ID726 STA ERRATUM

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Apixaban for the treatment and secondary prevention of deep vein thrombosis and/or pulmonary embolism

ID 726

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 14/15/01

Completed 21st January 2015

CONTAINS CIC/AIC

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

This document contains erratum in respect of the ERG report: Apixaban for the treatment and secondary prevention of deep vein thrombosis and/or pulmonary embolism. This document is provided in response to the factual error check submitted by the company.

1.6 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG is satisfied with the search strategy employed by the company to identify cost effectiveness studies, and is reasonably confident that no other relevant published articles exist.

A single economic model, constructed in MS Excel, has been submitted by the company. However, the decision problem encompasses two distinct assessments based on the results from two clinical trials in two quite different populations over two different time periods. The ERG considers that two separate models should have been developed to avoid compromising either assessment.

The company failed to provide sufficient information to allow the ERG to validate many of the company model parameters. The company only provided two of the 12 K-M analyses that were requested as part of the clarification process. Analyses of these two data sets identified parameter estimation errors and a mis-match between the model and an important trial result. There are, therefore, concerns that parameter values derived from the outstanding ten K-M analyses that were not carried out by the company may be similarly flawed.

The ERG is concerned about the use, within the model, of a number of the RR values generated by the company's secondary prevention period NMA. These RR values are unrealistic and lead to the generation of ICERs per QALY gained that appear to unduly favour treatment with apixaban.

In addition, the company model does not incorporate the impact of an accumulating history of recurrent VTE events on future risks; an omission that may lead to unreliable estimates of long-term health gains for apixaban. Furthermore, the ERG considers that background mortality is inadequately represented, treatment costs have been underestimated, incorrect UK utility values have been employed, and an excessive mean body weight estimate for a UK population has been applied.

1.7 End of life criteria

The company has not put forward a case for apixaban to be considered under NICE's end of life criteria. The ERG considers this to be appropriate.

- correction factors. This approach resulted in unrealistically high estimates of RR for some relevant comparisons; the company then used these values in the economic model. The ERG did not agree with the company's approach and asked the company to re-run the NMAs using a vague prior to allow the ERG to calculate revised ICERs
- There were small numbers of events for many trial outcomes and this resulted in wide confidence intervals for many of the RR estimates in the company's NMAs

Cost effectiveness

- The decision problem encompasses two distinct assessments based on results from two clinical trials in two quite different populations. The company submitted a single model which risks compromising both assessments
- The company failed, both within the CS and during the clarification process, to provide sufficient information to allow the ERG to validate many of the parameters in the company's submitted model
- The use of unrealistic RR values (generated by the company's secondary prevention period NMA) leads to the generation of ICERs per QALY gained which appear to unduly favour treatment with apixaban
- The company model does not incorporate the impact of an accumulating history of recurrent VTE events on future risks; an omission that may lead to unreliable estimates of long-term health gains for apixaban
- The ERG considers that, within the company model, background mortality is inadequately represented, treatment costs have been underestimated, incorrect UK utility values have been employed and an excessive mean body weight estimate for a UK population has been applied

1.9 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG applied six specific model changes using the ERG's preferred alternative parameter values or formulae. These changes were in relation to age/sex mortality rates, treatment costs, age-varying utility, annual discounting, use of a body weight value based on an English population (rather than trial data) and rebasing of the model for assessment of treatment in the prevention period. The impact of each of these changes on the company's long-term base-case results for the comparisons between apixaban and either LMWH/VKA, rivaroxaban or LMWH/dabigatran were small, with changes to the ICERs per QALY gained of ±£2,700 in the most extreme cases. The impact of the ERG's changes on the company's short-term base-case results for the comparisons between apixaban and all three of the comparators was, again, modest with changes in the ICERs per QALY gained being less than ±£300. The ERG also generated model results using parameter values that had been generated by three different alternative methods of calculating RRs from the secondary prevention period NMA. In addition, the ERG carried out a single stand-alone scenario analysis. The results from all of the ERG's analyses should be viewed with caution due to the underlying uncertainty with regard to the magnitude of the various model parameter values that the ERG has been unable to verify.

Apixaban is formulated as film-coated tablets of either 2.5mg or 5mg. The recommended dose of apixaban for the treatment of initial VTE is 10mg taken orally twice daily for the first 7 days followed by 5mg taken orally twice daily.[39]

The recommended dose of apixaban for the secondary prevention of recurrent VTE is 2.5mg taken orally twice daily. When prevention of recurrent VTE is indicated, the 2.5mg twice daily dose should be initiated following completion of 6 months of treatment with either apixaban 5mg twice daily or with another anticoagulant.[39]

The Summary of Product Characteristics[39] (SPC) states that whether a short duration of treatment (at least 3 months) is required the decision should be based on the presence of transient risk factors (e.g., recent surgery, trauma, immobilisation). The SPC[39] further states that the duration of overall therapy should be individualised after careful assessment of the treatment benefit against the risk of bleeding.

Apixaban should be used with caution in patients with severe renal impairment (creatinine clearance 15-29 mL/min). It is not recommended in patients with creatinine clearance <15ml/min or in patients undergoing dialysis. [39]

3.3 Comparators

The final scope[16] issued by NICE states that the comparators to apixaban are:

- initial treatment with a LMWH or
- fondaparinux and continued VKA or
- rivaroxaban

3.3.1 Initial treatment with a LMWH

In the CS,[1] the evidence for the clinical effectiveness of treatment with apixaban versus LMWH and continued VKA is derived from the AMPLIFY[36] trial. Patients in the AMPLIFY[36] trial were treated for 6 months. The LMWH used in the AMPLIFY[36] trial is enoxaparin and the VKA used is warfarin. The CS[1] states that enoxaparin and warfarin are the standard treatments used in clinical practice in the NHS in England and Wales. Clinical advice to the ERG agrees with the company's statement.

Sensitivity analyses

For the primary efficacy analysis, no imputation methods were used in the case of missing data[redacted]

Secondary outcomes

The results for the secondary outcomes are provided in **Error! Reference source not found.**; secondary outcomes are categorised into efficacy, composite, or death outcomes. Analyses were performed in order to detect the superiority of apixaban in relation to enoxaparin/warfarin in terms of each of the secondary outcomes.

The company states that the results describing the secondary outcomes are consistent with the findings of the primary efficacy analysis; apixaban was not shown to be superior in comparison to enoxaparin/warfarin for the primary efficacy outcome ([redacted]) and similarly, apixaban was not shown to be superior for the secondary outcomes, recurrent VTE or all-cause death and recurrent VTE or cardiovascular (CV) related death. Superiority of apixaban was demonstrated for the composite outcomes of recurrent VTE, VTE-related death or major bleeding (p=0.001), and recurrent VTE, CV-related death, MI, stroke, major bleeding or CRNM bleeding ([redacted]).

Table 1 Base-case analysis for the secondary prevention period NMA

Outcome	Apixaban 2.5mg bd vs dabigatran 150mg od RR (95% CrI)	Apixaban 2.5mg bd vs rivaroxaban 20mg od RR (95% CrI)	Apixaban 2.5mg bd vs VKA (INR 2.0–3.0) RR (95% CrI)	Apixaban 2.5mg bd vs aspirin 100mg od RR (95% CrI)	Apixaban 2.5mg bd vs placebo RR (95% CrI)
Recurrent VTE or VTE-related death*					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Recurrent VTE-related death					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Major or CRNM bleeding					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Major bleeding					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRNM bleeding					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Overall treatment discontinuation					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All-cause mortality					
Original analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis requested by ERG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

bd=twice daily; CrI=credible interval; CRNM=clinically relevant non-major; ERG=Evidence Review Group; od=once daily; RR=relative risk; vs=versus; VKA=vitamin K antagonist; VTE=venous thromboembolism

*VTE or VTE-related death outcome is composed of the sum of the incidences of non-fatal PE, DVT and VTE-related death

Source: Company clarification response.

included in the secondary prevention period NMA have been produced using sufficiently similar assumptions; the ERG considers that there are likely to be more events in studies with longer treatment periods and follow-up times than in studies with shorter treatment periods and follow-up times. This is a fundamental flaw in the secondary prevention period NMA and the ERG considers that the efficacy (and safety) of apixaban versus any active comparator as secondary prevention is not yet confirmed and that the results of the secondary prevention period NMA analysis should be interpreted with caution. Therefore, the ERG considers that only the direct clinical effectiveness evidence for apixaban from the AMPLIFY-EXT[37] trial may be used to provide evidence for the use of apixaban as secondary prevention for recurrent VTE.

In addition, the ERG has concerns about the methodology used for the initial treatment period and secondary prevention period NMAs. The company applied continuity correction factors for instances of zero events for certain outcomes. This led to high estimates of RRs for some comparisons, which in turn were used in the economic model and so affected the size of the ICERs. For example, the company calculated the RR for major bleeding for apixaban versus rivaroxaban to be [redacted] in the secondary prevention period NMA, which the ERG considered to be extremely small.

In order to investigate the robustness of the company's estimates, the ERG asked for both the initial treatment period and secondary prevention period NMAs to be re-run using an alternative vague prior for the trial effect, and for the analysis to treat the trial effect as random rather than fixed. The company provided the requested analyses, and the ERG is satisfied that the results of the re-analysis yield less extreme estimates of underlying treatment effect than those provided in the company's original analyses, i.e. less likely to be overestimating or underestimating treatment effect. The new estimate for the RR for major bleeding for apixaban versus rivaroxaban in the secondary prevention period ([redacted]) is more acceptable than that provided in the company's original analysis. The results from the requested analyses were used in the ERG's calculations of ICERs.

A minor concern that the ERG has in relation to the initial treatment period NMA is that the description of the NMA data and analysis is unclear. Tables 58, 61 and 63 in the CS[1] all show RRs for 'aspirin' and 'no treatment' versus apixaban for the initial treatment period, although further clarification (via the factual error check) reveals that 'aspirin' and 'no treatment' were not included in the treatment period NMA and that these RRs were derived from the prevention period NMA. The ERG has tried to replicate the company's

findings from the initial treatment period NMA using the network of evidence as provided in Figure 10 of the CS,[1] but was unable to obtain the same results. In addition, the ERG is uncertain about how the estimates from the initial treatment period NMA were obtained and used in the economic evaluation.

4.9 Conclusions of the clinical effectiveness section

The main source of clinical evidence described in the CS[1] is derived from two key RCTs that included two distinct patient populations. Patients in the AMPLIFY[36] trial required initial treatment for a VTE event. The AMPLIFY[36] trial compared apixaban versus enoxaparin/warfarin in adult patients with symptomatic proximal DVT or PE (with or without DVT) over a 6-month period. Patients recruited to the AMPLIFY-EXT[37] trial have symptomatic, proximal DVT or PE (with or without DVT) and were treated with standard anticoagulation therapy for 6 to 12 months without VTE recurrence. Patients in AMPLIFY-EXT[37] were considered to be at clinical equipoise meaning that the risks and benefits of extended anticoagulation treatment were uncertain.

The results of the AMPLIFY[36] trial demonstrate that for the initial treatment period of VTE apixaban is non-inferior to, but not superior to, treatment with enoxaparin/warfarin for the primary efficacy outcome of recurrent VTE or VTE-related death (p<0.001). This finding is consistent across a range of secondary outcomes and pre-specified subgroup analyses. A statistically significant reduction in major bleeding events is reported for apixaban-treated patients compared with enoxaparin/warfarin-treated patients (RR 0.31; 95% CI 0.17 to 0.55; p<0.001). This finding is consistent across a range of secondary outcomes and pre-specified subgroup analyses. The rate of AEs between the two arms of the trial is similar.

The results of the AMPLIFY-EXT[37] trial demonstrate that for patients at clinical equipoise, apixaban significantly reduces the risk of recurrent VTE or all-cause death compared to placebo/no treatment (RR 0.83; 95% CI 0.22 to 0.48; [redacted]). This finding is consistent across a range of secondary outcomes and pre-specified subgroups. For the outcome of major bleeding, apixaban is not statistically significantly different compared to placebo/no treatment and this finding is consistent across a range of secondary outcomes and pre-specified subgroup analyses. The rate of AEs between the two arms of the trial is similar.

No HRQoL data were collected during either of the trials and therefore patients' experience of treatment with apixaban versus enoxaparin/warfarin or apixaban versus placebo/no treatment remains unknown.

Due to the absence of direct data for the comparisons of apixaban versus rivaroxaban and apixaban versus dabigatran for the initial treatment period of VTE, the company conducted

5 COST EFFECTIVENESS

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by Bristol Myers Squibb Pharmaceuticals Ltd and Pfizer Ltd in support of the use of apixaban for the treatment of DVT and PE, and the prevention of recurrent DVT and PE in adults. The two key components of the economic evidence presented in the CS[1] are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company also provided an electronic copy of their economic model (developed in Microsoft Excel).

5.2 ERG comment on company's review of cost effectiveness evidence

5.2.1 Objective of the company's cost effectiveness review

The company updated an existing systematic review conducted for the rivaroxaban NICE submission (TA287).[24] The TA287[24] review considered articles published between 1 July 2011 and 25 October 2012 and was an update and refinement of a review carried out to inform NICE CG144[20] guidance (Venous thromboembolic diseases: the management of venous thromboembolic diseases and the role of thrombophilia testing). The objective of the CG144[20] review was to identify "health economic evidence within published literature relevant to the review questions". The time period for the CG144[20] search had no lower limit and included articles published up until 1 August 2011.

Details of the cost effectiveness search strategies employed by the company are included in Appendix 10.33 of the CS.[1] Medline (via Ovid), Medline R-In Process (Ovid), EMBASE (Ovid), EconLit (EBSCO) and The Cochrane Library (incorporating the NHS Economic Evaluations Database (NHS EED) and the Health Technology Assessment (HTA) database were searched for published economic evaluations. The time horizon for the search was 25 October 2012 to 15 July 2014. Searches of grey literature and completed and on-going studies were also carried out. In addition, the NICE website was searched for potentially relevant company and ERG reports.

For patients who had discontinued treatment, the risks of recurrent VTE events beyond the end of the AMPLIFY-EXT[37] trial or, in the treatment duration scenario beyond the end of the AMPLIFY[36] trial, were based on those reported in a prospective cohort study.[4]

Other deaths

In the base case, deaths from causes unrelated to VTE recurrences, VTE complications or bleeding were modelled as age- and sex-dependent background general mortality rates based on data from the Office of National Statistics 2010-2012.[71] These background rates were adjusted by a hazard ratio (4.41 [95% CI 3.63 to 5.36])[72] to reflect the excess mortality associated with a diagnosis of VTE. The risk of death was assumed to be equal across treatments. However, patients entering the non-fatal IC bleed and CTEPH semi-absorbing health states were modelled as having a higher risk of mortality, HR 2.6 (95% CI 2.2 to 3.0)[73] and HR 1.3 (95% CI 0.98 to 1.73)[74] respectively

Health related quality of life

No HRQoL data were collected in either the AMPLIFY[36] or AMPLIFY-EXT[37] trials. In the model a baseline utility (0.825) was applied to all patients. The company reports that this figure was based on a UK population-average score from Kind et al[75] and was updated upon the occurrence of an IC bleed and CTEPH (permanent changes). In addition, utility decrements were added to the baseline utility for PE, DVT, non-IC major bleed, CRNM bleed or severe PTS events. Model utility scores for baseline, IC bleed, and CTEPH are presented in **Error! Reference source not found.** and utility decrements associated with clinical events and those associated with the use of anticoagulants are presented in Table 3.

Table 2 Summary of model quality of life values for baseline, IC bleed and CTEPH

Utility scores for baseline, IC bleed, and CTEPH	Utility score	Utility measurement	Source	Duration	Source
Baseline utility	0.825	EQ-5D	Kind et al1998[75] (NICE TA261;[23] NICE TA287[24])		
IC bleed					
Acute care period	0.33	TTO	Locadia et al 2004[76] (NICE TA261;[23] NICE TA287[24])	91 days	Locadia et al 2004[76]
Post-acute care	0.61	EQ-5D	Pickard et al 2004;[77] (NICE TA287[24])	Beyond 91 days	
CTEPH					
Acute care period	0.65	EQ-5D	Ghofrani et al 2013[78]	30 days	Ghofrani et al 2013[78]
Post-acute care	0.65			Beyond 30 days	

AC=anticoagulant; CRNM=clinically relevant non-major; CTEPH=chronic thromboembolic pulmonary hypertension; DVT=deep vein thrombosis; IC=intracranial; LMWH=low molecular weight heparin; PE=pulmonary embolism; PTS=post-thrombotic syndrome; SG=standard gamble; TTO=time trade-off; VKA=vitamin K antagonist Source: CS[1] Table 67

Table 3 Model utility decrements associated with clinical events and use of anticoagulants

Utility decrements associated with clinical events and use of ACs	Utility decrement	Utility measurement	Source	Duration	Source
PE	-0.32	TTO	Locadia et al 2004[76] (NICE TA261[23]; NICE TA287[24])	30 days	Hogg et al 2013[79]
DVT	-0.11				
Bleeding					
Major non-IC bleed	-0.30	TTO	Locadia et al 2004[76] (NICE TA261[23]; NICE TA287[24])	30 days	Expert opinion
CRNM bleed	-0.0054	EQ-5D	Sullivan et al 2011[80]	2 days	Dorian et al 2014[81]
Severe PTS	-0.07	SG	Lenert et al 1997[82] (NICE TA261;[23] NICE TA287[24])	Throughout	Lenert et al 1997[82]
Anticoagulation					
Apixaban	-0.002	TTO	Gage et al 1996[83]	Whilst on treatment	Gage et al 1996[83]
Rivaroxaban	-0.002				
LMWH/dabigatran	-0.002				
LMWH/VKA	-0.013				
Aspirin	-0.002				

AC=anticoagulant; CRNM=clinically relevant non-major; CTEPH=chronic thromboembolic pulmonary hypertension; DVT=deep vein thrombosis; IC=intracranial; LMWH=low molecular weight heparin; PE=pulmonary embolism; PTS=post-thrombotic syndrome; SG=standard gamble; TTO=time trade-off; VKA=vitamin K antagonist
Source: CS,[1] Table 67

5.5 Detailed critique of company's economic model

5.5.1 Model structure and design

Trial populations

The design of the company model includes the initial treatment of the index VTE event for the first 3-6 months, as well as subsequent preventive treatment with apixaban, or no preventive treatment over the remaining lifetime of surviving patients. Model variables are calibrated with data taken from the AMPLIFY[36] and AMPLIFY-EXT[37] clinical trials. This approach to model design is problematic since the populations included in the two trials[36,37] are clearly dissimilar. The recruitment criteria for AMPLIFY-EXT[37] explicitly exclude patients surviving a recurrent VTE event during earlier treatment of their index event. Moreover, only a third of AMPLIFY-EXT[37] patients were recruited from AMPLIFY,[36] of whom, presumably, half had previously received apixaban (i.e. a sixth of the total cohort).

This has implications for modelling the cost effectiveness of long-term preventive treatment with apixaban, since the model begins with a pure AMPLIFY[36] population then applies long-term outcome effectiveness parameters derived from a very mixed population from which all patients surviving a prior recurrent VTE have been excluded. This essential incompatibility of the two parts of the model seriously compromises the results of the long-term preventive assessment of apixaban.

Ideally, two distinct decision models should be available, each based exclusively on a single trial; short-term use of apixaban versus comparators using AMPLIFY[36] data, and long-term use of apixaban versus no treatment using AMPLIFY-EXT[37] data. The ERG has devised a model modification (rebase prevention model) to improve the long-term assessment results by effectively excluding the first two 3-month model cycles of the submitted model. This is adequate to reflect the experience of those patients (one-third) who received 6 months initial treatment in the AMPLIFY[36] trial before joining AMPLIFY-EXT[37], but may not give accurate results for other AMPLIFY-EXT[37] patients who may have received active treatment for up to 12 months following their index VTE event.

Age, sex and index event

The submitted model uses a Markov state-transition structure to project the expected experience of a mixed cohort of male and female patients of varying ages (18-99 in the AMPLIFY[36] clinical trial) and who may have experienced either an index DVT alone or an

Recurrent VTE events

The model structure adopted by the company incorporates features intended to avoid complexity and duplication of features when projecting patients' long-term experience. The methods employed incorporate implicit assumptions which warrant careful examination.

At first sight two distinct populations are represented by parallel sub-models of very similar structure which distinguish between patients by their index event (PE+/-DVT versus DVT only). However, these sub-models are in fact linked; it is assumed that experiencing a PE event at any time has more serious consequences than experiencing a DVT event. To reflect this assumption, index DVT patients who suffer a subsequent PE event are removed from the DVT sub-model and permanently transferred to the PE sub-model. However, index PE patients suffering any number of subsequent DVT events remain in the PE sub-model indefinitely. Therefore, over time there is a steady transfer of patients between the separate cohorts, a factor which must be carefully considered when interpreting model results.

A closely related feature of the company model is that any patient who suffers a non-fatal VTE event without permanent AEs (i.e. bleeding, PTS or CTEPH) is returned to the pool of index PE or DVT patients after one cycle of event-related treatment. This implicitly assumes that the risk of a first recurrent VTE event is the same as the conditional risk of experiencing a second or third VTE event. No evidence is offered in support of this assumption. Published evidence of recurrence risks with long-term follow-up is rare. However, in 1995 Beyth and colleagues[92] reported on 6 to 8 years follow-up of 124 DVT patients to ascertain the incidence of recurrent VTE events and mortality. Of the 124 patients, 18 (14.5%) suffered a first recurrent event (15 DVT and 3 PE), but six of these (33.3%) suffered at least one subsequent VTE event. The RR of second recurrence, compared with first recurrent, therefore increased significantly (RR 2.1, p=0.02). This finding was supported by the authors' Cox proportional hazards modelling that indicated that the risk of a first recurrence was significantly related to a baseline prior history of VTE (relative hazard 2.9, p=0.03).

The natural consequence of the company model design is, therefore, that long-term estimates of future VTE events (including deaths) will be increasingly underestimated, resulting in understated event costs and disutilities, and overestimated cohort survival. The company acknowledges this limitation in their model and considers that its effect is to understate the advantage from apixaban treatment through reduced risk of bleeding events. The ERG recognises this argument in relation to long-term apixaban treatment, but considers it of limited applicability to the assessment of short-term apixaban use where initial apixaban therapy would be replaced by conventional long-term therapy. However, there is

no obvious model modification available to the ERG that would correct this important weakness in model structure; the only reliable approach would be to redesign the model completely to incorporate progressively increasing VTE recurrence rates by introducing several additional patient states for those experiencing second and subsequent recurrent events. Populating such a model may prove challenging

5.5.2 Long-term mortality rates

In the submitted model, background mortality is applied to reflect deaths due to all causes other than VTE-related events as a function of age and gender. Based on a study describing the survival of patients suffering venous embolism by Flinterman et al,[72] a single overall risk ratio has been estimated by the ERG and applied to uplift all general population mortality risks irrespective of age and gender, to reflect increased mortality risks from causes unrelated to VTE.

However, this does not take into account the fact that the risks of recurrent VTE events and of mortality from all causes increase progressively with each additional VTE event. A paper by Klok et al[93] compared outcomes for PE and non-PE patients for up to 5.5 years. This showed that in the long-term (3+ years) mortality rates in all patients increased and accelerated over time. Patients with PE showed higher mortality rates than non-PE patients, and the long-term mortality hazard ratio (PE versus non-PE) increased by 2.7% per year.

Beyth et al[92] reported an analysis of mortality in non-cancer DVT patients indicating a statistically significant higher long-term mortality risk for patients aged 75 years or over, and for those with prior history of stroke (long-term hazard ratio 5.4). In addition, the long-term risk of recurrent VTE was three times greater for patients younger than 65 years who had a previous history of VTE.

Also, stroke and IC haemorrhage events can result in the development or worsening of disability, which is known to be associated with increased subsequent mortality. In the NICE appraisal of clopidogrel and modified-release dipyridamole for the prevention of occlusive vascular events (review of Technology Appraisal No. 90[94]) it was found that risks of all types of stroke, MIs, other vascular deaths and non-vascular deaths (i.e. background deaths) increased with age, and were also increased if the patient had suffered prior serious stroke-related disability (modified Rankin score 3+). The risk of suffering a new haemorrhagic stroke multiplied by 1.93 due to existing disability, and the risk of non-vascular death multiplied by 1.87. In addition, it was found that the odds ratio for fatality from MI and stroke events increased by 2.44 for a second event and by 5.56 for a third event. There is,

therefore, good reason to expect that the RR risk applied to standard life-table mortality risks may increase over time as additional VTE events accumulate among the 'at risk' population.

The impact on the model of increasing recurrent event incidence, fatality rates and general mortality rates over time would be to increase the net cost of further events while also reducing the average survival estimates in both trial arms. The net effect on cost effectiveness of incorporating these time-varying risk parameters into the submitted economic model is difficult to anticipate without much more detailed analysis of the AMPLIFY[36] and AMPLIFY-EXT[37] trial data together with data from longer-term trials of apixaban and comparator products.

5.5.3 Treatment costs

The cost of anti-thromboembolic therapies is calculated in the company's model for each 3-month cycle based on the average number of patients alive and on treatment during the cycle (i.e. the number of patients incurring costs is estimated as the average of those patients on treatment at the beginning of the cycle and those remaining on treatment at the end of the cycle). This method of calculation is inaccurate for both the first cycle and for subsequent cycles.

The first cycle cost includes the initial treatment of the DVT/PE event which generally lasts only a few days (5, 7-8 or a maximum of 21 days); all VTE patients entering the model incur this cost in full. After this initial phase, patients move to continuing therapy for the rest of the first cycle, using prescribed medication sufficient to treat them for the remainder of the first cycle. However, the company's model assumes that those patients discontinuing treatment at any time during the first cycle incur this cost only for half a cycle. The ERG considers that this understates the true cost since oral medications prescribed early in the cycle cannot be returned and must be discarded. The cost incurred for treatment in the first 3 months is more accurately estimated using the full number of patients who begin treatment.

For second and subsequent cycles of extended treatment, the ERG considers it reasonable to assume that all patients remaining on long-term medication would receive repeat prescriptions quarterly. The only exception would be for patients receiving LMWH monotherapy who require regular injections. In this case it would be more appropriate to calculate medication and administration costs averaged across each cycle (including the first cycle).

5.5.7 Revised relative risks

The company was asked during the clarification process to undertake alternative NMA calculations for selected model variables (see Clarification letter). These are discussed in detail in Section 4.8 of this ERG report. The resultant RR values estimated by the company for VTE recurrence and related deaths, major bleeding and CRNM have been applied as model revisions by the ERG. However, the ERG has continuing concerns relating to the estimation of RR parameters for major bleeding and CRNM derived from the secondary prevention period NMA. Model results are especially sensitive to these variables, and some parameter estimates in the company's base-case model appear excessively large (e.g. RR of [redacted] for major bleeding for rivaroxaban versus apixaban). As previously discussed, this may be related to the wide range of follow-up periods in the included trials. The ERG carried out two additional analyses using a Poisson assumption (relating risk to the time of follow-up exposure), using Bayesian and non-Bayesian methods (Appendix 7 of this ERG report).

The company's base-case ICER is reliant on the original company model RR parameters. Using the results of the analysis method requested by the ERG in the clarification process and the two Poisson-based methods, a range of alternative parameter value sets are available. The impact of these alternatives on cost effectiveness estimates are compared in Section 6.

5.5.8 Time-to-event analysis and projective model coefficients

The submitted model uses parameter values derived from the AMPLIFY[36] and AMPLIFY-EXT[37] clinical trials to represent the various AEs suffered by VTE patients during the trial follow-up periods (i.e. recurrent VTE events, VTE related fatality, death from other causes, major bleeding events, CRNM events, IC bleeding, CTPH and PTS), as well as treatment duration.

As neither the trial publications[36,37] nor the related CSRs[47,52] provide relevant information, the ERG submitted a set of 12 clarification requests for K-M time-to-event analysis relating to the main model variables used to populate the model. The ERG intended to use this information to address three important issues:

- to validate the model parameter values against the original data

- to assess whether temporal trends in the risk of key events are consistent with the modellers' general assumption that risks evident in the last 3-month cycle of AMPLIFY-EXT[37] follow-up can be projected indefinitely at a constant level
- to formulate alternative projective hazard functions, where varying hazard trends are evident in any of the time-to-event variables, as a basis for providing revised ICER per QALY gained estimates

Of the 12 analyses requested by the ERG, the company supplied only two. The two analyses relate to overall survival and time to treatment discontinuation in the AMPLIFY trial.[36] As the ERG received the analyses 8 days prior to the submission date for the ERG report, the time available for the ERG to assess the data was limited. A brief discussion of the data is presented in this section.

Time to treatment discontinuation

The ERG's comparison of the time to treatment discontinuation results from the AMPLIFY[36] trial with the corresponding Markov trace model results has revealed important discrepancies. After 90 days treatment [redacted] of apixaban patients had discontinued treatment, whereas the company model indicates only 7.7% discontinuations. For patients assigned to conventional therapy the trial data show [redacted] of patients had stopped treatment compared to 13.0% estimated by the model.

This finding confirms the ERG belief that all similar model parameters should be independently validated against the original trial data. Without access to the requested K-M analyses, it is not possible for the ERG to determine whether or not similar discrepancies are present in other time-to-event model variables, which may render the calculated ICERs unreliable.

5.5.9 Cost effectiveness summary

The ERG has reviewed the decision model submitted by the company and has identified a number of areas of concern:

- The decision problem encompasses two distinct assessments based on results from two clinical trials[36,37] in quite different populations. The model design attempts to encompass both assessments within one structure, despite the evident incompatibility of the two sets of evidence. It is the ERG's view that two separate models should have been developed to avoid compromising either assessment.
- The company acknowledges that no attempt has been made to incorporate the impact of accumulating history of recurrent VTE events through increasing future risks (including background mortality), as is attested in the literature. This may lead to unreliable estimates of long-term health gains for apixaban.
- The model structure operates on the assumption that all patients begin with the same average age, and that background mortality rates are adequately represented by a simple weighted average of life-table values. The ERG has demonstrated that this is inappropriate and leads to biased results which tend to overstate estimated ICERs per QALY gained.
- In the model, treatment costs have been underestimated, incorrect UK utility values have been employed, and an excessive mean body weight estimate for a UK population has been used.
- The ERG notes that the company's INR monitoring costs are in line with previous NICE appraisals of rivaroxaban[23,24] and dabigatran[25] and do not significantly impact the size of the ERG's estimated ICERs.
- The method used to derive hazard ratios for the long-term secondary preventive treatment period comparisons has been shown to yield unrealistic values in some instances, tending to improve the estimated ICERs per QALY gained in favour of apixaban. Alternative approaches considered by the ERG have been shown to produce very different ICERs per QALY gained in some cases, indicating substantial unresolved uncertainty.
- The company only provided two of the requested 12 time-to-event analyses requested by the ERG. One of these revealed a significant error in calculating the risk of treatment discontinuation, and the other showed a significant difference between survival gain estimated by the model and that shown in the AMPLIFY[36] trial results. The ERG considers that it is likely that other discrepancies and errors may remain undetected in the model but this cannot be confirmed

without access to the other ten trial results. The ERG considers that it is likely that other discrepancies and errors may remain undetected in the model but this cannot be confirmed without access to the other ten analyses of trial data (from both the AMPLIFY[36] and AMPLIFY-EXT[37] studies) requested by the ERG.

- Results of sensitivity analyses show that the main drivers of cost effectiveness are baseline utility and starting age (6-month model) and baseline utility and risk of major bleed (lifelong model). Starting age should be a parameter for which there is general agreement. The ERG considers that there is uncertainty around the most plausible baseline utility as no HRQOL data were collected during the trials.[36,37]

In summary, the ERG considers that the company model (even with the amendments/corrections proposed by the ERG) does not provide a reliable basis for estimating the relative cost effectiveness of apixaban for the treatment and secondary prevention of DVT and/or PE.

The ERG, however, is not confident that the methods used in the secondary prevention period NMA are robust and considers that the results from this NMA are not credible.

8.2.1 Issues relating to NMA methodology

The ERG was unable to reproduce the results of either the company's initial treatment period or the secondary prevention period NMA using the information included in the CS.[1] The ERG notes that the information about the comparators included in the initial treatment period NMA presented in the clinical section of the CS[1] is not the same as the information presented in the economics section. Specifically there are results for two additional comparators (no treatment and aspirin) listed in Table 58 of the economics section. The legend for Table 58 is not sufficiently detailed to allow understanding of any additional assumptions/analyses that have been carried out by the company to explain this discrepancy. Further clarification (via the company's factual error check) has revealed that 'aspirin' and 'no treatment' were not included in the calculations for the treatment period NMA.

There were small numbers of events for many of the outcomes reported in the AMPLIFY[36,37] trials and this led to wide confidence intervals for many of the RR estimates resulting from both initial treatment period NMA and the secondary prevention period NMA.

The ERG has an additional major concern about the methodology and assumptions underpinning the secondary prevention period NMA. There is marked variation in the lengths of the included trial treatment periods (before recruitment into the trial, the trial treatment period and also of the observation period after the end of treatment). However, the number of events, and hence RRs, for these different periods were used in the analyses as if all of the trials had the same follow-up period. To illustrate, in AMPLIFY-EXT[37] apixaban is used for a 12-month treatment period following an initial treatment period of 6 to 12 months; in other trials[2-4,6,8,15,36,37,78,82] the initial treatment period varies from 3 to 18 months and the trial treatment period varies from 6 months to longer than 2 years. The advice given to the company was that the RRs between different treatment options were not likely to differ over time. However, recent evidence from Martinez et al[17] demonstrates that the risk of recurrence is highest in the first 6 months after an index event but then falls progressively between 6 months and 4 years, at which point it plateaus. The trials included in the secondary prevention period NMA[37,58-66] evaluate the effects of different therapies at different time-points during this decline in risk and the ERG is not aware of any data to justify the assumption that a comparison of the results of these trials is valid.

8.2.2 Economic model

The ERG's critique of the company's submitted economic model was severely hampered by a lack of information. This lack of information meant that the ERG was unable to verify key model parameters. Although the company was asked, via the clarification process, to carry out a series of 12 K-M analyses, only two analyses were provided. Examination of these data revealed issues with both the company's calculation of the risk of treatment discontinuation, and a significant difference between the survival gain estimated by the model and that shown by the AMPLIFY[36] trial results. The fact that there are outstanding issues relating to the parameters that had been estimated from both of the provided data sets raises concern that there may be other issues relating to the ten analyses carried for which no results were made available to the ERG.

The company created one model, combining the risks and benefits resulting from treatment with apixaban over an initial treatment period of 6 months with those associated with long-term secondary prevention. Risks of VTE-related events calculated from AMPLIFY[36] data were used to estimate risks in the initial 6-month period, whilst risks for VTE-related events calculated from AMPLIFY-EXT[37] data were used to estimate long-term risks beyond 6 months. As previously discussed, the population in AMPLIFY-EXT[37] only represents patients for whom there is uncertainty regarding the risks and benefits of treatment as patients who had a recurrent VTE within 6 months of an index event were excluded from this trial. The ERG considers that the risks of recurrent events for the whole population receiving apixaban are, therefore, likely to have been underestimated in the company's economic model.

Furthermore, it is not clear whether the RRs for the comparator treatments described in the CS[1] are valid for patients who definitely require lifelong treatment (i.e. patients at overall higher risk of VTE recurrence). From the information in the CS,[1] it appears that the RRs for the comparator treatments are mainly derived from data relating to the clinical equipoise population (AMPLIFY-EXT[37] and EINSTEIN-EXT[58]). This is an important consideration as these RRs are used in the economic model to generate the (lifelong) ICERs per QALY gained. This means that the company's economic model may yield an ICER that is not relevant to all patients who require lifelong treatment, including those with persistent factors who are at higher risk of recurrence.

The results from the company's secondary prevention period NMA were used as the basis for assessing lifelong use of apixaban in the economic model despite some values being unrealistic and despite the follow-up data for apixaban only being available for 12 months. The ERG also had minor concerns about the modelling of background mortality risk, treatment costs, utility and the discounting methods used in the model.

The ERG acknowledges that their adjustments to the company's model had minimal effects on the sizes of the ICERs per QALY gained. However, the ERG is very concerned that their critique was restricted by the failure of the company to provide requested data. Unfortunately, the currently available cost effectiveness estimates cannot be considered robust unless the ERG is given the opportunity to check the parameter estimates and assumptions used in the company model.

8.3 Implications for research

Due to the paucity of long-term evidence for both the continued benefit and safety of treatment with NOACs, close monitoring and surveillance of patients undergoing treatment with NOACs is necessary.

In addition, access to, support and best use of currently available clinical data from patient registries and hospital audits should be encouraged in order to explore the substantial uncertainty around key clinical issues for patients with VTE.

A robust study of quality of life in patients treated with NOACs compared to patients treated with warfarin would be welcome.