clinical effectiveness section of the MS to provide clinical data on apixaban. Based on the final scope issued by NICE,[30] the ERG considers that of the two trials only ARISTOTLE met the inclusion criteria for this STA. This is because AVERROES included an additional comparator (aspirin) to those listed in the final scope.[30] The ERG consider that apixaban versus aspirin was not a comparison of interest specified in the NICE final scope.[30] However, based on clinical advice, the ERG acknowledges that aspirin is utilised in clinical practice for some patients in the UK. Full details of AVERROES are presented in Appendix 9.3. The ERG has critiqued and reported on ARISTOTLE within the main body of this report, in Sections 4.2.1 and 4.2.2.

The ERG considers that the inclusion and exclusion criteria for ARISTOTLE were acceptable to address the trial's objectives. In addition, the ERG notes that the baseline characteristics of the randomised populations of ARISTOTLE appeared to be well balanced between trial arms. The intervention was apixaban, which is the focus of this STA, and the comparator warfarin.

The ERG considers that the outcome data reported from ARISTOTLE appeared to be consistent with the data collected in the trial. However, the ERG notes that TIA and HRQoL data requested in the NICE final scope[30] were not collected in the trial.

In terms of follow-up and statistical data analysis, the ERG considers that the duration of follow-up in ARISTOTLE was acceptable for the outcomes assessed. The ERG also considers that the statistical analysis plan was suitable.

4.3 Summary of results of ARISTOTLE trial

4.3.1 ARISTOTLE treatment compliance and discontinuations

A total of 20,098 patients were enrolled in ARISTOTLE and 18,201 were randomised to treatment. The ITT population for the apixaban treatment group consisted of 9,120 patients and the warfarin group 9,081. The manufacturer reports in the MS that 9,088 patients were treated in the apixaban group and 9,052 in the warfarin group; i.e. they received at least one dose of study drug and thus were included in the safety analyses.

TTR is a measure of warfarin control and not necessarily the same as treatment compliance but nonetheless important in assessing warfarin treatment efficacy. The target INR for ARISTOTLE was an INR in the range of 2.0–3.0. INR values during the first 7 days following randomisation and during study-drug interruptions were excluded from the TTR calculations in ARISTOTLE. The mean TTR for patients in the warfarin arm of ARISTOTLE was 62.2% and the median TTR was 66.0%. Based on the advice of clinicians regarding the TTR expected in a UK patient population, the ERG considers that the mean TTR for ARISTOTLE was acceptable.

In response to clarification questions the manufacturer provided details on the numbers of study drug interruptions that occurred in ARISTOTLE. The manufacturer reported that study interruptions were counted only if they lasted 5 consecutive days or more. A total of [redacted] of subjects in the apixaban and [redacted] subjects in the warfarin arm met this definition of study drug interruption. [redacted]). The durations of study drug interruptions were not reported in the clinical study report (CSR). However, the ERG notes that in general [redacted] patients in the [redacted] group experienced study drug interruptions and therefore any bias introduced by treatment interruption was likely to favour [redacted].

Regarding discontinuations, significantly fewer (defined as p<0.5) patients permanently discontinued treatment in the apixaban group compared with the warfarin group (25.3% vs 27.5% respectively; p=0.001). The reasons for discontinuations, in order of frequency were as follows:

- subject request;
- adverse event;
- death;
- other reasons.

The loss to follow-up in ARISTOTLE was low with only 51 patients in the apixaban arm and 39 in the warfarin arm lost to follow-up. However, the total number of patients at the end of the study whose vital status was unknown was slightly higher (apixaban arm 180 patients, warfarin arm 200 patients) for various reasons including some patients withdrawing consent for study participation.

4.3.2 ARISTOTLE treatment effectiveness results

The primary objective of ARISTOTLE was to prove the non-inferiority of apixaban versus warfarin in the prevention of stroke and systemic embolism. Non-inferiority was proven based on the definition of the upper bound of the two-sided 95% CI for the HR <1.38 in the [redacted]ITT populations (HR 0.79; 95% CI 0.66–0.95). In addition, superiority was proven as the HR for the primary efficacy end point of stroke and systemic embolism was 0.79 and the upper bound of the 95% confidence interval was <1 (95%CI

0.66–0.95). This suggests that apixaban was associated with significantly fewer stroke and systemic emboli when compared with warfarin (p=0.01). In addition, apixaban resulted in fewer events for each of the individual stroke or SE outcomes when compared with warfarin (Table 13). However, it should be noted that the only statistically significant difference in events between apixaban and warfarin was for the individual outcome of haemorrhagic stroke (HR 0.51; 95% CI 0.35–0.75; p<0.001). For the outcomes of ischaemic stroke (or uncertain type) and SE there was no statistically significant difference in treatment effect between apixaban and warfarin (p=0.42 and p=0.70, respectively). The ERG notes that haemorrhagic stroke is a bleeding-related outcome and thus an adverse effect of treatment with anticoagulants such as apixaban and warfarin. The significant reduction in bleeding events such as haemorrhagic strokes with apixaban is thus beneficial and is discussed further in Section 4.3.3 (along with the other safety data from ARISTOTLE).

From a clinical perspective it is also important to note that the incidence of fatal or disabling stroke was significantly lower in the apixaban group compared with the warfarin group (HR 0.71; 95% CI 0.54–0.94).

Table 13: Summary of primary efficacy outcome – randomised subjects (reproduced from MS; Table 17; pg 54)

	Apixaban N=9120		**Warfarin N=9081**		**Hazard ratio (95% CI)**	**p value**
	Pts with event	**Event rate**	**Pts with event**	**Event rate**		
	no.	**%/yr**	**no.**	**%/yr**		
Primary outcome: stroke or SE	212	1.27	265	1.60	0.79 (0.66–0.95)	0.01
Stroke	199	1.19	250	1.51	0.79 (0.65–0.95)	0.01
Ischaemic or uncertain type	162	0.97	175	1.05	0.92 (0.74–1.13)	0.42
Haemorrhagic stroke	40	0.24	78	0.47	0.51 (0.35–0.75)	<0.001
SE	15	0.09	17	0.10	0.87 (0.44–1.75)	0.70
Abbreviations used in table: pts, patients; SE, systemic embolism; yr, year; CI, confidence interval						

The results of the secondary efficacy outcomes are presented in Table 14. Similar to the primary efficacy outcome, compared with warfarin, apixaban was associated with fewer secondary efficacy outcome events. Apixaban was associated with significantly fewer all-cause deaths than warfarin (HR 0.89; 95% CI 0.80–0.99; p=0.047). However, the differences in MI and pulmonary embolism (PE) or deep vein thrombosis (DVT) are not statistically significant (p=0.37 and p=0.63, respectively). The ERG notes that MI, PE, DVT, ischaemic stroke and SE are all thrombotic events and their incidence is aimed to be reduced by the use of anticoagulants such as apixaban and warfarin. Apixaban demonstrated a non-significant reduction in all of these events when compared with warfarin.

Table 14: Summary of secondary efficacy outcomes – randomised subjects (reproduced from MS; Table 18; pg 55)

	Apixaban N=9120		Warfarin N=9081		Hazard ratio (95% CI)	p value
	Pts with event	Event rate	Pts with event	Event rate		
	no.	%/yr	no.	%/yr		
Key secondary outcome						
Death from any cause	603	3.52	669	3.94	0.89 (0.80–0.998)	0.047
Other secondary outcomes						
Stroke, SE, or death from any cause	752	4.49	837	5.04	0.89 (0.81–0.98)	0.02
MI	90	0.53	102	0.61	0.88 (0.66–1.17)	0.37
Stroke, SE, MI, or death from any cause	810	4.85	906	5.49	0.88 (0.80–0.97)	0.01
PE or DVT	7	0.04	9	0.05	0.78 (0.29–2.10)	0.63
Abbreviations used in table: DVT, deep vein thrombosis; MI, myocardial infarction; PE, pulmonary embolism; pts, patients; SE, systemic embolism; yr, year; CI, confidence interval						

4.3.3 ARISTOTLE safety and adverse events

The adverse events and safety analyses for ARISTOTLE were reported using the treated population, which consisted of all subjects who received at least one dose of study medication. The ERG notes that this population is slightly smaller than the ITT population (18,201 patients in the ITT population vs 18,140 patients in the treated population). However, the ERG acknowledges that the safety analyses for other novel oral anticoagulant (NOAC) trials (e.g. ROCKET-AF) were also limited to subjects who received at least one dose of study medication.

The primary safety outcome of ARISTOTLE was ISTH major bleeding and apixaban was proved to be superior to warfarin in reducing these bleeding events (HR 0.69; 95% CI 0.60–0.80; $p<0.01$). The ERG notes that apixaban resulted in fewer bleeding events than warfarin for all of the ISTH major bleeding and major bleeding or CRNM bleeding events reported in the MS (Table 15). Furthermore, the difference in bleeding events was statistically significant ($p<0.05$) for all of the bleeding outcomes except gastrointestinal (GI) bleeding ($p=0.37$). However, apixaban was associated with a non-significant reduction in GI bleeding events when compared with warfarin (HR 0.89; 95% CI 0.70–1.15; $p=0.37$).

Table 15: Bleeding outcomes and net clinical outcomes – treated patients (reproduced from MS; Table 27; pg 82)

	Apixaban (N=9,088)		Warfarin (N=9,052)		Hazard ratio (95% CI)	p value
	Pts with event	Event rate	Pts with event	Event rate		
	no.	%/yr	no.	%/yr		
Primary safety outcome: ISTH major bleeding	327	2.13	462	3.09	0.69 (0.60–0.80)	<0.001
Intracranial	52	0.33	122	0.80	0.42 (0.30–0.58)	<0.001
Other location	275	1.79	340	2.27	0.79 (0.68–0.93)	0.004
Gastrointestinal	105	0.76	119	0.86	0.89 (0.70–1.15)	0.37
Major or CRNM bleeding	613	4.07	877	6.01	0.68 (0.61–0.75)	<0.001
GUSTO severe bleeding	80	0.52	172	1.13	0.46 (0.35–0.60)	<0.001
GUSTO moderate or severe bleeding	199	1.29	328	2.18	0.60 (0.50–0.71)	<0.001
TIMI major bleeding	148	0.96	256	1.69	0.57 (0.46–0.70)	<0.001
TIMI major or minor bleeding	239	1.55	370	2.46	0.63 (0.54–0.75)	<0.001
Any bleeding	2356	18.1	3060	25.8	0.71 (0.68–0.75)	<0.001
Net clinical outcomes						
Stroke, SE, or major bleeding	521	3.17	666	4.11	0.77 (0.69–0.86)	<0.001
Stroke, SE, major bleeding or death from any cause	1009	6.13	1168	7.20	0.85 (0.78–0.92)	<0.001

Abbreviations used in table: CI, confidence interval; CRNM, clinically relevant non-major; GUSTO, Global use of strategies to open occluded coronary arteries; HR, hazard ratio; ISTH, International society on thrombosis and haemostatsis; no., number; SE, systemic embolism; TIMI, Thrombolysis In Myocardial Infarction; yr, year

The non-bleeding adverse events reported in the MS were limited to liver function test (LFT) and aggregated data for total adverse events, serious adverse events, bleeding adverse events and discontinuations due to adverse events. The aggregated adverse event data demonstrated that apixaban was associated with fewer adverse events than warfarin (Table 16). The manufacturer reported that no serious adverse event (SAE) occurred in >5% of the patients in either group. In addition, it was reported in the MS that the SAEs that occurred in >1% of patients in either treatment group occurred with similar frequency in each group. Regarding the hepatic safety of apixaban, the ERG notes that the rates of abnormalities detected in LFT results were similar between apixaban and warfarin (Table 17). The ERG considers that the adverse event profile of apixaban was generally comparable with that of warfarin.

Table 16: Summary of adverse events (treated population) (reproduced from MS; Table 29; pg 85)

Adverse events **Number (%) subjects**	**Apixaban** **(N=9,088)**	**Warfarin** **(N=9,052)**
AE	7406 (81.5)	7521 (83.1)
SAE	3182 (35.0)	3302 (36.5)
Bleeding AE	2288 (25.2)	2961 (32.7)
Discontinuation due to AEs	688 (7.6)	758 (8.4)
Abbreviations used in table: AE, adverse event; SAE, serious adverse event. Notes: AE, includes all serious or non-serious adverse events with onset from first dose through 2 days (for non-serious AEs) or 30 days (for serious AEs) after the last dose of blinded study drug; SAE, includes all serious adverse events with onset from first dose through 30 days after the last dose of blinded study drug; Bleeding AE, includes all serious or non-serious bleeding-related adverse events with onset from first dose through 2 days after the last dose of blinded study drug; Discontinuations due to AE, includes all serious or non-serious adverse events with onset from first dose of blinded study drug and with action taken regarding study drug (drug discontinued)		

Table 17: Summary of hepatic safety (treated population) (reproduced from MS; Table 30; pg 86)

Adverse events **Number (%) subjects**	**Apixaban** **(N=9,088)**	**Warfarin** **(N=9,052)**
ALT or AST >3x ULN and total bilirubin >2x ULN	30/8788 (0.2)	31/8756 (0.4)
ALT or AST >3x ULN and total bilirubin >2x ULN and alkaline phosphatase <2x ULN	17/8786 (0.2)	19/8755 (0.2)
ALT elevation		
3x ULN	100/8790 (1.1)	89/8759 (1.0)
5x ULN	45/8790 (0.5)	47/8759 (0.5)
10x ULN	16/8790 (0.2)	20/8759 (0.2)
20x ULN	8/8790 (<0.1)	12/8759 (0.1)
Abbreviations used in table: ALT, alanine aminotransferase; AST, aspartate aminotransferase; ULN, upper limit of normal		

4.3.4 ARISTOTLE subgroup analyses

In the MS, the manufacturer reported that the reduction of stroke and SE with apixaban was consistent across the majority of the 21 pre-specified subgroups (Table 18). In addition it was reported that statistical tests for within subgroup interaction were not significant (p>0.10).

Table 18: ARISTOTLE subgroup analysis results for primary efficacy outcome (stroke or SE) (adapted from manufacturer's response to clarification questions; Table 25; pg 29)

Subgroup	Total no. of patients	Apixaban, n (N)	Warfarin, n (N)	HR (95% CI)
Prior use of warfarin or other VKA				
Yes	10401	102 (5208)	138 (5193)	0.73 (0.57, 0.95)
No	7800	110 (3912)	127 (3888)	0.86 (0.66, 1.11)
Age				
<65 yr	5471	51 (2731)	44 (2740)	1.16 (0.77, 1.73)
65 to <75 yr	7052	82 (3539)	112 (3513)	0.72 (0.54, 0.96)
≥75 yr	5678	79 (2850)	109 (2828)	0.71 (0.53, 0.95)
Sex				
Male	11785	132 (5886)	160 (5899)	0.82 (0.65, 1.04)
Female	6416	80 (3234)	105 (3182)	0.74 (0.56, 1.00)
Weight				
≤60 kg	1985	34 (1018)	52 (967)	0.63 (0.41, 0.97)
>60 kg	16154	177 (8070)	212 (8084)	0.83 (0.68, 1.01)
Type of AF				
Permanent or persistent	15412	191 (7744)	235 (7668)	0.80 (0.66, 0.97)
Paroxysmal	2786	21 (1374)	30 (1412)	0.72 (0.41, 1.25)
Prior stroke or TIA				
Yes	3436	73 (1694)	98 (1742)	0.76 (0.56, 1.03)
No	14765	139 (7426)	167 (7339)	0.82 (0.65, 1.03)
Diabetes mellitus				
Yes	4547	57 (2284)	75 (2263)	0.75 (0.53, 1.05)
No	13654	155 (6836)	190 (6818)	0.81 (0.65, 1.00)
Heart failure				
Yes	5541	70 (2784)	79 (2757)	0.86 (0.63, 1.19)
No	12660	142 (6336)	186 (6324)	0.76 (0.61, 0.94)
$CHADS_2$ score				
≤1	6183	44 (3100)	51 (3083)	0.85 (0.57, 1.27)
2	6516	74 (3262)	82 (3254)	0.90 (0.66, 1.23)
≥3	5502	94 (2758)	132 (2744)	0.70 (0.54, 0.91)
Level of renal impairment				
Severe or moderate	3017	54 (1502)	69 (1515)	0.79 (0.56, 1.13)
Mild	7587	87 (3817)	116 (3770)	0.74 (0.56, 0.97)
No impairment	7518	70 (3761)	79 (3757)	0.88 (0.64, 1.21)
Apixaban dose				
2.5 mg BD or placebo	831	12 (428)	22 (403)	0.50 (0.25, 1.02)
5 mg BD or placebo	17370	200 (8692)	243 (8678)	0.82 (0.68, 0.98)
Geographic region				
North America	4474	42 (2249)	56 (2225)	0.75 (0.51, 1.13)
Latin America	3468	43 (1743)	52 (1725)	0.81 (0.54, 1.21)
Europe	7343	75 (3672)	77 (3671)	0.96 (0.70, 1.32)

Subgroup	Total no. of patients	Apixaban, n (N)	Warfarin, n (N)	HR (95% CI)
Asia pacific	2916	52 (1456)	80 (1460)	0.65 (0.46, 0.92)
Aspirin use at randomisation				
Yes	5632	70 (2859)	94 (2773)	0.72 (0.53, 0.98)
No	12569	142 (6261)	171 (6308)	0.83 (0.67, 1.04)
Abbreviations used in table: AF, Atrial fibrillation; BD, twice daily; TIA, transient ischaemic attack; HR, Hazard ratio; vitamin K antagonist				

The ERG notes that the subgroup for age <65 years demonstrated a non-significant benefit in favour of treatment with warfarin rather than with apixaban. The ERG also notes that the subgroups for geographical region suggest a trend towards patients in Europe experiencing less benefit with apixaban treatment compared with patients from North America, Latin America or Asia Pacific regions.

In addition, the ERG notes that for the primary safety end point of reduction in ISTH major bleeding events apixaban was associated with fewer bleeding events compared with warfarin in all of the subgroups reported in the MS (Table 19). However, the ERG notes that the statistical test for subgroup interaction was significant ($p<0.10$) for the subgroups of diabetes (yes or no) and level of renal impairment (none, mild and moderate or severe). Patients without diabetes and those with moderate or severe renal impairment received a greater reduction in major bleeding events compared with those with diabetes or with mild or no renal impairment. The ERG notes that the definition of statistical significance for subgroup interaction reported in the MS and primary publication for ARISTOTLE[28] was $p<0.10$. Based on this definition of statistical significance, the ERG notes that there is also a significant interaction for the outcome of major bleeding events in the gender subgroup ($p=0.08$). Female patients received a greater benefit with apixaban compared with male patients, although both subgroups were associated with statistically significant reductions in major bleeding events with apixaban when compared with warfarin. Other within subgroup trends suggesting less of a reduction in major bleeding events with apixaban included the following subgroups:

- patients <65 years (vs those >65 years);
- patients in the Europe or North America regions (vs those in Latin America or Asia Pacific).

The ERG notes that these subgroups were not stratified at randomisation and were not statistically powered and thus the drawing of conclusions from the subgroup results is inappropriate.

Table 19: ARISTOTLE subgroup analysis results for primary safety outcome (major bleeds) (adapted from manufacturer's response to clarification questions; Table 26; pg 31)

Subgroup	Total no. of patients	Apixaban, n (N)	Warfarin, n (N)	HR (95% CI)
Prior use of warfarin or other VKA				
Yes	10376	185 (5196)	274 (5180)	0.66 (0.55, 0.80)
No	7764	142 (3892)	188 (3872)	0.73 (0.59, 0.91)
Age				
<65 yr	5455	56 (2723)	72 (2732)	0.78 (0.55, 1.11)
65 to <75 yr	7030	120 (3529)	166 (3501)	0.71 (0.56, 0.89)
≥75 yr	5655	151 (2836)	224 (2819)	0.64 (0.52, 0.79)
Sex				
Male	11747	225 (5868)	294 (5879)	0.76 (0.64, 0.90)
Female	6393	102 (3220)	168 (3173)	0.58 (0.45, 0.74)
Weight				
≤60 kg	1978	36 (1013)	62 (965)	0.55 (0.36, 0.83)
>60 kg	16102	290 (8043)	398 (8059)	0.72 (0.62, 0.83)
Type of AF				
Permanent or persistent	15361	283 (7715)	402 (7646)	0.68 (0.59, 0.80)
Paroxysmal	2776	44 (1371)	60 (1405)	0.73 (0.50, 1.08)
Prior stroke or TIA				
Yes	3422	77 (1687)	106 (1735)	0.73 (0.54, 0.98)
No	14718	250 (7401)	356 (7317)	0.68 (0.58, 0.80)
Diabetes mellitus				
Yes	4526	112 (2276)	114 (2250)	0.96 (0.74, 1.25)
No	13614	215 (6812)	348 (6802)	0.60 (0.51, 0.71)
Heart failure				
Yes	5527	87 (2777)	137 (2750)	0.61 (0.47, 0.80)
No	12613	240 (6311)	325 (6302)	0.73 (0.61, 0.86)
$CHADS_2$ score				
≤1	6169	76 (3093)	126 (3076)	0.59 (0.44, 0.78)
2	6492	125 (3246)	163 (3246)	0.76 (0.60, 0.96)
≥3	5479	126 (2749)	173 (2730)	0.70 (0.56, 0.88)
Level of renal impairment				
Severe or moderate	3005	73 (1493)	142 (1512)	0.50 (0.38, 0.67)
Mild	7565	157 (3807)	199 (3758)	0.76 (0.62, 0.94)
No impairment	7496	96 (3750)	119 (3746)	0.79 (0.61, 1.04)
Apixaban dose				
2.5 mg BD or placebo	826	20 (424)	37 (402)	0.50 (0.29, 0.86)
5 mg BD or placebo	17314	307 (8664)	425 (8650)	0.71 (0.61, 0.82)
Geographic region				
North America	4463	106 (2244)	137 (2219)	0.77 (0.60, 1.00)
Latin America	3460	60 (1739)	94 (1721)	0.60 (0.44, 0.84)
Europe	7313	110 (3657)	135 (3656)	0.80 (0.62, 1.02)

Subgroup	Total no. of patients	Apixaban, n (N)	Warfarin, n (N)	HR (95% CI)
Asia pacific	2904	51 (1448)	96 (1456)	0.52 (0.37, 0.74)
Aspirin use at randomisation				
Yes	5608	129 (2846)	164 (2762)	0.75 (0.60, 0.95)
No	12532	198 (6242)	298 (6290)	0.66 (0.55, 0.79)
Abbreviations used in table: AF, Atrial fibrillation; BD, twice daily; TIA, transient ischaemic attack; HR, Hazard ratio; vitamin K antagonist				

During the clarification stage the ERG requested further data for some of the subgroups reported in the MS ($CHADS_2$ scores [≤1, =2, ≥3], Western Europe, 2.5 mg apixaban, 5 mg apixaban and age <65 years subgroups) as well as some additional subgroups (individual $CHADS_2$ scores, 2.5 mg apixaban and age >80 years subgroups). However, the manufacturer did not provide all of the data requested by the ERG. The manufacturer provided data for the subgroups reported in the MS ($CHADS_2$ scores, Western Europe, 2.5 mg apixaban, 5 mg apixaban and age <65 years subgroups). However, the outcome data were limited to stroke or SE, any bleeding and major bleeding for the $CHADS_2$, 2.5 mg apixaban, 5 mg apixaban and age <65 years subgroups. The Western Europe subgroup data were limited to only stroke or SE and major bleeding. The manufacturer reported that subgroup analyses were conducted only for these outcomes in the CSR and not for any other secondary outcomes. The manufacturer also presents an argument that post-hoc analyses on other secondary outcomes would be underpowered and run the risk of producing spurious results. The ERG considers that the additional outcome data requested would have been useful to investigate the consistency of the treatment effect of apixaban within subgroups but acknowledges that the subgroups would have been underpowered.

The ERG notes that for the subgroups by apixaban dose received (2.5 mg or 5 mg), the results for each dose [redacted] (Table 20).

Table 20. Results of ARISTOTLE by apixaban dose received compared with the whole trial population

Event	2.5 mg apixaban	5 mg apixaban	Whole trial population
	HR (95% CI)	HR (95% CI)	HR (95% CI)
Stroke or systemic embolism	[redacted]	[redacted]	0.79 (0.66, 0.95)
Any bleeding	[redacted]	[redacted]	0.71 (0.68, 0.75)
Major bleeding	[redacted]	[redacted]	0.69 (0.60, 0.80)
Abbreviations used in table: CI, confidence interval; HR, hazard ratio			

During the clarification question stage, the ERG requested subgroup data for the primary efficacy and safety outcomes of ARISTOTLE based on individual $CHADS_2$ score (i.e. 0, 1, 2, 3, 4, 5 and 6). However, the manufacturer only provided data for the $CHADS_2$ =2 subgroup. Data for the remaining $CHADS_2$ scores were provided only at an aggregate level within the following subgroups: $CHADS_2$ ≤1, and $CHADS_2$ ≥3 (Table 21). [redacted]. However, the lack of detailed individual $CHADS_2$ score data, particularly for the higher $CHADS_2$ scores (i.e. 3, 4, 5 and 6) limits the ability of the ERG to comment on any potential variation in apixaban treatment effect for these subgroups.

Table 21. Results of ARISTOTLE by $CHADS_2$ score compared with the whole trial population.

Event	≤1	2	≥3	Whole trial population
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Stroke or systemic embolism	[redacted]	[redacted]	[redacted]	0.79 (0.66, 0.95)
Any bleeding	[redacted]	[redacted]	[redacted]	0.71 (0.68, 0.75)
Major bleeding	[redacted]	[redacted]	[redacted]	0.69 (0.60, 0.80)
Abbreviations used in table: CI, confidence interval; HR, hazard ratio				

The results of ARISTOTLE for the post-hoc subgroup analysis of Western Europe demonstrated [redacted] (Table 22). The study was not powered to detect differences in efficacy and safety in the different geographical subgroups. However, the ERG notes that for the Western Europe subgroup, [redacted]

Table 22. Results of ARISTOTLE for Western Europe subgroup compared with the whole trial population.

Event	Western Europe	Whole trial population
	HR (95% CI)	HR (95% CI)
Stroke or systemic embolism	[redacted]	0.79 (0.66, 0.95)
Major bleeding	[redacted]	0.69 (0.60, 0.80)
Abbreviations used in table: CI, confidence interval; HR, hazard ratio		

During the clarification stage, the ERG requested details on the %TTR for each of the regional subgroups of ARISTOTLE. This was intended to enable an assessment by the ERG of any potential relationship between TTR and the efficacy of apixaban. However, the manufacturer provided these data at country rather than regional level; the manufacturer stated that patient level data were not available within the CSR. The ERG notes that based on the country level data the range of %TTR in the warfarin group for each region were as follows:

[redacted] Based on the country level %TTR results, the ERG considers that a potential explanation [redacted] However, without weighted region level %TTR data, the ERG is unable to comment further on any potential relationship between TTR and region subgroup results.

The results for the centre TTR subgroup analyses reported in ARISTOTLE suggested that apixaban was associated with fewer events than warfarin regardless of INR control, although this difference was not statistically significant (Table 23). The ERG notes that in the warfarin group there is a reduction in events as the %TTR improves. This trend is in keeping with what would be expected in clinical practice; i.e. with better warfarin control (or %TTR) it would be expected that there would be fewer strokes and SEs. The ERG also notes that a similar trend is seen in the apixaban group and the reason for this is unclear. Apixaban efficacy is not known to be related to INR control. The patients in the apixaban group and from the centres with the worst TTR (<58.0%) received less benefit from apixaban when compared with the patients on apixaban from the centres with the best TTR (>72.2%). The results of the centre TTR subgroup analysis thus suggest that there may be factors other than INR control that are impacting on the efficacy of apixaban and warfarin. However, it is unclear what these factors would be.

Table 23: Relative risks of stroke or SE with apixaban compared with warfarin, according to cTTR subgroup (reproduced from MS; Table 19; pg 58)

	Apixaban		Warfarin		
cTTR (%)	**Events**	**Rate/100 person years**	**Events**	**Rate/100 person years**	**HR (95% CI)**
<58.0	70	1.75	88	2.28	0.77 (0.56–1.06)
58.0–65.7	54	1.30	68	1.61	0.80 (0.56–1.15)
65.7–72.2	51	1.21	65	1.55	0.79 (0.54–1.13)
>72.2	36	0.83	44	1.02	0.81 (0.52–1.26)
Abbreviations used in table: CI, confidence interval; cTTR, centre time in therapeutic range; HR, hazard ratio					

The results of the primary safety analysis (major bleeding) for the centre TTR subgroups show that there are fewer major bleeding events with apixaban compared with warfarin regardless of centre TTR (Table 24). The ERG notes that this difference in major bleeding events between apixaban and warfarin is statistically significant in three out of the four centre TTR quartiles. However, the ERG notes that the study was not powered for the subgroup analyses and thus the absence of a significant difference in the remaining centre TTR quartile subgroup is not necessarily an indicator of no difference in treatment effect.

Table 24: Relative risks of major bleeding with apixaban compared with warfarin, according to cTTR subgroup (reproduced from MS; Table 28; pg 25)

cTTR (%)	**Apixaban**		**Warfarin**		
	Events	**Rate/100 person years**	**Events**	**Rate/100 person years**	**HR (95% CI)**
<58.0	64	1.75	115	3.34	0.53 (0.39–0.72)
58.0–65.7	61	1.60	102	2.68	0.60 (0.43–0.82)
65.7–72.2	103	2.68	109	2.89	0.93 (0.71–1.21)
>72.2	98	2.49	136	3.46	0.72 (0.55–0.93)
Abbreviations used in table: CI, confidence interval; cTTR, centre time in therapeutic range; HR, hazard ratio					

4.3.5 Summary of ARISTOTLE results

- Apixaban was demonstrated to be non-inferior to warfarin in the prevention of stroke and systemic embolism.
- For the primary efficacy end point of reduction in stroke and systemic embolism, apixaban demonstrated superiority compared with warfarin.
- Apixaban was associated with statistically fewer haemorrhagic strokes compared with warfarin and although there was no statistically significant difference in the individual outcomes of ischaemic stroke and systemic embolism.
- For the primary safety end point of reduction in ISTH major bleeding, apixaban demonstrated superiority compared with warfarin.
- The overall adverse event and safety profile of apixaban was comparable or better when compared with warfarin for the outcomes reported in the MS.
- Subgroup analyses suggest that European patients may have derived slightly less benefit with apixaban for both the efficacy and safety outcomes when compared with the whole trial population.
- The results of the subgroup analyses by centre TTR suggested that the safety and efficacy of apixaban compared with warfarin were independent of the level of warfarin control i.e. %TTR.

4.4 Description and critique of network meta-analysis

As a result of the absence of head-to-head trials, the manufacturer carried out two network meta-analyses (NMAs). The aim of these NMAs was to compare apixaban with dabigatran and rivaroxaban, as specified in the final scope issued by NICE.[30] NMA 1 consisted of patients suitable for treatment with warfarin and compared apixaban, warfarin, dabigatran and rivaroxaban. NMA 2 was reported in the MS to be in a population of patients unsuitable for VKA therapy and compared apixaban, dabigatran, rivaroxaban and aspirin. In the MS, the manufacturer stated that although aspirin was not included in the NICE final scope[30] it remained a relevant comparator in VKA unsuitable patients. The manufacturer's rationale for this was that aspirin is recommended in CG36 for patients at moderate to high risk of stroke or SE who are unsuitable for VKA therapy. In addition, the manufacturer stated that aspirin is widely used in clinical practice in England and Wales.

The ERG considers that the trials included in NMA 2 may be unsuitable for the intended analysis as a result of differences in the patient populations. The ERG notes that NMA 2 aimed to estimate the relative efficacy of apixaban, aspirin, dabigatran and rivaroxaban in a population of patients unsuitable for VKA therapy. However, the ERG notes that the majority (three out of four) of the trials included patients who were suitable for treatment with warfarin (ARISTOTLE, RE-LY and ROCKET-AF). In fact, AVERROES was the only trial included in NMA 2 that consisted of patients deemed unsuitable for treatment with warfarin. The ERG acknowledges that, at present, there are no published trials assessing the efficacy of dabigatran or rivaroxaban in patients unsuitable for VKA therapy. However, the ERG also notes that manufacturer does not offer any rationale to suggest that the efficacy and safety of rivaroxaban and dabigatran would be equivalent in people suitable and

unsuitable for VKA therapy. Therefore, the ERG considers it inappropriate to combine results from trials in VKA suitable patient populations with trial results for VKA intolerant patient populations as similar efficacy for NOACs has yet to be demonstrated. Thus the ERG does not consider NMA 2 to appropriately address the efficacy of dabigatran or rivaroxaban in patients unsuitable for VKA therapy. The focus of this report is on NMA 1. However, the network diagram and results of NMA 2 are reported in Appendix 9.4.

4.4.1 Methods

The trials included in NMA 1 were identified in the RCT evidence SR described in Sections 4.1.1 and 4.1.2. A total of three RCTs were identified as suitable for inclusion in NMA 1; these are detailed in Section 4.4.3.

For each outcome of interest, event rates from each RCT were used in the NMA to calculate HRs. The event rate was defined as the total number of events across all patients, divided by the total patient-years exposed. The ERG notes that if the event rate was not reported in published sources the manufacturer estimated event rates from the number of patients experiencing an outcome. The ERG is unable to provide any comment on the likely impact of this calculation as the true event rates are unknown. It is thus impossible to establish whether the calculated event rate could be over or under estimating the true event rate. However, the manufacturer reported that this method of calculating event rates accurately predicted event rates in the studies where data were available for both event rates and the number of patients with each event.

Where possible, to enable a degree of consistency in the way the populations were analysed, ITT data were used from each trial in NMA 1. In the MS, the manufacturer reported that the ITT data for ROCKET-AF were identified from a slide set on the FDA website and the rivaroxaban SPC. However, the manufacturer highlighted that there was an absence of published ITT outcomes data for fatal stroke, disabling stroke and non-disabling stroke from ROCKET-AF. Therefore, data from the "on-treatment" population of ROCKET-AF were used instead. The "on-treatment" population was defined as patients who received at least one dose of a study drug and followed for events until 2 days after discontinuation of blinded study drug treatment. In addition, the ERG notes that the manufacturer conducted two sensitivity analyses based on alternative datasets. Sensitivity analysis 1 used updated ITT efficacy data for RE-LY. Sensitivity analysis 2 used "on-treatment" data from the primary publication of ROCKET-AF in place of ITT data used in the main analysis.

The manufacturer used a Bayesian Markov Chain Monte Carlo simulation in WinBUGS to conduct the NMA. In the MS, the manufacturer reported that both fixed and random effects models were fitted to the data; the model with the best fit was chosen for the reporting of the results. Model fit was determined using the deviance information criterion (DIC) and residual deviance for each outcome

assessed. The manufacturer reported that there was little difference in model fit between the fixed and random effects models and all outcomes were reported using a fixed effects model. The manufacturer's rationale for the use of a fixed rather than random effects model was built around the small number of studies in the network (three studies). The manufacturer considered that as a result of the small number of included studies, a random effects model would produce poor estimates of the variation in between-study treatment effects. The manufacturer also cites text from the Cochrane Systematic Review Handbook[41] that recommends that at least 10 studies are used in the calculations to investigate heterogeneity. The ERG notes that while a random effects model incorporates heterogeneity it does not investigate heterogeneity. Consequently, the ERG does not consider the number of studies in the network to be sufficient reason to choose a fixed effects model over a random effects model. Rather, the ERG considers that the best fitting model should be chosen. However, the ERG acknowledges that NMA 1 is a "star-shaped" network; i.e. it has a single "common comparator" that links the five treatments of interest together. The between study heterogeneity generated using the random effects model reflects the prior value inputted into the model as there are insufficient trial data to further inform this estimate. The ERG thus considers the manufacturer's use of a fixed effects model to be a reasonable choice given the limited data set.

The ERG notes that within the MS, the manufacturer did not report the DIC or residual deviance values for either the fixed or random effects models. However, upon request, the manufacturer supplied the residual deviance values for NMA 1 during the clarification stage. Although the ERG notes that values were not supplied for all of the outcomes reported in the base case. For the outcomes where the residual deviance was supplied, the ERG agrees with the manufacturer's assessment that both the fixed and random effects models fit the data well. However, it is unclear why the residual deviance for some of the outcomes was not provided by the manufacturer.

4.4.2 Outcomes reported in network meta-analysis

The following safety and efficacy outcomes were reported in NMA 1:

- stroke + SE;
- any stroke;
- SE;
- haemorrhagic stroke;
- ischaemic stroke;
- MI;
- all-cause mortality;
- fatal stroke;
- disabling stroke;

- non-disabling stroke;
- discontinuations;
- ICH;
- major bleeding;
- GI bleeding;
- other major bleed (calculated by subtracting ICH from total major bleeding events);
- CRNM bleeding;
- any bleeding.

The ERG considers that these outcomes were generally consistent with those reported for ARISTOTLE. Furthermore, with the exception of TIA and HRQoL, the outcomes considered met the requirements of the final scope issued by NICE;[30] the absence of data on apixaban for the outcomes of TIA and HRQoL is discussed in Section 3.4. The absence of data on apixaban for TIA and HRQoL prevented the comparison of apixaban for these outcomes with the other treatments specified in the final scope issued by NICE.[30] Therefore, the ERG considers that these outcomes are appropriately excluded from the MS.

The ERG notes that for some outcomes, the manufacturer made some assumptions to facilitate calculation of event-rate data. These outcomes and assumptions were as follows:

- any bleed – event rates for ROCKET-AF AF were calculated by adding major/CRNM bleeding to minimal bleeding events;
- disabling stroke – for RE-LY and ARISTOTLE, fatal strokes were subtracted from the number of fatal or disabling strokes;
- non-disabling stroke – event rates in ARISTOTLE were calculated by subtracting disabling or fatal stroke events from total strokes;
- other major bleed – calculated by subtracting ICH from major bleed events.

The ERG notes that for some of these outcomes, the event calculation may have resulted in the double-counting of events. This is because a patient may have experienced more than one event; e.g. a minimal bleed and a major or CRNM bleed. The calculations could thus have resulted in:

- an over estimation of the number of 'any bleeds' in ROCKET-AF AF;
- underestimation of the number of disabling strokes in RE-LY or ARISTOTLE;
- underestimation of the number of non-disabling strokes in ARISTOTLE;
- underestimations in the numbers of 'other major bleeds'.

The ERG is unable to comment further on the likely impact of these potential biases on the results of NMA 1.

4.4.3 Studies included in the network

NMA 1 consisted of the following three RCTs (Figure 2):

- ARISTOTLE: apixaban 5 mg twice daily (BD) vs warfarin dosed to achieve a target INR 2.0–3.0;
- RE-LY: dabigatran 110 mg BD vs dabigatran 150 mg BD vs warfarin dosed to achieve a target INR 2.0–3.0;
- ROCKET-AF: rivaroxaban 20 mg once daily (OD) vs warfarin dosed to achieve a target INR 2.0–3.0.

The three trials were all conducted in people suitable for treatment with warfarin at the time of randomisation. Further background details on each of the trials are presented in Table 25. The trials were combined in an NMA to enable comparisons to be made between apixaban and dabigatran (110 mg and 150 mg doses), and apixaban and rivaroxaban and so address the comparisons requested in the final scope issued by NICE.[30]

Figure 2: Network diagram for warfarin-suitable population (NMA 1) (reproduced from MS; Figure 9; pg 68)

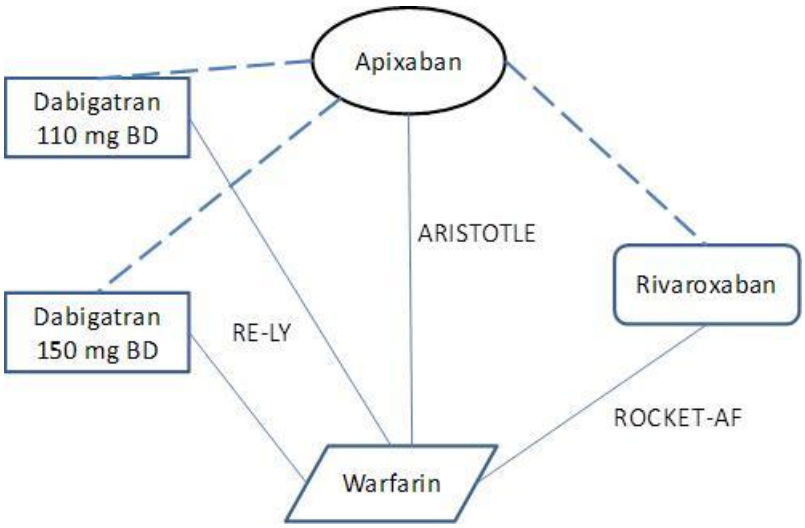


Table 25: Summary of the trials used to conduct NMA 1 (adapted from MS; Table 22; pg 66)

Trial name (primary ref)	Treatment	Dose	Trial design	Patient population	Mean age	% Male	Length follow-up	Mean % TTR	Efficacy and safety populations as reported in the primary publication
ARISTOTLE[28]	Apixaban	5 mg BD*	Randomised, double-blind, double-dummy	Subjects with AF and a $CHADS_2$ score ≥1	69.1	64.4	1.8 years	–	Efficacy: ITT – all randomised patients Safety: OT – all patients who received ≥1 dose of study drug
	Dose-adjusted warfarin	INR 2.0–3.0			64.5	65.0		62%	
RE-LY[26]‡	Dabigatran 110 mg	100 mg BD	Randomised, two doses of dabigatran administered in a blinded fashion, open-label use of warfarin	Subjects with AF and a $CHADS_2$ score ≥1	71.4	64.3	2 years	–	Efficacy: ITT Safety: ITT
	Dabigatran 150 mg	150 mg BD			71.5	63.2		–	
	Dose-adjusted warfarin	INR 2.0–3.0			71.6	63.3		64%	

Trial name (primary ref)	Treatment	Dose	Trial design	Patient population	Mean age	% Male	Length follow-up	Mean % TTR	Efficacy and safety populations as reported in the primary publication
ROCKET–AF[27]	Rivaroxaban	20 mg OD	Randomised, double-blind, double-dummy	Subjects with non-valvular AF and a $CHADS_2$ score ≥2	71.2	60.3	1.9 years	–	Efficacy: PP – all patients who received ≥1 dose of study drug, did not have a major protocol violation. OT population was used to test for superiority in the event non-inferiority was achieved on the PP population. ITT population analysed for the primary outcome only Safety: OT – all patients who received ≥1 dose of study drug regardless of adherence to protocol
	Dose-adjusted warfarin	INR 2.0–3.0			71.2	60.3		55%	
Abbreviations used in table: BD, twice daily; INR, international normalised ratio; OD, once daily; TTR, time in therapeutic range; ‡A later publication of the RE-LY trial was identified by the systematic review that reported additional primary efficacy and safety outcome events recorded during routine clinical site closure visits after the database was locked (Connolly et al. 2010).[42] Data from the 2010 publication were used in sensitivity analyses. *2.5 mg BD was used in small subpopulations.									

The ERG considers it important to highlight the following differences between the three trials included in NMA 1:

- ARISTOTLE and ROCKET-AF were double-blind, double dummy trials whereas treatment allocation to dabigatran or warfarin was not concealed in RE-LY;
- ROCKET-AF consisted of a higher stroke/SE risk population based on baseline $CHADS_2$ score compared with ARISTOTLE and RE-LY (mean baseline $CHADS_2$ 3.6, 2.1 and 2.1, respectively);
- Mean %TTR was lower in ROCKET-AF compared with ARISTOTLE and RE-LY (55%, 62% and 64%, respectively).

4.4.4 Heterogeneity

Within the MS, the manufacturer did not present any statistical assessments of heterogeneity for NMA 1. This is because, each pair-wise comparison was informed by single studies and therefore an assessment of heterogeneity within pair-wise comparisons could not be carried out. Similarly, as the network of studies was "star shaped" there was no opportunity to assess between study heterogeneity. The between study heterogeneity in the random effects model was informed by the prior value of tau included in the analysis. That is, there were insufficient data in the analysis to inform the between study heterogeneity. The ERG thus considers the manufacturer's use of a fixed effects model to be a reasonable choice given the limited data set.

In addition, the manufacturer commented on some potential sources of clinical heterogeneity within the network; these included the following observations:

- ROCKET-AF baseline population was at higher risk of stroke or SE ($CHADS_2$ higher than for RE-LY or ARISTOTOLE;
- Statistically significant difference in MI at baseline between treatment groups in ROCKET-AF;
- Open-label design of RE-LY vs double blind design of ARISTOTLE and ROCKET-AF;
- "On treatment" population of ROCKET-AF AF was used for some outcomes analysed in NMA 1 because of the absence of published ITT data.

The manufacturer stated that because of the presence of only one study for each comparator, no studies were excluded from any of the analyses. This is because exclusion of any single study would have resulted in the exclusion of one of the treatments from the analysis.

In addition, the manufacturer considered the exploration of the potential presence of heterogeneity within NMA 1 using covariate analysis to be inappropriate. However, subgroup analyses were conducted "to explore the consistency of treatment effects across stroke risk severity ($CHADS_2$) and centre-level TTR (cTTR) patient subgroups in accordance with the NICE scope". These subgroup analyses were only carried out for the outcomes of stroke or SE and major bleeds. This is because

there were insufficient data available for the other outcomes of interest reported in the base case NMA 1.

The ERG notes that for the $CHADS_2$ subgroup analyses no data were available from ROCKET-AF AF in the $CHADS_2 \leq 1$ subgroup because of the exclusion of such patients from ROCKET-AF AF. The ERG also notes that the subgroups of the cTTR analyses were defined differently in each of the included trials. Therefore, the ERG does not consider the subgroup analyses to be directly comparable (Table 26).

Table 26: Centre-TTR quartile subgroups across the three trials in NMA 1 (reproduced from MS; Table 26; pg 77)

cTTR quartile	ARISTOTLE	RE-LY	ROCKET
Lowest	<58.0%	<57.1%;	<50.6%;
2nd lowest	58.0–65.7%	57.1–65.5%	50.7–58.5%
2nd highest	65.7–72.2%	65.5–72.6%	58.6–65.7%
Highest	>72.2%	>72.6%	>65.7%
Abbreviations used in table: cTTR, centre-level time in therapeutic range.			

4.4.5 Results of the network meta-analysis

The results of NMA 1 are presented in Table 27 and the results of the sensitivity and subgroup analyses are presented in Tables 28, 29 and 30.

Table 27: NMA 1 (warfarin suitable population) base case analysis (HR<1 favours apixaban, HR>1 favours comparator) (reproduced from MS; Table 24; pg 74)

Outcome	**Hazard ratio [95% CrI]**			
	Apixaban vs dabigatran 150 mg	**Apixaban vs dabigatran 110 mg**	**Apixaban vs rivaroxaban**	**Apixaban vs warfarin**
Stroke + SE	[redacted]	[redacted]	[redacted]	[redacted]
Any stroke	[redacted]	[redacted]	[redacted]	[redacted]
SE	[redacted]	[redacted]	[redacted]	[redacted]
Haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]
MI	[redacted]	[redacted]	[redacted]	[redacted]
All-cause mortality	[redacted]	[redacted]	[redacted]	[redacted]
Fatal stroke	[redacted]	[redacted]	[redacted]	[redacted]
Disabling stroke	[redacted]	[redacted]	[redacted]	[redacted]
Non-disabling stroke	[redacted]	[redacted]	[redacted]	[redacted]
ICH	[redacted]	[redacted]	[redacted]	[redacted]
Major bleeding	[redacted]	[redacted]	[redacted]	[redacted]
GI bleeding	[redacted]	[redacted]	[redacted]	[redacted]
Other major bleed	[redacted]	[redacted]	[redacted]	[redacted]
CRNM bleeding	NR†	NR†	[redacted]	[redacted]
Any bleeding	[redacted]	[redacted]	[redacted]	[redacted]
Discontinuations	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations used in table: CrI, credibility interval; CRNM, clinically relevant non-major; GI, gastrointestinal; ICH, intracranial haemorrhage; MI, myocardial infarction; SE, systemic embolism.
Results shown in bold are significantly different; †Data for this outcome not reported for the RE-LY trial;

The base case results of NMA 1 suggested that apixaban was associated with a significantly lower incidence of MI compared with dabigatran 150 mg or 110 mg. [redacted].

For the bleeding safety outcomes, apixaban was associated with significantly fewer events for the following outcomes:

- ICH compared with rivaroxaban [redacted];
- major bleeding compared with rivaroxaban, with dabigatran 150 mg [redacted];
- GI bleeding compared with rivaroxaban and with dabigatran 150 mg;

- other major bleeding compared with rivaroxaban, with dabigatran 150 mg [redacted];
- CRNM bleeding compared with rivaroxaban [redacted];
- any bleeding compared with rivaroxaban, with dabigatran 150 mg, with dabigatran 110 mg [redacted].

In addition, apixaban was associated with significantly fewer discontinuations compared with dabigatran 150 mg, dabigatran 110 mg, rivaroxaban [redacted]

Table 28: NMA 1 sensitivity analysis 1 (RE-LY 2010 and ROCKET-AF ITT data) (reproduced from MS; Table 112; pg 249)

Outcome	Hazard ratio [95% CrI]			
	Apixaban vs dabi 150 mg	Apixaban vs dabi 110 mg	Apixaban vs rivaroxaban	Apixaban vs warfarin
Stroke + SE	[redacted]	[redacted]	[redacted]	[redacted]
Any stroke	[redacted]	[redacted]	[redacted]	[redacted]
SE	NR	NR	NR	NR
Haemorrhagic stroke	NR	NR	NR	NR
Ischaemic stroke	NR	NR	NR	NR
MI	[redacted]	[redacted]	[redacted]	[redacted]
All-cause mortality	NR	NR	NR	NR
Fatal stroke	NR	NR	NR	NR
Disabling stroke	NR	NR	NR	NR
Non-disabling stroke	NR	NR	NR	NR
ICH†	[redacted]	[redacted]	[redacted]	[redacted]
Major bleeding†	[redacted]	[redacted]	[redacted]	[redacted]
GI bleeding	[redacted]	[redacted]	[redacted]	[redacted]
Other major bleed	[redacted]	[redacted]	[redacted]	[redacted]
CRNM bleeding	NR	NR	NR	NR
Any bleeding	[redacted]	[redacted]	[redacted]	[redacted]
Discontinuations	NR	NR	NR	NR

Abbreviations used in table: CrI, credible interval; CRNM, clinically relevant non-major; dabi, dabigatran; GI, gastrointestinal; ICH, intracranial haemorrhage; ITT, intention-to-treat; MI, myocardial infarction; NR, not reported; SE, systemic embolism.
†RE-LY 2010 and ROCKET-AF OT data

Sensitivity analysis 1 (SA1) utilised post-hoc data from RE-LY published as an updated analysis in 2010.[(42)] The results are generally consistent with those of the base case NMA with the exception of the MI outcome. In the base case there is a statistically significant reduction in MI with apixaban compared with both dabigatran 150 mg and 110 mg doses. Whereas in SA1, the reduction in MI with apixaban for either dose of dabigatran is not statistically significant.

Table 29: NMA 1 sensitivity analysis 2 (RE-LY 2009 and ROCKET-AF "on-treatment" (OT) data) (reproduced from MS; Table 113; pg 250)

Outcome	Hazard ratio [95% CrI]			
	Apixaban vs dabi 150 mg	Apixaban vs dabi 110 mg	Apixaban vs rivaroxaban	Apixaban vs warfarin
Stroke + SE	[redacted]	[redacted]	[redacted]	[redacted]
Total stroke	[redacted]	[redacted]	[redacted]	[redacted]
SE	[redacted]	[redacted]	[redacted]	[redacted]
Haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]
MI	[redacted]	[redacted]	[redacted]	[redacted]
All-cause mortality	[redacted]	[redacted]	[redacted]	[redacted]
Fatal stroke	NR	NR	NR	NR
Disabling stroke	NR	NR	NR	NR
Non-disabling stroke	NR	NR	NR	NR
ICH[†]	NR	NR	NR	NR
Major bleeding	NR	NR	NR	NR
GI bleeding	NR	NR	NR	NR
CRNM bleeding	NR	NR	[redacted]	[redacted]
Any bleeding	NR	NR	NR	NR
Discontinuations	NR	NR	NR	NR

Abbreviations used in table: CrI, credible interval; CRNM, clinically relevant non-major; dabi, dabigatran; GI, gastrointestinal; ICH, intracranial haemorrhage; ITT, intention-to-treat; MI, myocardial infarction; NR, not reported; SE, systemic embolism.
[†]Data included in NMA 1 sensitivity analysis 1

- [redacted] Sensitivity analysis 2 (SA 2) utilised the safety on treatment dataset from ROCKET-AF instead of the ITT data, which are used in the base case NMA. [redacted]

Table 30: NMA 1 (warfarin suitable population) subgroup analyses based on $CHADS_2$ score and cTTR (reproduced from MS; Table 124; pg 262)

Subgroup	Hazard ratio [95% CrI]			
	Apixaban vs dabi 150 mg	**Apixaban vs dabi 110 mg**	**Apixaban vs rivaroxaban**	**Apixaban vs warfarin**
Stroke or SE				
$CHADS_2$ ≤1	[redacted]	[redacted]	–	[redacted]
$CHADS_2$ =2	[redacted]	[redacted]	[redacted]	[redacted]
$CHADS_2$ ≥3	[redacted]	[redacted]	[redacted]	[redacted]
TTR lowest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR 2nd lowest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR 2nd highest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR highest quartile	[redacted]	[redacted]	[redacted]	[redacted]
Major bleeding				
$CHADS_2$ ≤1	[redacted]	[redacted]	–	[redacted]
$CHADS_2$ =2	[redacted]	[redacted]	[redacted]	[redacted]
$CHADS_2$ ≥3	[redacted]	[redacted]	[redacted]	[redacted]
TTR lowest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR 2nd lowest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR 2nd highest quartile	[redacted]	[redacted]	[redacted]	[redacted]
TTR highest quartile	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations used in table: CrI, credible interval; SE, systemic embolism; TTR, time in therapeutic range				

As discussed in Section 4.4.4, the ERG notes that for the $CHADS_2$ subgroup analyses there were no data from ROCKET-AF AF in the $CHADS_2$ ≤1 subgroup (because of the exclusion of such patients from ROCKET-AF AF). In addition, the ERG would have been interested to see the analyses by $CHADS_2$ scores broken down by the individual $CHADS_2$ scores; i.e. 0, 1, 2, 3, 4, 5 and 6. Furthermore, the ERG is concerned that the number of patients within each $CHADS_2$ score category for each of the trials may be disproportionately different. Therefore, aggregation of multiple scores may provide misleading results.

[redacted]

[redacted]

The second subgroup analyses presented by the manufacturer for NMA 1 were based on cTTR. The ERG considers that the cTTR subgroups defined by the manufacturer are misleading. This is because the cTTR quartiles were different for each of the trials included in the analysis (see Section 4.4.4 for further details). In particular, the ERG notes that the %cTTR is substantially lower for each of the cTTR quartiles in ROCKET-AF AF compared with the %cTTR for each of the cTTR quartiles in RE-LY and ARISTOTLE. The ERG considers that the results of the manufacturer's aggregation of the data for the cTTR subgroup analysis should be interpreted with caution.

[redacted]

4.4.6 Apixaban direct pair-wise results compared with the network meta-analysis

The ERG notes that the direct pair-wise results for apixaban versus warfarin from ARISTOTLE are in keeping with the apixaban versus warfarin results generated by NMA 1 (Table 31). The manufacturer highlighted that there was a difference in all-cause mortality between ARISTOTLE and NMA 1 but attributed this to rounding differences between calculations. In ARISTOTLE, apixaban resulted in a statistically significant reduction in all-cause mortality compared with warfarin (HR 0.89; 95% CI 0.80–0.99; p=0.047). However in NMA 1, the result for this outcome was not statistically significant [redacted]. The ERG considers the manufacturer's rationale for this difference in results to be reasonable.

Table 31: Comparison of ARISTOTLE and NMA 1 results for apixaban versus (vs) warfarin.

Outcome	**Hazard ratio [95% CrI]**	
	NMA 1 (apixaban vs warfarin)	**ARISTOTLE (apixaban vs warfarin)**
Stroke + SE	[redacted]	0.79 (0.66–0.95)
Any stroke	[redacted]	0.79 (0.65–0.95)
SE	[redacted]	0.87 (0.44–1.75)
Haemorrhagic stroke	[redacted]	0.51 (0.35–0.75)
Ischaemic stroke	[redacted]	0.92 (0.74–1.13)
MI	[redacted]	0.88 (0.66–1.17)
All-cause mortality	[redacted]	0.89 (0.80–0.998)
Fatal stroke	[redacted]	–
Disabling stroke	[redacted]	–
Non-disabling stroke	[redacted]	–
ICH	[redacted]	0.42 (0.30–0.58)
Major bleeding	[redacted]	0.69 (0.60–0.80)
GI bleeding	[redacted]	0.89 (0.70–1.15)
Other major bleed	[redacted]	–
CRNM bleeding	[redacted]	–
Any bleeding	[redacted]	0.71 (0.68–0.75)
Discontinuations	[redacted]	–

Abbreviations used in table: CRNM, clinically relevant non-major; ICH, intracranial haemorrhage; GI, gastrointestinal; MI, myocardial infarction; NMA, network meta-analyses; SE, systemic embolism; vs, versus.

4.5 Conclusions of the clinical effectiveness section

4.5.1 Summary of clinical results

- Apixaban was demonstrated to be non-inferior and also superior to warfarin in the prevention of stroke and SE.
- Apixaban was associated with statistically fewer haemorrhagic strokes compared with warfarin and although there was no statistically significant difference in the individual outcomes of ischaemic stroke and systemic embolism.
- For the primary safety endpoint of reduction in ISTH major bleeding, apixaban demonstrated superiority compared with warfarin.
- The overall adverse event and safety profile of apixaban was comparable or better when compared with warfarin for the outcomes reported in the MS.
- [redacted]The results of the subgroup analyses by centre TTR suggested that the safety and efficacy of apixaban compared with warfarin were independent of the level of warfarin control; i.e. %TTR.
- [redacted]
- The base case results of NMA 1 suggested that apixaban was associated with a significantly lower incidence of MI compared to dabigatran 150 mg or 110 mg. [redacted]
- For the bleeding safety outcomes assessed in NMA 1[redacted]

4.5.2 Clinical issues

- UK marketing authorisation is not currently held for the use of apixaban for stroke prevention in AF; although, approval is expected to be granted in December 2012.
- In addition to the comparators listed in the final scope issued by NICE[30], the manufacturer has included evidence for apixaban versus aspirin in people for whom warfarin is unsuitable. The key trial presented by the manufacturer within the MS to provide the evidence for this comparison is the AVERROES trial comparing apixaban with aspirin.
- The manufacturer presents data from a network meta-analysis (NMA 2) for the comparisons of apixaban versus rivaroxaban, and apixaban versus dabigatran in people for whom warfarin is unsuitable. The ERG considers this network meta-analysis to be potentially flawed as the population of the trials of rivaroxaban and dabigatran informing the network consist of patients suitable for treatment with warfarin.
- TIA was listed as an outcome of interest in the final scope issued by NICE[30]. However, no data on TIA were collected in ARISTOTLE or AVERROES and thus no data on TIAs were presented in the MS.
- HRQoL data presented within the MS was limited to generic AF HRQoL data identified from a systematic review of the literature presented in the cost-effectiveness section of the MS.

- The lack of detailed individual $CHADS_2$ score data, particularly for the higher $CHADS_2$ scores (i.e. 3, 4, 5 and 6) limits the ability of the ERG to comment on any potential variation in apixaban treatment effect for these subgroups.
- There is potential clinical heterogeneity in NMA 1 due to the differences between the trials included. However, due to the presence of only one study for each comparator, the exclusion of any single study would have resulted in the exclusion of one of the treatments from the analysis.
- The centre TTR subgroups were defined differently in each of the included trials in NMA 1 and thus the ERG does not consider the manufacturer's aggregation of the data for the centre TTR subgroups to be directly comparable.

5 COST EFFECTIVENESS

5.1 Introduction

This single technology appraisal (STA) considers the cost-effectiveness of treatment with apixaban for the prevention of stroke and systemic embolism (SE) in patients with non-valvular atrial fibrillation (NVAF). The following sections provide a summary and critique of the economic evidence submitted by the manufacturer of apixaban in support of this STA. Table 32 summarises the location of key economic information with the manufacturer's written submission. In addition to the evidence provided in the MS, the manufacturer supplied a Microsoft® EXCEL® based economic model and additional report. The additional report provided details of a targeted literature review into the relative risk of mortality in a patient population with atrial fibrillation (AF).

Table 32. Summary of key information within the manufacturer's submission

Information	Section (MS)
Details of the systematic review of the economic literature	7.1
Population	7.2.1
Model structure	7.2.2 to 7.2.6
Technology	7.2.7
Treatment continuation rules	7.2.8
Clinical parameters and variables	7.3
Measurement and valuation of health effects and adverse events	7.4
Resource identification, valuation and measurement	7.5
Sensitivity analysis	7.6
Results	7.7
Validation	7.8
Subgroup analysis	7.9
Interpretation of economic evidence	7.10
Strengths and weaknesses of economic evaluation	7.10.3
Abbreviations used in table: MS, manufacturer's submission.	

5.2 ERG comment on manufacturer's review of cost-effectiveness evidence

A systematic review of the cost-effectiveness literature was carried out by the manufacturer. The objective of the manufacturer's review was to identify economic evaluations of interventions for the prevention of stroke and/or SE in adult patients with AF. Standard databases (NHS EED, Embase, EconLit, Medline In-Process & Other Non-Indexed Citations and OVID MEDLINE) were searched, from 1990 until 12th December 2011. In addition hand-searches of: manufacturer databases, the Cost-Effectiveness Analysis (CEA) Registry, conference proceedings and health technology assessment (HTA) submissions to the National Institute for Health and Clinical Excellence (NICE), were carried

out. The evidence review group (ERG) notes that the search terms used were reasonable and both inclusion and exclusion criteria were explicitly stated; the ERG considers it unlikely that relevant publications were excluded.

The manufacturer's review identified 27 potentially relevant publications: 20 full text economic evaluations, five conference abstracts and two NICE HTA submissions. Of the 20 full text publications identified, the manufacturer considered five as relevant to the review objective, on the grounds that they "evaluated currently available pharmacological interventions in an active comparator setting, and reported an [incremental cost-effectiveness ratio] ICER" (MS; pg 101). Following the completion of the manufacturer's review, a cost-utility analysis of apixaban versus aspirin in an AF population was published. The manufacturer included this study in their review; details of this and other studies considered relevant by the manufacturer (including the two HTA submissions) are provided in Table 33. In addition, the ERG identified a study published subsequent to the manufacturer's submission, which evaluates apixaban versus warfarin in an AF patient population. Details of this study are also presented in Table33.

The two HTA submissions[43],[44] included in the manufacturer's review used a Markov modelling approach with a United Kingdom (UK) National Health Service (NHS) perspective. Dabigatran or rivaroxaban were compared with warfarin and aspirin (as a secondary analysis) in a vitamin K antagonist (VKA) suitable and VKA unsuitable population, respectively. However, the ERG notes that neither submission was based on reliable data for the efficacy of dabigatran or rivaroxaban in a VKA unsuitable population. Of the remaining studies identified (either by the manufacturer or the ERG), five used a Markov modelling approach (Freeman 2011,[45] Gage 1995,[46] Shah 2011,[47] Lee 2012,[48] Kamel 2012[49]), one used a semi-Markov approach[50] and one used a discrete event simulation.[51] The cost-effectiveness of the following interventions were evaluated, dabigatran,[45;51] warfarin,[45;51] aspirin,[46;48] aspirin plus clopidogrel[47] and apixaban.[48],[49] Four of the studies used a lifetime,[45],[49] one a 10 year,[46] one a 20 year[47] and one[48] both a one year and 10 year time horizon. Five of the studies focused on the United States of America (USA),[45]-[49] one on the UK[51] and one on Canada.[50]

Table 33. Summary of relevant cost-utility studies (adapted from MS;Table 108; pg 222)

Study, Country	Model framework, Population & time horizon	Intervention and comparators	Effectiveness data source & outcomes collected	Study results and conclusions
HTA submissions				
Dabigatran STA,[43] UK	Markov model NVAF patients at risk of SE or stroke and eligible for anticoagulation treatment Mean age 71 and mean $CHADS_2$ score 2.1 Lifetime horizon	Dabigatran 110 mg BD. Dabigatran 150 mg BD. Warfarin Aspirin+ clopidogrel Aspirin No therapy	Ischaemic stroke, intracranial haemorrhage, haemorrhagic stroke, extracranial bleeds, SE, TIA and acute MI Data from RE-LY	*Manufacturer results* Vs aspirin (incremental cost per QALY) for all patients Dabigatran 150 mg: £4,434 Dabigatran 110 mg: £9,397 Warfarin: £2,493 Aspirin + clopidogrel: Dominant (less costly and more effective) Vs warfarin (incremental cost per QALY) for all patients Dabigatran 150 mg: £6,264 Dabigatran 110 mg: £18,691 *ERG most plausible ICER* Vs warfarin (incremental cost per QALY) for all patients Dabigatran 150 mg: £24,173 Aspirin, aspirin + clopidogrel and dabigatran 110 mg: dominated by warfarin (more costly and less effective)
Rivaroxaban STA,[44] UK	Markov model NVAF patients at risk of SE or stroke and eligible for anticoagulation treatment Mean age 73 and mean $CHADS_2$ score 3.5 30 year time horizon	Rivaroxaban Warfarin Dabigatran Aspirin	Major and minor stroke, SE, major and minor extra-cranial bleeding, intracranial bleeding, MI and death. Data from ROCKET-AF and NMA conducted by the manufacturer	*Manufacturer results* Vs aspirin (incremental cost per QALY) for all patients Rivaroxaban: £2,083 Vs warfarin (incremental cost per QALY) for all patients Rivaroxaban: £18,883 *ERG most plausible ICER* Vs warfarin (incremental cost per QALY) for all patients Rivaroxaban: £33,758
Studies identified in the manufacturer's systematic review				
Freeman 2011,[45] USA	Markov Model NVAF patients, aged ≥65 with $CHADS_2$ score >0 Lifetime horizon	Dabigatran 110 mg BD Dabigatran 150 mg BD Warfarin	Ischaemic stroke (fatal, moderate to severe, mild, or reversible); haemorrhagic stroke (fatal, moderate to severe intracranial, mild intracranial, major or minor non-cerebral); MI Data from RE-LY	Vs warfarin (incremental cost per QALY) Dabigatran 110 mg: $51,229 Dabigatran 150 mg: $45,372 Results were sensitive to cost of dabigatran, risk of stroke and intracranial haemorrhage
Gage 1995,[46] USA	Markov Model. 65 year-old US NVAF patients, and good candidates for warfarin and aspirin therapy 10 year time	Warfarin Aspirin No treatment	Ischaemic stroke and haemorrhage (fatal, moderate to severe, mild, or reversible) Data from published clinical trials in NVAF	Incremental cost per QALY: *High risk of stroke:* Warfarin dominates (is less expensive and more effective) aspirin and no therapy *Medium risk of stroke:* Warfarin vs aspirin: $8,000 Warfarin dominates no therapy (is

Study, Country	Model framework, Population & time horizon	Intervention and comparators	Effectiveness data source & outcomes collected	Study results and conclusions
	horizon			less expensive and more effective) *Low risk of stroke:* Warfarin vs aspirin: $370,000 Warfarin vs no therapy: $14,000
Pink 2011,[51] UK	Discrete event simulation Cohort of 50,000 simulated patients at moderate to high risk of stroke with mean baseline $CHADS_2$ score of 2.1 Lifetime horizon	Dabigatran 110 mg BD Dabigatran 150 mg BD Warfarin	Stroke and MI (fatal, permanent, reversible); pulmonary embolism (fatal, reversible); bleed (fatal, minor, major, ICH); TIA; congestive heart failure Data from RE-LY and published meta-analysis for thromboembolic events with aspirin	Incremental cost per QALY: Dabigatran 110 mg vs warfarin: £43,074 Dabigatran 150 mg vs warfarin: £23,082 Dabigatran 150 mg vs dabigatran 110 mg: Dabigatran 150 mg dominant strategy (less costly and more effective) Dabigatran 150 mg vs warfarin was considered cost effective for all subgroups of patients except centres with good control of INR (£42,386 per QALY)
Shah 2011[47] USA	Markov Model Hypothetical cohort of 70-year-old patients based on the RE-LY clinical trial with moderate risk of stroke 20 year time horizon	Dabigatran 110 mg BD Dabigatran 150 mg BD Warfarin Aspirin+ clopidogrel Aspirin No therapy	Ischaemic stroke, TIA, intracranial haemorrhage, major and minor non-cerebral haemorrhage, MI, dyspepsia Data from RE-LY, a meta-analysis of published clinical trials, and observational data	Vs aspirin (incremental cost per QALY) Dabigatran 150 mg: $50,000 Dabigatran 110 mg: $66,000 Warfarin: $12,500 Aspirin+ clopidogrel: $99,000 Vs warfarin (incremental cost per QALY) Dabigatran 150 mg: $86,000 Dabigatran 110 mg: $150,000 Aspirin+ clopidogrel: dominated (more costly and less effective)
Sorensen 2011[50] Canada	Semi-Markov Model Canadian AF patients based on the RE-LY trial, with mean $CHADS_2$ score 2.1 and mean age 69 years Lifetime horizon	Dabigatran 150 mg BD Dabigatran 110 mg BD Warfarin Aspirin	Ischaemic stroke, intracranial haemorrhage and haemorrhagic stroke (fatal, disability: moderate, independent, dependent); SE, acute MI and extracranial haemorrhage (fatal, non-fatal); TIA; minor bleed Data from a meta-analysis and RE-LY	Incremental cost per QALY: Dabigatran sequential dosing vs 'trial-like' warfarin: CAN$10,440 Dabigatran vs 'real-world prescribing' (warfarin, aspirin or no treatment): CAN$3,962 Dabigatran 150 mg vs 'trial-like' warfarin: CAN$9,041 Dabigatran 110 mg vs 'trial-like' warfarin: CAN$29,994
Studies identified subsequent to the manufacturer's systematic review				
Lee 2012[48] USA	Markov model 1 and 10 year time horizons Hypothetical cohort of 70 year old patients with AF, a $CHADS_2$ score of 2 and a low risk of	Apixaban Aspirin	Ischaemic stroke and intracranial haemorrhage (fatal, minor, major, reversible); MI Data from historic data and AVERROES	One-year model Apixaban inferior strategy (most costly but no more effective) 10-year model Apixaban dominant (less costly and more effective) Results were sensitive to baseline stroke rate, monthly cost of stroke and prior VKA use in both models

Study, Country	Model framework, Population & time horizon	Intervention and comparators	Effectiveness data source & outcomes collected	Study results and conclusions
	bleeding			
Kamel 2012,[49] USA	Markov model Lifetime time horizon Hypothetical cohort of 70-year-old patients with no contraindication to anticoagulation and a history of stroke or TIA from nonvalvular AF	Apixaban Warfarin	Ischaemic stroke and intracranial haemorrhage (fatal, minor, major, reversible); MI Data from ARISTOTLE	Apixaban provided a gain of 0.28 quality-adjusted life-years (QALYs) at an additional cost of $3,200, resulting in an incremental cost-effectiveness ratio of $11,400 per QALY. The findings were robust in univariate sensitivity analyses varying model inputs across plausible ranges. In Monte Carlo analysis, apixaban was cost-effective in 62% of simulations using a threshold of $50,000 per QALY and 81% of simulations using a threshold of $100,000 per QALY.
Abbreviations used in table: AF, atrial fibrillation; BD., twice daily; CEAC, cost-effectiveness acceptability curve; $CHADS_2$, clinical prediction rule for estimating the risk of stroke in AF patients; ICER, incremental cost-effectiveness ratio; MI, myocardial infarction; NMA, network meta-analysis; NR, not reported; NVAF, non-valvular atrial fibrillation; PSA, probabilistic sensitivity analysis; QALY, quality adjusted life year; RCT, randomised controlled trial; RE-LY, Randomized Evaluation of Long-Tern Anticoagulation Therapy; SE, systemic embolism; SA, sensitivity analysis; TIA, transient ischaemic attack; UK, United Kingdom; USA, United States of America; VKA, vitamin-K antagonist; vs, versus				

5.3 Summary of manufacturer's de novo economic evaluation

The manufacturer constructed a Markov model to evaluate the long- and medium-term consequences of apixaban for the prevention of stroke or SE in AF patients. The model captured the clinical and economic consequences of treatment by modelling the movement of patients between discrete health states over a lifetime time horizon.

5.3.1 Interventions and comparators

The landscape of standard care for AF patients at moderate to high risk of stroke has changed recently with the approval of two novel oral anticoagulants (NOACs). Rivaroxaban (20 mg OD) and dabigatran (110 mg twice daily or 150 mg BD) were approved by NICE for the prevention of stroke or SE in VKA suitable or unsuitable AF patients in March 2012 and May 2012, respectively.[43],[44] Consequently, the final scope issued by NICE stipulated the comparison of apixaban with dabigatran and with rivaroxaban. This was in addition to the historical comparison against warfarin as the long standing prophylactic treatment of AF patients who are suitable for VKA therapy.[30] For patients who are unsuitable for VKA therapy the final scope specified the comparison of apixaban with dabigatran and with rivaroxaban.[30] However, the manufacturer presented two fully incremental analyses of:

- apixaban, dabigatran blend (150 mg BD switching to 110 mg at the age of 80 years), dabigatran 110 mg, rivaroxaban and warfarin in a VKA suitable population;
- apixaban, dabigatran blend (150 mg BD switching to 110 mg at the age of 80 years), dabigatran 110 mg, rivaroxaban and aspirin in a VKA unsuitable population.

Although, aspirin was not listed as a comparator in the NICE final scope, the manufacturer stated that aspirin was a relevant comparator in VKA unsuitable AF patients. The manufacturer's rationale for including aspirin as a comparator was that aspirin is at present recommended[23] and widely used to treat VKA unsuitable patients in England and Wales.[7][52][53] The ERG agrees with the manufacturer that aspirin is currently recommended and widely used. However, the recent approval of the more effective NOACs (dabigatran and rivaroxaban) in VKA suitable and VKA unsuitable patients lessens the importance of the comparison with aspirin.[43][44] Therefore, as discussed in Section 3.3, the focus of the ERG report is on the comparison of apixaban with dabigatran, rivaroxaban and warfarin. However, the results of apixaban compared with aspirin in the VKA unsuitable population are presented in Appendix 9.6.

5.3.2 Population

VKA suitable and VKA unsuitable patient populations were considered separately within the manufacturer's model. However, the baseline characteristics of both populations were assumed to be equivalent to the characteristics reported in the UK General Practice Research Database (GPRD).[52] Data from the manufacturer's network meta-analysis (NMA) 1 (based on data from RE-LY, ROCKET-AF and ARISTOTLE) and NMA 2 (based on data from RE-LY, ROCKET-AF, ARISTOTLE and AVERROES) were used to inform the clinical effectiveness of treatments in the VKA suitable and VKA unsuitable populations, respectively. The ERG considers that NMA 1 provided a reliable estimate of the relative treatment effect of apixaban versus dabigatran and versus rivaroxaban. However (as discussed in Section 4.4), the viability of a comparison of apixaban with dabigatran or with rivaroxaban was limited; as a result of an absence of efficacy data for dabigatran or rivaroxaban in a VKA unsuitable patient population. Consequently, the ERG notes that NMA 2 may not have provided a reliable relative estimate of apixaban versus dabigatran or rivaroxaban in a VKA unsuitable population. Moreover, the difference between the patient populations of AVERROES and those included in NMA 1 further increased the uncertainty around the estimates of relative treatment effect in NMA 2. Therefore, the results of the model assuming a VKA suitable population are the focus of the ERG report. However, the results of the model assuming a VKA unsuitable patient population are presented in Appendix 9.5.

In addition to the main analysis carried out in the VKA suitable population, the manufacturer also presented subgroup analyses by:

- level of international normalised ratio (INR) control;
- $CHADS_2$ score.

Time in therapeutic range (TTR) was used as a measure of INR control; however, trial data were not stratified by TTR. Therefore, the manufacturer endeavoured to preserve a level of randomisation in these subgroup analyses by grouping centres by median TTR (cTTR). However, as a result of limited cTTR data available from RE-LY and ROCKET-AF, economic subgroup analyses by cTTR were viable only in the comparison of apixaban with warfarin. The manufacturer did carry out clinical analyses of cTTR across all the relevant comparators; however, the applicability of these analyses was limited by the inconsistencies in the data (see Section 4.3.1). Four ranges of cTTR were considered in accordance with the quartiles of cTTR observed in ARISTOTLE:

- cTTR <52.38%;
- 52.38% ≤ cTTR < 66.02%;
- 66.02% ≤ cTTR < 76.51%;
- cTTR ≥76.51%.

Subgroup analyses by $CHADS_2$ score were carried out for all comparators in the following $CHADS_2$ score categories:

- $CHADS_2$ score = 1;
- $CHADS_2$ score = 2;
- $CHADS_2$ score = 3–6.

Trial data were not stratified by $CHADS_2$ score, however, $CHADS_2$ score categories were chosen based on the number of ARISTOTLE patients in each category (i.e. to ensure sufficient patient numbers in each category).

Sections 5.3.7 and 5.3.15 provide details of the methods used and results of these subgroup analyses, respectively.

5.3.3 Perspective, time horizon and discounting

The manufacturer's economic evaluation was undertaken from the perspective of the NHS and Personal Social Services (PSS) in England and Wales. A lifetime time horizon was used and both costs and benefits were discounted at 3.5% per annum after the first year.

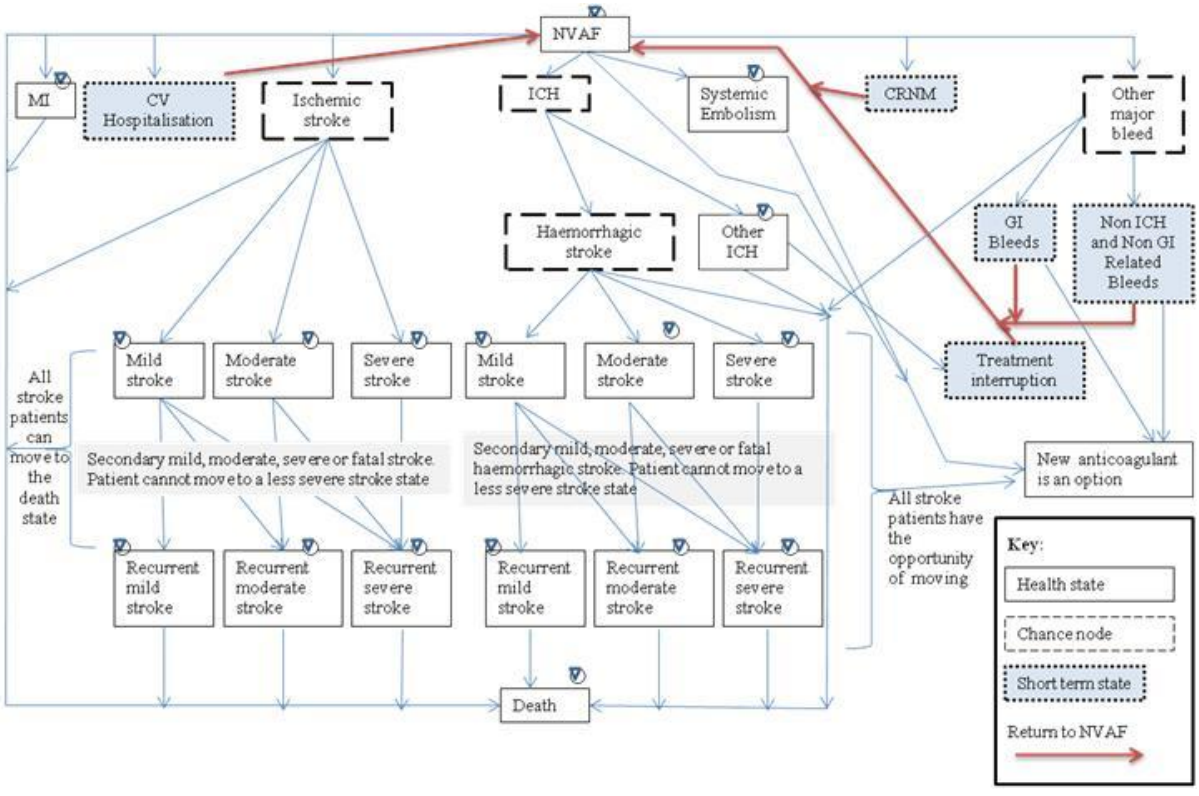
5.3.4 Model structure

The Markov model submitted by the manufacturer consisted of 18 health states, including the absorbing state of death. Patients transitioned between health states in cycles of 6 weeks with only one

clinical event permitted per cycle. The ERG notes that, given the influence of individual patient characteristics on outcomes in AF, a discreet event simulation rather than a Markov cohort modelling approach may have been more appropriate. However, the ERG acknowledges that a well-constructed Markov model may be sufficient to capture the mean differences in costs and consequences associated with prophylactic treatments in AF. Furthermore, the ERG notes that the use of Markov modelling techniques is consistent with previous novel oral anticoagulant (NOAC) HTA submissions.[43][44] By contrast, the cycle length used in the manufacturer's model was shorter than the cycle length used in the dabigatran and rivaroxaban submissions (6 weeks versus 3 months, respectively).[43][44] However, it has been previously noted that a cycle length of 3 months may compromise the reasonableness of the assumption of one clinical event per cycle. Therefore, the ERG considers the manufacturer's use of a shorter (6 week) cycle length to be appropriate.

The model allowed a maximum of two lines of therapy (see Section 5.3.5 for further details on treatment algorithms). The model structure for patients on first-line therapy is displayed in Figure 3.

Figure 3. Model structure for patients on first-line therapy (reproduced from MS figure 15 pg105)

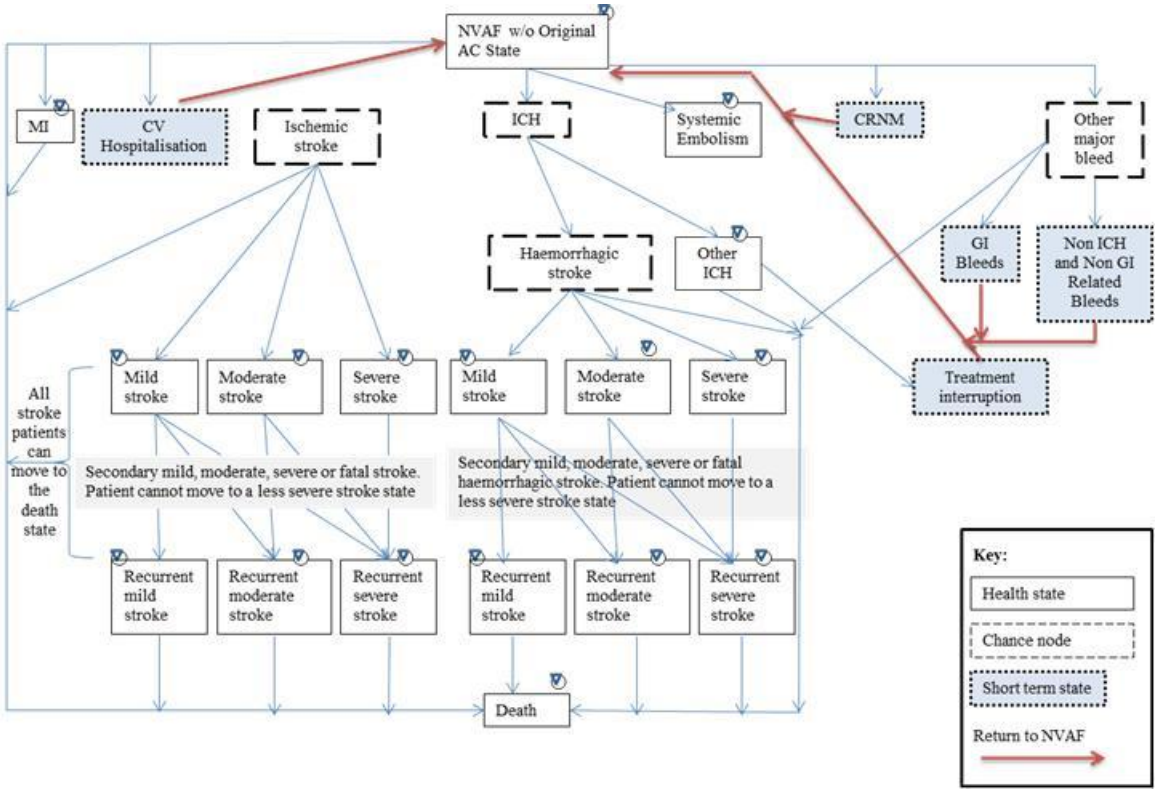


Patients entered the model in the "NVAF" health state. Patients in the "NVAF" health state were assumed to have stable AF and to have received any one of the considered interventions; patients in the "NVAF" health state had not yet experienced any clinical event. The model assumed that patients either remained in-state until death or experienced one of the following events:

- ischaemic stroke (mild/moderate/severe/fatal);[1]
- haemorrhagic stroke (mild/moderate/severe/fatal);[1]
- non-fatal or fatal other intracranial haemorrhage (ICH i.e non-haemorrhagic stroke);
- non-fatal or fatal SE;
- non-fatal or fatal other major bleeds (gastrointestinal (GI) bleeds or other non-ICH and non-GI related bleeds);
- clinically relevant non-major (CRNM) bleeds;
- non-fatal or fatal myocardial infarction (MI);
- other cardiovascular (CV) hospitalisations (i.e. CV hospitalisations unrelated to stroke or MI).

In addition to the transitions listed above, patients in the "NVAF" health state were able to switch from their initial anticoagulation therapy; a switch to second-line therapy was either a consequence of certain events or a result of other causes (see Section 5.3.5 for further details of treatment switching and discontinuation). Following a switch to second-line therapy, patients transitioned into the "NVAF w/o original AC" health state. Henceforth, patients were at risk of the same events (with the exception of switch to second-line therapy) as patients in the "NVAF" health state. Figure 4 summarises the transitions of patients from the "NVAF without original AC" health state.

Figure 4. Model structure for patients on second-line therapy (reproduced from MS figure 16 pg106)



[1] Stroke (ischaemic or haemorrhagic) is categorised into three severity levels based on the modified Rankin scale (mRS). The three severity levels considered are mild (mRS 0-2), moderate (mRS 3-4) and severe (mRS 5).

The ERG notes that in addition to thromboembolic events, the manufacturer's model captured clinically relevant adverse events. In particular, bleeding events (the most important safety outcomes of treatment to prevent stroke or SE in AF patients), which were captured in distinct health states. In addition to adverse bleeding events, the model captured the impact of dyspepsia as a result of anticoagulation therapy and renal complications associated with dabigatran (see Section 5.3.11 for details). However, the ERG notes that transient ischaemic attack (TIA), an outcome specified in the NICE final scope[30] was not included in the manufacturer's model; the manufacturer stated that TIA data were not collected in the ARISTOTLE trial (MS; pg 32). Although, following expert clinical advice, the ERG notes that TIA events may have been indirectly accounted for within the outcome of mild stroke. Therefore, overall the ERG is satisfied that the model captured all clinically relevant events.

Permanent and temporary events

The events captured in the manufacturer's model were categorised as either permanent (ischaemic stroke, haemorrhagic stroke, SE, MI) or temporary (other ICH, other major bleeds, CRNM bleeds and other CV hospitalisations) events, in accordance with the long-term impact associated with the event. Patients who experienced a temporary event transitioned into a temporary health state where they incurred the acute cost and disutility associated with the event; the duration of each temporary health state was user-defined. In the base case, patients who experienced an (non-fatal) "other ICH" or major bleed remained in the event-related temporary health state for 6 and 2 weeks, respectively. Similarly, patients who experienced a CV hospitalisation or CRNM bleed were assumed to remain in-state for 6 and 2 days, respectively. Following a (non-fatal) temporary event, all patients were assumed to recover to their previous health status (i.e. transitioned back to the "NVAF" or "NVAF without original AC" health states).

However, patients who experienced permanent events (stroke, SE or MI) accrued both acute and long-term maintenance costs and were not assumed to recover to their previous level of health (i.e. experienced long-term disutility). Therefore, patients who experienced a permanent event transitioned into a health state representative of the longer term costs and quality of life (QoL) associated with the event they had experienced. Acute costs were applied per episode and were dependent upon user-specified episode durations; in the manufacturer's base case the acute aspect of all permanent events was assumed to endure for 2 weeks. Long-term maintenance costs were applied for as long as the patient remained in the permanent health state and included such costs as the costs of drugs, monitoring and patient care.

Based on expert clinical advice, the ERG considers the manufacturer's categorisation of events to be appropriate.

Further transitions

Following a permanent event, the future transitions of patients were limited, that is, patients were not exposed to the risk of all events. In particular, the health states of MI and SE were assumed to be semi-absorbing; following entrance to the MI or SE health states the only risk patients were exposed to was that of death. However, patients who experienced a (non-fatal) stroke were exposed to the risk of one and only one recurrent stroke, in addition to the risk of death. Recurrent strokes were assumed to be of the same type (ischaemic or ICH) as the initial event; however, the severity of a recurrent stroke was assumed to be independent of initial stroke severity. The resource use and disutility associated with a recurrent stroke was assumed to be equal to that of the most severe stroke experienced (e.g. if a patient experienced a moderate stroke followed by a mild stroke, the resource use and disutility accrued by the patient would be that of a moderate stroke).

The manufacturer's rationale for limiting the future risks of patients was a paucity of data to inform further event risks. The implication of the manufacturer's assumptions regarding future transitions in terms of potential model bias is discussed in Section 5.4.6.

5.3.5 Treatment switching and discontinuation

In the manufacturer's model, discontinuation from first-line treatment was assumed to be either a result of particular temporary clinical events (event-related discontinuation) or other causes (other-cause discontinuation). In the base-case, second-line therapy with aspirin was assumed to follow either event-related or other-cause discontinuation. However, cessation of treatment (i.e. no treatment second-line) was considered in a sensitivity analysis (see Section 5.3.14 for sensitivity analysis results). A switch to second-line treatment was assumed to affect a patients risk profile, with the risk of clinical events being dependant on therapy received (see Section 5.3.7 for full details of treatment specific event risks).

In addition to a switch from first- to second-line therapy, other treatment switches were permitted following certain "permanent" events. In particular, patients who had experienced an MI or haemorrhagic stroke were assumed to discontinue anticoagulation therapy. Similarly, patients who were receiving therapy with aspirin (e.g. patients on second-line therapy) were assumed to switch to warfarin following either an ischaemic stroke or SE. However, the effect of these treatment switches was limited to costs accrued; i.e. did not affect patients risk profiles. This is because the risk of events for patients in "permanent" health states was assumed to be independent of treatment received.

Event-related discontinuation

Patients who experienced an ICH that was not a haemorrhagic stroke ("other ICH"), a GI bleed or other major bleed were exposed to a risk of treatment discontinuation (event-related discontinuation).

Following an "other ICH" the manufacturer assumed that 56% of patients switched to second-line therapy; the remaining 44% of patients were assumed to experience a 6-week treatment interruption. These assumptions were based on evidence presented in a study by Claassen et al.[54] which considered a cohort of patients from a single centre in the US (n=52). Similarly, following a major bleed (GI or non-GI related) 25% of patients were assumed to switch to second-line therapy. The proportion of patients who switched therapy following a major bleed was taken from a cost-effectiveness analysis by Sorensen et al.[55] The study by Sorensen et al. considered trial versus "real-world" data for warfarin in the prevention of stroke or SE in AF patients.[55]

The ERG notes that data for event-related discontinuation was based on data associated with warfarin treatment. However, in ARISTOTLE, rates of discontinuation due to adverse events were lower for patients treated with apixaban than for patients treated with warfarin (7.6% versus 8.4%, respectively) (MS; pg 85). Therefore, the ERG considers that the use of warfarin data to inform event-related discontinuation may have been a conservative assumption (i.e. any bias is likely to be against apixaban).

Furthermore, the ERG notes that following a major bleed, patients that remained on their original anticoagulation therapy (rather than switched treatment) were not assumed to experience a treatment interruption. However, based on expert clinical opinion, the ERG notes that in clinical practice a treatment interruption may occur for patients who experience a major bleed. In addition, the ERG notes that patients who do experience a treatment interruption may be at higher risk of thromboembolic events. However, the ERG understands that there is a paucity of data to inform treatment interruptions following a major bleed, or risk profiles following a treatment interruption. In addition, the ERG notes that treatment interruption was assumed to last for 6 weeks. Therefore, it is unlikely that any changes in risk profile or costs over a period of 6 weeks will affect the model's lifetime cost-effectiveness results.

Other-cause discontinuation

By contrast to event-related discontinuation, the rate of other-cause treatment discontinuation was dependant on treatment received. Data from the ARISTOTLE trial and the manufacturer's NMA were used to inform the rate of other-cause discontinuation used in the manufacturer's model (see Section 5.3.7 for full details). Table 34 summarises the treatment specific rate of other-cause discontinuation used in the manufacturer's model.

Table 34, Rate of other-cause discontinuation associated with each treatment (per 100 patient years)

Treatment	Rate
Apixaban	[redacted]
Rivaroxaban	[redacted]
Dabigatran 110 mg	[redacted]
Dabigatran 150 mg	[redacted]
Warfarin	[redacted]
Abbreviations used in table: mg, milligram	

5.3.6 Summary of model parameters

Table 35 summarises all parameters used in the manufacturer's base case model in theVKA suitable population.

Table 35: Summary of (VKA suitable) base case model parameters

Parameter	Treatment					
	Apixaban	Warfarin	Dabigatran 110mg	Dabigatran 150mg	Rivaroxaban	Aspirin (2nd line)
Baseline $CHADS_2$ score						
0	9.80%					N/A
1	30.10%					
2	29.60%					
3	17.90%					
4	8.50%					
5	4.10%					
6	0.00%					
Ischaemic stroke						
Baseline rate (per 100 PY)	[redacted]	N/A				[redacted]
HR	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
Annual risk adjustment factor	1.40					N/A
Proportion mild stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Proportion moderate stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Proportion severe stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Case fatality rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Annual rate of recurrence	0.04					
Bleeding						
Baseline ICH rate (per 100 PY)	0.33	N/A				[redacted]
HR for ICH	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
ICH annual risk adjustment factor	2.0					N/A
Proportion haemorrhagic stroke	[redacted]	[redacted]	64.00%	41.00%	57.00%	[redacted]
Proportion mild haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Proportion moderate haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Proportion severe haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Haemorrhagic stroke case fatality rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Annual recurrent haemorrhagic stroke rate	0.03					
Other ICH case fatality	[redacted]					
Other ICH treatment interruption proportion	44.0%					
Baseline Other major bleeding rate (per 100 PY)	1.79	N/A				[redacted]
HR for other major bleed	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
Other major bleed annual risk adjustment factor	1.97					N/A
Other major bleed	[redacted]					

case fatality						
Major bleed no treatment change proportion	75.00%					
Proportion GI bleed	[redacted]	[redacted]	41.00%	49.00%	45.00%	[redacted]
Baseline CRNM bleeding rate (per100 PY)	[redacted]	N/A				[redacted]
HR for CRNM bleed	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
CRNM bleed annual risk adjustment factor	1.97					N/A
Other events						
Baseline MI rate (per 100 PY)	0.53	N/A				[redacted]
HR for MI	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
MI male case fatality rate	0.11					
MI female case fatality rate	0.16					
MI annual risk adjustment factor	1.30					N/A
Baseline CV hospitalisation rate (per 100 PY)	[redacted]	N/A				[redacted]
HR for CV hospitalisation	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
Baseline Tx discontinaution rate (per 100 PY)	[redacted]	N/A				
HR for Tx discontinuation	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
Baseline SE rate (per 100 PY)	[redacted]	N/A				[redacted]
HR for SE	N/A	[redacted]	[redacted]	[redacted]	[redacted]	N/A
SE case fatality rate	[redacted]					
Other-cause mortality						
Other death risk trial period	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AF baseline death HR	[redacted]					
HR following mild stroke (isc/haem)	3.18					
HR following moderate stroke (isc/haem)	5.84					
HR following severe stroke (isc/haem)	15.75					
Male HR following MI	2.56					
Female HR following MI	4.16					
HR following SE	1.34					
Costs (£)						
Monthly management cost (dyspepsia/anticoagulant)	0.04	0.04	0.13	0.13	0.04	0.04
Mild ischaemic stroke acute care (per episode)	3,515.64					
Mild ischaemic stroke maintenance care (per month)	183.91					
Moderate ischaemic stroke acute care (per episode)	18,341.08					
Moderate ischaemic	358.78					

stroke maintenance care (per month)	
Severe ischaemic stroke acute care (per episode)	25,050.88
Severe ischaemic stroke maintenance care (per month)	544.76
Fatal ischaemic stroke cost (per episode)	3,162.11
Mild haemorrhagic stroke acute care (per episode)	10,236.81
Mild haemorrhagic stroke maintenance care (per month)	183.91
Moderate haemorrhagic stroke acute care (per episode)	26,299.60
Moderate haemorrhagic stroke maintanence care (per month)	358.78
Severe haemorrhagic stroke acute care (per episode)	44,486.65
Severe haemorrhagic stroke maintanence care (per month)	544.76
Fatal haemorrhagic stroke cost (per episode)	1,645.66
SE acute care cost (per episode)	4,077.98
SE acute care cost (per month)	183.91
Other ICH (excluding haemorrhagic stroke)	3,010.00
Other major bleed GI (excluding ICH)	1,493.68
Other major bleed non-GI (excluding ICH)	3,947.92
CRNM bleed	1,133.93
MI acute care (per episode)	2,018.84
MI maintanence care (per month)	6.65
Other CV hospitalisation	1,570.89
Utility value	
AF baseline	0.8
Mild ischaemic stroke	0.8
Moderate ischaemic stroke	0.4
Severe ischaemic stroke	0.1
Mild haemorrhagic stroke	0.8
Moderate haemorrhagic stroke	0.4
Severe haemorrhagic stroke	0.1

SE	0.7
Utility decrement	
Other ICH	0.107
Other major Bleed	0.107
CRNM bleed	0.1
MI	0.7
Other CV hospitalisation	0.1
Apixaban	0.002
Dabigatran	0.002
Rivaroxaban	0.002
Aspirin	0.002
Warfarin	0.013
Abbreviations used in table: AF, atrial fibrillation; mg, milligram; $CHADS_2$, Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score; CRNM, clinically relevant non-major; CV, cardiovascular; GI, gastrointestinal; HR, hazard ratio; ICH, intracranial haemorrhage; MI, myocardial infarction; N/A, not applicable; PY, patient years; SE, systemic embolism; Tx, treatment; VKA, vitamin K antagonist.	

5.3.7 Treatment effectiveness

The manufacturer's model was structured to allow treatment to affect a variety of parameters, in particular:

- event risks;
- distribution of stroke severity;
- distribution of bleed type.

Each of these parameters were informed by data from the ARISTOTLE trial and/or the manufacturer's NMA. The appropriateness of assuming that each of these parameters was specific to treatment regimen is discussed in Section 5.4.5. However, the focus of this section is to describe (for first-line and second-line treatment) the data sources for and implementation of event risks, stroke severity distribution and distribution of bleed type. In addition, this section details the treatment effectiveness parameters used in the manufacturer's subgroup analyses.

Event risks

The risk of occurrence of each modelled event varied depending on which treatment a patient received. In addition, the risk of stroke was adjusted for baseline $CHADS_2$ score distribution and the risks of stroke, ICH, MI, other major bleeds and CRNM bleeds were adjusted for age.

The treatment specific relative event rates were derived from NMA 1 (in the VKA suitable population) carried out by the manufacturer. For the purposes of the economic model, apixaban was used as the reference treatment (as oppose to warfarin used as the reference treatment for the results of NMA 1 reported in Section 4.4). Therefore, the baseline event rates in the model were those associated with apixaban treatment. The event rates associated with other treatments considered in the manufacturer's model were calculated by applying a relative treatment effect to the baseline rate for

each event; the manufacturer's NMA used hazard ratios as a measure of relative treatment effect. For each event the treatment specific hazard ratio was applied multiplicatively to the baseline event rate; and the resultant rate converted into a cycle specific probability using standard formulae:

$$O_r = B_r.HR$$

Where: O_r is the overall rate (treatment and event specific); B_r is the baseline event rate (with apixaban); HR is the event and treatment specific hazard ratio.

$$r = \frac{O_r / d_Y}{n}$$

Where: r is the daily rate; d_Y is the number of days in a year (i.e. 365.25); n is the number of patients.

$$p = 1 - \exp(-rt)$$

Where: p is the cycle length probability; r is the daily rate and t is the cycle length (i.e. 42 days).

Table 36 summarises the baseline and relative event rates used in the manufacturer's model.

Table 36. Baseline and relative event rates used in the manufacturer's model (VKA suitable population only) HR>1 favours apixaban; HR<1 favours comparator

Event	Baseline rate (per 100 PY)	Treatment specific hazard ratio (95%CrI)[a]			
		Rivaroxaban	Dabigatran 110 mg	Dabigatran 150 mg	Warfarin
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICH	0.330[c]	[redacted]	[redacted]	[redacted]	[redacted]
Other major bleeds	1.790 [c]	[redacted]	[redacted]	[redacted]	[redacted]
CRNM bleeds	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MI	0.530[c]	[redacted]	[redacted]	[redacted]	[redacted]
Other cardiovascular hospitalisations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other treatment discontinuations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations used in table: CrI, credible interval; CRNM, clinically relevant non-major; HR, hazard ratio; ICH, intracranial haemorrhagic; mg, milligram; MI, myocardial infarction; PY, patient years; SE, systemic embolism; VKA, vitamin K antagonist.

[a]Source: manufacturer NMA 1 (warfarin suitable patient population); [b]calculated from a weighted average of ischaemic stroke risk by $CHADS_2$ score. [c]Source: ARISTOTLE; [d]Source: secondary analysis of ARISTOTLE data.

The baseline risk of ischaemic stroke was calculated as a weighted average of stroke risk by $CHADS_2$ score; weighted by the proportion of patients in each $CHADS_2$ score category at baseline. Table 37 summarizes the risk of stroke by $CHADS_2$ score used in the manufacturer's model. The $CHADS_2$ specific stroke risks were derived from a secondary analysis of ARISTOTLE trial data.

Table 37. Stroke risk by $CHADS_2$ score for patients receiving apixaban

$CHADS_2$ score	Stroke rate (per 100 PY)	Proportion of patients (%)
0–1	[redacted]	39.90
2	[redacted]	29.60
3–6	[redacted]	30.50
Average stroke risk	[redacted]	

Abbreviations used in table: CHADS, cardiac failure, hypertension, age, diabetes, and stoke; PY, patient years.

In addition to the modifying effect of treatment on event risks (and $CHADS_2$ score on stroke risk), the effect of patient age on certain event risks was also accounted for. Risk adjustment factors identified from the reference list of Freeman et al.[45] (The study by Freeman et al. was identified in the manufacturer's review of the cost-effectiveness literature) were applied to the treatment specific event risks. The per decade risk adjustment factors used are summarised in Table 38.

Table 38. Risk adjustment factors used in the manufacturer's model (adapted from MS; Table 43; pg 115)

Event	Risk adjustment factor (per decade)	Source
Ischaemic stroke	1.400	Pooled data from 5 RCTs[56]
ICH	1.970	Systematic review by Ariesen et al.[57]
Other major bleeds	1.970	
CRNM bleeds	1.970	
MI	1.300	Freeman et al.[45]
Abbreviations used in table: CRNM, clinically relevant non-major; ICH, intracranial haemorrhage; MI, myocardial infarction; RCT, randomised controlled trial.		

To allow the risk of events to gradually increase with patients increasing age, each factor was converted into a per cycle adjustment factor as follows:

$$R_C = R_D^{\left(\frac{t}{10.d_Y/d_W}\right)}$$

Where: R_C is the per cycle risk adjustment factor; R_D is the per decade risk adjustment factor; t is the cycle length in weeks; d_Y is the number of days in a year and d_W is the number of days in a week.

Recurrent stroke

As discussed in Section 5.3.4, patients who experienced a stroke event were exposed to the risk of a single recurrent stroke. The risk of subsequent stroke was independent of treatment and patient characteristics (such as age or $CHADS_2$ score). The annual rates of recurrent ischaemic and haemorrhagic stroke used in the manufacturer's model were 0.04 and 0.03, respectively. These data were taken from a study by Mohan et al.[58] Mohan et al. considered the recurrence of stroke in patients of the South London Stroke Registry for up to 10 years after their initial stroke event.

Stroke severity distribution

Stroke severity has been classified by the manufacturer according to the Modified Rankin Scale (mRS) with:

- mild stroke = mRS 0–2;
- moderate stroke = mRS 3–4;
- severe stroke = mRS 5;
- fatal stroke = mRS 6.

The likelihood that an initial stroke event was mild, moderate, severe or fatal was assumed to depend upon the treatment a patient received. Table 39 summarises the treatment specific ischaemic and haemorrhagic stroke severity distributions used in the manufacturer's model. The severity of recurrent

stroke was assumed to be independent of treatment received and assumed to be equivalent to the apixaban distribution of first stroke severity.

Table 39. Treatment specific stroke severity distribution

Treatment	Severity				Source
	Mild	Moderate	Severe	Fatal	
Ischaemic stroke					
Apixaban	[redacted]	[redacted]	[redacted]	[redacted]	Secondary analysis of ARISTOTLE data
Warfarin	[redacted]	[redacted]	[redacted]	[redacted]	
Dabigatran (110 mg)	[redacted]	[redacted]	[redacted]	[redacted]	Connolly et al.[26]
Dabigatran (150 mg)	[redacted]	[redacted]	[redacted]	[redacted]	
Rivaroxaban	[redacted]	[redacted]	[redacted]	[redacted]	Patel et al.[27]
Haemorrhagic stroke					
Apixaban	[redacted]	[redacted]	[redacted]	[redacted]	Secondary analysis of ARISTOTLE data
Warfarin	[redacted]	[redacted]	[redacted]	[redacted]	
Dabigatran (110 mg)	[redacted]	[redacted]	[redacted]	[redacted]	Connolly et al.[26]
Dabigatran (150 mg)	[redacted]	[redacted]	[redacted]	[redacted]	
Rivaroxaban	[redacted]	[redacted]	[redacted]	[redacted]	Patel et al.[27]
Abbreviations used in table: mg, milligram.					

However, the stroke severity distribution of RE-LY and ROCKET-AF was not reported in precisely the same way as the stroke severity distribution of ARISTOTLE. In particular, no distinction was made in RE-LY or ROCKET-AF between ischaemic and haemorrhagic stroke. Furthermore, the proportion of patients experiencing a severe (mRS 5) stroke was not disaggregated from the proportion of patients experiencing a moderate (mRS 3–4) stroke. Therefore, the manufacturer disaggregated moderate and severe stroke as reported in RE-LY and ROCKET-AF by weighting the data with proportions reported in ARISTOTLE. For example:

$$Mod_{Dabi} = Mod_{ARIS} .Mod / Sev_{RE-LY}$$

Where: Mod_{Dabi} is the proportion of patients treated with dabigatran who experienced a moderate stroke; Mod_{ARIS} is the proportion of patients (from those who experienced a moderate or severe stroke) who experienced a moderate stroke reported in ARISTOTLE and Mod/Sev_{RE-LY} is the proportion of patients who experienced a moderate or severe stroke reported in RE-LY. The ERG considers that the manufacturer's approach was reasonable and pragmatic, however, the ERG notes that the assumption of treatment dependant stroke severity may not be appropriate (discussed in Section 5.4.5).

Distribution of bleed type

In addition to haemorrhagic stroke, the manufacturer's model also accounted for "other ICH" (i.e., intracranial bleeds that were not haemorrhagic strokes). The distribution of ICH type was assumed to be dependent on the treatment regimen received. Similarly, the model accounted for two distinct types of (non-ICH) major bleeding: GI bleeding and non-GI bleeding. The distribution of type of major bleed was also assumed to depend on the treatment a patient was receiving. Table 40 summarises the treatment specific distribution of bleed type that was incorporated into the manufacturer's model.

Table 40. Distribution of bleed type by treatment regimen

Treatment	ICH			Major bleeds		
	Haemorrhagic stroke	Other ICH	Source	GI Bleeds	Non-GI related bleeds	Source
Apixaban	[redacted]	[redacted]	Secondary analysis of ARISTOTLE data	[redacted]	[redacted]	ARISTOTLE
Warfarin	[redacted]	[redacted]		[redacted]	[redacted]	Secondary analysis of ARISTOTLE data
Dabigatran (110 mg)	64%	36%	Connolly et al.[26]	41%	59%	Connolly et al.[26]
Dabigatran (150 mg)	41%	59%		49%	51%	
Rivaroxaban	57%	43%	Patel et al.[27]	45%	55%	Patel et al.[27]
Abbreviations used in Table: GI, gastrointestinal; ICH, intracranial haemorrhage; mg, milligram.						

Second-line treatment

As discussed in Section 5.3.5, patients were eligible to switch to second-line therapy (with either aspirin or no treatment) following certain bleeding events (other ICH and major bleeding) or as a result of other-cause treatment discontinuation. Additional treatment switches were also permitted within the model; however, these did not result in an alteration of a patient's risk profile. A switch to second-line therapy was associated with an alteration of a patients risk profile as similar to first-line therapy: event risks; stroke severity and bleed type were associated with treatment received. Tables 41 to 43 summarise the: absolute event rate, initial stroke severity distribution and distribution of bleed type for patients on second-line therapy.

Table 41. Absolute event rate (per 100 patient-years) for patients on second-line therapy (adapted from MS; Table 50; pg 120)

Event	Aspirin (2nd line)	Source	No Treatment	Source
Ischaemic stroke	[redacted]	Secondary analysis of AVERROES data	4.186	Mandema et al.[59]
ICH	[redacted]		0.000	Assumption
Other major bleeds	[redacted]		0.000	Assumption
CRNM bleeds	[redacted]		0.000	Assumption
MI	[redacted]		1.003	Mandema et al.[59]
SE	[redacted]	Assumption, AVERROES CSR	0.959	
Other CV hospitalisation	[redacted]	Secondary analysis of AVERROES data	16.506	
Abbreviations used in Table: CSR, clinical study report; CV, cardiovascular; ICH, intracranial haemorrhage; MI, myocardial infarction; SE, systemic embolism.				

With the exception of SE, the event risks for patients on second-line aspirin therapy were derived from secondary analyses of AVERROES data. The secondary analyses were carried out on data from patients who had failed first-line therapy with warfarin; the event rate of SE was assumed to be equivalent to that observed in the overall population of AVERROES.

Furthermore, by contrast to patients on first-line therapy, patients on second-line therapy were assumed to be exposed to a constant risk of events. That is, the risk of events was assumed to be the same regardless of age, duration of second-line treatment, prior anticoagulation therapy or other patient characteristics (e.g. $CHADS_2$ score). The potential impact of this assumption is discussed in Section 5.4.5. In addition, patients who ceased to receive anticoagulation therapy (i.e. switched to no treatment) were no longer exposed to the risk of bleeding events.

Table 42. Treatment specific stroke severity distribution (second-line therapy)

Treatment	Severity				Source
	Mild	Moderate	Severe	Fatal	
Ischaemic stroke					
Aspirin	[redacted]	[redacted]	[redacted]	[redacted]	Secondary analysis of ARISTOTLE data
No treatment	22%	37%	36%	5%	Hylek et al.[60]
Haemorrhagic stroke					
Aspirin	[redacted]	[redacted]	[redacted]	[redacted]	Secondary analysis of ARISTOTLE data
No treatment*	N/A				

Abbreviations used in table: N/A, not applicable.
*Patients who do not receive anticoagulation therapy are assumed not to be at risk of bleeding events.

Table 43. Distribution of bleed type by treatment regimen (second-line therapy)

Treatment	ICH			Major bleeds		
	Haemorrhagic stroke	Other ICH	Source	GI Bleeds	Non-GI related bleeds	Source
Aspirin	[redacted]	[redacted]	Secondary analysis of AVERROES data	[redacted]	[redacted]	Secondary analysis of AVERROES data
No treatment*	N/A					

Abbreviations used in Table: GI, gastrointestinal; ICH, intracranial haemorrhage; N/A, Not applicable.
*Patients who do not receive anticoagulation therapy are assumed not to be at risk of bleeding events.

Subgroup analyses

As discussed in Section 5.3.2, the manufacturer carried out subgroup analyses by level of INR control and $CHADS_2$ score categories. In the VKA suitable population, subgroup analyses by $CHADS_2$ score categories were carried out for all comparators. These analyses were enabled by a function within the manufacturer's model that allowed user-specification of the baseline $CHADS_2$ score distribution; the relative effect of treatment was assumed to be constant across $CHADS_2$ score categories. Based on evidence presented in the rivaroxaban STA,[26][61;62] the ERG considers that the assumption of constant relative treatment effect across $CHADS_2$ score categories was reasonable.

As a consequence of limited data available from RE-LY and ROCKET-AF, subgroup analyses by level of INR control were limited to comparison of apixaban with warfarin (in a VKA suitable population); TTR was used as a measure of INR control. In order to preserve a degree of randomisation TTR was assessed by centre (cTTR) with centres being grouped by median TTR. Four cTTR ranges, representing the quartiles of cTTR observed in ARISTOTLE were considered:

- cTTR <52.38%;
- 52.38% ≤ cTTR < 66.02%;
- 66.02% ≤ cTTR < 76.51%;
- cTTR ≥76.51%.

Secondary analyses of ARISTOTLE trial data were carried out to obtain absolute event rates for apixaban and warfarin in each cTTR range (Table 44).

Table 44. Absolute event rates for apixaban and warfarin by cTTR range.

cTTR range	Apixaban	Warfarin
Ischaemic Stroke		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
Intracranial haemorrhage		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
Other major bleed		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
CRNM bleed		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
Abbreviations used in table: CRNM, clinically relevant non-major; cTTR, centre time in therapeutic range		

These data were then used to adjust the baseline event rates associated with apixaban and warfarin. In order for the model to be enabled to consider any (user-specified) cTTR distribution the adjustment of baseline event rates was carried out as follows:

1) For each cTTR range a treatment specific relative rate was calculated for each event; relative to the absolute event rate for the cTTR range 52.38% ≤ cTTR < 66.02% (the range containing the average TTR from ARISTOTLE) (see Table 45).

Table 45. Relative event rates for apixaban and warfarin by cTTR range

cTTR range	Apixaban relative rate (95% CI)	Warfarin relative rate (95% CI)
Ischaemic stroke		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
ICH		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
Other major bleed		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
CRNM bleed		
cTTR <52.38%	[redacted]	[redacted]
52.38% ≤ cTTR < 66.02%	[redacted]	[redacted]
66.02% ≤ cTTR < 76.51%	[redacted]	[redacted]
cTTR ≥76.51%	[redacted]	[redacted]
Abbreviations used in table: CI, confidence interval; CRNM, clinically non-major; cTTR, centre time in therapeutic range.		

2) The relative event rates (Table 45) for each cTTR range were then weighted by the (user specified) proportion of patients in each cTTR range to determine an overall treatment and event specific "average relative rate".

   For example, consider the following cTTR relative event rates of ischaemic stroke for patients treated with apixaban:

   - cTTR <52.38% = [redacted];
   - 52.38% ≤ cTTR < 66.02% = [redacted];
   - 66.02% ≤ cTTR < 78.61% = [redacted];
   - cTTR >78.61% = [redacted].

   Now, assuming the user-specified distribution of cTTR was:

   - cTTR <52.38% = 10%;
   - 52.38% ≤ cTTR < 66.02% = 15%;
   - 66.02% ≤ cTTR < 78.61% = 25%;
   - cTTR >78.61% = 50%.

Then for patients treated with apixaban the "average relative rate" of ischaemic stroke would be:

$ARR_{USER\text{-}SPECIFIED}$ = ([redacted]*10%) + ([redacted]*15%) + ([redacted]*25%) + ([redacted]*50%) = [redacted]

Where $ARR_{USER\text{-}SPECIFIED}$ is the "average relative rate" based on the user-specified cTTR distribution

3) Then for each treatment the "relative rate for cTTR adjustment" was calculated by comparing the "average relative rate" for each event with the "average relative rate" for each event had the cTTR distribution specified been akin to that of ARISTOTLE (i.e. 25% of patients in each cTTR range). For example, the "relative rate for cTTR adjustment" for ischaemic stroke in patients receiving therapy with apixaban would be:

$ARR_{ARISTOTLE}$ = ([redacted]*25%) + ([redacted]*25%) + ([redacted]*25%) + ([redacted]*25%) = [redacted]

$$RR_{cTTR} \frac{ARR_{ARISTOTLE}}{ARR_{USER-SPECIFIED}}$$

Where: $ARR_{ARISTOTLE}$ is the "average relative rate" assuming an ARISTOTLE cTTR distribution, $ARR_{USER\text{-}SPECIFIED}$ is the "average relative rate" based on the user-specified cTTR distribution and RR_{cTTR} is the relative rate for cTTR adjustment.

4) The treatment specific "relative rate for cTTR adjustment" for each event is then used to adjust the baseline event rates. Table 46 presents the baseline event rates of apixaban and warfarin used in the cTTR subgroup analyses. It is important to note that the baseline event rates for warfarin differ from those used in the manufacturer's main analysis. This is because the main analysis calculated baseline rates for warfarin by applying a hazard ratio (HR, derived from the NMA) to the baseline event rates for apixaban. Whereas, for the purpose of these subgroup analyses the baseline event rates for warfarin were derived directly from the ARISTOTLE trial.

Table 46. Baseline event rates (per 100 PY) used in cTTR subgroup analysis

Event	Treatment	
	Apixaban[a]	Warfarin[b]
Ischaemic stroke	[redacted]	[redacted]
ICH	0.330	0.800
Other major bleeds	1.790	2.270
CRNM bleed	[redacted]	[redacted]
MI	0.530	0.610
Other CV hospitalisation	[redacted]	10.460
Other treatment discontinuations	[redacted]	[redacted]
SE	[redacted]	[redacted]

Abbreviations used in table: CRNM, clinically relevant non-major; cTTR, centre time in therapeutic range; CV, cardiovascular; ICH, intracranial haemorrhage; MI, myocardial infarction; PY, patient years; SE, systemic embolism.
[a]same as main analysis, [b]derived from ARISTOTLE trial.

The results of the manufacturer's subgroup analyses are presented in Section 5.3.15.

5.3.8 Mortality

Within the manufacturer's model, patient mortality was either a consequence of modelled events or other causes, i.e. all-cause mortality minus event-related mortality.

Other-cause mortality

Mortality that was not a consequence of modelled events (i.e. other-cause mortality) was disaggregated across the within-trial and post-trial periods. During the within-trial period (1.2 years, based on the duration of follow-up for other-cause mortality in ARISTOTLE; MS: pg 121), other-cause mortality was assumed to depend on treatment received. Treatment specific estimates of other-cause mortality were calculated by applying HRs to the baseline rate of other-cause mortality observed in the apixaban arm of ARISTOTLE [redacted] A HR of [redacted] was calculated for warfarin from a secondary analysis of ARISTOTLE trial data. However, dabigatran and rivaroxaban were assumed to have the same other-cause mortality profile as apixaban (i.e. HR = 1). The ERG's opinion on and impact of assuming other-cause mortality was treatment specific for the within-trial period is discussed in Section 5.4.3.

During the post-trial period, other-cause mortality was assumed to be equivalent across all considered treatments. Estimates of other-cause mortality for the general population were taken from UK life tables and a Gompertz distribution fitted. The Gompertz distribution was then used to interpolate annual data to 6-week data. The ERG notes that before selecting the Gompertz distribution, the manufacturer assessed a number of other distributions. Distributions were assessed by considering fit, predicted life expectancy and predicted median survival against observed data (Table 47).

Table 47. Summary of the manufacturer's assessments of distributional fit (predicted versus observed) for other-cause mortality (reproduced from MS; Table 129; pg 266)

Distribution	Males			Females		
	Predicted LE (Observed=78.6)	Predicted Median Survival (Observed=81.8)	SSE	Predicted LE (Observed=82.8)	Predicted Median (Observed=85.9)	SSE
Exponential	38.5	26.7	23.69	46.0	31.9	22.30
Weibull	104.5	98.9	5.23	128.1	120.8	7.07
Gompertz	78.1	80.6	0.035	82.1	84.6	0.067
Log-Logistic	87.6	69.9	3.80	113.0	87.3	3.91
Log-Normal	60.9	56.6	7.23	66.4	61.7	6.91
Abbreviations used in table: LE, life expectancy; SSE, sum of squared errors						

Based on the information presented in Table 47, the ERG accepts the manufacturer's choice of distribution for interpolation.

Furthermore, based on a systematic review of the literature (submitted as an additional report in support of this STA) the manufacturer highlighted that the risk of death from other causes is higher in the AF than in the general population.[10][63] Therefore, the manufacturer adjusted the general population other-cause mortality rate using an HR that captured the increased risk of other-cause mortality in an AF population. The HR used was derived from a study by Friberg et al.[10] identified as part of the manufacturer's additional systematic review. The study by Friberg et al. was carried out in Stockholm and considered mortality rates in AF (paroxysmal, persistent or permanent) patients compared with a matched general population. The HR for other-cause mortality in patients with AF used in the manufacturer's model was 1.34; calculated as follows:

1. Annual mortality rates for the general and AF population were extracted from Friberg et al. along with data on cause of death (Table 48);

Table 48. General population and AF mortality data (by subgroup) extracted from Friberg et al.[10]

Population	Annual mortality rate (%)	Number of patients in subgroup	All Cause death (n)	Death due to MI)	Death due to all stroke	Death due to stroke and MI
General population	5	N/A	–	–	–	–
Paroxysmal AF	7	888	111	18	11	29
Persistent AF	3	618	27	6	2	8
Permanent AF	14	1186	291	50	24	74
Abbreviations used in table: AF, atrial fibrillation; MI*, myocardial infarction; N/A, not applicable.						

2. The all-cause mortality rates for each AF subtype were converted into all-cause mortality risks as follows:

$$p = 1 - e^{-rt}$$

Where: p is the annual all-cause mortality risk; r is the annual all-cause mortality rate and t is time in years.

3. The all-cause mortality risk for each AF subtype was adjusted by the percentage of deaths as a result of stroke or MI (i.e. events accounted for elsewhere in the manufacturer's model) observed in that subtype.

$$p_{OC} = p.(1 - p_{MI/stroke})$$

Where: p_{OC} is the other-cause mortality risk; p is the all-cause mortality risk and $p_{MI/stroke}$ is the proportion of patients dying as a result of MI or stroke.

4. The other-cause mortality risk for each AF subtype is then converted back into a rate as follows:

$$r_{OC} = -\ln(1 - p_{OC})$$

Where r_{OC} is the other-cause mortality rate and p_{OC} is the other-cause mortality risk.

Table 49 displays the other-cause mortality rates calculated for each AF subtype.

Table 49. Adjusted other-cause mortality rates derived from data reported in Friberg et al.[10]

Population	Annual mortality risk (%)	% (n/N) of deaths attributed to Stroke and MI	Adjusted risk of death	Adjusted rate of death
Paroxysmal AF	6.76	26.13(29/111)	6.76%*(1-26.13%)= 4.99%	-ln(1-4.99%)= **5.12**%
Persistent AF	2.96	29.63(8/27)	2.96%*(1-29.63%)= 2.08%	-ln(1-2.08%)= **2.10**%
Permanent AF	13.06	25.43(74/291)	13.06%*(1-25.43%)= 9.74%	-ln(1-9.74%)= **10.25**%
Abbreviations used in table: AF, atrial fibrillation.				

5. A weighted average of other-cause mortality rates is then calculated; weighted by the proportion of patients in each AF subtype as follows:

$$R_{OC} = \frac{(5.12\%.888) + (2.10\%.618) + (10.25\%.1186)}{(888 + 618 + 1186)}$$

Where R_{OC} is the average other-cause mortality rate

6. The average other-cause mortality rate for the AF population was compared with the overall mortality rate in the general population (see Table 48) to generate the hazard ratio of 1.34. (6.69%/5.00%)

The ERG notes that the Friberg et al. study was carried out in a Swedish rather than an UK population. However, the ERG accepts the manufacturer's rationale for selecting this study to calculate the relative other-cause mortality in an AF population (Box 7).

Box 7. Rationale for selecting Friberg et al.[10] to inform the relative rate of "other-cause" mortality used in the manufacturer's model (reproduced from the manufacturer's additional report; pg 10)

This study was selected for the following reasons:

- The death rate for a matched general population was provided enabling a comparison to the exact model requirements;
- It provided number of deaths for both stroke and MI, specifically, as well as all-cause death;
- The HR was calculated by adjusting the death rate for AF patients to exclude deaths from MI and stroke and comparing to the general population death rate;
- It has sufficient patient population and follow-up duration.

Event related mortality

Stroke, SE, MI, other ICH and major bleeding (GI or non-GI) were associated with a mortality risk. For the temporary events of other ICH and major bleeding there was an incident risk of death but no increase in long-term mortality risk. However, the permanent events of stroke, SE and MI were associated with an incident risk of death and an increase in long-term mortality risk. Table 50 summarises the event-related incident risks of death used in the manufacturer's model.

Table 50 Event-related incident risks of death used in the manufacturer's model.

Event	Treatment related case fatality					Source
	Apixaban	Dabigatran 150 mg	Dabigatran 110 mg	Rivaroxaban	Warfarin	
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[a]Secondary analysis of ARISTOTLE data [b]Connolly 2009[26] [c]Patel 2009[27]
Haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
SE	[redacted]					ARISTOTLE
Other ICH	[redacted]					Secondary analysis of ARISTOTLE and AVERROES data (assumed to apply to all treatments)
Major bleeding	[redacted]					
MI	10.8% for males and 15.6% for females					Scarborough 2010 (CHD stats)[64]
Abbreviations used in table: CHD, coronary heart disease; ICH, intracranial haemorrhage, mg, milligram; MI, myocardial infarction; SE, systemic embolism.						

The increase in long-term mortality risk associated with permanent events was implemented in the model using HRs applied to UK life table data. Table 51 displays the HRs used for each permanent event.

Table 51. Hazard ratios used to adjust post-event mortality (reproduced from MS; Table 55; pg 122)

Event	HR	Source
Stroke (excluding haemorrhagic stroke)		
Mild	3.18	Bronnum-Hansen et al.,[65] Henriksson et al.,[66] Huybrechts et al.[67]
Moderate	5.84	
Severe	15.75	
Haemorrhagic Stroke		
Mild	3.18	Bronnum-Hansen et al.,[65] Henriksson et al.,[66] Huybrechts et al.[67]
Moderate	5.84	
Severe	15.75	
MI		
Males	2.56	Bronnum-Hansen et al.[68]
Females	4.16	
SE	1.34	Assumption. Same as AF[10]
Abbreviations used in table: AF, atrial fibrillation; HR, hazard ratio; MI, myocardial infarction; SE, systemic embolism.		

It is important to note that the HR applied to patients in the SE health state is equal to the hazard ratio applied to patients in the "NVAF"/"NVAF w/o original AC" health states. That is, the long-term mortality risk of patients who have experienced an SE is the same as at model entry. The implication of this assumption is discussed in Section 5.4.6.

5.3.9 Extrapolation

The manufacturer's model extrapolates the risk of events and treatment discontinuation from the within-trial into the post-trial period. The effect of treatment on the risk of events was assumed to remain constant for the full model time horizon; i.e. the effect of treatment was assumed to be maintained. The manufacturer's rationale for this assumption was the absence of evidence to the contrary. The ERG notes that an absence of evidence is not evidence of an absence, for example of a change in treatment effect. However, the ERG acknowledges that the manufacturer's assumption of constant treatment effect was consistent with assumption used in other NOAC HTA submissions[43][44]

As discussed in Section 5.3.5, discontinuation may be event-related or a result of other causes. Event-related discontinuation was assumed to be the same in the within-trial and post-trial periods. However, discontinuation as a result of other causes was assumed to be treatment specific and treatment independent in the within-trial and post-trial periods, respectively. Furthermore, the manufacturer carried out sensitivity analyses around the extrapolation of other-cause discontinuation in the post-trial period, assuming that rates:

- remained the same as the within trial period;
- were nil for all treatments.

The results of the manufacturer's sensitivity analyses are discussed in Section 5.3.15.

5.3.10 Health-related quality of life

Health-related quality of life data were not collected as part of the ARISTOTLE or AVERROES clinical trials. Therefore, the manufacturer carried out a systematic review to identify health state utility value (HSUV) studies relevant to the health states considered in the model. The focus of the manufacturer's review was HSUV studies that used the EQ-5D method of elicitation in accordance with NICE methods guidance.[69] Standard databases (Medline/Medline (R) In-Process, EMBASE, Econlit and NHS EED) were searched and supplemented by hand searches of conference proceedings, the CEA registry and reference lists of identified economic evaluations and related NICE technology appraisals.

Initially, 24 HSUV studies were identified by the manufacturer's review, 21 of these studies reported EQ-5D data in an AF population. The remaining three "reported EQ-5D values for a variety of chronic conditions after controlling for co-morbidities". However, following the initial search and appraisal utility values had not been identified for all health states. Therefore, the manufacturer relaxed the EQ-5D inclusion criterion and re-appraised originally excluded studies. A further nine studies were identified that reported utility data in an AF population using non-EQ-5D (Time trade off or standard gamble) elicitation methods; 33 HSUV studies in total.

Of the 33 HSUV studies identified in the manufacturer's review, 18 reported utility values applicable to health states included in the model. However, only six of these studies were used to inform the manufacturer's model. Table 52 summarises the sources used to inform the utilities associated with model health states and disutilites associated with temporary events and treatment regimen.

Table 52. Utility and disutility values used in the manufacturer's model (adapted from MS; Table 61; pg 130).

Health state		Utility value	Elicitation method	Source	Rationale for source used
Health states					
"NVAF"/"NVAF w/o original AC"		0.7800	EQ-5D (TTO)	Khan et al. 2004[70]	The only UK based EQ-5D HSUV study identified
Stroke (ischaemic or haemorrhagic)	Mild	0.7600	TTO	Gage et al. 1996[71]	The only HSUV study identified to report utility by mild, moderate and severe stroke severity for as defined by the mRS
	Moderate	0.3900			
	Severe	0.1100			
SE		0.6795	EQ-5D (TTO)	Sullivan et al. 2011[72)]	Only source identified
MI		0.6830	EQ-5D (TTO)	Lacey et al. 2003[73]	
Temporary events*					
Other ICH (applied upon occurrence for a duration of 6 weeks)		-0.1070	Standard gamble	Thomson et al. 2000[74]	Only source identified
Other major bleeds (applied upon occurrence for a duration of 2 weeks)		-0.1070		Thomson et al. 2000[74]	
CRNM bleed (applied upon occurrence for a duration of 2 days)		-0.0582	EQ-5D (TTO)	Sullivan et al. 2011[72]	
Other CV hospitalisation (applied upon occurrence for a duration of 6 days)		-0.0970	EQ-5D (TTO)	Lacey et al. 2003[73]	Assumed to be equivalent to the disutility associated with an MI
Treatment					
Apixaban		-0.0020	TTO	Gage et al. 1996[71]	Only source identified
Aspirin		-0.0020			
Aspirin (2nd line)		-0.0020			
Warfarin		-0.0130			
Dabigatran (110 mg)		-0.0020			
Dabigatran (150 mg)		-0.0020			
Rivaroxaban		-0.0020			

Abbreviations used in table: AC, anticoagulation; CRNM, clinically relevant non-major; CV, cardiovascular; HSUV, health state utility value; ICH, intracranial haemorrhage; MI, myocardial infarction; NVAF, non-valvular atrial fibrillation; SE, systemic embolism; TTO, Time trade off; UK, United Kingdom; w/o, without.
*Duration of temporary events was based on expert opinion.

The ERG is satisfied that the manufacturer carried out a comprehensive review of the literature and that it is unlikely that any relevant studies have been missed. Furthermore, the ERG accepts the sources selected to inform the utility and disutility values of the manufacturer's model. However, the ERG notes that the manufacturer has assumed a disutility associated with treatment. Furthermore, the disutility associated with warfarin treatment is higher than the disutility associated with the NOAC treatment. In addition, the ERG notes that utility is not adjusted for age; i.e. does not change as the model cohort ages. The ERG's opinion on and impact of these assumptions is discussed in Section 5.4.4.

5.3.11 Resources and costs

The manufacturer's model adopted an NHS and PSS perspective and, therefore, wherever possible unit costs were taken from NHS Reference Costs 2010/11.[75] The manufacturer stated that, where available, Healthcare Resource Group (HRG) codes specified in the costing report for AF from NICE CG36[76] were used. However, in instances where "procedures do not have HRG codes or the codes are not sufficiently disaggregated, such as long term care and type and severity of stroke, unit costs have been identified from the published literature" (MS; pg 133). However, the manufacturer highlighted that the published literature used to supplement NHS reference costs was not identified through a systematic review. Instead references used in previous NICE technology appraisals (TA249 and TA256) or identified as part of the manufacturer's review of the cost-effectiveness or quality of life literature were used.

Resource use and costs used in the manufacturer's model can be categorised into:

- intervention costs;
- health state costs;
- adverse event costs.

Where appropriate costs were inflated to 2010/11 prices using the Pay & Prices Index.[77]

Intervention costs

The cost of interventions considered in the manufacturer's model was comprised of drug acquisition and monitoring costs. The acquisition costs of each intervention are presented in Table 53 as costs per day for the given dose.

Table 53. Drug acquisition costs (reproduced from the MS; Table 62; pg 133)

Anticoagulation	Tablet size (mg)	Cost per tablet	Average daily dose (mg)	Average daily cost	Source
Apixaban	5	£1.10	10	£2.20	BMS/Pfizer
Warfarin	0.5, 1.0, 3.0	£0.40	4.5	£0.12	Electronic Drug Tariff‡
Dabigatran 110 mg	110	£1.10	220	£2.20	MIMS†
Dabigatran 150 mg	150	£1.10	300	£2.20	MIMS†
Rivaroxaban	20	£2.10	20	£2.10	MIMS†
Abbreviations used in table: BMS, Bristol-Myers Squibb; mg, milligram. ‡Electronic Drug Tariff, August 2012, Department of Health by the NHS Business Services Authority, NHS Prescription Services, http://www.ppa.org.uk/ppa/edt_intro.htm; †MIMS, August 2012.					

In addition to drug acquisition costs, patients treated with warfarin or dabigatran accrued the cost of INR or renal monitoring, respectively. The cost of INR monitoring calculated by the ERG for TA249[40] was used (inflated to £248.19 per year) and applied to all warfarin patients. The ERG considers that the use of a previously approved INR monitoring cost rather than an INR cost calculated from first principles was a pragmatic and acceptable approach.

The cost of renal monitoring was applied to 19.4% of patients treated with dabigatran (150 mg or 110 mg). The manufacturer cited the trial publication by Connolly et al.[26] as justification for the proportion of dabigatran patients that accrued the cost of renal monitoring. The ERG was unable to validate this figure; however, given the low cost (£3) of renal monitoring it is unlikely that the inclusion of renal monitoring for dabigatran would have had a significant impact on model results. The NHS reference cost of direct access pathology services (code DAP283) was used to inform the cost (£3) of renal monitoring.

Health state costs

The health states captured in the manufacturer's model were categorised as either permanent or temporary according to the events they represented. Permanent health states were associated with acute and long-term costs, whereas, patients who entered temporary health states accrued only acute costs. Tables 54 and 55 summarise the costs associated with each modelled permanent and temporary health states, respectively.

Table 54. Acute and long-term costs associated with permanent health states

Health state		Acute cost (£ per episode)	Reference	Long-term cost (£ per month)	Reference
Ischaemic stroke	Mild	3,515.64	Luengo-Fernandez et al. 2012[78]	183.91	Luengo-Fernandez et al. 2012[78]
	Moderate	18,341.08		358.78	
	Severe	25,050.88		544.76	
	Fatal	3,162.11		N/A	
Haemorrhagic stroke	Mild	10,236.81		183.91	
	Moderate	26,299.60		358.78	
	Severe	44,486.65		544.76	
	Fatal	1,645.66		N/A	
SE		4,077.98		183.91	
MI		2,018.84	Comprised of: • EB10Z: Acute or suspected MI;[75] • Cardiac rehabilitation;[79] • Coronary revasculation.[75]	6.65	A weighted average of the monthly cost of treatment with beta-blocker (atenolol), ACE inhibitor (ramipril) and statin (simvastatin). Costs taken from EDR.[80] Weights taken from prescribing data[81]
Abbreviations used in table: ACE, Angiotensin-converting enzyme; EDR, Electronic Drug Tariff; MI, myocardial infarction; N/A, Not applicable; SE, systemic embolism.					

With the exception of MI, the acute and long-term costs associated with each permanent health state (stroke or SE) were taken from a study by Luengo-Fernandez et al.[78] The study by Luengo-Fernandez et al. was a population-based assessment of the acute and long-term costs associated with stroke in patients with atrial fibrillation. Within the study, costs are broken down by stroke type (ischaemic, haemorrhagic or unknown) and/or severity (mild, moderate, severe and fatal). However, the ERG notes that the manufacturer used the costs (acute and long-term) associated with all mild stroke to inform the costs of SE. The ERG considers this approach to be reasonable for long-term costs. However, the ERG notes that the acute cost of SE was approximately double the acute costs used in the rivaroxaban and dabigatran submissions.[43][44] The impact of the higher acute cost of SE on the cost-effectiveness results is discussed in Section 5.4.7.

The acute cost of MI was comprised of:

- the average cost associated with HRG code EB10Z (acute or suspected MI);[75]
- the total cost per patient of cardiac rehabilitation;[79]
- the average cost associated with coronary revascularization assessment (service code 320 Cardiology).[75]

Based on NICE's costing template from CG48 (MI secondary prevention)[82] the cost of cardiac rehabilitation and coronary revascularisation assessment were applied to 56% and 78% of patients,

respectively. The ERG considers that the elements included in the calculation of acute MI were reasonable.

The long-term cost of MI was assumed to be purely associated with pharmacological costs. A weighted average of the cost of treatment with a beta-blocker (atenolol), an ACE inhibitor (ramipril) and a statin (simvastatin) was calculated. Drug costs were taken from Electronic Drug Tariff[80] and weighted by the proportion of patients receiving each available dose; weights were taken from national prescribing data.[81] Based on expert clinical opinion, the ERG considers that the long-term costs of MI were acceptable; clinical experts fed back that patients would be expected to recover well from an MI and not require costly long-term care.

Table 55. Acute costs associated with temporary health states

Health state	Acute cost (£ per episode)	Reference Description
Other ICH	3,010.00	Weighted average of the following HRG codes: • AA23A: Haemorrhagic Cerebrovascular Disorders with CC; • AA23B: Haemorrhagic Cerebrovascular Disorders without CC.
GI bleeds	1,493.68	Weighted average of the following HRG codes: • FZ38D: GI Bleed with length of stay 2 days or more with Major CC; • FZ38E: GI Bleed with length of stay 2 days or more without Major CC; • FZ38F: GI Bleed with length of stay 1 day or less.
Other major bleeds	3,947.92	Weighted average of the following HRG codes: • HC28B: Spinal Cord Conditions with CC; • HC28C: Spinal Cord Conditions without CC; • HD24A: Non-Inflammatory Bone or Joint Disorders with Major CC; • BZ24A: Non-Surgical Ophthalmology with length of stay 2 days or more; • PA23A: Cardiac Conditions with CC; • FZ12D: General Abdominal – Very Major or Major Procedures 19 years and over with Major CC; • FZ12F: General Abdominal – Very Major or Major Procedures 19 years and over without CC.
CRNM bleed	1,133.93	Weighted average of the following HRG codes: • FZ38F: GI Bleed with length of stay 1 day or less; • CZ13Y: Intermediate Nose Procedures 19 years and over without CC; • LB38B: Unspecified Haematuria without Major CC.
CV hospitalisations	1,570.89	Weighted average of HRG codes: AA29A, AA29B, PA22Z, QZ20Z, EB03H, EB03I, QZ17A, QZ17B, QZ17C and EB01Z.
Abbreviations used in table: CC, complication and co-morbidity; CRNM, clinically relevant non-major; CV, cardiovascular; GI, gastrointestinal; ICH, intracranial haemorrhage.		

The acute cost of each temporary event was calculated as a weighted average of appropriate HRG codes;[75] the average cost associated with each relevant code was weighted by the level of activity reported for that code. The ERG notes that the cost of temporary events in the manufacturer's submission differs to those used in the dabigatran and rivaroxaban submissions; in particular, the cost

associated with CRNM bleeds (apixaban submission, £1,133.93; dabigatran submission, £84; rivaroxaban submission, £126).[43][44] However, the ERG considers that the codes included in the calculation of each temporary event were reasonable and accepts the costs used in the manufacturer's submission.

Adverse event costs

Dyspepsia was the only adverse event that was not explicitly modelled as a health state (permanent or temporary). An additional cost of dyspepsia management was applied to all patients who experienced this adverse event. Table 56 summarises the proportion of patients assumed to experience dyspepsia for each considered intervention.

Table 56. Treatment specific proportions of patients experiencing dyspepsia

Treatment	Proportion of patients assumed to experience dyspepsia (%)	Source
Apixaban	[redacted]	ARISTOTLE
Warfarin	[redacted]	ARISTOTLE
Dabigatran (110 mg)	3.69	RE-LY
Dabigatran (150 mg)	3.53	RE-LY
Rivaroxaban	1.67	Assumption*
*[redacted].		

The yearly cost of dyspepsia management was assumed to be £27.60 and was comprised of:

- endoscopy costs – £6.12 based on a cost of £612 (HRG code FZ42Z]) applied to 1% of patients (NICE CG17);[83]
- GP visits – £1.80 based on a cost of £36 (personal social services research unit costs)[77] applied to 5% of patients (NICE CG17);[83]
- the weighted average cost of treatment with omeprazole and lansoprazole (omeprazole and lansoprazole accounted for 95.7% of all proton pump inhibitors prescribed);[81] costs were taken from Electronic Drug Tariff[80] and weighted by the proportion of patients receiving each available dose; weights were taken from national prescribing data.[81]

5.3.12 Model validation and face validity check

Within the MS, the manufacturer stated that the model was assessed for internal (verification) and external (validation) validity. Verification was carried out by two independent economists and used extreme value analysis to identify any flawed algorithms or irregularities. Validation was carried out by assessment of the face validity of the model with clinicians and comparison of the model results against published results.

Model results were compared with the cost-effectiveness results reported in the dabigatran and rivaroxaban submissions.[43][44] A higher ICER (£13,648 versus £6,264) was estimated for dabigatran versus warfarin, whereas a lower ICER (£14,071 versus £18,883) was estimated for rivaroxaban

versus warfarin. The manufacturer highlighted that the difference in model results was likely to be because of differences in key parameters such as costs associated with stroke events, INR monitoring costs (between rivaroxaban and apixaban submissions) and HRQoL. Furthermore, the manufacturer stated that the parameters used to inform the economic analysis of apixaban were more robust because "The present analysis assumes conservative estimates of the costs of INR monitoring. HRQL inputs were identified following a full systematic review. Relative treatment effects were identified following a comprehensive systematic review and network meta-analysis".

The robustness of assumptions around event costs and HRQoL are discussed in Sections 5.4.7 and 5.4.4, respectively. However, based on the cost of INR monitoring approved by NICE in TA256, the ERG agrees with the manufacturer that the cost of INR monitoring used in this submission was conservative.

5.3.13 Cost-effectiveness results

The manufacturer presented the base case model results as a fully incremental analysis between all considered interventions. The results were generated deterministically rather than probabilistically; i.e. mean values rather than distributions were used for each parameter. Therefore, the ERG assessed the manufacturer's deterministic results against those estimated by the manufacturer's probabilistic sensitivity analysis (PSA, see Section 5.3.14 for full details of the PSA). The costs, life years gained (LYG) and quality adjusted life years (QALYs) gained estimated from the deterministic analysis were compared with the average costs, LYG and QALYs gained estimated from the PSA. Table 57 displays the manufacturer's deterministic and probabilistic incremental results (VKA suitable population).

Table 57. Base case incremental results – VKA suitable population (adapted from MS; Table 79; pg 146)

Treatment	Total			Incremental†			ICER (£/QALY)	
	Costs (£)	LYG	QALY	Costs (£)	LYG	QALY	Versus warfarin	Incremental
Deterministic								
Warfarin	7,188	7.469	5.696	–	–	–	–	–
Dabigatran (150/110 mg)	8,437	7.537	5.788	1,248	0.068	0.091	13,648	Extendedly dominated
Dabigatran (110 mg)	8,684	7.503	5.756	247	-0.034	-0.032	25,308	Strictly Dominated
Rivaroxaban	8,778	7.553	5.809	95	0.050	0.054	14,071	Extendedly dominated
Apixaban	8,983	7.614	5.860	205	0.06	0.05	11,008	11,008
Probabilistic								
Warfarin	5,331	6.869	5.303	–	–	–	–	–
Dabigatran (150/110 mg)	6,737	6.921	5.342	1,406	0.05	0.04	36,450	Extendedly dominated
Dabigatran (110 mg)	6,832	6.899	5.321	95	-0.02	-0.02	83,628	Strictly Dominated
Rivaroxaban	7,070	6.943	5.366	237	0.04	0.05	27,565	Extendedly dominated
Apixaban	7,228	7.002	5.416	159	0.06	0.05	16,852	16,852

Abbreviations used in table: LYG, life year gained; mg, milligram; QALY, quality adjusted life year; VKA, vitamin K antagonist.
†Versus the next least costly technology.

In both the deterministic and probabilistic incremental results, dabigatran 110 mg is strictly dominated by (i.e. was less costly and less effective than) dabigatran blend (150 mg/110 mg) and may therefore be excluded from the analysis. Furthermore, apixaban extendedly dominated (i.e. resulted in a lower ICER versus warfarin despite having higher total costs) rivaroxaban and dabigatran blend. Apixaban had an ICER versus warfarin of £11,008 and £16,852 in the deterministic and probabilistic incremental analyses, respectively.

Tables 58 and 59 summarise the QALYs and costs gained for each treatment disaggregated by health state. In addition, Table 60 presents model outcomes compared with the clinical results of ARISTOTLE.

Table 58. QALYs gained by health state

Health state	Total QALYs				
	Apixaban	**Warfarin**	**Rivaroxaban**	**Dabigatran (150/110 mg)**	**Dabigatran (110 mg)**
NVAF	5.458	5.282	5.388	5.363	5.317
Ischaemic stroke					
Mild	0.151	0.143	0.151	0.136	0.146
Moderate	0.040	0.045	0.041	0.045	0.049
Severe	0.002	0.002	0.002	0.002	0.002
Recurrent stroke*	0.015	0.014	0.015	0.014	0.015
Haemorrhagic stroke					
Mild	0.010	0.015	0.023	0.007	0.008
Moderate	0.005	0.005	0.004	0.002	0.003
Severe	0.000	0.001	0.000	0.000	0.000
Recurrent stroke*	0.000	0.000	0.001	0.000	0.000
Other events					
SE	0.059	0.061	0.060	0.063	0.063
MI	0.122	0.129	0.127	0.157	0.154
Other temporary events**	-0.001	-0.002	-0.002	-0.002	-0.001
Total QALYs	**5.860**	**5.696**	**5.809**	**5.788**	**5.756**

Abbreviations used in table: CRNM, clinically relevant non-major; CV, cardiovascular; ICH, intracranial haemorrhage; mg, milligram; MI, myocardial infarction; NVAF, non-valvular atrial fibrillation; SE, systemic embolism; QALY, quality adjusted life year.

*Mild, moderate and severe recurrent stroke health states have been aggregated as QALY numbers were small and recurrent stroke is independent of treatment.

**Other temporary events includes (Other ICH, Other major bleeds, CRNM bleeds and CV hospitalisation).

Table 59. Costs accrued by health state

Health state	Total costs (£)				
	Apixaban	**Warfarin**	**Rivaroxaban**	**Dabigatran (150/110 mg)**	**Dabigatran (110 mg)**
Anticoagulation	3,347	252	2,891	2,657	2,716
Routine care	0	0	0	0	0
Monitoring	72	977	80	90	88
Ischaemic stroke					
Mild	651	620	652	594	631
Moderate	1341	1487	1394	1522	1627
Severe	605	648	618	683	722
Fatal	61	58	74	75	78
Recurrent stroke*	195	201	200	194	210

Haemorrhagic stroke					
Mild	63	98	144	44	49
Moderate	195	175	143	94	113
Severe	143	230	128	115	122
Fatal	12	25	11	8	8
Recurrent stroke*	19	24	28	10	13
Other events					
SE	251	263	257	271	268
MI	114	120	119	142	140
Other temporary events**	1,912	2,013	2,040	1,929	1,891
Total costs	**8,983**	**7,188**	**8,778**	**8,437**	**8,684**

Abbreviations used in table: CRNM, clinically relevant non-major; CV, cardiovascular; ICH, intracranial haemorrhage; mg, milligram; MI, myocardial infarction; NVAF, non-valvular atrial fibrillation; SE, systemic embolism; QALY, quality adjusted life year.

*Mild, moderate and severe recurrent stroke health states have been aggregated as QALY numbers were small and recurrent stroke is independent of treatment.

**Other temporary events includes (Other ICH, Other major bleeds, CRNM bleeds, CV hospitalisation and dyspepsia).

Table 60. Model outcomes compared with the clinical results of ARISTOTLE (reproduced from MS; Table 77; pg 145)

Outcome	**ARISTOTLE Events**			**Model events***		
	Apixaban (N=9,120)	**Warfarin (N=9,081)**	**Incremental events on Warfarin**	**Apixaban (N=9,120)**	**Warfarin (N=9,081)**	**Incremental events on Warfarin**
Primary outcome: stroke or SE	212	265	53	260§	307§	47
Stroke	199	250	51	240§	286§	46
Ischaemic or uncertain type	162	175	13	199	207	8
Haemorrhagic	40	78	38	41	79	38
SE	15	17	2	20	21	1
Death – any cause	603	669	66	593	665	72

Abbreviations used in table: SE, systemic embolism.

* Approximation estimated at 1.84 years using patient characteristics from ARISTOTLE.

The number of events predicted by the manufacturer's model was higher than that observed in ARISTOTLE. However, the relative difference between events modelled for apixaban and warfarin was lower than that observed in ARISTOTLE. Therefore, in the comparison of apixaban with warfarin, the manufacturer's model may be considered as conservative (biased against apixaban).

5.3.14 Sensitivity analysis

In support of the apixaban submission, the manufacturer carried out PSAs, deterministic sensitivity analyses and scenario analyses. This section summarises the methods and results of each analysis.

Probabilistic sensitivity analyses

The sensitivity of the model to parameter uncertainty while not accounted for in the base case model results (discussed in Section 5.3.13) has been explored in probabilistic sensitivity analyses. Parameters were assigned a probability distribution from which estimates were simultaneously sampled for 2,000 runs. Table 61 summarises the type of distribution used for each group of parameters considered within the sensitivity analyses; Figures 5 and 6 and Table 62 present the probabilistic results.

Table 61. Probability distributions used for model parameters

Parameter type	Distribution(s) used	Manufacturer's rationale
Probabilities	Beta	Probabilities that were based on the proportion of observed outcomes (i.e. probability of event is 1-probability of non-event) could be assumed to follow a binomial distribution. Therefore the beta distribution was used as it is the conjugate of the binomial distribution and is bounded by 0 and 1.
	Dirichlet	For probabilities that described the distribution of patients across different categories (e.g. stroke severity) a Dirichlet distribution was applied, as the Dirichlet distribution is a generalisation of the beta distribution for multiple events. The Dirichlet distribution was applied using the normalised sum of independent gamma or normal variable as described in Briggs et al. 2003.[84]
Costs	Gamma or lognormal	Either the Gamma or lognormal distribution was considered suitable as "both distributions can be highly skewed to reflect the natural skew in costs". The distribution chosen for each individual cost parameter was dependant on the ability of that distribution to reproduce the inputted 95% confidence interval or standard error. Note: where 95% confidence intervals or standard errors were not available from the literature a 25% level of variation was assumed.
HRs	Lognormal or Gamma	The lognormal or gamma distribution was used depending on the ability of the distribution to replicate the "real-world" confidence intervals (with the exception of the relative risks of death which assumed a 25% level of variation).
	Uniform	The uncertainty around the HR of other-cause mortality for AF patients was assumed to follow a uniform distribution. The manufacturer's rationale for assuming a uniform distribution was that "The paper calculates the increase risk in death for AF patients including factors like stroke. The model already incorporates the increased risk in mortality due to strokes and bleeds therefore this risk was used as the absolute upper bound."
Utilities	Beta	The beta distribution was chosen based on the (0,1) boundary imposed by this distribution.
Abbreviations used in table: AF, atrial fibrillation; HR, hazard ratio.		

Generally, the ERG accepts the distributions and rationale provided for the choice of distribution used for each group of parameters. However, the ERG notes that the manufacturer's rationale for choosing the uniform distribution to sample HRs for the risk of other-cause mortality is unclear. The manufacturer stated that the uniform distribution was chosen to reflect the fact that the increased risk of death included death from modelled events such as stroke. However, the ERG notes that the manufacturer calculated an HR that accounted for the increased risk of death over and above that of modelled events (see Section 5.3.8). Therefore, the ERG considers that the HR for other-cause

mortality should be sampled from a similar distribution to those used for other HRs in the model (lognormal or Gamma). The ERG carried out an exploratory analysis in which the Gamma distribution, with an assumed variation of 20% was used to sample the other-cause mortality HR. The probabilistic incremental results of this exploratory analysis are presented in Appendix 9.7. The ERG notes that the probabilistic results of the ERG's exploratory analysis are more representative of the manufacturer's deterministic incremental results.

Figure 5. Scatter plot results of probabilistic sensitivity analysis of apixaban versus (running clockwise from top left) warfarin, rivaroxaban, dabigatran blend (150 mg/110 mg) and dabigatran 110 mg in the VKA suitable population

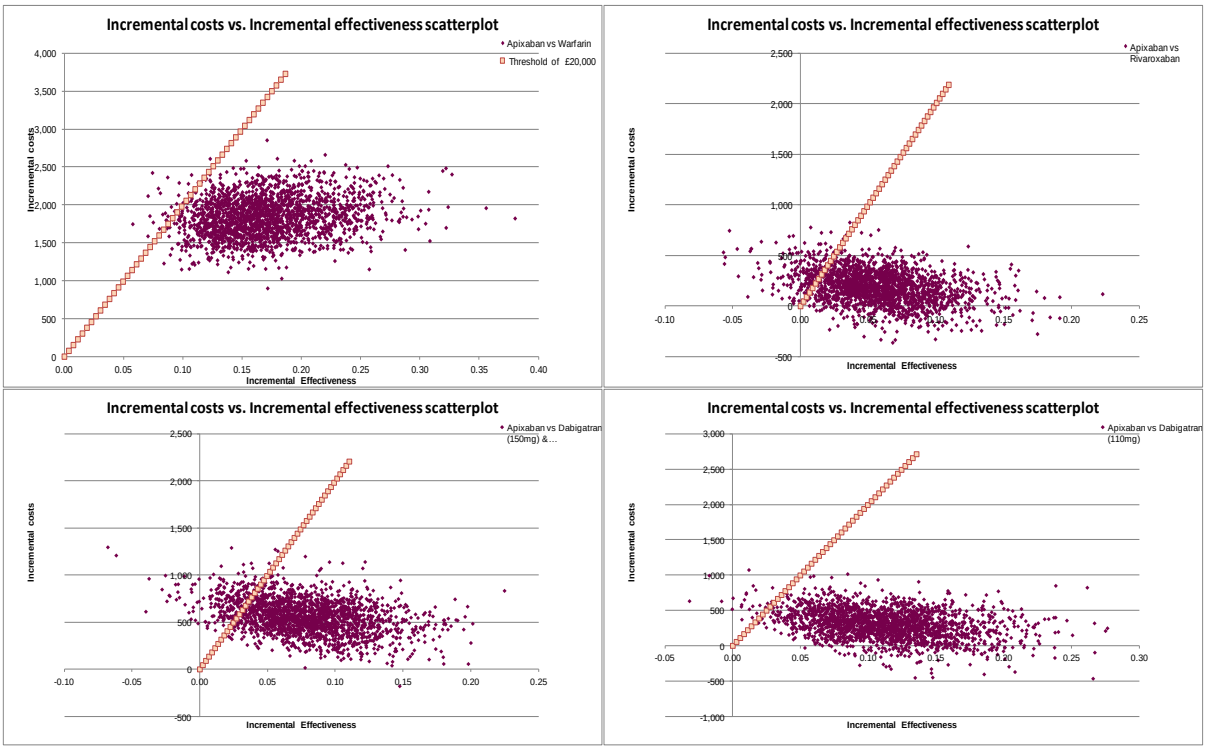


Figure 6. Cost-effectiveness acceptability curves in the VKA suitable population (reproduced from MS; Figure 33; pg 157)

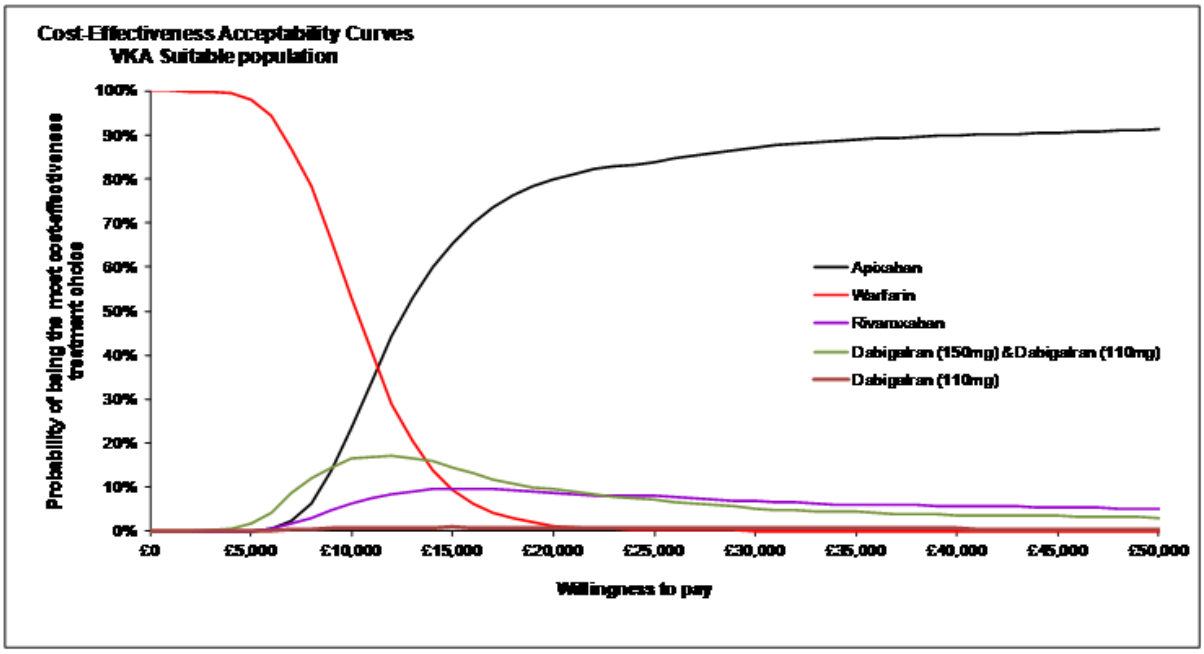


Table 62. Probability of cost-effectiveness at different willingness-to-pay thresholds – VKA suitable population (adapted from MS; Table 81; pg 157)

Intervention	£20,000	£30,000
Apixaban	80%	87%
Dabigatran (110 mg & 150 mg)	10%	5%
Rivaroxaban	9%	7%
Warfarin	1%	0%
Dabigatran (110 mg)	1%	1%
Abbreviations used in table: mg, milligram.		

Deterministic analysis

Within the deterministic sensitivity analyses, the manufacturer assessed the univariate sensitivity of the model to a total of 117 parameters. Each parameter was alternately assigned a low and high value estimated from the 95% confidence intervals associated with that parameter; where confidence intervals were not available or could not be derived, variation was assumed to be either 10% or 25% of the mean. Figures 7 to 10 present the deterministic sensitivity analysis results for the 13 most influential parameters in each comparison. In each comparison, the HRs (comparator vs apixaban) of ischaemic stroke, ICH and other-cause mortality have an influential effect on the ICER. Similarly, the ICER of apixaban versus warfarin is affected by the:

- disutility associated with warfarin use;
- cost of an INR monitoring visit;
- discount rate applied to QALYs.

Furthermore, the ICERs of apixaban versus dabigatran or rivaroxaban are affected by:

- the absolute stroke risk for apixaban;
- second line stroke risk for aspirin.

However, all of the ICERs calculated in the manufacturer's deterministic sensitivity analyses were below £20,000. This suggests that within the confines of the model structure the manufacturer's cost-effectiveness results were robust to parameter variation.

Figure 7. Tornado diagram of top 13 deterministic sensitivity analysis results for apixaban versus warfarin – VKA suitable (reproduced from MS; Figure 17; pg 148)

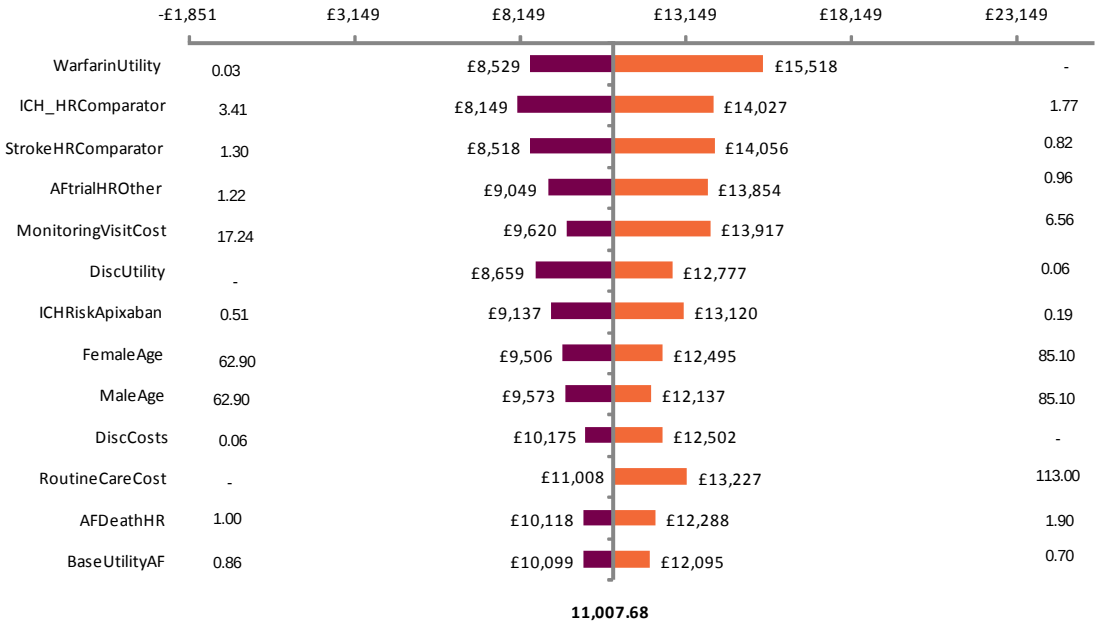


Figure 8. Tornado diagram of top 13 deterministic sensitivity analysis results for apixaban versus dabigatran (150/110 mg) – VKA suitable (reproduced from MS; Figure 19; pg 150)

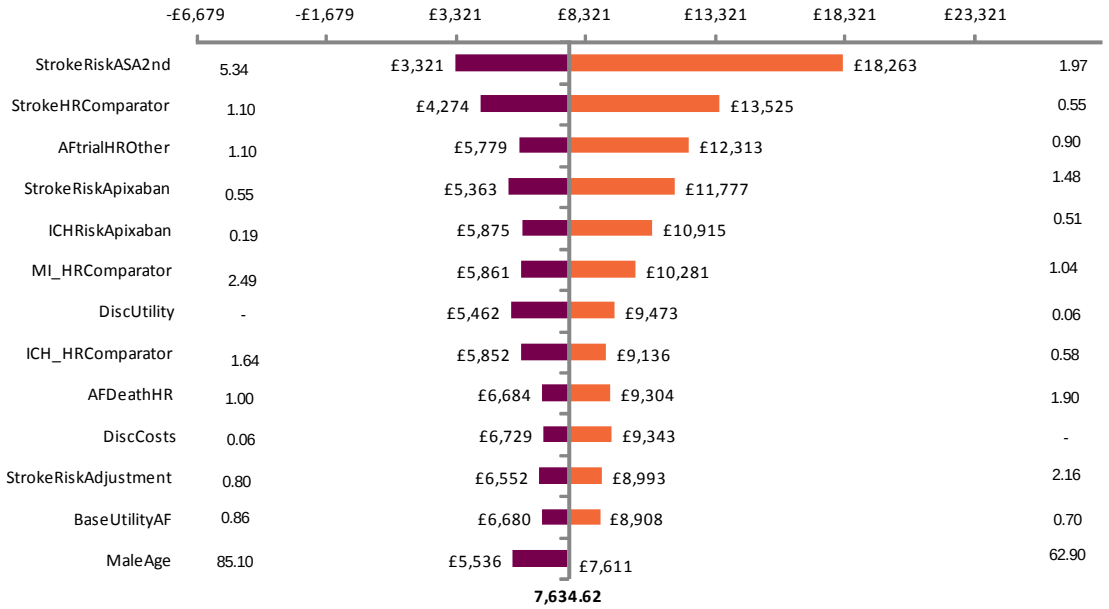


Figure 9. Tornado diagram of top 13 deterministic sensitivity analysis results for apixaban versus dabigatran (110 mg) – VKA suitable (reproduced from MS; Figure 20; pg 150)

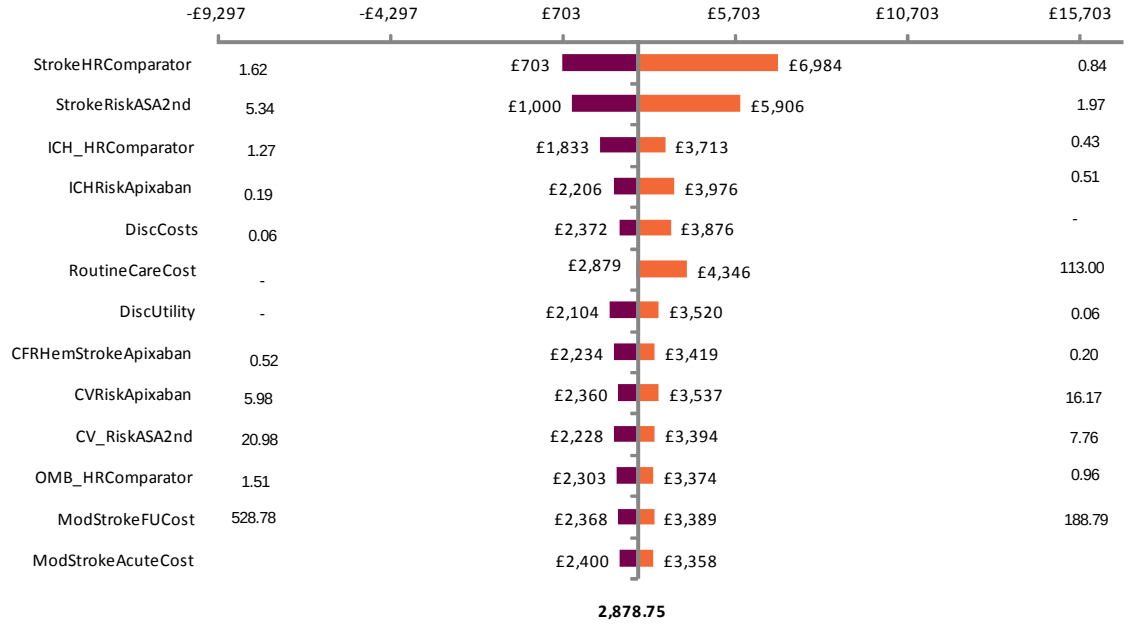
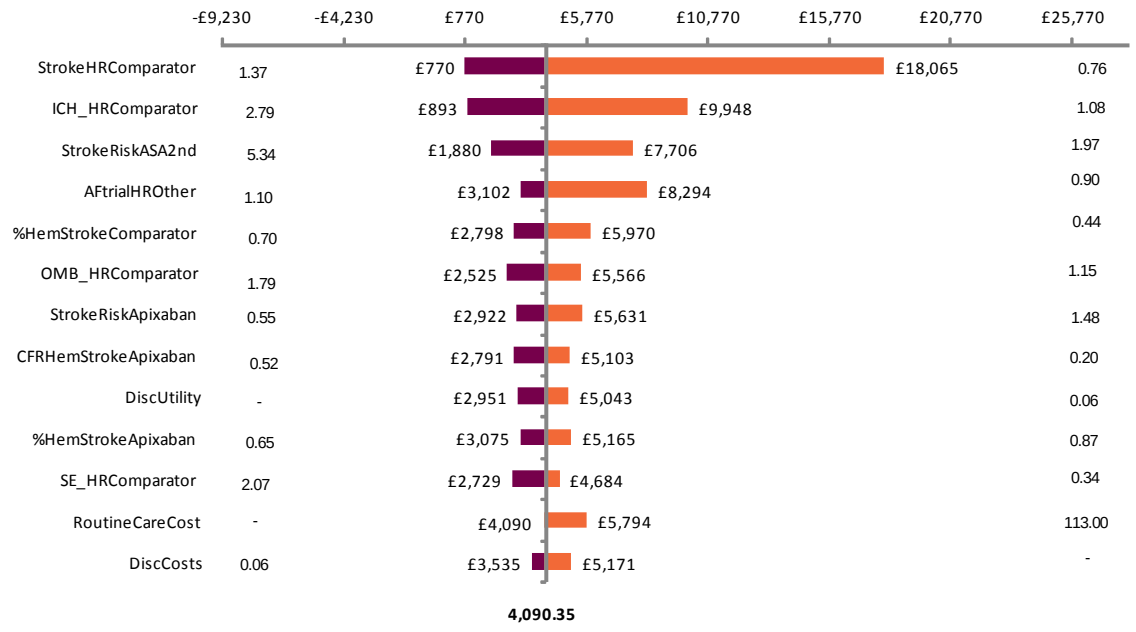


Figure 10. Tornado diagram of top 13 deterministic sensitivity analysis results for apixaban versus rivaroxaban – VKA suitable (reproduced from MS; Figure 21; pg 151)



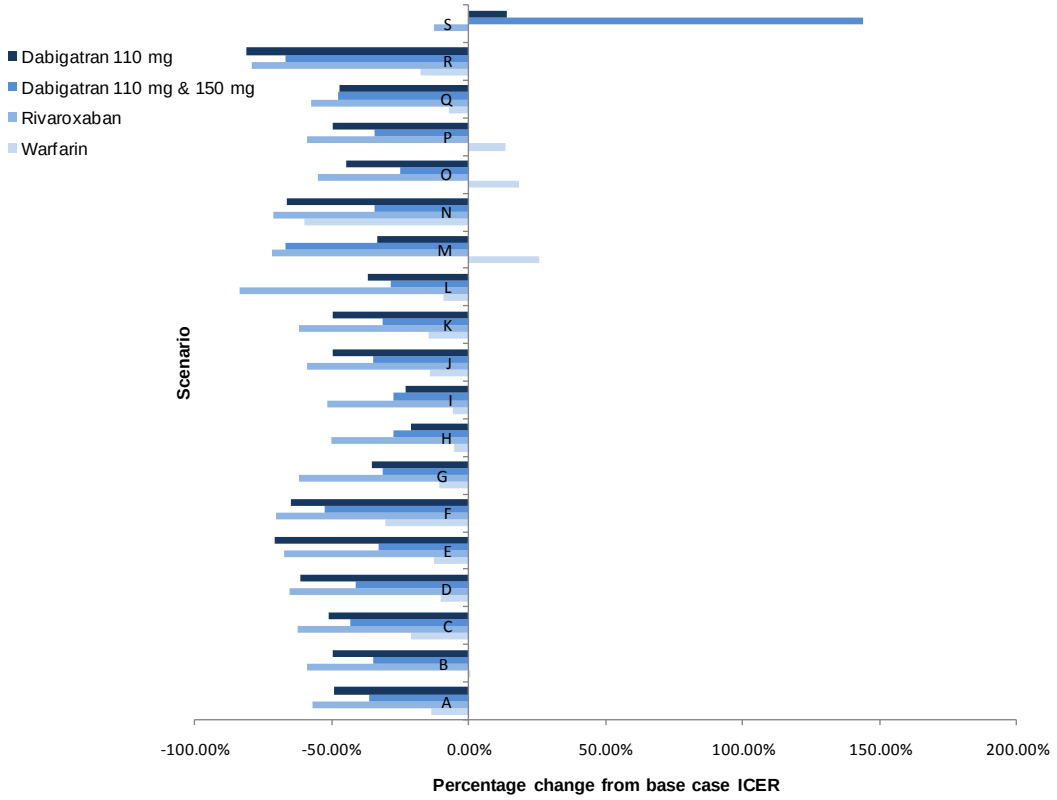
Scenario analysis

In total, the manufacturer carried out 19 scenario analyses (in the VKA suitable population) around various model assumptions, Table 63 gives the details of each scenario analysis carried out. The results of each analysis with respect to impact on the base case ICER are displayed in Figure 11; impact is assessed by percentage change from the base case ICER (for apixaban versus each comparator). The majority of scenario analyses decrease the base case ICER (apixaban versus comparator), suggesting that the assumptions of the base case model were conservative (i.e. any bias was likely to be against apixaban).

Table 63. Scenario analysis carried out by the manufacturer (adapted from MS; Table 76; pg 141)

Scenario label	Scenario description
A	Recurrent stroke (ischaemic and haemorrhagic) switched off
B	Trial specific other-cause mortality switched off (adjusted UK life-table data used for full model time horizon)
C	Long term mortality based on general public (AF correction switched off, HR=1)
D	Other-cause discontinuation set equal to apixaban for all comparators
E	Discontinuation rates set the same as apixaban
F	Discount costs and benefits at 6% and 1.5%, respectively
G	Set mild, moderate and severe acute stroke costs (ischaemic and haemorrhagic), SE cost and long-term maintenance costs equal to Youman et al. (2002) (inflated to 2010/11 costs)*
H	Set mild, moderate and severe acute stroke costs (ischaemic and haemorrhagic) and SE cost equal to NHS reference cost of stroke (estimated as £2,952 based on weighted average of AA04A, AA04B, A10A, AA10B, AA16A, AA16B,AA22A, AA22B, AA23A, AA23B; cost of fatal stroke cost=£0)
I	Set mild, moderate and severe acute stroke costs (ischaemic and haemorrhagic) and SE cost equal to PBR Tariff costs of stroke (estimated as £4,231 based on weighted average of AA04Z, AA10Z, AA16Z, AA22Z, AA23Z, cost of fatal stroke=£0)
J	Reduce health state utility decrements for Other ICH, Other Major Bleeds and CRNM bleeds by 25%
K	Reduce utility values for ischaemic and haemorrhagic stroke and SE health states by 25%
L	Assume same (apixaban) stroke severity distribution for all interventions (mild, moderate, severe and fatal)
M	Age = 80, risks calculated using cTTR specific data, 100% of patients have cTTR >76.51%, all drugs have same stroke severity distribution, trial mortality off, no cost for fatal strokes, NHS reference costs used for stroke and SE, utility decrements associated with bleeding reduced by 25%
N	Age = 70, risks calculated using cTTR specific data, 100% of patients have cTTR <52.38%, costs of stroke and SE inflated by 15%
O	Apply warfarin disutility of 0.013 to all NOACs
P	Apply disutility of 0.0 to all anticoagulants
Q	Gallagher et al. (2008)[52] baseline characteristics
R	Treatment Choice Post Other ICH/Other Major Bleeds – No treatment
S	Treatment Choice Post Other ICH/Other Major Bleeds – Warfarin
Abbreviations used in table: AF, atrial fibrillation; cTTR, centre time in therapeutic range; CRNM, clinically relevant non-major; HR, hazard ratio; ICH, intracranial haemorrhage; NOAC, novel oral anticoagulant; NHS, national health service; PBR, payment by results; SE, systemic embolism; UK, united kingdom. *Reference not provided.	

Figure 11. Impact of scenario analysis on base ICER (apixaban versus comparator)



5.3.15 Subgroup analyses

As discussed in Sections 5.3.2 and 5.3.7, the manufacturer carried out subgroup analyses by level of INR control and by categories of baseline $CHADS_2$ scores.

INR control

Subgroup analysis by level of INR control was limited to the comparison of apixaban with warfarin and based on data from the ARISTOTLE trial (full details of effectiveness parameters used are given in Section 5.3.7). Trial centres were grouped by median level of TTR (cTTR) and subgroup analysis carried out for the quartiles (i.e. 100% of patients were assumed to be in one quartile per analysis):

- cTTR <52.38%;
- 52.38% ≤ cTTR < 66.02%;
- 66.02% ≤ cTTR < 76.51%;
- cTTR ≥76.51%.

Results of the subgroup analyses by level of INR control are summarised in Table 64.

Table 64. Cost-effectiveness results by cTTR subgroup

cTTR subgroup	Technologies	Total			Incremental†			ICER (£/QALY)
		Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
<52.38%	Warfarin	7,508	7.38	5.62	–	–	–	–
	Apixaban	8,895	7.60	5.85	1,387	0.22	0.23	6,077
≥52.38% and <66.02%	Warfarin	7,202	7.47	5.69	–	–	–	–
	Apixaban	9,156	7.55	5.80	1,954	0.08	0.11	18,102
≥66.02% and <76.51%	Warfarin	7,107	7.49	5.72	–	–	–	–
	Apixaban	9,003	7.63	5.87	1,896	0.14	0.15	12,286
≥76.51%	Warfarin	7,037	7.51	5.73	–	–	–	–
	Apixaban	8,875	7.68	5.92	1,838	0.17	0.19	9,889

Abbreviations used in table: cTTR (centre time in therapeutic range; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality adjusted life years.

$CHADS_2$ score

The manufacturer carried out fully incremental subgroup analysis by $CHADS_2$ score categories (1, 2, 3–6), the results of which are presented in Tables 65 to 67.

Table 65. Incremental cost-effectiveness results for $CHADS_2$ score of 1 (adapted from MS; Table 89; pg 164)

Technologies	Total			Incremental†			ICER (£) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	6,930	7.536	5.756	–	–	–	–	–
Dabigatran 150 mg & 110 mg	8,297	7.593	5.835	1,367	0.056	0.079	17,233	Extendedly dominated
Dabigatran 110 mg	8,450	7.579	5.822	153	-0.013	-0.013	23,068	Strictly Dominated
Rivaroxaban	8,596	7.623	5.869	146	0.044	0.047	14,794	Extendedly dominated
Apixaban	8,745	7.685	5.921	149	0.06	0.05	11,010	11,010

Abbreviations used in table: ICER, incremental cost-effectiveness results; LYG, life years gaineg; mg, milligram; QALYs, quality adjusted life years.

Table 66. Incremental cost-effectiveness results for $CHADS_2$ score of 2 (adapted from MS; Table 91; pg 164)

Technologies	Total			Incremental†			ICER (£) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7,184	7.470	5.697	–	–	–	–	–
Dabigatran 150 mg & 110 mg	8,434	7.538	5.789	1,250	0.068	0.091	13,697	Extendedly dominated
Dabigatran 110 mg	8,680	7.504	5.757	246	-0.034	-0.032	25,269	Strictly Dominated
Rivaroxaban	8,776	7.554	5.810	95	0.050	0.054	14,083	Extendedly dominated
Apixaban	8,979	7.615	5.860	204	0.06	0.05	11,008	11,008
Abbreviations used in table: ICER, incremental cost-effectiveness results; LYG, life years gaineg; mg, milligram; QALYs, quality adjusted life years.								

Table 67. Incremental cost-effectiveness results for $CHADS_2$ score of 3-6 (adapted from MS; Table 93; pg 165)

Technologies	Total			Incremental†			ICER (£) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	£7,517	7.383	5.621	–	–	–	–	–
Dabigatran 150 mg & 110 mg	£8,615	7.465	5.727	£1,098	0.083	0.106	£10,321	–
Dabigatran 110 mg	£8,981	7.406	5.671	£366	-0.059	-0.056	£29,042	Strictly Dominated
Rivaroxaban	£9,011	7.464	5.734	£29	0.058	0.063	£13,178	Extendedly dominated
Apixaban	£9,286	7.524	5.781	£275	0.06	0.05	£10,998	£10,998
Abbreviations used in table: ICER, incremental cost-effectiveness results; LYG, life years gaineg; mg, milligram; QALYs, quality adjusted life years.								

The cost-effectiveness results of the manufacturer's subgroup analyses by $CHADS_2$ score categories do not vary substantially from the base case results. However, the results of subgroup analyses by level of INR control, while remaining below £20,000 per QALY gained, showed more variation from the manufacturer's base case. In addition, as highlighted by the manufacturer the results of cTTR subgroup analyses were "surprising". (MS; pg 165) That is, the ICER of apixaban versus warfarin did not increase as expected with increasing level of INR control. By contrast, the ICER for patients treated at centres with median cTTR of at least 76.51% was lower than the base case ICER. The manufacturer highlighted that this result was driven by the "the lower number of ischaemic and haemorrhagic strokes experienced by patients on both medications, resulting in a better incremental QALY gain for patients on apixaban compared with warfarin". (MS; pg 165) The ERG considers that

subgroup analyses results suggest that other factors may be more influential in the manufacturer's analyses than the level of INR control (e.g. the ability of centres to manage AF, different patient risk profiles, socioeconomic factors, etc.). Therefore, the ERG considers that the manufacturer's subgroup analyses were not an accurate or reliable reflection of the cost-effectiveness of apixaban versus warfarin by level of INR control.

5.4 Critique of manufacturer's de novo economic evaluation

The manufacturer provided an economic model constructed in Microsoft© EXCEL© with Visual Basic for Applications, along with a written submission of the economic evaluation. In addition, following the clarification requests of the ERG, the manufacturer provided a revised model which incorporated:

- age adjustment of utility;
- a scenario analysis allowing patients who have experienced an SE to be at risk of further ischaemic stroke events;
- a time horizon truncated when patients reach 100 years of age (26 years).

The ERG considers that the manufacturer's models were well constructed, transparent, accurate and easy to navigate. However, the ERG notes that some of the formulae were more complex than was perhaps necessary and the probabilistic analysis took almost 3 hours to run.

The following sections outline the key points of the ERG's critique of the manufacturer's submission. Points of critique are grouped by theme (e.g. relating to model structure or treatment effectiveness) and are ordered by potential impact on the manufacturer's cost-effectiveness results. In addition, an ERG revised base case is presented in Section 6, along with the ERG's rationale for which elements of the critique are included in the revised base case.

5.4.1 NICE reference case checklist (TABLE ONLY)

Tables 68 and 69 summarise the ERG's assessment of the manufacturer's economic evaluation against the NICE reference case and Philips checklists, respectively. The ERG notes that the manufacturer's base case economic evaluation satisfied many of the requirements set out in the 'Guide to the Methods of Technology Appraisal'.[69] and criteria specified in the checklist by Philips et al.[85] The ERG notes that the outcome of TIA specified in the NICE scope was excluded from the manufacturer's economic evaluation as a result of an absence of data from ARISTOTLE or AVERROES. The ERG considers that the modelled pathways following an MI or SE may have been over simplified; however, the impact of this simplification on the model results is not substantial (see Section 5.4.6). In addition, the ERG notes that the uniform distribution was inappropriately used in the PSA.

Table 68. NICE reference case[69]

Attribute	Reference case	Does the *de novo* economic evaluation match the reference case?
Decision problem	The scope developed by NICE	With the exception of assessment of TIA as an outcome, the manufacturer followed the NICE scope in full. In addition, the manufacturer included aspirin as a comparator in a VKA unsuitable patient population. The ERG acknowledges that TIA was not recorded in either ARISTOTLE or AVERROES trial. Furthermore, based on expert clinical opinion the ERG considers that the exclusion of TIA will not change overall model results.
Comparator(s)	Alternative therapies routinely used in the NHS	Yes
Perspective costs	NHS and Personal Social Services	Yes
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost-utility analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes; however, a shorter time horizon (100 years vs 123 years) would have provided better face validity given the life expectancy of AF patients.
Synthesis of evidence on outcomes	Systematic review	Yes
Outcome measure	QALYs	Yes
Health states for QALY	Described using a standardised and validated instrument	Utility values were obtained from the literature using EQ-5D data where available.
Benefit valuation	Time-trade off or standard gamble	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	Yes
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Sensitivity analysis	Probabilistic sensitivity analysis	Yes. The manufacturer carried out deterministic sensitivity analysis, scenario analysis and probabilistic sensitivity analysis.
Abbreviations used in table: AF, atrial fibrillation; EQ-5D, EuroQol 5 dimensions questionnaire; NICE, National Institute for Health and Clinical Excellence; NHS, National Health Service; QALY, quality adjusted life year; TIA, transient ischaemic attack; VKA, vitamin K antagonist.		

Table 69. Philips[85] checklist

Dimension of quality	Comments
Structure	
S1: Statement of decision problem/objective	Clearly stated.
S2:Statement of scope/perspective	The scope and perspective of the model were clearly stated. With the exception of the exclusion of TIA (for which there was no trial data collected), the manufacturer fully followed the NICE scope. Over and above the NICE scope, the manufacturer assessed the cost-effectiveness of aspirin in a VKA unsuitable population.
S3: Rationale for structure	The ERG notes that the manufacturer's rationale for the structure of the model was based on previous publications of related technology appraisals. The ERG considers the model structure to be appropriate and well-constructed.
S4: Structural assumptions	The structural assumptions were transparent. In addition, a number of scenario and sensitivity analyses were undertaken to test the robustness of the different assumptions.
S5: Strategies/comparators	All relevant comparators were evaluated. In addition aspirin was evaluated in a VKA unsuitable patient population.
S6: Model type	Correct, cost-utility analysis.
S7: Time horizon	Lifetime, in accordance with NICE methods guide.[69] In addition, shorter time horizons (1, 3 and 5 year) were assessed.
S8: Disease states/pathways	With the exception of pathways following an MI or SE, the ERG agrees with the pathways/health states modelled. However, the exclusion of subsequent events following an MI or SE may bias the model against treatments that are more effective at preventing MI or SE, respectively.
S9: Cycle length	The ERG considers 6 weeks to be a reasonable cycle length to capture the consequences of model events. Half-cycle correction was included in the model.
Data	
D1: Data identification	Data were taken from ARISTOTLE and AVERROES trials. The manufacturer also carried out NMAs. Utility data were identified through a systematic literature review.
D2: Pre-model data analysis	Correct formulae were used in all pre-model data analyses.
D2a: Baseline data	Baseline data were taken from ARISTOTLE and AVERROES trials. Conversion of yearly rates to quarterly probabilities was carried out using standard formulae.
D2b: Treatment effects	Treatment effects for each outcome were estimated from an NMA where apixaban data from ARISTOTLE was used as the baseline. The model used hazard ratios as the relative treatment effect. Extrapolation of treatment effects is clearly described and justified.
D2d: Quality of life weights (utilities)	ARISTOTLE and AVERROES trials did not collect QoL data. Utility data used in the model was obtained from literature and appropriately referenced.
D3: Data incorporation	The manufacturer clearly described how data were used in the model, all sources were referenced and copies of referenced papers were provided. Standard distributions were used for different outcomes (e.g. the gamma distribution for costs).
D4: Assessment of uncertainty	The assessment of sensitivity was thorough and robust. Probabilistic, one-way sensitivity analysis and various scenario analyses were satisfactorily reported.
D4a: Methodological	Appropriate analytical methods were used, and were supported with sensitivity and scenario analyses to test the robustness of the chosen base case approach.
D4b: Structural	The manufacturer described deterministic sensitivity analysis and scenario analysis in detail.
D4c: Heterogeneity	Heterogeneity was partially addressed by the analysis of different subgroups of patients.

D4d: Parameter	Probabilistic sensitivity analysis was carried out to a generally high standard. However, the ERG notes that the uniform distribution was inappropriately used for other-cause mortality HR.
Consistency	
C1: Internal consistency	The model seems to be mathematically sound with no obvious inconsistencies. The manufacturer reported that the model was validated by independent economists.
C2: External consistency	The model results are intuitive and conclusions are valid given the data presented.
Abbreviations used in table: ERG, Evidence Review Group; HR, hazard ratio; MI, myocardial infarction; MS, manufacturer's submission; NICE, National Institute for Health and Clinical Excellence; NMA, network meta-analysis; QoL, quality of life; SE, systemic embolism; TIA, transient ischaemic attack; VKA, vitamin K antagonist.	

5.4.2 Treatment switching and discontinuation

Within the manufacturer's model, discontinuation of first-line therapy was either a result of other ICH or major bleeding events (event-related discontinuation) or other causes (other-cause discontinuation). In the base case, aspirin was assumed to be the second-line treatment following either event-related or other-cause discontinuation. The manufacturer's rationale for choosing aspirin as second-line treatment was based upon recommendations made in CG36 (MS; pg 127).

However, based on expert clinical input, there is uncertainty regarding the second-line treatment choice for patients who discontinue therapy with apixaban, dabigatran or rivaroxaban; it was considered that some of these patients (depending on the reason for discontinuation) may be eligible for treatment with warfarin or possibly a different NOAC. Therefore, as part of the clarification process the ERG asked for further clarification of the manufacturer's rationale for choosing aspirin as the second-line treatment.

The manufacturer stated that:

- discontinuation (on clinical grounds) from one anticoagulation therapy would mean that alternative anticoagulants would be contraindicated;
- using aspirin as second-line treatment allows fair comparison against warfarin (i.e. the same treatment sequence can be used);
- the same approach was employed in TA249.

The ERG notes that consistency with other NOAC HTA submissions can only be achieved by considering a treatment sequence with aspirin as second-line therapy. However, based on the current uncertainty over clinical practice, the ERG considers it useful to note the impact of other treatment sequences on the cost-effectiveness results.

Therefore, the ERG considered the following incremental analyses:

- apixaban, dabigatran blend (150 mg BD switching to 110 mg at the age of 80 years), dabigatran 110 mg and rivaroxaban: using warfarin as second-line treatment;
- apixaban, dabigatran blend (150 mg BD switching to 110 mg at the age of 80 years), dabigatran 110 mg and warfarin: using rivaroxaban as second-line treatment;
- apixaban, rivaroxaban and warfarin using dabigatran blend as second-line treatment.

The results of these analyses are presented in Appendix 9.7

To summarise, when second-line treatment was assumed to be warfarin, dabigatran 110 mg and rivaroxaban were strictly dominated by dabigatran blend. The ICER of apixaban versus dabigatran blend was £28,695. When rivaroxaban was used as second-line therapy, dabigatran 110 mg was strictly dominated by warfarin treatment, the ICERs of dabigatran blend and apixaban versus warfarin were £9,923 and £11,637, respectively. An incremental ICER of apixaban versus dabigatran blend was £60,366. When second-line treatment was assumed to be dabigatran 110 mg, rivaroxaban was extendedly dominated by apixaban and the ICER of rivaroxaban versus warfarin was £287.

The ERG considers it important to note that the risks patients were exposed to on second-line treatment were constant (see Section 5.3.7 and 5.4.5). Therefore, caution should be used when interpreting these results. However, the main driver of the higher ICERs seen in the analyses around second-line treatment choice (e.g. apixaban versus dabigatran blend) was discontinuation. That is, patients who discontinued treatment fared far better than in the base case. Therefore, treatments with higher discontinuation rates (e.g. dabigatran) appeared more effective than in the manufacturer's base case.

In addition, based on expert clinical advice, the ERG notes that there is some uncertainty regarding the relative other-cause discontinuation rates of apixaban and dabigatran. This is because, by contrast to ARISTOTLE, RE-LY was an open label trial. Expert clinical advice was that within open label trials, unexplained new symptoms may be associated with the novel therapeutic and treatment stopped. Therefore, it is possible that some of the higher level of discontinuation observed with dabigatran versus warfarin in RE-LY may be attributable to this phenomenon. The ERG carried out an exploratory analysis to investigate the impact of other-cause discontinuation on the manufacturer's cost-effectiveness results. In the exploratory analysis, the ERG assumed there was no difference in other-cause discontinuation between apixaban and dabigatran (both doses). The results of this exploratory analysis are presented in Appendix 9.7. To summarise, rivaroxaban and dabigatran 110 mg were strictly dominated by dabigatran blend. However, apixaban no longer extendedly dominated dabigatran blend, rather the ICER for the comparison of apixaban with dabigatran blend became £14,456.

5.4.3 Mortality

Mortality was accounted for in the manufacturer's model by categorising as event-specific and other-cause mortality. Other-cause mortality was further disaggregated between the within-trial and post-trial phases of the model. During the within-trial period, the rate of other-cause mortality was [redacted] for patients treated with warfarin than for patients treated with an NOAC (HR=[redacted]). Expert clinical advice received by the ERG suggested that patients treated with warfarin may be at a higher risk of death as a result of adverse bleeding events. However, it is not expected that patients receiving treatment with warfarin would be at a [redacted] risk of other-cause mortality. The ERG notes that the manufacturer carried out a scenario analysis that assumed equivalent other-cause mortality rates between treatments for the full model time horizon. The incremental cost-effectiveness results were unaffected by assuming equal rates of other-cause mortality across treatments; i.e. dabigatran and rivaroxaban continued to be extendedly dominated (see Appendix 9.6). However, the ICER of apixaban versus warfarin increased from £11,008 to £12,829.

5.4.4 Health-related quality of life

Treatment-related disutilities

By contrast with the dabigatran and rivaroxaban STA submissions,[43][44] the manufacturer submitted a model that accounted for disutility associated with treatment. A study by Gage et al. identified in the manufacturer's systematic literature review was used to inform the treatment related disutilities used in the model.[71] Gage et al. examined the impact of stroke and prophylactic treatment for stroke on patients quality of life. The mean disutility of patients treated with warfarin was reported as 0.01 (perfect health minus utility when treated with warfarin), whereas the mean disutility associated with aspirin treatment was 0.002.[71] The manufacturer assumed that the disutility associated with all of the NOACs would be equivalent to that associated with aspirin therapy.

The ERG notes that in ARISTOTLE, apixaban had a more favourable discontinuation profile than warfarin (MS; pg 51). Therefore, based on ARISTOTLE and on expert clinical advice, the ERG considers that treatment with warfarin might incur a higher level of disutility than treatment with apixaban. However, the ERG responsible for considering the dabigatran STA noted that dabigatran was associated with a higher level of disutility than patients treated with warfarin.[40] In addition, no evidence of treatment related disutility is available from ROCKET-AF;[27] however, the ERG responsible for evaluating the rivaroxaban STA noted that the level of discontinuation with rivaroxaban was higher than that observed with warfarin.[86] Therefore, the ERG considers that the assumption of equivalent disutility between the NOACs may not be robust with any resultant bias likely to have been against apixaban. However, the manufacturer carried out a sensitivity analysis that investigated the effect of treatment-related disutility on the cost-effectiveness results; no disutility was

assumed for any treatment. The incremental results were unaffected by the removal of treatment-related disutility (dabigatran and rivaroxaban continued to be extendedly dominated, see Appendix 9.6); however, the ICER of apixaban versus warfarin increased from £11,008 to £14,530.

Age adjustment of utility

In addition to the use of treatment-related disutilites, the manufacturer assumed that health-related quality of life would remain constant over time. That is, a patient's quality of life would be affected by events experienced but not by increasing age (i.e. utilities were not age adjusted). As part of the clarification process, the ERG requested an updated model in which utility was adjusted for age. The manufacturer implemented an age adjustment of -0.00029 per year and provided an updated model in the clarification response. A study by Sullivan et al. of UK EQ-5D scores was used to inform the utility decrement used by the manufacturer.[72] Sullivan et al. applied UK general population preferences to EQ-5D data derived from the US-based Medical Expenditure Panel Survey. The incremental cost-effectiveness results were unaffected by the addition of age adjustment (dabigatran and rivaroxaban continued to be extendedly dominated, see Appendix 9.6); however, the ICER of apixaban versus warfarin increased from £11,008 to £11,227.

5.4.5 Treatment effectiveness

Uncertainty around MI event risks

Within, the manufacturer's model event risks were assumed to be dependent on treatment received. The ERG considers that this assumption was reasonable, but notes that there is uncertainty surrounding the relative effect of apixaban and dabigatran on the risk of MI. As discussed in Section 4.4.5, the manufacturer carried out a sensitivity analysis of NMA 1 (based on data from RE-LY, ROCKET-AF and ARISTOTLE). The manufacturer's sensitivity analysis substituted efficacy data from RE-LY with updated efficacy data from an analysis of RE-LY carried out in 2010.[42] The ERG notes that the results of this sensitivity analysis indicate that the reduction in MI associated with apixaban versus dabigatran (both doses) was not statistically significant. The ERG carried out an exploratory analysis to investigate the impact of MI event risk on the manufacturer's cost-effectiveness results. In the exploratory analysis, the ERG assumed there was no difference in the risk of MI for patients treated with apixaban or dabigatran (both doses). The results of this exploratory analysis are presented in Appendix 9.6. To summarise, dabigatran 110 mg continued to be strictly dominated by dabigatran blend. However, apixaban no longer extendedly dominated dabigatran blend, rather the ICER for the comparison of apixaban with dabigatran blend was £11,191.

Treatment specific stroke severity and bleed distributions

The ERG notes that amongst the parameters in the manufacturer's model that were assumed to be treatment dependant were the distributions of stroke severity (ischaemic and ICH). However, based on expert clinical opinion, the ERG notes that it may have been inappropriate to assume that the severity of a stroke event (ischaemic or haemorrhagic) depends upon treatment received. Moreover, stroke severity was not a pre-specified outcome in ARISTOTLE (or AVERROES), RE-LY or ROCKET-AF; the severity distributions used in the model were derived from a secondary analysis of ARISTOTLE data and weighting of the data reported in RE-LY and ROCKET-AF publications (see Section 5.3.7). Therefore, the difference in stroke severity profiles observed between apixaban and warfarin may have been a result of random chance or other factors not accounted for in the analysis.

In addition, the manufacturer's model assumed that treatment affected the proportion of:

- patients experiencing a haemorrhagic stroke versus any other ICH;
- patients experiencing a major bleed that is a GI bleed versus a non-GI bleed.

For similar reasons to those given for stroke severity, the ERG notes that it may have been inappropriate to assume that bleed type was dependant on treatment received.

However, the ERG notes that the manufacturer carried out a scenario analysis that assumed equivalent stroke severity distributions (mild, moderate, severe and fatal) for all interventions; assumed to be equivalent to the distribution of stroke severity observed in ARISTOTLE for patients treated with apixaban. The ERG carried out a similar scenario analysis for the distribution of bleed type; assuming the distribution of bleed type is equivalent to that observed with apixaban for all treatments. These scenario analyses had little impact on the manufacturer's incremental cost-effectiveness results; dabigatran and rivaroxaban continued to be extendedly dominated (see Appendix 9.6 for incremental results). However, the ICERs of apixaban versus warfarin increased (from £11,008 to £12,277) and decreased (from £11,008 to £9,771) for the scenario analyses of stroke severity and bleed type distribution, respectively.

Event-risks for patients on second-line therapy

Furthermore, the ERG notes that the risk profile of patients on second-line therapy was not adjusted for patient characteristics such as age or $CHADS_2$ score. No rationale for this assumption was provided in the MS; however, the ERG understands that the task of adjusting for patient characteristics in second-line treatment may be beyond the reasonable scope of a Markov model. Moreover, the ERG assessed the sensitivity of the model to increased event-risks over time for patients on second-line therapy. The same risk adjustment factors used for patients receiving first-line therapy were used to adjust the risk of stroke, ICH, major bleeding, CRNM bleeding and MI in

patients receiving second-line therapy. The effect of this sensitivity analysis on the incremental results is displayed in Appendix 9.6. Dabigatran and rivaroxaban persisted in being extendedly dominated; however, the ICER for apixaban versus warfarin fell slightly from £11,008 to £10,779.

5.4.6 Model structure

Within the manufacturer's model, the risks of events following a permanent event (stroke, SE or MI) were limited. In particular, patients who experienced a stroke event (ischaemic or haemorrhagic) were exposed to the risk of one and only one subsequent stroke event (see Section 5.3.4). Moreover, patients who experienced an MI or SE entered a semi-absorbing health state and were not exposed to any further event risks (except death).

The manufacturer's rationale for limiting the potential number of strokes was that data on recurrent stroke were not available for any of the interventions considered. In addition, the manufacturer stated that patients who experienced an MI transitioned into the semi-absorbing MI health state "to keep the model simple but to incorporate important sequelae of AF" (MS; pg 108). No rationale was provided for the assumption that SE was a semi-absorbing event.

Based on expert clinical advice, the ERG accepts the risk limitation applied to patients who experienced a stroke event. However, following expert clinical input, the ERG notes that patients who experienced an MI or SE would remain at risk of further events (in particular ischaemic stroke). Furthermore, the ERG notes that the study by Mohan et al.[58] (from which the manufacturer derived the risk of recurrent stroke, see Section 5.3.7) reports cumulative probabilities of recurrent stroke following an MI. These are 11.3, 22.6 and 41.1 for 1, 5 and 10 years post-MI, respectively.

The ERG notes that the risk and severity distribution of subsequent stroke is independent of treatment. Therefore, any bias introduced assuming that MI and SE are semi-absorbing health states will depend entirely on the difference (in cost, utility and long-term mortality risk) between the permanent health states. Table 70 summarises the costs, utilities and long-term mortality risks associated with the MI, SE and recurrent stroke health states.

Table 70. Long-term costs and utilities associated with MI, SE or ischaemic stroke health states in the manufacturer's model

Health state	Long-term cost (£)	Health state utility	Long-term mortality risk
MI	6.65	0.6830	2.56 (males) and 4.16 (females)
SE	183.91	0.6795	1.34
Recurrent ischaemic stroke			
Mild	183.91	0.7600	3.18
Moderate	358.78	0.3900	5.84
Severe	544.76	0.1100	15.75
Abbreviations used in table: MI, myocardial infarction; SE, systemic embolism			

The costs, utilities and long-term mortality risks associated with recurrent stroke are generally less favourable than those experienced by patients in the MI or SE health states. Therefore it might be expected that the assumption that MI and SE are semi-absorbing events may:

- bias the model against the treatment that is most effective at preventing MI as MI patients are not at risk of moving into health states with poorer outcomes;
- bias the model against the treatment that is most effective at preventing SE as SE patients are not at risk of moving into health states with poorer outcomes.

As part of the clarification process, the ERG requested that the manufacturer provide a scenario analysis around the future risks associated with SE. The scenario analysis requested was to apply the risks of recurrent ischaemic stroke to patients who experienced an SE; using the same risks as those applied to patients who had experienced an ischaemic stroke. The manufacturer provided the scenario analysis with the caveat that no data were available "specifically for SE patients to support this modelling" (Manufacturer's clarification response; pg 33).

In addition, the ERG carried out a scenario analysis that assumed that patients who had experienced an MI were at risk of subsequent stroke events. The annual rate of recurrent stroke employed in this analysis was 0.053; calculated from the 10 year cumulative risk (41.1%) of recurrent stroke in patients who had experienced an MI, reported in Mohan et al.[58]

These scenario analyses did not alter the overall incremental results, in that dabigatran and rivaroxaban continued to be extendedly dominated (see Appendix 9.6). However, the ICER of apixaban versus warfarin decreased from £11,008 to ~£10,980 for both scenario analyses. For each analysis, the decrease in the ICER was driven by an increase in the relative number of mild recurrent strokes; i.e. patients treated with warfarin experienced more recurrent mild strokes than patients treated with apixaban. The increase in the number of mild strokes is in turn driven by the distribution

of recurrent stroke severity; assumed to be equivalent to that observed in ARISTOTLE for patients treated with apixaban for both scenario analyses is (i.e. ~[redacted]% of recurrent strokes will be mild).

5.4.7 Resources and costs

As discussed in Section 5.3.12, differences exist between the ICERs estimated by the manufacturer's model and those estimated in the dabigatran and rivaroxaban submissions.[43][44] The manufacturer attributed these differences to differences in the key parameters such as costs associated with stroke events. Furthermore, the ERG notes that the acute cost of SE is approximately double the acute cost used in other NOAC submissions (apixaban submission, £4,077.98; dabigatran submission, £2,772 [fatal and non-fatal acute costs]; rivaroxaban, £1,658.12). Moreover, the acute cost associated with SE used in the dabigatran and rivaroxaban submissions was derived from NHS reference costs. Therefore, the ERG considers the costs used in the other NOAC submissions to be more plausible than those employed in the manufacturer's model. The ERG carried out sensitivity analyses to assess the impact of using the lower costs for acute SE. The incremental results did not change and dabigatran and rivaroxaban continued to be extendedly dominated by apixaban (Appendix 9.6). However, the ICER of apixaban versus warfarin increased from £11,008 to £11,012 and £11,016 when the costs used in the dabigatran and rivaroxaban submission were used, respectively.

5.4.8 Perspective, time horizon and discounting

The model submitted by the manufacturer adopted a lifetime time horizon where patients were followed for 49 years (from 74 to 123 years of age). Given that the observed life expectancy of the general population was approximately 79 years (MS; pg 271), the ERG considers the maximum modelled age of 123 years to lack face validity. In addition, the ERG notes that within the model 99.7% of patients had died after 26 years (by 100 years of age). Therefore, in line with current good research practices[87] the ERG recommends using a time horizon of 26 years (74 to 100 years of age). As part of the clarification process the ERG requested a revised model with a 26 year time horizon. To which the manufacturer provided a model with a truncated time horizon. Upon implementation of the shorter time horizon the incremental model results remained the same, with dabigatran and rivaroxaban being extendedly dominated (see Appendix 9.6). However, the ICER of apixaban versus warfarin increased by £6, from £11,007 to £11.013.

5.5 Conclusions of the cost-effectiveness section

The manufacturer submitted an economic evaluation in the VKA suitable AF patient population that was robust and generally conservative; i.e. generally, bias potentially associated with model assumptions was likely to be against apixaban. However, the ERG identified some areas where model face validity and accuracy could have been improved; although, the ERG acknowledges that the impact of these on the cost-effectiveness results was limited (see Section 6). Overall, the ERG considers that the sources and assumptions used in the manufacturer's economic model were systematically identified and appropriate.

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

The ERG carried out numerous sensitivity analyses, in response to points of critique, which are detailed throughout Section 5.4. Some of the sensitivity analyses carried out have been combined to form a revised ERG base case. The ERG's revised base case is displayed in Table 71 and consisted of the following amendments:

- other-cause mortality assumed to be independent of treatment received;
- utility adjusted for age;
- stroke severity distribution assumed to be independent of treatment received;
- bleed type assumed to be independent of treatment received;
- SE patients assumed to be at risk of recurrent stroke;
- MI patients assumed to be at risk of recurrent stroke;
- acute cost of SE assumed to be equal to the cost used in the rivaroxaban HTA submission[44](chosen over dabigatran as the more conservative cost);
- time horizon assumed to be 26 years.

However, the ERG notes that the incremental results were unaffected by any of the above amendments (both individual and combined amendments); i.e. dabigatran and rivaroxaban continued to be extendedly dominated by apixaban. Therefore, only the comparison of apixaban versus warfarin was considered in the ERG's revised base case (Table 71).

In addition, some of the sensitivity analyses carried out by the ERG were not included in the ERG's revised base case, these were:

- age adjustment of event risks for patients on second-line therapy;
- the removal of treatment-related disutility;
- the use of alternative second-line treatments.

The age adjustment of event risks for patients on second-line therapy was a sensitivity analysis carried out to investigate the effect of assuming constant risk (see Section 5.4.5). However, the ERG acknowledges that age is not the only patient characteristic that event risks (for patients on second-line therapy) would need to be adjusted for; duration of treatment and $CHADS_2$ score may also be important. Therefore, the ERG did not include age adjustment of the second-line treatment risk profile into the revised base case, as it was considered that this analysis did not accurately reflect the impact of patient characteristics on second-line treatment event risks.

The ERG considered that the assumption that all NOACs had the same treatment-related disutility as patients who received treatment with aspirin was unfounded (see Section 5.4.4). However, the impact

of this assumption was only important when apixaban was compared with other NOACs. Therefore, the ERG decided not to include the removal of treatment-related disutility into the revised base case as only the comparison of apixaban with warfarin was considered. Similarly, the use of alternative second-line treatments was not relevant to the comparison of apixaban versus warfarin; therefore, these were not included in the ERG's revised base case.

Table 71. Incremental impact of the ERG's revised base case amendments

Analysis	Tx	Total costs (£)	Total QALYs	Inc. costs (£)	Inc. QALYs	Inc ICER (£/QALYs)	Cumulative ICER (£/QALYs)
Manufacturer's base case	**Warfarin**	**7,188.49**	**5.70**	**–**	**–**	**–**	**11,007.68**
	Apixaban	**8,983.07**	**5.86**	**1,794.58**	**0.163**	**11,007.68**	
Other-cause mortality assumed to be independent of Tx	Warfarin	7,166.80	5.68	–	–	–	12,829.47
	Apixaban	8,917.83	5.82	1,751.02	0.14	12,829.47	
Utility adjusted for age	Warfarin	7,188.49	5.59	–	–	–	13,081.07
	Apixaban	8,983.07	5.75	1,794.58	0.16	11,226.73	
Stroke severity assumed to be independent of Tx	Warfarin	6,576.61	5.73	–	–	–	14,788.43
	Apixaban	8,485.48	5.89	1,908.87	0.16	12,276.64	
Bleed Type assumed to be independent of Tx	Warfarin	7,264.28	5.67	–	–	–	12,805.14
	Apixaban	9,018.76	5.85	1,754.48	0.18	9,771.15	
SE patients assumed to be at risk of stroke	Warfarin	7,201.36	5.69	–	–	–	12,774.04
	Apixaban	8,994.67	5.85	1,793.31	0.16	10,982.25	
MI patients assumed to be at risk of stroke	Warfarin	7,310.59	5.69	–	–	–	12,739.74
	Apixaban	9,104.35	5.86	1,793.76	0.16	10,981.19	
Acute SE costs from rivaroxaban HTA submission[44] used	Warfarin	7,152.47	5.70	–	–	–	12,748.90
	Apixaban	8,948.37	5.86	1,795.90	0.163	11,015.79	
26 year time horizon	Warfarin	7,185.87	5.70	–	–	–	12,757.14
	Apixaban	8,980.27	5.86	1,794.40	0.163	11,013.85	
Event-risk on second-line Tx adjusted for patient characteristics	Impact unknown, however analysis of impact from age adjustment of second-line treatment event risks suggested that impact may favour apixaban						
ERG revised base case	Warfarin	**6,733.32**	**5.57**	**–**	**–**	**–**	**12,757.14**
	Apixaban	**8,556.42**	**5.71**	**1,823.17**	**0.14**	**12,757.14**	

Abbreviations used in table: ERG, Evidence Review Group; HTA, health technology assessment; ICER, incremental cost-effectiveness ratio; MI, myocardial infarction; QALY, quality adjusted life year; SE, systemic embolism; Tx, treatment

7 OVERALL CONCLUSIONS

The manufacturer presented data in the MS for the use of apixaban in patients with AF and at least one additional risk factor for stroke.

The manufacturer presents the case for the use of apixaban compared with adjusted-dose warfarin for the prevention of stroke and systemic embolism in patients with AF based on data from the ARISTOTLE trial. The Evidence Review Group (ERG) considers ARISTOTLE to be a good quality, well conducted RCT. The ERG notes that there is an absence of direct evidence regarding the efficacy of apixaban in the low risk AF population ($CHADS_2$ <1). However, the ERG acknowledges that the proposed licensed indication for the use of apixaban in AF would not include patients with a $CHADS_2$ score of 0.

The primary objective of ARISTOTLE, to prove non-inferiority of apixaban versus warfarin in the prevention of stroke and SE was met. In addition, superiority was proven; i.e. apixaban was associated with significantly fewer strokes and systemic emboli when compared with warfarin (HR 0.79; 95% CI 0.66–0.95; p=0.01). The primary safety outcome of ARISTOTLE was ISTH major bleeding and apixaban was proven superior to warfarin in reducing the risk of these bleeding events (HR 0.69; 95% CI 0.60–0.80; p<0.01). ARISTOTLE also demonstrated that apixaban was associated with significantly fewer permanent treatment discontinuations compared with warfarin (25.3% vs 27.5% respectively; p=0.001). The ERG considers that the overall adverse event and safety profile of apixaban was comparable or better when compared with warfarin for the outcomes reported in the MS.

The ERG acknowledges that ARISTOTLE was not powered appropriately to detect differences in efficacy and safety in the subgroup analyses. However, the ERG notes that the post hoc subgroup analysis of Western Europe demonstrated [redacted]. In addition, the results of the centre TTR subgroup analyses reported in ARISTOTLE suggested that, regardless of INR control, apixaban was associated with fewer stroke or SE and fewer major bleeding events than warfarin. However, this difference was not statistically significant. The final subgroup that the ERG considers important to highlight is [redacted].

The clinical data presented in the MS also included a trial (AVERROES) comparing apixaban with aspirin in patients who had failed VKA therapy, or had been deemed as unsuitable for VKA therapy (i.e. warfarin unsuitable). The ERG notes that aspirin was not listed as a comparator of interest in the final scope issued by NICE. The ERG thus does not consider AVERROES to meet the inclusion criteria for this STA.

The manufacturer conducted a NMA to compare apixaban with dabigatran, rivaroxaban and warfarin in patients suitable for VKA therapy; NMA 1. The ERG considers that there is potential clinical heterogeneity within the network of included trials. However, the ERG acknowledges that each treatment in the analysis was informed by only one study. The exclusion of a study from the analysis to explore the potential heterogeneity would thus have resulted in the exclusion of a treatment from the network.

The base case results of NMA 1 suggested that apixaban was associated with a significantly lower incidence of MI compared to dabigatran 150 mg or 110 mg. [redacted].

For the bleeding safety outcomes, apixaban was generally associated with fewer events compared with rivaroxaban, dabigatran 150 mg, dabigatran 110 mg [redacted] In addition, apixaban was associated with significantly fewer discontinuations compared with dabigatran 150 mg, dabigatran 110 mg, rivaroxaban [redacted]

[redacted]

The ERG notes that the subgroups of the cTTR analyses were defined differently in each of the included trials and the [redacted]. Therefore, the ERG does not consider the subgroup analyses to be directly comparable. The ERG considers that the subgroup results should be interpreted with caution.

The manufacturer presented a robust and predominantly conservative (direction of bias generally more likely to be against rather than towards apixaban) economic evaluation of apixaban versus

warfarin, dabigatran 110 mg, dabigatran blend (150 mg moving to 110 mg at age 80) and rivaroxaban in the VKA suitable AF patient population. The manufacturer presented fully incremental cost-effectiveness results. These indicated that:

- dabigatran 110 mg was strictly dominated by (i.e. is less costly and less effective than) dabigatran blend;
- rivaroxaban and dabigatran blend were extendedly dominated (i.e. resulted in a lower incremental cost-effectiveness ratio (ICER) versus warfarin despite having higher total costs) by apixaban;
- apixaban had an ICER versus warfarin of £11,008 per quality adjusted life year (QALY).

The ERG carried out several sensitivity analyses to investigate uncertainty around the model's base case assumptions. It is important to note that none of these sensitivity analyses altered the incremental cost effectiveness results. Furthermore, the ICER of apixaban versus warfarin generated by the ERG's revised base case (£12,757) remained relatively consistent with the manufacturer's base case ICER (£11,008).

In addition to sensitivity analyses, the ERG carried out exploratory analyses around the choice of second-line treatment. Exploratory analyses around second-line treatment options were prompted by expert clinical input regarding uncertainty in clinical practice. The results of these analyses were highly variable, with incremental ICERs for apixaban varying between £287 (versus warfarin when dabigatran 110 mg was chosen as second-line treatment) and £60,366 (versus dabigatran blend, when rivaroxaban was chosen as second-line treatment). However, the ERG notes that within the manufacturer's model patients on second-line treatment were exposed to a constant risk of events. Therefore, the results of these analyses should be interpreted with caution as the main driver of the ICERs was discontinuation rates associated with first-line therapy. That is, patients who discontinued treatment fared far better than in the base case. Consequently, treatments with higher discontinuation rates (e.g. dabigatran) appeared more effective than in the manufacturer's base case.

In addition, the ERG carried out exploratory analyses around the level of discontinuation and risk of MI associated with dabigatran. Exploration of the impact of treatment discontinuation associated with dabigatran was prompted by expert clinical input. Clinical opinion was that the manufacturer's analysis of treatment specific discontinuation rates may have been biased by the open-label trial used to inform the treatment effect of dabigatran. Regarding the risk of MI, the manufacturer's first sensitivity analysis of NMA 1 revealed uncertainty in the significance of the reduction of MI risk associated with apixaban versus dabigatran. However, it is important to note that the analyses carried out by the ERG were extreme value analyses which assumed no difference between apixaban and dabigatran for treatment discontinuation and MI risk. These analyses resulted in ICERs of £11,191 and £14,456, respectively.

7.1 Implications for research

The ERG considers that there is a need for further research into the safety and clinical benefit of apixaban compared with dabigatran etexilate and rivaroxaban in patients who are not suitable for warfarin or other VKAs. In addition, the ERG considers that head-to-head trials comparing apixaban, dabigatran etexilate and rivaroxaban in patients whom are suitable for treatment with VKAs would be informative for service provision within the NHS. Finally, the ERG notes that there is an absence of HRQoL studies in people taking apixaban; HRQoL studies would be useful for further informing the clinical effectiveness of apixaban.

8 REFERENCES

(1) Simpson J FIDMMD. Antithrombotic therapy and atrial fibrillation in Scotland: Results of a national audit. *J R Coll Physicians Edinb* 2012;42 (Suppl 18); 64.

(2) National Institute for Health and Clinical Excellence. The management of atrial fibrillation. NICE. 2006. [last accessed 2012 October]; Available from: http://guidance.nice.org.uk/CG36

(3) NHS. The Information Centre. Quality of outcomes framework (QOF) for April 2009-March 2010, England. NHS. The Information Centre. 2012. [last accessed 2012 October]; Available from: http://www.ic.nhs.uk/webfiles/QOF/2009-10/Prevalence%20tables/QOF0910_National_Prevalence.xls.

(4) Health in Wales. QOF prevalence charts 2009-10. Health in Wales. 2010. [last accessed 2012 October]; Available from: www.wales.nhs.uk/sites3/docopen.cfm?orgid=480&id=168155

(5) Stewart S, Hart CL, Hole DJ, McMurray JJ. Population prevalence, incidence, and predictors of atrial fibrillation in the Renfrew/Paisley study. *Heart* 2001 Nov;86**(5)**:516-21.

(6) de LS, van VJ, Hague N, Thana L, Dzregah B, Chan T. Preventing stroke in people with atrial fibrillation: a cross-sectional study. *J Public Health (Oxf)* 2005 Mar;27**(1)**:85-92.

(7) DeWilde S, Carey IM, Emmas C, Richards N, Cook DG. Trends in the prevalence of diagnosed atrial fibrillation, its treatment with anticoagulation and predictors of such treatment in UK primary care. *Heart* 2006 Aug;92**(8)**:1064-70.

(8) Majeed A, Moser K, Carroll K. Trends in the prevalence and management of atrial fibrillation in general practice in England and Wales, 1994-1998: analysis of data from the general practice research database. *Heart* 2001 Sep;86**(3)**:284-8.

(9) Murphy NF, Simpson CR, Jhund PS, Stewart S, Kirkpatrick M, Chalmers J, *et al*. A national survey of the prevalence, incidence, primary care burden and treatment of atrial fibrillation in Scotland. *Heart* 2007 May;93**(5)**:606-12.

(10) Friberg L, Hammar N, Pettersson H, Rosenqvist M. Increased mortality in paroxysmal atrial fibrillation: report from the Stockholm Cohort-Study of Atrial Fibrillation (SCAF). *Eur Heart J* 2007 Oct;28**(19)**:2346-53.

(11) Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991 Aug;22**(8)**:983-8.

(12) Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, *et al*. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010 Oct;31**(19)**:2369-429.

(13) Frost L, Engholm G, Johnsen S, Moller H, Henneberg EW, Husted S. Incident thromboembolism in the aorta and the renal, mesenteric, pelvic, and extremity arteries after discharge from the hospital with a diagnosis of atrial fibrillation. *Arch Intern Med* 2001 Jan 22;161**(2)**:272-6.

(14) Lamassa M, Di CA, Pracucci G, Basile AM, Trefoloni G, Vanni P, *et al*. Characteristics, outcome, and care of stroke associated with atrial fibrillation in Europe: data from a

multicenter multinational hospital-based registry (The European Community Stroke Project). *Stroke* 2001 Feb;32**(2)**:392-8.

(15) Jorgensen HS, Nakayama H, Reith J, Raaschou HO, Olsen TS. Acute stroke with atrial fibrillation. The Copenhagen Stroke Study. *Stroke* 1996 Oct;27**(10)**:1765-9.

(16) Seet RC, Zhang Y, Wijdicks EF, Rabinstein AA. Relationship between chronic atrial fibrillation and worse outcomes in stroke patients after intravenous thrombolysis. *Arch Neurol* 2011 Nov;68**(11)**:1454-8.

(17) Marini C, De SF, Sacco S, Russo T, Olivieri L, Totaro R, *et al*. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke: results from a population-based study. *Stroke* 2005 Jun;36**(6)**:1115-9.

(18) Lin HJ, Wolf PA, Kelly-Hayes M, Beiser AS, Kase CS, Benjamin EJ, *et al*. Stroke severity in atrial fibrillation. The Framingham Study. *Stroke* 1996 Oct;27**(10)**:1760-4.

(19) Hilari K, Northcott S, Roy P, Marshall J, Wiggins RD, Chataway J, *et al*. Psychological distress after stroke and aphasia: the first six months. *Clin Rehabil* 2010 Feb;24**(2)**:181-90.

(20) Luengo-Fernandez R, Gray AM, Rothwell PM. Population-based study of determinants of initial secondary care costs of acute stroke in the United Kingdom. *Stroke* 2006 Oct;37**(10)**:2579-87.

(21) National Institute for Health and Clinical Excellence. TA249 Dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation. NICE. 2012. [last accessed 2012 October]; Available from: http://www.nice.org.uk/guidance/TA249

(22) National Institute for Health and Clinical Excellence. TA256 Atrial fibrillation (stroke prevention) - rivaroxaban: guidance. NICE. 2012. [last accessed 2012 October]; Available from: http://www.nice.org.uk/guidance/TA256.

(23) National Collaborating Centre for Chronic Conditions. Atrial fibrillation: National clinical guideline for management in primary and secondary care. London: Royal College of Physicians. 2006.

(24) NHS Improvement Programme. Heart and stroke improvement. Commissioning for stroke prevention in primary care - the role of atrial fibrillation. NHS Improvement Programme. 2009. [last accessed 2012 October]; Available from: http://www.improvement.nhs.uk/heart/Portals/0/documents2009/AF_Commissioning_Guide_v2.pdf

(25) Connolly SJ, Eikelboom J, Joyner C, Diener HC, Hart R, Golitsyn S, *et al.* Apixaban in patients with atrial fibrillation. *N Engl J Med* 2011 Mar 3;364**(9)**:806-17.

(26) Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, *et al*. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009 Sep 17;361**(12)**:1139-51.

(27) Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, *et al*. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011 Sep 8;365**(10)**:883-91.

(28) Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, *et al.* Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011 Sep 15;365**(11)**:981-92.

(29) Gamra H, Naditch-Brule L, Chiang CE, Lewalter T, Murin J, Rosenqvist M. Management of valvular atrial fibrillation in real-life practice: Insights from the RealiseAF survey (abstract). *Circulation*. 2011;124: A12992.

(30) National Institute for Health and Clinical Excellence. Stroke and systemic embolism (prevention, non-valvular atrial fibrillation) - apixaban: final scope - June 2012. NICE. 2012.

(31) Committee for Medicinal Products for Human Use (CHMP). Eliquis (apixaban): Summary of opinion (post authorisation). European Medicines Agency. 2012. [last accessed 2012 October]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/002148/WC500132869.pdf

(32) Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* 2012;**(3)**:80-97.

(33) Waldo AL. Treatment of atrial flutter. *Heart* 2000 Aug;84*(2)*:227-32.

(34) Bristol-Myers Squibb Company and Pfizer. Clinical Study Report for Study CV185030 - A phase 3, active (Warfarin) controlled, randomized, double-blind, parallel arm study to evaluate the efficacy and safety of apixaban in preventing stroke and systemic embolism in subjects with non-valvular atrial fibrillation. Data on file 2011.

(35) Bristol-Myers Squibb Company, Pfizer. Clinical Study Report for Study CV185048 - Apixaban versus acetylsalicylic acid (ASA) to prevent stroke in atrial fibrillation patients who have failed or are unsuitable for vitamin K antagonist treatment: A randomized double-blind trial. Data on file 2011.

(36) Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare. Centre for Reviews and Dissemination. 2011. [last accessed 2012 October]; Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

(37) National Institute for Health and Clinical Excellence. Single Technology Appraisal (STA): Specification for manufacturer/sponsor submission of evidence. National Institute for Health and Clinical Excellence. 2012. [last accessed 2012 October]; Available from: http://www.nice.org.uk/media/D6D/E2/SpecificationForManufacturerSponsorSubmissionEvidence.doc

(38) Bristol-Myers Squibb (BMS), Pfizer. Protocol for: Granger CB, Alexander JH, McMurray JJV, *et al*. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981-92. [last accessed 2012 October]; Available from: http://www.nejm.org/doi/suppl/10.1056/NEJMoa1107039/suppl_file/nejmoa1107039_protocol.pdf

(39) Snapin SM. Noninferiority trials. *Curr Control Trials Cardiovasc Med* 2000;1:19-21.

(40) Spackman E, Burch J, Faria R, Corbacho B, Fox D, Woolacott N. Dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation: Evidence Review Group Report. 2011.

(41) Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011. [last accessed 2012 October]; Available at: http://www.cochrane-handbook.org/

(42) Connolly SJ, Ezekowitz MD, Yusuf S, Reilly PA, Wallentin L. Newly identified events in the RE-LY trial. *N Engl J Med* 2010 Nov 4;363**(19)**:1875-6.

(43) Boehringer Ingelheim. Atrial fibrillation - dabigatran etexilate: Single Technology Appraisal, manufacturer submission. 2009 Oct.

(44) Bayer plc. Atrial fibrillation (stroke prevention) - rivaroxaban: Single Technology Appraisal, manufacturers submission. 2011 Aug.

(45) Freeman JV, Zhu RP, Owens DK, Garber AM, Hutton DW, Go AS, *et al*. Cost-effectiveness of dabigatran compared with warfarin for stroke prevention in atrial fibrillation. *Ann Intern Med* 2011 Jan 4;154**(1)**:1-11.

(46) Gage BF, Cardinalli AB, Albers GW, Owens DK. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. *JAMA* 1995 Dec 20;274**(23)**:1839-45.

(47) Shah SV, Gage BF. Cost-effectiveness of dabigatran for stroke prophylaxis in atrial fibrillation. *Circulation* 2011 Jun 7;123**(22)**:2562-70.

(48) Lee S, Anglade MW, Meng J, Hagstrom K, Kluger J, Coleman CI. Cost-effectiveness of apixaban compared with aspirin for stroke prevention in atrial fibrillation among patients unsuitable for warfarin. *Circ Cardiovasc Qual Outcomes* 2012 Jul 1;5**(4)**:472-9.

(49) Kamel H, Easton JD, Johnston SC, Kim AS. Cost-effectiveness of apixaban vs warfarin for secondary stroke prevention in atrial fibrillation. *Neurology* 2012 Oct 2;79**(14)**:1428-34.

(50) Sorensen SV, Kansal AR, Connolly S, Peng S, Linnehan J, Bradley-Kennedy C, *et al*. Cost-effectiveness of dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation: a Canadian payer perspective. *Thromb Haemost* 2011 May;105**(5)**:908-19.

(51) Pink J, Lane S, Pirmohamed M, Hughes DA. Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses. *BMJ* 2011;343:d6333.

(52) Gallagher AM, Rietbrock S, Plumb J, van Staa TP. Initiation and persistence of warfarin or aspirin in patients with chronic atrial fibrillation in general practice: do the appropriate patients receive stroke prophylaxis? *J Thromb Haemost* 2008 Sep;6**(9)**:1500-6.

(53) Cowan C. The management of AF. Presentation at Improving the management of atrial fibrillation: the role of GRASP-AF. 2011 Sep.

(54) Claassen DO, Kazemi N, Zubkov AY, Wijdicks EF, Rabinstein AA. Restarting anticoagulation therapy after warfarin-associated intracerebral hemorrhage. *Arch Neurol* 2008 Oct;65**(10)**:1313-8.

(55) Sorensen SV, DeWilde S, Singer DE, Goldhaber SZ, Monz BU, Plumb JM. Cost-effectiveness of warfarin: trial versus "real-world" stroke prevention in atrial fibrillation. *Am Heart J* 2009 Jun;157**(6)**:1064-73.

(56) Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. *Arch Intern Med* 1994 Jul 11;154**(13)**:1449-57.

(57) Ariesen MJ, Claus SP, Rinkel GJ, Algra A. Risk factors for intracerebral hemorrhage in the general population: a systematic review. *Stroke* 2003 Aug;34**(8)**:2060-5.

(58) Mohan KM, Crichton SL, Grieve AP, Rudd AG, Wolfe CD, Heuschmann PU. Frequency and predictors for the risk of stroke recurrence up to 10 years after stroke: the South London Stroke Register. *J Neurol Neurosurg Psychiatry* 2009 Sep;80**(9)**:1012-8.

(59) Mandema J. Meta-analysis of placebo, ASA and warfarin controlled studies in AF. 2011.

(60) Hylek EM, Go AS, Chang Y, Jensvold NG, Henault LE, Selby JV, *et al*. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med* 2003 Sep 11;349**(11)**:1019-26.

(61) Connolly S, Yusuf S, Budaj A, Camm J, Chrolavicius S, Commerford PJ, et al. Rationale and design of ACTIVE: the atrial fibrillation clopidogrel trial with irbesartan for prevention of vascular events. *Am Heart J* 2006 Jun;151**(6)**:1187-93.

(62) European Medicines Agency. Summary of product characteristics: Pradaxa (dabigatran etexilate). European Medicines Agency. 2011. [last accessed 2012 October]; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000829/human_med_000981.jsp&mid=WC0b01ac058001d124

(63) Boggon R, Lip GY, Gallagher AM, van Staa TP. Resource utilization and outcomes in patients with atrial fibrillation: a case control study. *Appl Health Econ Health Policy* 2012 Jul 1;10**(4)**:249-59.

(64) Scarborough P, Bhatnagar P. Coronary heart disease statistics 2010 edition; British Health Foundation Health Promotion Research Group, Department of Public Health, University of Oxford. 2012.

(65) Bronnum-Hansen H, Davidsen M, Thorvaldsen P. Long-term survival and causes of death after stroke. *Stroke* 2001 Sep;32**(9)**:2131-6.

(66) Henriksson KM, Farahmand B, Johansson S, Asberg S, Terent A, Edvardsson N. Survival after stroke--the impact of CHADS2 score and atrial fibrillation. *Int J Cardiol* 2010 May 14;141**(1)**:18-23.

(67) Huybrechts KF, Caro JJ, Xenakis JJ, Vemmos KN. The prognostic value of the modified Rankin Scale score for long-term survival after first-ever stroke. Results from the Athens Stroke Registry. *Cerebrovasc Dis* 2008;26**(4)**:381-7.

(68) Bronnum-Hansen H, Jorgensen T, Davidsen M, Madsen M, Osler M, Gerdes LU, *et al*. Survival and cause of death after myocardial infarction: the Danish MONICA study. *J Clin Epidemiol* 2001 Dec;54**(12)**:1244-50.

(69) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. National Institute for Health and Clinical Excellence. 2008. [last accessed 2012 October]; Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

(70) Khan TI, Kamali F, Kesteven P, Avery P, Wynne H. The value of education and self-monitoring in the management of warfarin therapy in older patients with unstable control of anticoagulation. *Br J Haematol* 2004 Aug;126**(4)**:557-64.

(71) Gage BF, Cardinalli AB, Owens DK. The effect of stroke and stroke prophylaxis with aspirin or warfarin on quality of life. *Arch Intern Med* 1996 Sep 9;156**(16)**:1829-36.

(72) Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 2011 Nov;31**(6)**:800-4.

(73) Lacey EA, Walters SJ. Continuing inequality: gender and social class influences on self perceived health after a heart attack. *J Epidemiol Community Health* 2003 Aug;57**(8)**:622-7.

(74) Thomson R, Parkin D, Eccles M, Sudlow M, Robinson A. Decision analysis and guidelines for anticoagulant therapy to prevent stroke in patients with atrial fibrillation. *Lancet* 2000 Mar 18;355**(9208)**:956-62.

(75) Department of Health. NHS reference costs 2010-2011. Department of Health. 2011.

(76) National Institute for Health and Clinical Excellence. Atrial fibrillation: the management of atrial fibrillation. Costing report. NICE Clinical Guideline no. 36. NICE. 2006. [last accessed 2012 October]; Available from: http://www.nice.org.uk/guidance/CG36

(77) Curtis L. Unit costs of health and social care 2011. Canterbury, UK: Personal Social Services Research Unit. 2011.

(78) Luengo-Fernandez R, Yiin GS, Gray AM, Rothwell PM. Population-based study of acute- and long-term care costs after stroke in patients with AF. *Int J Stroke* 2012 May 9.

(79) Beswick AD, Rees K, Griebsch I, Taylor FC, Burke M, West RR, *et al*. Provision, uptake and cost of cardiac rehabilitation programmes: improving services to under-represented groups. *Health Technol Assess* 2004 Oct;8**(41)**:iii-x, 1.

(80) Electronic Drug Tariff. Department of Health by the NHS Business Services Authority, NHS Prescription Services. 2012 Aug.

(81) Health and Social Care Information Centre. Prescribing and Primary Care. Prescription cost cnalysis: England 2011 (Excel Format). Health and Social Care Information Centre. 2012.

(82) National Institute for Health and Clinical Excellence. NICE Clinical Guideline 48. MI: secondary prevention. Costing template: Implementing NICE guidance. NICE. 2007. [last accessed 2012 October]; Available from: http://www.nice.org.uk/guidance/CG48

(83) National Institute for Health and Clinical Excellence. NICE guidance CG17. Dyspepsia: Managing dyspepsia in adults in primary care. NICE. 2012. [last accessed 2012 October]; Available from: http://www.nice.org.uk/guidance/CG17

(84) Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Making* 2003 Jul;23**(4)**:341-50.

(85) Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al*. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004 Sep;8**(36)**:iii-xi, 1.

(86) Edwards S, Hamilton V, Nherera L, Trevor N, Barton S. Rivaroxaban for the prevention of stroke and systemic embolism in people with atrial fibrillation: Evidence Review Group report. 2011.

(87) Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health* 2012 Sep;15**(6)**:812-20.

9 APPENDICES

9.1 PRISMA diagrams for RCT evidence SR and non-RCT evidence SR

Figure 12. PRISMA diagram for RCT evidence SR (reproduced from MS; Figure 35; pg 202)

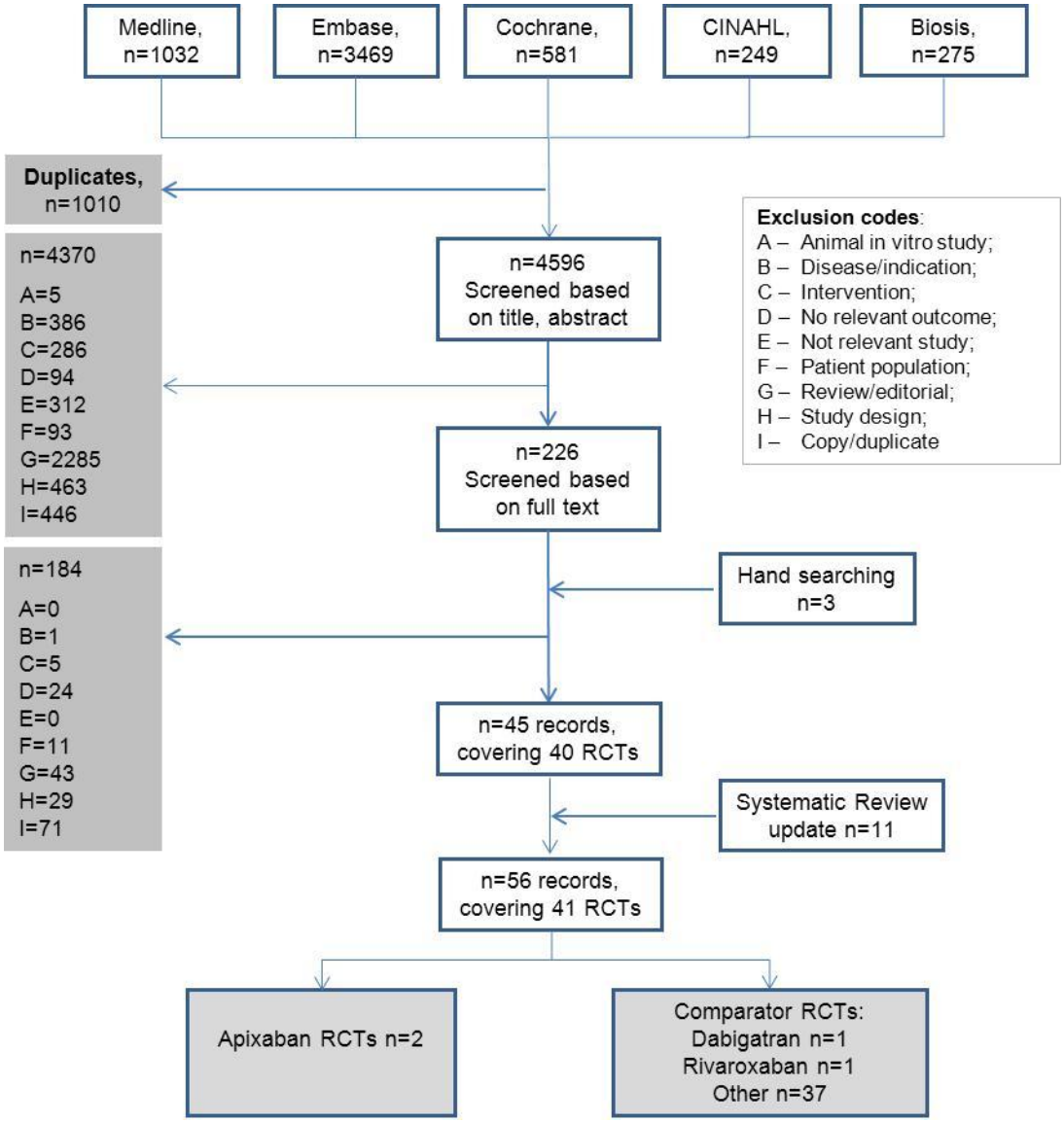
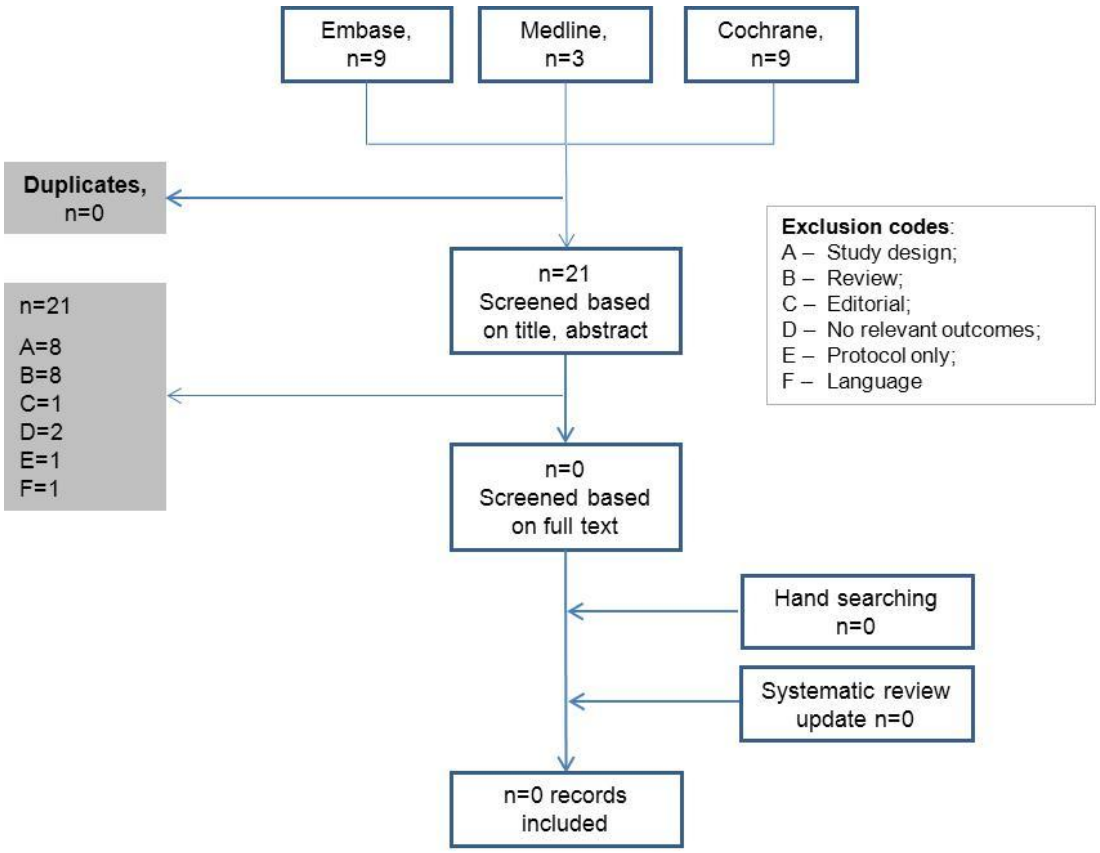


Figure 13. PRISMA diagram for non-RCT evidence SR (reproduced from MS; Figure 36; pg 212)



9.2 Quality assessment of RCTs

Table 72 Quality assessment of ARISTOTLE (reproduced from MS; Table 105; pg 203)

ARISTOTLE[28]		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/ not clear/NA)**
Was randomisation carried out appropriately?	At the time of enrolment, each subject was assigned a unique sequential subject number via IVRS. Subjects were randomised 1:1 to apixaban or warfarin via IVRS. Randomisation was stratified by investigative site and prior warfarin/VKA status (experienced or naive). Subjects were randomised in blocks of 2.	Yes
Was the concealment of treatment allocation adequate?	Study medications were prepared in a double dummy design using placebo matching the active treatments. Dosing for warfarin/warfarin-placebo was based on INR monitoring using a blinded, encrypted, point-of-care INR device. An algorithm was provided to guide the adjustment of the warfarin dose	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	The two treatment groups were well balanced with respect to both baseline demographic and disease characteristics.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Subjects, investigators, administrative/adjudication committees, and the Sponsor's staff conducting the study were blind to treatment assignments.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Fewer subjects discontinued study drug permanently in the apixaban group (25.3%) than in the warfarin group (27.5%) (p=0.001). The most common reasons for discontinuation in both treatment arms were subject's request to discontinue study treatment and AEs.	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in the clinical study report.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The primary and secondary efficacy analyses included all patients who underwent randomisation (ITT). The analyses of bleeding events included all patients who received at least one dose of a study drug. This was considered appropriate.	Yes

Table 73 Quality assessment of AVERROES (reproduced from MS; Table 106; pg 204)

AVERROES[25]		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/ not clear/NA)**
Was randomisation carried out appropriately?	At the time of enrolment, each subject was assigned a unique sequential subject number via IVRS. Subjects were randomised 1:1 to apixaban or aspirin via IVRS. Randomisation was stratified by study site. The subjects were randomised in blocks of 4.	Yes
Was the concealment of treatment allocation adequate?	Study medications were prepared in a double dummy design using placebo matching the active treatments.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	The treatment groups were well balanced for the baseline characteristics and physical measurements with no clinically relevant differences for randomised subjects.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Subjects, investigators, administrative/adjudication committees, and the Sponsor's staff conducting the study were blind to treatment assignments.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Fewer subjects discontinued study drug permanently in the apixaban group (19.9%) than in the aspirin group (23.3%). The most common reasons for discontinuation in both treatment arms were subject's request to discontinue study treatment and AEs.	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in the clinical study report.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	All primary efficacy and safety analyses were based on the intention-to-treat (ITT) principle. This was considered appropriate.	Yes

9.3 AVERROES trial

9.3.1 Description of AVERROES trial

AVERROES was a phase III, active-controlled, randomised, double-blind, double-dummy study. The primary objective of AVERROES was to determine whether apixaban 5 mg BD was superior to aspirin (81–324 mg OD) for preventing the composite outcome of stroke or SE. The patient population of AVERROES had AF and at least one additional risk factor for stroke. In addition, patients included in AVERROES had failed on or were unsuitable for VKA therapy.

AVERROES population

AVERROES was conducted in 36 countries and included 18 UK sites. The inclusion and exclusion criteria for AVERROES are listed in Table 74 and the baseline characteristics of the resulting randomised trial population are reported in Table 75. The ERG agrees with the manufacturer's assessment that the treatment groups in AVERROES appear well balanced for the baseline characteristics and physical measurements.

Table 74: Inclusion/exclusion criteria of AVERROES (adapted from MS; Table 9; pg 40)

Trial	Inclusion criteria	Exclusion criteria
AVERROES	Male or females ≥50 years of age, with documented permanent, paroxysmal or persistent AF, presenting with ≥1 risk factor for stroke, and not currently receiving VKA therapy. Risk factors for stroke: Prior stroke or transient ischaemic attack Age ≥75 years Arterial hypertension on treatment Diabetes mellitus Heart failure (NYHA class 2 or higher at time of enrolment) Left ventricular ejection fraction of 35% or less Documented peripheral arterial disease	Presence of conditions other than AF for which the patient required long-term anticoagulation Valvular disease requiring surgery A serious bleed in the previous 6 months or a high risk of bleeding Current alcohol or drug abuse or psychosocial issues Life expectancy of less than 1 year Severe renal insufficiency (a serum creatinine level of >2.5 mg/dL or a calculated creatinine clearance of <25 mL/min) Alanine aminotransferase or aspartate aminotransferase level >2x ULN or a total bilirubin >1.5x ULN Allergy to aspirin
Abbreviations: AF, atrial fibrillation; ECG, electrocardiogram; MI, myocardial infarction; NYHA, New York Heart Association; SE, systemic embolism; ULN, upper limit of normal; VKA, vitamin K antagonist		

Table 75: Characteristics of participants in AVERROES across randomised groups (reproduced from MS; Table 11; pg 43)

	Apixaban (N=2808)	**Aspirin (N=2791)**
Age (years), mean±SD	69.7±9.44	70.0±9.71
Gender male, n (%)	1660 (59.1)	1617 (57.9)
Region n (%)		
North America	408 (15)	396 (14)
Latin America	589 (21)	596 (21)
Western Europe	625 (22)	633 (23)
Eastern Europe	639 (23)	611 (22)
Asia and South Africa	547 (19)	555 (20)
Mean BMI (kg/m^2)	28.4	28.2
Systolic blood pressure (mm Hg), mean ±SD	132±16	132±16
Baseline electrocardiographic findings n (%)		
Atrial fibrillation	1923 (68)	1894 (68)
Atrial flutter	19 (1)	20 (1)
Sinus rhythm	707 (25)	730 (26)
Paced or other rhythm	147 (5)	139 (5)
Left ventricular hypertrophy	490 (17)	498 (18)
Classification of atrial fibrillation n (%)		
Paroxysmal	760 (27)	752 (27)
Persistent	587 (21)	590 (21)
Permanent	1460 (52)	1448 (52)
Use of VKA within 30 days before screening n (%)	401 (14)	426 (15)
Use of aspirin within 30 days before screening n (%)	2137 (76)	2081 (75)
Risk factors for stroke n (%)		
Prior stroke or transient ischaemic attack	390 (14)	374 (13)
Hypertension, receiving treatment	2408 (86)	2429 (87)
Heart failure	1118 (40)	1053 (38)
NYHA class 1 or 2	932 (33)	878 (31)
NYHA class 3 or 4	186 (7)	175 (6)
Left ventricular ejection fraction ≤35%	144 (5)	144 (5)
Peripheral artery disease	66 (2)	87 (3)
Diabetes, receiving treatment	537 (19)	559 (20)
Mitral stenosis	64 (2)	50 (2)
$CHADS_2$ score at enrolment, n (%)		
0 or 1	1004 (36)	1022 (37)
2	1045 (37)	954 (34)
≥3	758 (27)	812 (29)
Mean score	2.0±1.1	2.1±1.1
Medication use at baseline n (%)		
ACE inhibitor or ARB	1790 (64)	1786 (64)
Verapamil or diltiazem	251 (9)	248 (9)

Beta-blocker	1563 (56)	1534 (55)
Digoxin	821 (29)	754 (27)
Amiodarone	298 (11)	328 (12)
Statin	883 (31)	879 (31)
Study dose of aspirin or aspirin-placebo		
81 mg	1816 (65)	1786 (64)
162 mg	718 (26)	750 (27)
243 mg	73 (3)	60 (2)
324 mg	193 (7)	184 (7)
Data not available	7 (<1)	11 (<1)
Study dose of 2.5 mg BD apixaban (or placebo)	179 (6)	182 (7)
Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin-receptor blocker; BD, twice daily; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; TIA, transient ischaemic attack; VKA, vitamin K antagonist		

The ERG requested further breakdown of the $CHADS_2$ distribution of patients at baseline during the clarification stage which the manufacturer provided (Table 76). The ERG notes that the majority of patients in AVERROES had a $CHADS_2$ score of [redacted]. Therefore, the ERG considers that the $CHADS_2$ score distribution of AVERROES is comparable with the UK population for whom apixaban treatment would be expected to be considered.

Table 76: Baseline $CHADS_2$ scores for AVERROES (adapted from manufacturer's response to clarification questions; Table 11; pg 13)

$CHADS_2$ score	**AVERROES**	
	Apixaban (N=2,807)	**Aspirin (N=2,791)**
	N (%)	**N (%)**
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
Abbreviations used in table: $CHADS_2$, Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score.		

The ERG considers it important to highlight that the AVERROES population was comprised of patients for whom warfarin was unsuitable. In addition, the ERG notes that this population were not specified in the NICE final scope.[30] The reasons why the patients in AVERROES were considered unsuitable for treatment with warfarin are listed in Table 77. The ERG notes that 40% of the

randomised patients had previously received but discontinued VKA therapy. In addition, for approximately 15% of the patients in each trial arm the only reason given for being VKA unsuitable was "patient's refusal". Therefore, the ERG considers that the population of AVERROES was not exclusively comprised of patients for whom treatment with warfarin or other VKAs was contraindicated. However, clinical advisors to the ERG highlighted that the reasons for VKA unsuitability given in ARISTOTLE were consistent with reasons given in UK clinical practice.

Table 77: Reasons for unsuitability of VKA therapy† (reproduced from MS; Table 12; pg 44)

Reason for unsuitability n (%)	Apixaban (N=2808)	Aspirin (N=2791)	Previous use of VKA (N=2216)	No previous use of VKA (N=3383)
Assessment that INR could not be maintained in the therapeutic range	465 (17)	468 (17)	932 (42)	–
AE not related to bleeding during VKA therapy	86 (3)	94 (3)	180 (8)	–
Serious bleeding event during VKA therapy	92 (3)	82 (3)	173 (8)	–
Assessment that INR could not or was unlikely to be measured at requested intervals	1196 (43)	1191 (43)	827 (37)	1560 (46)
Expected difficulty in contacting patient for urgent change in dose of VKA	322 (11)	331 (12)	167 (8)	486 (14)
Uncertainty about patient's ability to adhere to instructions regarding VKA therapy	437 (16)	405 (15)	262 (12)	580 (17)
Concurrent medications that could alter activity of VKA	50 (2)	53 (2)	33 (1)	70 (2)
Concurrent medications whose metabolism could be affected by VKA	35 (1)	46 (2)	19 (1)	62 (2)
Assessment that patient would be unable or unlikely to adhere to restrictions	134 (5)	141 (5)	127 (6)	148 (4)
Hepatic disease	13 (<1)	9 (<1)	4 (<1)	18 (1)
Mild cognitive impairment	85 (3)	86 (3)	56 (3)	115 (3)
Heart failure or cardiomyopathy	179 (6)	188 (7)	95 (4)	272 (8)
Other factors that could be associated with increased risk of VKA use	96 (3)	123 (4)	121 (5)	98 (3)
$CHADS_2$ score of 1 and VKA therapy not recommended by physician	590 (21)	605 (22)	458 (21)	737 (22)
Other characteristics indicating risk of stroke too low to warrant treatment with VKA	55 (2)	40 (1)	32 (1)	63 (2)
Patient's refusal to take VKA	1053 (38)	1039 (37)	819 (37)	1273 (38)
Other reasons	184 (7)	189 (7)	249 (11)	124 (4)
$CHADS_2$ score of 1 as only reason for unsuitability of VKA therapy	313 (11)	336 (12)	216 (10)	433 (13)
Patient's refusal to take VKA as only reason for unsuitability	421 (15)	394 (14)	199 (9)	616 (18)
Multiple reasons for unsuitability of VKA therapy	1444 (51)	1440 (52)	1436 (65)	1448 (43)
†The reason for unsuitability was missing for one patient in the apixaban group				

AVERROES intervention and comparator

Patients in AVERROES were randomised 1:1 to treatment with apixaban or aspirin via IVRS. The study medications were administered using a double dummy design with placebo tablets matched to the active treatments. In the MS, the manufacturer reported that "subjects, investigators, administrative/adjudication committees, and the Sponsor's staff conducting the study were blind to treatment assignments" (MS section 6.3.1).

The intervention under investigation in AVERROES was apixaban 5 mg BD (or 2.5 mg BD for patients with an increased risk of bleeding meeting the criteria for the lower dose) plus aspirin placebo tablet(s). The aspirin placebo dose was at the discretion of the investigator in keeping with the active treatment dose.

The comparator in AVERROES was aspirin at a dose of 81–342 mg a day (consisting of between one to four 81 mg tablets) plus apixaban placebo tablets. The aspirin dose was at the discretion of the investigator.

Patients taking a thienopyridine at baseline were excluded from AVERROES, although patients not taking a thienopyridine at baseline could be prescribed one during the study if an indication emerged. The use of potent inhibitors of CYP3A4, and other antithrombotic agents, were prohibited in AVERROES. In addition, study investigators were encouraged to discontinue any non-study aspirin that patients were taking prior to randomisation. During the clarification stage the manufacturer provided additional information on the number of patients who remained on non-study aspirin during randomised treatment. The manufacturer reported that at baseline [redacted] in the apixaban arm and [redacted] in the aspirin arm were taking aspirin 30 days prior to study start. On the day of randomisation, [redacted] and [redacted] of patients were still on non-study aspirin (acetylsalicylic acid [ASA]) in the apixaban and aspirin arms respectively. However, the ERG also notes that the manufacturer reported post study start, only [redacted] of patients in each arm took non-study aspirin for more than [redacted] of the time during the study.

AVERROES outcomes

The safety and efficacy outcomes reported in AVERROES were as follows:

- the composite of time to first occurrence of stroke (ischaemic or haemorrhagic) or SE during the treatment period;
- the composite of days from randomisation to first occurrence of stroke, SE, MI or vascular death;
- days from randomisation to first occurrence of all-cause death;
- the composite of days from first dose of study drug to first occurrence of major or CRNM bleeding;

- days from randomisation to first occurrence of any bleeding;
- occurrence of major bleeding.

Definitions of each outcome were the same as those used in ARISTOTLE (section 4.2.1). The ERG notes that data from the individual components of each of the outcomes were also reported in the MS. The ERG considers the inclusion of these outcomes to be appropriate as the individual rather than the composite outcomes are specified in the NICE final scope.[30]

The ERG notes that the efficacy and safety outcomes in AVERROES were adjudicated on the basis of pre-specified criteria by a clinical events committee. The clinical events committee was blinded to the patients' study-group assignments. The ERG considers that this has reduced the risk of investigator bias affecting the results in terms of outcome assessment.

The ERG acknowledges that data on the outcomes of TIA and HRQoL, specified in the final scope issued by NICE,[30] were not collected in AVERROES and thus could not be presented in the MS. The ERG considers that all other outcomes specified in the final scope issued by NICE[30] were captured in AVERROES and reported appropriately in the MS for this STA.

AVERROES subgroup analyses

A large number of subgroup analyses were specified *a priori* and carried out on the AVERROES trial data (Table 78).

Table 78: Subgroups assessed for primary efficacy and safety end points (adapted from MS; Table 15; pg 49)

Characteristic	Subpopulations in AVERROES
VKA unsuitable	Demonstrated; Expected
Reason VKA unsuitable	Subject refused treatment with VKA (only reason); $CHADS_2$ score =1 and physician does not recommend VKA (only reason); All others
Apixaban dose	2.5 mg BD or matching placebo; 5 mg BD or matching placebo
Aspirin dose	81 mg; 162 mg; 243 mg; 324 mg
Geographic region	North America; Latin America; Europe; Asia/Pacific
Age	<65 years; ≥65 to <75 years; ≥75 years
Gender	Male; Female
Race	White; Black or African American; Asian; Other
Ethnicity	Hispanic/Latino; Not Hispanic/Latino
Weight	≤60 kg; >60 kg
Body mass index	≤28 kg/m^2; >28 to 33 kg/m^2; >33 kg/m^2
Level of renal impairment	Severe or moderate: ≤50 mL/min; Mild >50 to 80 mL/min; Normal >80 mL/min
Number of risk factors	≤1; ≥2
$CHADS_2$ score	≤1; 2; ≥3
Prior stroke or TIA	Yes; No
Age ≥75 years	Yes; No
Diabetes mellitus	Yes; No
Hypertension requiring pharmacological treatment	Yes; No
Heart failure	Yes; No
Abbreviations: AF, atrial fibrillation; BD, twice daily; TIA, transient ischaemic attack; VKA, vitamin K antagonist	

The ERG notes that the subgroup analyses in AVERROES were limited to the primary efficacy and safety outcomes. In addition, the ERG notes that patients were not stratified at randomisation for any of the characteristics assessed within the subgroup analyses and that AVERROES was not statistically powered to draw conclusions for any of the subgroup analyses reported.

AVERROES follow-up

The double-blind treatment period of the AVERROES was anticipated to be completed after at least 226 subjects had a primary efficacy end point. However, the ERG notes that AVERROES was terminated early following a recommendation by the data and safety monitoring committee (DMC) because of "the superior efficacy of apixaban". This recommendation was made based on a review of the results from the first planned interim analysis of efficacy and the results of a further confirmatory analysis. The first analysis was reviewed by the DMC on 19th February 2010 and showed a treatment benefit in favour of apixaban for the primary outcome that exceeded four standard deviations. The

results of the confirmatory analysis were reviewed on 28th May 2010 and the p value was 0.000002. AVERROES was thus terminated on 28th May 2010 with all events occurring up to this date included in the primary analyses resulting in a mean duration of follow-up of 1.1 years.

9.3.2 Description and critique of statistical approaches used in AVERROES

The primary objective of AVERROES was to determine whether apixaban was superior to aspirin in the prevention of the composite outcome of stroke or SE.

In the protocol for AVERROES it was stated that 226 primary outcome events would be required for the study to have at least 90% power to detect a 35% relative risk reduction (RRR) of apixaban vs aspirin at the one-sided α=0.025. It was estimated that this would be achieved via an average 1.6 years of follow-up of at least 5,600 randomised subjects assuming a 1% loss to follow-up.

Formal interim analyses in AVERROES were planned when 50% and 75% of the primary efficacy events had accrued; the total primary efficacy outcomes calculated to be required for AVERROES was 226. Pre-specified stopping rules were defined and included modified Haybittle–Peto boundaries of 4 SD (log hazard ratio) for the primary outcome in the first half of the study. If this threshold was crossed, a confirmatory analysis was to be performed 3 months later, and if that analysis also crossed the specified boundary, the DMC could recommend that the trial be terminated. This occurred in AVERROES and resulted in the early termination of the study (i.e. prior to the occurrence of 226 primary efficacy outcome events).

Similar to ARISTOTLE, a hierarchical testing strategy was pre-specified in AVERROES to control the overall type 1 error in the study. The strategy was comprised of sequential testing provided each previous hypothesis was proven true for the following analyses:

- superiority of apixaban relative to aspirin for the primary efficacy end point;
- superiority of apixaban relative to aspirin for the secondary efficacy end point;
- superiority of apixaban relative to aspirin for the end point for all-cause death.

All of the primary efficacy analyses in AVERROES were based on the ITT population and the safety analyses were conducted on the population of treated-subjects.

Cox proportional-hazards modelling and log-rank testing were used for both the efficacy and safety analyses in AVERROES with hazard ratios the preferred outcome measure reported in the study publications and MS.

The ERG considers the manufacturer's approach to the statistical analysis of the data in AVERROES to be appropriate.

9.3.3 Results of AVERROES trial

AVERROES treatment compliance and discontinuations

There were 6,421 patients enrolled in AVERROES and 5,598 of these were randomised; 2,808 patients were randomised to treatment with apixaban and 2,791 to treatment with aspirin. The ERG notes that treatment with apixaban was associated with significantly fewer discontinuations compared with aspirin (17.9% in the apixaban group vs 20.5% in the aspirin group [HR 0.88; 95% CI: 0.78-0.99, p=0.03]).

AVERROES treatment effectiveness results

Apixaban demonstrated superiority to aspirin (p<0.001) for the prevention of stroke or SE in subjects with AF and at least one additional risk factor for stroke who were deemed to be unsuitable for VKA treatment (HR 0.45; 95% CI: 0.32–0.62; p<0.001). The results of AVERROES for the primary efficacy outcome of stroke or SE and its individual components are presented in Table 79.

Table 79: Summary of primary efficacy outcome – randomised subjects (reproduced from MS; Table 20; pg 60)

	Apixaban N=2,808		Aspirin N=2,791		Hazard ratio (95% CI)	p value
	Pts with event	Event rate†	Pts with event	Event rate†		
	no.	%/yr	no.	%/yr		
Primary outcome: stroke or SE	51	1.6	113	3.7	0.45 (0.32–0.62)	<0.001
Stroke	49	1.6	105	3.4	0.46 (0.33–0.65)	<0.001
Ischaemic	35	1.1	93	3.0	0.37 (0.25–0.55)	<0.001
Haemorrhagic	6	0.2	9	0.3	0.67 (0.24–1.88)	0.45
Unspecified	9	0.3	4	0.1	2.24 (0.69–7.27)	0.18
Disabling or fatal	31	1.0	72	2.3	0.43 (0.28–0.65)	<0.001
Systemic embolism	2	0.1	13	0.4	0.15 (0.03–0.68)	0.01

Abbreviations: Pts, patients; SE, systemic embolism; yr, year; CI, confidence interval
†The percent per year is the rate per 100 patient-years of follow-up. All analyses were based on the time to a first event; patients could have more than one event

The results of the secondary efficacy outcomes for AVERROES are presented in Table 80. Apixaban resulted in a statistically significant reduction in all of the composite outcomes, including that of stroke, SE, MI, or vascular death relative to aspirin (HR 0.66; 95% CI: 0.53–0.83, p=0.003). Apixaban also resulted in a reduction of MIs, all-cause deaths, vascular deaths and hospitalisations for a CV cause compared with aspirin although none of these results reached statistical significance.

Table 80: Summary of secondary efficacy outcomes – randomised subjects (reproduced from MS; Table 21; pg 61)

	Apixaban N=2808		Aspirin N=2791		Hazard ratio (95% CI)	p value
	Pts with event	Event rate†	Pts with event	Event rate†		
	no.	%/yr	no.	%/yr		
Stroke, SE, or death	143	4.6	223	7.2	0.64 (0.51–0.78)	<0.001
Stroke, SE, MI or death from vascular cause	132	4.2	197	6.4	0.66 (0.53–0.83)	<0.001
Stroke, SE, MI, death from vascular cause, or major bleeding event	163	5.3	220	7.2	0.74 (0.60–0.90)	0.003
MI	24	0.8	28	0.9	0.86 (0.50–1.48)	0.59
Death from any cause	111	3.5	140	4.4	0.79 (0.62–1.02)	0.07
Death from vascular cause	84	2.7	96	3.1	0.87 (0.65–1.17)	0.37
Hospitalisation for CV cause	367	12.6	455	15.9	0.79 (0.69–0.91)	<0.001

Abbreviations: CV, cardiovascular; MI, myocardial infarction; Pts, patients; SE, systemic embolism; yr, year; CI, confidence interval

†The percent per year is the rate per 100 patient-years of follow-up. All analyses were based on the time to a first event; patients could have more than one event

AVERROES safety and adverse events

Similar to ARISTOTLE, the adverse event data for AVERROES were reported for the treated population. However, the ERG notes that the primary study publication for AVERROES reported safety data based on the ITT population and that the manufacturer included these data in the MS (MS, Appendix 15).

The ERG notes that based on the treated population, apixaban was associated with more major bleeding, CRNM bleeding, minor bleeding and all bleeding events compared with aspirin (Table 81). In addition, the ERG notes that these differences were statistically significant (defined as $p<0.05$) for the outcomes of major or CRNM bleeding and all bleeding.

Table 81: Summary of bleeding outcomes – treated subjects (reproduced from MS; Table 31; pg 87)

	Apixaban (N=2,798)		Aspirin (N=2,780)		Hazard ratio (95% CI)	P Value
	Pts with event	Event rate[†]	Pts with event	Event rate		
	no.	%/yr	no.	%/yr		
Major bleeding	45	1.41	29	0.92	1.54 (0.96–2.45)	0.07
Fatal Intracranial	5	–	5	–	–	–
Bleeding into a critical site	22	–	12	–	–	–
Intracranial	11	–	11	–	–	–
Intraarticular	2	–	1	–	–	–
Intraocular	6	–	0	–	–	–
Pericardial	1	–	0	–	–	–
Intramuscular	1	–	0	–	–	–
Retroperitoneal	1	–	0	–	–	–
Major or CRNM bleeding	140	4.46	101	3.24	1.38 (1.07–1.78)	0.01
CRNM bleeding	98	–	74	–	–	–
Minor bleeding	200	–	153	–	–	–
All bleeding	325	10.85	250	8.82	130 (1.10–1.53)	0.002
Abbreviations: CI, confidence interval; CRNM, clinically relevant non-major; Pts, patients; no., number; yr, year						

The results of AVERROES suggest that the overall safety profile of apixaban is similar to that of aspirin based on the incidence of AEs (65.5% vs 69.2%, respectively) (Table 82). The most common AEs (defined by the manufacturer as occurring in >5% of subjects in either treatment group) were dizziness and dyspnoea. In addition, it is reported in the MS that most AEs were mild to moderate in severity in both treatment groups and that the incidence of AEs leading to study discontinuation was lower in the apixaban group compared with the aspirin group (9.5% vs 13.0%, respectively).

The ERG notes that the incidence of SAEs was also lower in the apixaban group (23.5%) compared with the aspirin group (28.9%). In addition, the manufacuturer reported that the most common SAEs occurred in the system organ classes of Cardiac Disorders, Infections and Infestations, and Nervous System Disorders. The SAEs generally occurred in similar frequencies in the two treatment groups with the exception of Nervous System Disorders; where the frequency was lower in the apixaban group compared with the aspirin group (3.0% vs 6.5%).

The manufacturer also provided data on liver function test (LFT) abnormalities in AVERROES. The ERG notes that the overall frequency of LFT elevations was low and that the incidence of LFT abnormalities was similar for the apixaban and aspirin treatment groups.

Table 82: Summary of adverse events – treated subjects (reproduced from MS; Table 32; pg 90)

Adverse events **Number (%) subjects**	**Apixaban** **(N=2,798)**	**Aspirin** **(N=2,780)**
AE	1833 (65.5)	1925 (69.2)
SAE	657 (23.5)	804 (28.9)
Bleeding AE	281 (10.0)	259 (9.3)
Discontinuation due to AEs	266 (9.5)	362 (13.0)
Abbreviations: AE, adverse event; SAE, serious adverse event		

AVERROES subgroup analyses

The results of the subgroup analyses of AVERROES for the primary efficacy and safety outcomes are presented in Tables 83 and 84, respectively. The manufacturer reports in the MS that "overall, the results within each subgroup were consistent with the primary efficacy results for the study". The manufacturer also highlights that "although the study was not designed to ensure adequate power for subgroup analyses, the upper bounds of the 95% CIs for the HR were <1 (when estimable) for most of the subgroup categories" (MS section 6.5.2).

In the MS it was reported that there were no statistically significant interactions between the treatment effects and patients characteristics in the subgroup analyses conducted. However, in response to clarification questions the manufacturer reported that there was a significant treatment-by-subgroup interaction ($p<0.05$) for the ethnicity and weight subgroups for the primary efficacy outcome (0.03 and 0.02, respectively). However, the manufacturer reported there were no significant subgroup interactions for major bleeding.

Table 83: AVERROES subgroup analysis results for the primary efficacy outcome (stroke or SE) (reproduced from manufacturer's response to clarification questions; Table 20; pg 20)

Subgroup	**Total no. of patients**	**Apixaban (n/N)**	**Aspirin (n/N)**	**HR (95% CI)**
Age				
<65 yr	1720	7/855	19/865	0.38 (0.16, 0.89)
65 to <75 yr	1987	24/1049	29/938	0.73 (0.43, 1.25)
≥75 yr	1891	20/903	65/988	0.34 (0.20, 0.56)
Sex				
Male	3277	26/1660	49/1617	0.52 (0.32, 0.83)
Female	2321	25/1147	64/1174	0.40 (0.25, 0.63)
Estimated GFR (ml/min)				
<50	Not reported	Not reported	Not reported	Not reported
50 to <80	Not reported	Not reported	Not reported	Not reported
≥80	Not reported	Not reported	Not reported	Not reported

$CHADS_2$ score				
≤1	2142	12/1066	19/1076	0.63 (0.31, 1.30)
2	1973	23/1037	43/936	0.49 (0.29, 0.81)
≥3	1483	16/704	51/779	0.35 (0.20, 0.61)
Prior stroke or TIA				
Yes	764	10/390	33/374	0.29 (0.14, 0.60)
No	4834	41/2417	80/2417	0.51 (0.35, 0.74)
Study aspirin dose				
<162 mg daily	3602	39/1816	85/1786	0.45 (0.31, 0.65)
≥162 mg daily	1978	12/984	27/994	Not reported
Previous VKA use				
Yes	2215	17/1108	52/1107	0.33 (0.19, 0.56)
No	3383	34/1699	61/1684	0.55 (0.36, 0.84)
Patient refused VKA				
Yes	2092	16/Not reported	40/Not reported	Not reported
No	3506	35/Not reported	73/Not reported	Not reported
Heart failure				
Yes	1810	19/920	35/890	0.52 (0.30, 0.91)
No	3788	32/1887	78/1901	(0.27, 0.62)

Abbreviations used in table: HR, Hazard ratio; 95% CI, 95% confidence interval; VKA, vitamin K antagonist; $CHADS_2$, Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score; yr, year; TIA, transient ischaemic attack.

Table 84: AVERROES subgroup analysis results for the primary safety outcome (major bleeding) using the treated population data set.

Subgroup	Total number of patients	Apixaban (n/N)	Aspirin (n/N)	HR (95% CI)
VKA unsuitable				
Demonstrated	[redacted]	[redacted]	[redacted]	[redacted]
Expected	[redacted]	[redacted]	[redacted]	[redacted]
Reason VKA unsuitable				
Subject refused	[redacted]	[redacted]	[redacted]	[redacted]
$CHADS_2$=1 / physician not required	[redacted]	[redacted]	[redacted]	[redacted]
All other reasons	[redacted]	[redacted]	[redacted]	[redacted]
Apixaban/apixaban placebo dose				
2.5 mg BD	[redacted]	[redacted]	[redacted]	[redacted]
5 mg BD	[redacted]	[redacted]	[redacted]	[redacted]
Study aspirin/aspirin placebo dose				
91 mg OD	[redacted]	[redacted]	[redacted]	[redacted]
162 mg OD	[redacted]	[redacted]	[redacted]	[redacted]

243 mg OD	[redacted]	[redacted]	[redacted]	[redacted]
324 mg OD	[redacted]	[redacted]	[redacted]	[redacted]
Geographic region				
North America	[redacted]	[redacted]	[redacted]	[redacted]
Latin America	[redacted]	[redacted]	[redacted]	[redacted]
Europe	[redacted]	[redacted]	[redacted]	[redacted]
Pacific/Asia	[redacted]	[redacted]	[redacted]	[redacted]
Age				
<65 yr	[redacted]	[redacted]	[redacted]	[redacted]
≥65 to <75 yr	[redacted]	[redacted]	[redacted]	[redacted]
≥75 yr	[redacted]	[redacted]	[redacted]	[redacted]
Gender				
Male	[redacted]	[redacted]	[redacted]	[redacted]
Female	[redacted]	[redacted]	[redacted]	[redacted]
Race				
White	[redacted]	[redacted]	[redacted]	[redacted]
Black/African American	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]
Ethnicity				
Hispanic/latino	[redacted]	[redacted]	[redacted]	[redacted]
Not Hispanic/latino	[redacted]	[redacted]	[redacted]	[redacted]
Weight				
≤60kg	[redacted]	[redacted]	[redacted]	[redacted]
>60kg	[redacted]	[redacted]	[redacted]	[redacted]
BMI				
≤28 kg/m^2	[redacted]	[redacted]	[redacted]	[redacted]
>28g/m^2 to 33 kg/m^2	[redacted]	[redacted]	[redacted]	[redacted]
>33kg/m^2	[redacted]	[redacted]	[redacted]	[redacted]
Level of renal impairment				
Severe/moderate	[redacted]	[redacted]	[redacted]	[redacted]
Mild	[redacted]	[redacted]	[redacted]	[redacted]
Normal	[redacted]	[redacted]	[redacted]	[redacted]
Number of risk factors for stroke				
≤1	[redacted]	[redacted]	[redacted]	[redacted]
≥2	[redacted]	[redacted]	[redacted]	[redacted]
$CHADS_2$ score				
≤1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥3	[redacted]	[redacted]	[redacted]	[redacted]
Prior stroke or TIA				

No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Age ≥75yr				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Diabetes mellitus				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension requiring treatment				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Heart failure				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations used in table: HR, Hazard ratio; 95% CI, 95% confidence interval; VKA, vitamin K antagonist; $CHADS_2$, Congestive heart failure, Hypertension, Age ≥75 [doubled], Diabetes, Stroke [doubled] risk score; BD, twice daily; OD, once daily; yr, year; TIA, transient ischaemic attack.				

The ERG would also have liked to investigate the efficacy of apixaban in patients for whom warfarin is contraindicated. This was because the reasons for patients in AVERROES being unsuitable for treatment with warfarin or other VKAs included patient refusal which the ERG does not consider to be a true contraindication to treatment with warfarin. However, insufficient information was supplied by the manufacturer for such analysis.

9.3.4 AVERROES Summary

AVERROES was one of two RCTs (ARISTOTLE and AVERROES) included in the clinical effectiveness section of the MS to provide clinical data on apixaban for this STA. The ERG considered only ARISTOTLE meets the inclusion criteria for this STA based on the final scope issued by NICE[30] and thus reporting of AVERROES is limited to the appendices of the ERG report.

The ERG considers the inclusion and exclusion criteria for AVERROES to be acceptable for addressing the trial's objectives. In addition, the ERG considers the baseline characteristics of the randomised populations in AVERROES appear well balanced between trial arms. The intervention was apixaban which is the focus of this STA, and the comparator aspirin. The ERG consider that apixaban vs aspirin was not a comparison of interest specified in the NICE final scope.[30] However, based on clinical advice, the ERG acknowledges that aspirin is utilised in clinical practice for some patients in the UK.

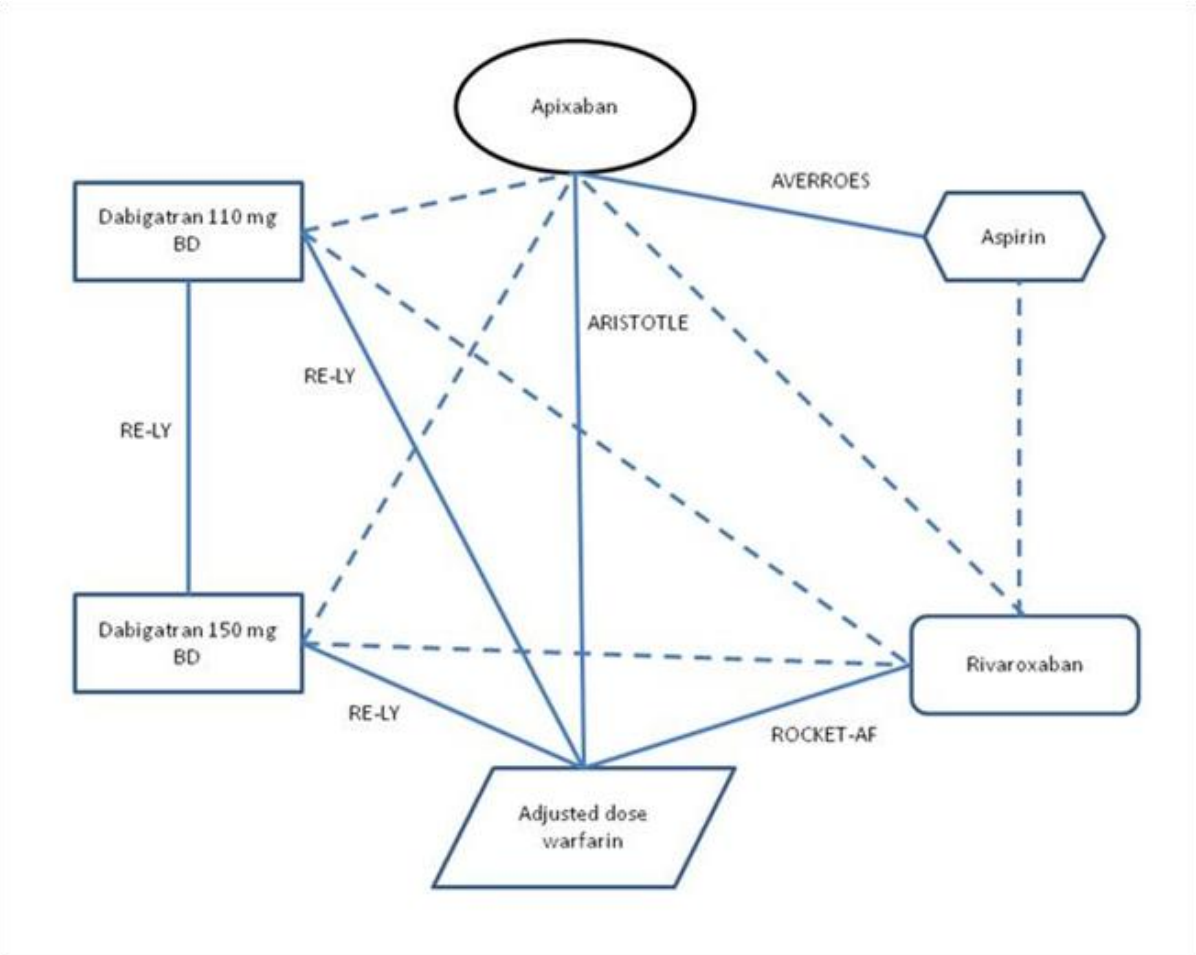
The ERG considers that the outcome data reported from AVERROES appears to be consistent with the data collected in the trial. However, the ERG note that the TIA and HRQoL data requested in the NICE final scope[30] were not collected in AVERROES.

In terms of follow-up and the statistical analysis of the data from the AVERROES, the ERG consider the duration of follow-up to be acceptable for the outcomes assessed and the statistical analysis plan to be suitable.

9.4 NMA 2

9.4.1 Network diagram for NMA 2

Figure 14: Network diagram for NMA 2 (warfarin suitable and unsuitable population) (reproduced from MS; Figure 10; pg 69)



9.4.2 Base case results of NMA 2

Table 85: NMA 2 (warfarin suitable and unsuitable population) base case analysis (reproduced from MS; Table 25; pg 76)

Outcome	Hazard ratio [95% CrI]				
	Apixaban vs dabigatran 150 mg	**Apixaban vs dabigatran 110 mg**	**Apixaban vs rivaroxaban**	**Apixaban vs warfarin**	**Apixaban vs aspirin**
Stroke + SE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Any stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Haemorrhagic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All-cause mortality	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatal stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disabling stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-disabling stroke	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICH	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Major bleeding	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GI bleeding	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other major bleed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRNM bleeding	NR[†]	NR[†]	[redacted]	[redacted]	[redacted]
Any bleeding	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinuations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Results shown in bold are significantly different; [†]Data for this outcome not reported for the RE-LY trial; Abbreviations: CrI, credibility interval; CRNM, clinically relevant non-major; ICH, intracranial haemorrhage; GI, gastrointestinal; MI, myocardial infarction; SE, systemic embolism

9.5 Manufacturer's incremental results in VKA unsuitable patient population

Table 86. Model outcomes compared with the clinical results of AVERROES (reproduced from MS; Table 78; pg 145)

Outcome	AVERROES events			Model events*		
	Apixaban (N=2,808)	Aspirin (N=2,791)	Incremental events on Aspirin	Apixaban (N=2,808)	Aspirin (N=2,791)	Incremental events on Aspirin
Primary outcome: stroke or SE	51	113	62	58§	131§	73
Stroke	49	105	56	55§	122§	67
Ischaemic	35	93	58	49	116	67
Haemorrhagic	6	9	3	6	6	0
SE	2	13	11	3	9	6
Death – any cause	111	140	29	112	138	26

Abbreviations: SE, systemic embolism.
* Approximation estimated at 1.15 years using patient characteristics from AVERROES.

Table 87. Base case incremental results - VKA unsuitable population (adapted from MS; Table 80; pg 146)

Treatment	Total			Incremental†			ICER (£/QALY)	
	Costs (£)	LYG	QALY	Costs (£)	LYG	QALY	Versus aspirin	Incremental
Aspirin	7,916	7.063	5.354	-	-	-	-	-
Dabigatran (150 & 110 mg)	8,228	7.357	5.635	312	0.294	0.281	1,111	1,111
Dabigatran (110 mg)	8,531	7.311	5.592	303	-0.046	-0.043	2,587	Strictly Dominated
Rivaroxaban	8,608	7.367	5.651	77	0.056	0.060	2,326	23,027
Apixaban	8,870	7.410	5.683	262	0.043	0.031	2,903	8,401

Abbreviations: LYG, life year gained; mg, milligram; QALY, quality adjusted life year; VKA, vitamin K antgonist.
†Versus the next least costly technology

9.6 Results of ERG sensitivity analyses

Treatment switching and discontinuation

Table 88. Incremental results following the assumption that warfarin would be given second-line

Technologies	Total			Incremental			ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Dabigatran 150 mg & 110 mg	7,816.57	7.84	6.03	-	-	-	-
Dabigatran 110 mg	8,081.00	7.79	5.99	264.43	-0.04	-0.04	Strictly dominated (by dabigatran blend)
Rivaroxaban	8,230.20	7.82	6.02	149.20	0.02	0.03	Strictly dominated (by dabigatran blend)
Apixaban	8,488.10	7.85	6.05	257.90	0.04	0.03	28,694.64
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.							

Table 89. Incremental results following the assumption that rivaroxaban would be given second-line

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	8,333.75	7.76	5.98	-	-	-	-	-
Dabigatran 150 mg & 110 mg	9,777.78	7.89	6.12	1,444.03	0.12	0.15	9,922.53	9,922.53
Dabigatran 110 mg	9,991.25	7.84	6.08	213.47	-0.04	-0.042	Strictly dominated	
Apixaban	10,086.95	7.89	6.13	95.69	0.05	0.05	11,637.41	60,365.68
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 90. Incremental results following the assumption that dabigatran 110 mg would be given second-line

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	8,420.30	7.73	5.95	-	-	-	-	-
Rivaroxaban	10,075.96	7.83	6.08	1,655.65	0.10	0.13	12,997.49	Extendedly dominated
Apixaban	10,170.40	7.87	6.10	94.45	0.03	0.02	286.78	286.78
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Mortality

Table 91. Incremental results following removal of higher "other-cause" mortality for warfarin patients during the within-trial period

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7166.80	7.44	5.68	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8375.41	7.48	5.74	1208.61	0.03	0.07	18444.94	Extendedly dominated
Dabigatran 110 mg	8621.96	7.45	5.71	246.56	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8714.98	7.50	5.77	93.01	0.05	0.05	17802.29	Extendedly dominated
Apixaban	8917.83	7.56	5.81	202.85	0.06	0.05	12829.47	12829.47
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Health-related quality of life

Table 92. Incremental results following removal of treatment-related disutility

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7188.49	7.47	5.75	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8436.53	7.54	5.81	1248.04	0.07	0.05	23774.16	Extendedly dominated
Dabigatran 110 mg	8683.84	7.50	5.77	247.31	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8778.37	7.55	5.83	94.53	0.05	0.05	21589.67	Extendedly dominated
Apixaban	8983.07	7.61	5.88	204.70	0.06	0.05	14530.26	14530.26
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 93. Incremental results following age adjustment of utilities

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7188.49	7.47	5.59	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8436.53	7.54	5.68	1248.04	0.07	0.09	13819.39	Extendedly dominated
Dabigatran 110 mg	8683.84	7.50	5.65	247.31	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8778.37	7.55	5.70	94.53	0.05	0.05	14283.12	Extendedly dominated
Apixaban	8983.07	7.61	5.75	204.70	0.06	0.05	11226.73	11226.73
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Treatment effectiveness

Table 94. Incremental results assuming stroke severity is independent of treatment received (equal to that observed with apixaban in ARISTOTLE)

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	6,712.06	7.50	5.74	-	-	-	-	-
Dabigatran 150 mg & 110 mg	7,878.42	7.58	5.84	1,166.36	0.08	0.10	11,663.60	Extendedly dominated
Dabigatran 110 mg	8,067.31	7.54	5.81	188.89	-0.04	-0.03	Strictly dominated	
Rivaroxaban	8,307.13	7.58	5.85	239.82	0.04	0.04	14,698.63	Extendedly dominated
Apixaban	8,477.45	7.63	5.89	170.32	0.05	0.04	11,863.08	11,863.08
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 95. Incremental results assuming bleed type is independent of treatment received (equal to that observed with apixaban in ARISTOTLE)

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7,264.28	7.44	5.67	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8,554.28	7.51	5.77	1,290.00	0.07	0.10	13,649.70	Extendedly dominated
Dabigatran 110 mg	8,753.10	7.50	5.74	198.82	-0.02	-0.02	Strictly dominated	
Rivaroxaban	8,903.21	7.53	5.79	150.11	0.04	0.05	14,118.78	Extendedly dominated
Apixaban	9,018.76	7.61	5.85	115.55	0.08	0.06	9,771.15	9,771.15
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 96. Incremental results following application of risk adjustment factors (per decade) to patients on second-line treatment

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7,264.28	7.44	5.67	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8,554.28	7.51	5.77	1,290.00	0.07	0.10	13,649.70	Extendedly dominated
Dabigatran 110 mg	8,753.10	7.50	5.74	198.82	-0.02	-0.02	Strictly dominated	
Rivaroxaban	8,903.21	7.53	5.79	150.11	0.04	0.05	14,118.78	Extendedly dominated
Apixaban	9,314.91	7.61	5.85	411.70	0.08	0.06	11,420.49	11,420.49
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Model structure

Table 97. Incremental results following application of future stroke risks for patients who have experienced an SE

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7201.36	7.46	5.69	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8449.75	7.53	5.78	1248.40	0.07	0.09	13680.92	Extendedly dominated
Dabigatran 110 mg	8696.82	7.49	5.75	247.07	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8790.65	7.54	5.80	93.84	0.05	0.05	14045.14	Extendedly dominated
Apixaban	8994.67	7.61	5.85	204.02	0.06	0.05	10982.25	10982.25
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 98. Incremental results following application of future stroke risks for patients who have experienced an MI

Technologies	Total			Incremental†			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7340.83	7.46	5.69	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8633.08	7.53	5.78	1292.25	0.07	0.09	14372.31	Extendedly dominated
Dabigatran 110 mg	8876.30	7.49	5.75	243.22	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8935.52	7.55	5.80	59.22	0.05	0.06	14100.12	Extendedly dominated
Apixaban	9134.42	7.61	5.85	198.90	0.06	0.05	10974.91	10974.91
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Resource use and costs

Table 99. Incremental results following application of acute SE costs from dabigatran submission[43]

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7169.07	7.47	5.70	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8416.36	7.54	5.79	1247.29	0.07	0.09	13640.24	Extendedly dominated
Dabigatran 110 mg	8663.90	7.50	5.76	247.54	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8759.28	7.55	5.81	95.38	0.05	0.05	14074.43	Extendedly dominated
Apixaban	8964.36	7.61	5.86	205.08	0.06	0.05	11012.05	11012.05
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Table 100. Incremental results following application of acute SE costs from rivaroxaban submission[44]

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7152.47	7.47	5.70	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8399.13	7.54	5.79	1246.66	0.07	0.09	13633.31	Extendedly dominated
Dabigatran 110 mg	8646.87	7.50	5.76	247.74	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8742.97	7.55	5.81	96.10	0.05	0.054	14076.94	Extendedly dominated
Apixaban	8948.37	7.61	5.86	205.40	0.06	0.05	11015.79	11015.79
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Perspective, time horizon and discounting

Table 101. Incremental results following application of a 26 year time horizon

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7185.87	7.47	5.70	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8433.87	7.53	5.79	1248.00	0.07	0.09	13650.31	Extendedly dominated
Dabigatran 110 mg	8681.21	7.50	5.75	247.34	-0.03	-0.03	Strictly dominated	
Rivaroxaban	8775.69	7.55	5.81	94.47	0.05	0.05	14076.13	Extendedly dominated
Apixaban	8980.27	7.61	5.86	204.59	0.06	0.05	11013.85	11013.85
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

9.7 Results of ERG exploratory analyses

Discontinuation

Table 102. Incremental results following the assumption that apixaban and dabigatran (150 mg/110 mg) are associated with the same level of other-cause discontinuation

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7,188	7.47	5.70	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8,708	7.59	5.84	1,519.63	0.12	0.14	10,552	10,552
Dabigatran 110 mg	8,778	7.55	5.81	70.25	-0.04	-0.03	14,071	Strictly dominated
Rivaroxaban	8,967	7.55	5.80	188.29	-0.01	-0.01	17,481	Strictly dominated
Apixaban	8,983	7.61	5.86	16.41	0.07	0.06	11,008	14,456
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.								

Risk of MI

Table 103. Incremental results following the assumption that apixaban and dabigatran (150 mg/110 mg) are associated with the same risk of MI

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	7,188	7.47	5.70	-	-	-	-	-
Dabigatran 150 mg & 110 mg	8,498	7.57	5.82	1,309.53	0.10	0.12	10,941	10,941
Dabigatran 110 mg	8,744	7.53	5.78	245.81	-0.04	-0.03	18,155	Strictly dominated
Rivaroxaban	8,778	7.55	5.81	34.54	0.02	0.03	14,071	Strictly dominated
Apixaban	8,983	7.61	5.86	204.70	0.06	0.05	11,008	11,191
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; MI, myocardial infarction; QALYs, quality adjusted life-years.								

Gamma distribution for HR of other-cause mortality

Table 104. Probabilistic incremental results following the use of Gamma distribution to sample the hazard ratio of other-cause mortality

Technologies	Total			Incremental			ICER (£/QALY) versus warfarin	ICER (£) Incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs		
Warfarin	5,535	5.75	4.38	-	-	-	-	-
Dabigatran 150 mg & 110 mg	6,661	5.80	4.46	1,126	0.05	0.08	14,923	Extendedly dominated
Dabigatran 110 mg	6,921	5.78	4.44	260	-0.01	-0.02	24,187	Strictly dominated
Rivaroxaban	6,966	5.80	4.46	46	0.02	0.02	17,455	Extendedly dominated
Apixaban	7,060	5.83	4.49	93	0.03	0.02	14,281	14,281

Abbreviations used in table: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mg, milligram; QALYs, quality adjusted life-years.