EXTERNAL ASSESSMENT REPORT

Watch BP Home A for diagnosing and monitoring hypertension and detecting atrial fibrillation

Produced by NUTH and YHEC EAC

DR IAIN WILLITS, Medical Technologies Evaluator, NUTH
CHRISTOPHER A. REAY, Clinical Scientist, NUTH
KIM KELTIE, Research Scientist, NUTH
JOYCE CRAIG, Project Director, YHEC
STEVEN DUFFY, Research Associate, YHEC
DR ANDREW SIMS, Head of Clinical Measurement and Engineering Unit, NUTH

JANUARY 2012
Correspondence to:  
Dr Iain Willits
Medical Technologies Evaluator
Regional Medical Physics Department
Newcastle upon Tyne Hospitals NHS Foundation Trust
Freeman Hospital
Newcastle upon Tyne
NE7 7DN

Tel: +44 (0) 191 213 7787 (direct)
Fax: +44 (0) 191 213 0290

Declared interests of the authors: None

Contribution of authors

IW wrote the clinical evidence and economic evidence sections with the help of JC and AJS. CAR and SD replicated the literature searches and performed additional searches. KK and AJS performed additional statistical analysis and modelling.

The views expressed in this report are those of the authors and not those of NICE. Any errors are the responsibility of the authors.
# Contents

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iii</td>
</tr>
</tbody>
</table>

## Section 1: Summary

1.1 Scope of the Sponsor’s Submission ....................................................... 1
1.2 Summary of Clinical Evidence Submitted by the Sponsor ......................... 1
1.3 Summary Critique of Clinical Evidence Submitted by the Sponsor ............... 2
1.4 Summary of Economic Evidence Submitted by the Sponsor .......................... 2
1.5 Summary Critique of Economic Evidence Submitted by the Sponsor .............. 3
1.6 External Assessment Centre Commentary on the Robustness of Evidence Submitted by the Sponsor .......................................................... 3
1.7 Summary of any Additional Work Carried out by the External Assessment Centre .......................................................... 3

## Section 2: Background

2.1 Overview and Critique of Sponsor’s Description of Clinical Context ............. 5
2.2 Overview of Sponsor’s Description of Ongoing Studies ............................... 9
2.3 Critique of Sponsor’s Definition of the Decision Problem .......................... 9

## Section 3: Clinical Evidence

3.1 Critique of the Sponsor’s Search Strategy ................................................. 12
3.2 Critique of the Sponsor’s Study Selection ................................................. 16
3.3 Included and Excluded Studies .................................................................. 17
3.4 Overview of Methodologies of all Included Studies .................................... 19
3.5 Overview and Critique of Sponsor’s Critical Appraisal for each Study .......... 24
3.6 Results .................................................................................................. 24
3.7 Does Each Relevant Study Include the Patient Population(s), Intervention(s), Comparator(s) and Outcomes as Defined in the Final Scope? .................... 27
3.8 Description of the Adverse Events Reported by the Sponsor ....................... 29
3.9 Description and Critique of Evidence Synthesis and Meta-Analysis Carried out by the Sponsor .......................................................... 29
3.10 Additional Work Carried out by the External Assessment Centre in Relation to Clinical Evidence .......................................................... 30
3.11 Conclusions on the Clinical Evidence ...................................................... 32

## Section 4: Economic Evidence

4.1 Critique of the Sponsor’s Search Strategy ................................................. 36
4.2 Critique of the Sponsor’s Study Selection ................................................. 36
4.3 De Novo Cost Analysis ............................................................................ 37
4.4 Model Validation .................................................................................... 52
4.5 Interpretation of Economic Evidence ...................................................... 53
4.6 Additional Work Undertaken by the External Assessment Centre in Relation to Economic Evidence .......................................................... 55
4.7 Conclusions on the Economic Evidence .................................................. 57
Section 5: Conclusions
5.1 Conclusions on the Clinical Evidence
5.2 Conclusions on the Economic Evidence
5.3 Differences of Opinion Between EAC and Sponsor

Section 6: Implications for Research

Appendices:

Appendix A: Calculation of prevalence rate in population
Appendix B: Replication of Sponsor’s de novo economic results
Appendix C: Sensitivity Analysis of de novo economic model

Bibliography
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF</td>
<td>Atrial Fibrillation</td>
</tr>
<tr>
<td>EAC</td>
<td>External Assessment Centre</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Clinical Excellence</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>GP</td>
<td>General Practitioner</td>
</tr>
<tr>
<td>PSS</td>
<td>Personal Social Services</td>
</tr>
<tr>
<td>ABPM</td>
<td>Ambulatory blood pressure measurement</td>
</tr>
<tr>
<td>BP</td>
<td>Blood pressure</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiograph</td>
</tr>
<tr>
<td>HBPM</td>
<td>Home Blood Pressure Measurement</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative predictive value</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
</tbody>
</table>
Section 1: Summary

1.1 SCOPE OF THE SPONSOR’S SUBMISSION

The Sponsor (Microlife) has presented evidence in their submission to support the use of the Watch BP Home A oscillometric HBPM (home blood pressure measurement) monitor for the measurement of blood pressure (BP) and the detection of atrial fibrillation (AF). The external assessment centre (EAC), NUTH/YHEC, identified some deviations from the scope outlined by NICE in both the clinical evidence and economic evidence submissions.

The AF detection system was considered to be the innovative function of the device and so was the main focus of the submission (the WatchBP Home A is fully validated for the detection of BP). However, when considering AF detection (or more correctly pulse irregularity), some extrapolation from the literature was required, as the clinical evidence identified did not always directly reflect the population, intervention, or comparators that were identified in the Scope.

The main comparator used in the clinical evidence submission (12-lead electrocardiograph [ECG] for the detection of AF) was different to the main comparator used in the economic submission (manual pulse palpation for the detection of AF) and was not the comparator specified in the Scope (confirmatory 12-lead ECG to detect AF is part of the clinical pathway common to both the intervention and the comparator).

The population identified in the de novo cost analysis of the economic submission (symptomatic patients and/or patients at high risk of AF) was not that specified by the Scope (patients with suspected or existing hypertension or those who are being screened for hypertension).

1.2 SUMMARY OF CLINICAL EVIDENCE SUBMITTED BY THE SPONSOR

The sponsor identified ten studies in total; seven of which were from their literature search and a further three were identified from unpublished studies. Of these studies, the EAC considered five were highly relevant to the decision problem (four published studies and one study awaiting publication). Four studies (three published and one unpublished) were not fully appraised as they were BP measurement validation studies, or measurement of BP was their sole outcome. One unpublished study was in the process of recruitment and no results were available.

Of the remaining five studies, three were cross-sectional diagnostic studies with the diagnostic accuracy of the equivalent AF detection algorithm of the WatchBP Home A compared with 12-lead ECG as their main outcome. These studies indicated that the sensitivity of the AF detection algorithm was 93-100%, specificity was 69-91%, PPV was 68%, and NPV was 98%. For comparison, a high-quality Health Technology Assessment
(HTA) reported a sensitivity of 87.2%, a specificity of 81.3%, a PPV of 30.1%, and an NPV of 98.6% when comparing GP or nurse-led pulse palpation with 12-lead ECG.

The two remaining studies investigated the AF detection algorithm in a home setting. One study was a small case series in discharged hospital patients who had previously been diagnosed with AF but were now in sinus rhythm. Over the course of days or months using the device, 10/19 patients recorded positive for AF, and three of these were false positives. The remaining unpublished study compared a device with the AF detection algorithm with a hand-held ECG over 30 days. In this home setting, the diagnostic accuracy of the device appeared to be similar to the device in a clinical setting. The device misdiagnosed 7/139 patients and identified 2/139 new true positive cases.

1.3 SUMMARY CRITIQUE OF CLINICAL EVIDENCE SUBMITTED BY THE SPONSOR

The sponsor’s literature search was poorly implemented and could not be replicated as described. However, there was no evidence of confirmation bias or missed studies. An unpublished study of a trial in progress was considered to be highly relevant to the decision problem but no results were reported.

The three studies investigating the diagnostic accuracy of the AF detection algorithm were variants of cross-sectional diagnostic studies with a gold standard (12-lead ECG), which were considered appropriate for the assessment of the diagnostic accuracy of screening techniques. These studies had minimal potential for operator or observer bias and were considered to have high internal validity. However, care is required in the interpretation of this data as the studies were set in secondary care environments where the AF status of the patients was either known, or the pre-test probability of the patients having AF was high. In addition, the devices used had differing measurement protocols (e.g. requiring one, two, or three positive measurements in succession for positive detection of AF) which may lead to different results to the WatchBP Home A, and add to the uncertainty of the diagnostic accuracy of the device. Thus, it was not clear how generalisable the results of these studies were to the community population envisaged in the scope, or the WatchBP Home A device itself.

One study in a home setting was a small case series and as such lacked the methodological quality and generalisability to inform the decision problem. The other study in a home setting was unpublished. Whilst giving an interesting insight into how the WatchBP Home A device could be used at home, this study was underpowered to provide the statistical analysis required to substantiate the results in terms of true and false positives. Overall there was considered to be a lack of evidence to support the home use of the WatchBP Home A for AF detection, rather than evidence of no benefit.

1.4 SUMMARY OF ECONOMIC EVIDENCE SUBMITTED BY THE SPONSOR

The sponsor did not provide any economic evidence specifically relating to the WatchBP Home A derived from the existing published literature. The sponsor presented a de novo
cost consequence model with a time horizon of one year that compared pulse palpation with the WatchBP Home A in a community healthcare setting. The focus of the analysis was the diagnosis of AF through the detection of irregular pulse and confirmation with 12-lead ECG. Costs of pulse palpation; ECG; anticoagulant and antiplatelet drugs; adverse effects of these drugs; and strokes prevented were monetised and used to generate an estimation of the net economic benefit of the device. Both pulse palpation and the WatchBP Home A incurred costs to the NHS; however, the sponsors estimated for the base case analysis the WatchBP Home A would free resources for the NHS in England and Wales equivalent to £11.6 million compared with pulse palpation. Deterministic sensitivity analysis identified that more savings would be made if the cost of stroke was assumed to be greater, and less savings would be made if a proportion of patients were assumed to asymptomatic.

1.5 SUMMARY CRITIQUE OF ECONOMIC EVIDENCE SUBMITTED BY THE SPONSOR

The EAC considered that the de novo cost consequence model used appropriate methodology and was generally populated using accepted reference sources. However, the population studied, patients symptomatic of AF, was not that of the Scope. There was also some uncertainty caused by the inputs used to populate the model. These included uncertainty in the model structure; estimation of prevalence of AF in the population; diagnostic accuracy of the WatchBP Home A in a community setting; and the cost of screening at the point of patient contact.

1.6 EXTERNAL ASSESSMENT CENTRE COMMENTARY ON THE ROBUSTNESS OF EVIDENCE SUBMITTED BY THE SPONSOR

The EAC considered that the studies that provided the clinical evidence lacked generalisability to the community setting that was relevant to the decision problem. As these studies informed the de novo economic evidence they cast doubt on the robustness of the model. Additionally, the EAC considered that some of the key costing assumptions were questionable, particularly concerning the opportunity cost involved with pulse palpation.

1.7 SUMMARY OF ANY ADDITIONAL WORK CARRIED OUT BY THE EXTERNAL ASSESSMENT CENTRE

The EAC carried out the following additional work:

- Repeating and attempting to replicate the sponsor’s literature search for clinical evidence.

- Literature search and reporting of studies concerning the diagnostic accuracy of pulse palpation.
- Multivariate sensitivity analysis of *de novo* economic model. Parameters that were varied included the prevalence of AF in the study population; the diagnostic accuracy of the WatchBP Home A and pulse palpation; and the cost of screening associated with pulse palpation.
2.1 OVERVIEW AND CRITIQUE OF SPONSOR’S DESCRIPTION OF CLINICAL CONTEXT

2.1.1 The technology

The WatchBP Home A is a conventional automated oscillometric blood pressure (BP) monitor with a heart beat (pulse) detection system that detects irregular heart rhythm during cuff deflation. The last ten pulse intervals are measured, with intervals differing by more than 25% from the mean (fixed threshold) being rejected as premature beats. The device calculates an “irregularity index” from remaining beats, and applies a fixed threshold (0.06) to the index to identify irregular rhythms.

The EAC considered that the algorithm was straightforward compared with state-of-the-art beat-based atrial fibrillation (AF) detectors, perhaps partly due to the restriction of only detecting beats during the cuff deflation phase, which lasts around 30 seconds in total and 15 seconds with sufficient pressure to detect pulses reliably. Fixed thresholds are used for rejecting beats and deciding irregularity based on a sample of ten successive beat intervals. Premature beats are excluded from the analysis.

The EAC also noted that the sponsor often used the terms “pulse irregularity” and “atrial fibrillation” synonymously. In fact there are several causes of irregular pulse (including atrial flutter, atrial extrasystoles, sinus tachycardia, supraventricular tachycardias, and multifocal atrial tachycardia). The sponsor acknowledges that all cases of irregular pulse detected with the WatchBP Home A should be confirmed by 12-lead electrocardiography (ECG) to either confirm AF or sinus rhythm, or detect a differential diagnosis.

2.1.2 Overview of Condition

The WatchBP Home A is a dual function device indicated for the measurement of BP and detection of AF.

Hypertension (high blood pressure) is one of the most preventable causes of premature morbidity and mortality in the UK, and is a major risk factor for stroke (ischaemic and haemorrhagic), myocardial infarction, heart failure, chronic kidney disease, peripheral vascular disease, cognitive decline and premature death. Hypertension is also a major risk factor for AF. It is estimated that about a quarter of the UK population are hypertensive with more than half of these being aged over 60 years (also around the age when the risk of AF significantly increases) (1).

Atrial fibrillation is an atrial tachyarrhythmia characterised by predominantly uncoordinated atrial activation with consequent deterioration of atrial mechanical function. It is classified into three types depending on how long episodes last. These are paroxysmal (multiple
episodes that cease within 7 days without treatment); persistent (episodes lasting longer than 7 days, or less when treated); and persistent or permanent (continuous AF which has occurred for more than one year) (2).

Atrial fibrillation predisposes to thrombus formation causing sufferers to have a five-fold increased risk of stroke and thromboembolism (3). Cardiac output may be compromised regardless of ventricular function, and AF can lead to structural heart disease and can precipitate critical cardiac ischaemia. Overall, AF is associated with an odds ratio (OR) for death of 1.5 for men and 1.9 in women, with most of the excess mortality occurring soon after the diagnosis of AF (2).

According to the 2011 Quality and Outcomes Framework (QOF) database, 914,634 people were registered as having AF in the UK, which accounts for about 1.5% of the population (4). The prevalence of AF advances with age, roughly doubling with each advancing decade of age (2). The annual incidence of AF in people aged over 65 years has been estimated to be as high as 1.04% (5).

Given the relationship between hypertension and AF, and the potential seriousness of both conditions alone or together, the EAC recognised that a single device that could measure, monitor, or detect hypertension and/or AF could be of significant value.

2.1.3 Relevant guidelines

The sponsor identified one guideline from National Institute for Health and Clinical Excellence (NICE) that they considered relevant to the decision problem. This was the NICE clinical guideline on Atrial Fibrillation: National clinical guideline for management in primary and secondary care (CG36) (2). The sponsors also used the costing report (6) associated with these guidelines.

The EAC considered that two NICE guidelines were directly relevant to the decision problem, and one was indirectly relevant.

- The NICE clinical guideline on Atrial Fibrillation: National clinical guideline for management in primary and secondary care (CG36) (2) was considered highly relevant. This guideline provides full recommendations on when to suspect AF, how to screen for AF, how to diagnose AF, as well as its ongoing management.

- The NICE clinical guideline on Hypertension: The clinical management of primary hypertension in adults (CG127, previously CG34) (1) was also considered highly relevant. This guideline provides advice on how to measure BP and when to screen for an irregular pulse. In addition, the guideline covers the diagnosis and monitoring of hypertension at home using a home blood pressure measurement (HBPM) device. There was no indication that this guideline was considered by the sponsor.

- The NICE clinical guideline on Lipid modification: Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of
cardiovascular disease (CG67) (7) was considered indirectly relevant. This covers the assessment and management of cardiovascular disease in general.

2.1.4 Current clinical care pathway

The current clinical pathway described by the sponsor is that described by the NICE clinical guidelines on Atrial Fibrillation (CG36) (2). These guidelines recommend targeted or opportunistic screening for AF in patients who have symptoms of the condition or are at particularly high risk. The associated symptoms of AF include breathlessness, dyspnoea, palpitations, syncope/dizziness or chest discomfort.

However, the EAC considers that the NICE clinical guidelines on Hypertension (CG127) are also very relevant to the current clinical care pathway (1). These recommend that, during consultation for the screening for hypertension, irregular pulse should be screened for during the measurement of BP in the office or clinic (this is mainly because irregular pulse can adversely affect the diagnostic accuracy of BP measurement devices). If an electronic oscillometric BP monitor is used, as is increasingly common, this should be done by using manual palpation of the ulnar or radial pulse. If an irregular pulse is detected, a manual measurement of BP should be performed, preferably using a non-mercury sphygmomanometer with auscultation of the brachial artery, and the patient should be referred for ECG.

The EAC noted that the failure of the sponsor to consider the pathway recommended by NICE for the measurement of BP was an omission from their analysis.

2.1.5 Clinical care pathway with WatchBP Home A

2.1.5.1 Sponsors claims for clinical pathway with WatchBP Home A

The sponsor considered that the WatchBP Home A would be used in patients presenting with symptoms of AF as a point of good practice, in accordance with NICE guidelines on Atrial fibrillation (2). The sponsor claimed that “both pulse irregularity and blood pressure are checked at the same time with high accuracy”. In addition, the sponsor asserted that the device would be useful for the detection of AF in asymptomatic patients, leading to the detection of at least 30% more cases of AF than is currently the case.

The sponsor correctly stated that currently all cases of irregular pulse detected by manual pulse palpation require further investigation with a 12-lead ECG, and acknowledged that, as the WatchBP Home A detects irregular pulse and not AF per se, a positive result with the device would not eliminate the need for ECG referral. Investigation with ECG is necessary because there is a need to confirm or rule out the diagnosis, or identify a differential diagnosis. However, the sponsor also claimed that “less unnecessary ECG’s are performed” because “AF detection system shows better results for detecting AF than manual pulse palpation and is not liable to observer bias”.

Additionally, the sponsor suggested that if the General Practitioner (GP) felt there was reason to suspect paroxysmal AF (e.g. the patient had described previous symptoms
compatible with AF), the patient could take the device home for regular self-monitoring, or for use when they were feeling symptomatic.

### 2.1.5.2 EAC evaluation of clinical pathway with WatchBP Home A

The EAC considered that the presence of an irregular pulse may mean it is not possible to accurately measure BP (1). After discussion with the sponsor, it was clarified that the WatchBP Home A has not been validated for the measurement of BP when an irregular pulse is present, so the claim of “high accuracy” in this context appears to be currently unsupported (although the sponsor stated that this is an issue that is currently undergoing investigation). This is an area that the NICE clinical guideline on Hypertension recommends is suitable for further research (Which automated blood pressure monitors are suitable for people with hypertension and atrial fibrillation?) (1). The accuracy of the device to measure BP in patients with AF is a significant consideration given the device purports to have this dual functionality.

The estimate that 30% more cases of AF would be picked up by use of the device was considered likely to be a significant overestimate, since the NICE guidelines on Hypertension recommend the manual pulse palpation of all patients who require BP measurement (1). Additionally, there is a great deal of uncertainty regarding the prevalence of asymptomatic and paroxysmal AF and it is dependent on the population screened (2).

The EAC noted that the sponsor claimed the WatchBP Home A would reduce the need for “unnecessary” ECGs. This implies that the device is more specific than manual pulse palpation, leading to fewer false positive results. This is further discussed in Section 3 (Clinical Evidence). This is an important aspect of the device, as the reduction in false positive results during screening could potentially free immediate resources within the NHS, as well as sparing patients unnecessary anxiety.

The EAC considered that the sponsor’s suggestion that the device could be used at home to monitor BP and (in patients in whom the GP suspects it) detect paroxysmal AF may be of potential benefit. However, in fact the population that is relevant to the final Scope is patients who have suspected hypertension in whom paroxysmal AF may be detected as an incidental finding during diagnosis of hypertension (including measurement at home). Additionally, ad hoc use of the device to detect AF during symptomatic episodes was not in the remit of the Scope.

### 2.1.6 Changes to current services

The sponsor claimed the WatchBP Home A would eliminate the need for manual palpation of the pulse and could reduce the need for unnecessary ECGs. The sponsor also considered that the WatchBP Home A could be used at home for the monitoring of BP and incidental detection of AF.

The EAC considered pulse palpation required no capital investment and limited time investment, as palpation is carried out just before BP measurement. In addition, if the
WatchBP Home A device does not accurately measure BP in patients with an irregular pulse, manual measurement of BP with auscultation would be required in any case. Acquisition of the WatchBP Home A for clinic or home use would also inevitably confer capital expenditure to the NHS through replacement of existing BP monitors (see Section 4, Economic Evidence).

2.2 OVERVIEW OF SPONSOR’S DESCRIPTION OF ONGOING STUDIES

The sponsor described details of three unpublished studies, one of which was ongoing ("Oxford") (8), and two which were awaiting publication (Wiesel, 2012 and Wessel, 2010) (9) (10). The sponsor supplied full manuscripts (in academic confidence) of the studies awaiting publication, and provided an abridged trial protocol for the ongoing study, as well as a patient information leaflet. The sponsor did not give details of dates of publication or completion of the trial. The sponsor did not describe the studies in detail or indicate how significant they might be to the decision problem.

The EAC’s consideration of the relevance and possible implications of these studies are fully described in Section 3 (Clinical Evidence).

2.3 CRITIQUE OF SPONSOR’S DEFINITION OF THE DECISION PROBLEM

The sponsor did not complete Section 1 of the Clinical Evidence submission, “Statement of the decision problem”. Therefore variations on the Scope, or rationale for these variations, have not been reported.

2.3.1 Population

Two populations were described in the Scope, reflecting the dual functionality of the technology. These were “People with suspected or existing hypertension or those who are being screened for hypertension” and “People in the above group in whom atrial fibrillation is present but undetected". It is understood that the second group of patients was included to reflect the fact that the WatchBP Home A’s novel function compared to other HBPM devices is the detection of AF, and this is the main focus of the submission.

The sponsor did not provide a definition of the population in relation to the decision problem, in several sections in the submission the population used by the sponsor did not match that of the Scope.

In their clinical evidence submission, the sponsor failed to identify specific populations or scenarios relevant to the Scope, and failed to extrapolate or interpret how the differences in the populations described in the included studies from their literature search might impact compared to the population described in the Scope (see section 3.7.1).

In their economic submission, the sponsor focussed on symptomatic patients being screened for AF in a clinical environment (typically a GP or nurses’s office). Whilst this was considered to be a relevant setting, patients with symptoms of AF were not included in the Scope. The use of the WatchBP Home A to detect AF in a home environment was not included in the sponsor’s de novo cost consequence analysis.
2.3.2 Intervention

The sponsor did not provide a definition of the intervention in relation to the decision problem. The definition provided by the final Scope is simply “Watch BP Home A”. Without further clarification, this has been taken to mean the use of the WatchBP Home A using “usual” or “diagnostic” mode within the context of the relevant population and comparators. However, throughout the submission the sponsor did not adequately describe these modes and their implications in the context of the decision problem. Problems with applying the available evidence to the WatchBP Home A in are discussed in Section 3 (Clinical Evidence). Consideration of the use of diagnostic mode is discussed in Section 4 (Economic Evidence).

2.3.3 Comparator(s)

The sponsor did not provide a definition of the comparators in relation to the decision problem. The Final Scope included a description of two comparators which were “Home or ambulatory blood pressure monitoring without an atrial fibrillation algorithm in patients with suspected hypertension” and “Clinical examination including pulse palpation, followed by a confirmatory ECG to diagnose atrial fibrillation”. These comparators were intended to reflect the dual functionality of the WatchBP Home A device.

The sponsor focussed on the use of 12-lead ECG as a comparator in the Clinical Evidence submission. In the Economic submission, the sponsor used pulse palpation as a comparator. Thus the comparators in each part of the submission were different, which made it difficult for the EAC to directly relate the clinical evidence with the economic evidence. The difficulties raised by this issue are highlighted throughout the EAC’s critique of the sponsor’s submission.

The use of a comparator in a home setting (i.e. HBPM or ABPM devices) was not considered in either the clinical or economic evidence submissions.

2.3.4 Outcomes

The sponsor did not provide a definition of the outcomes in relation to the decision problem.

The outcomes listed in the Scope are as follows:

- Diagnostic accuracy for hypertension.
- Diagnostic accuracy for AF.
- Incidence of AF-related stroke in patients with AF detected.
- Reduced mortality from AF-related stroke.
- Reduced disability from AF-related stroke.
• Device-related adverse events.

The EAC considers that these are correct and relevant outcomes to the decision problem. The sponsor reported on the first four of these outcomes in the two submissions. Reduced disability from AF-related stroke was not reported as an outcome in either document. Device-related adverse effects are of limited relevance to this class of devices (non-invasive BP monitors).

2.3.5 Cost Analysis

The sponsor did not provide a definition of the cost outcomes in relation to the decision problem.

The EAC considered that the de novo economic analysis provided in the economic evidence submission, which was a comparison of the potential cost savings of using the WatchBP Home A to detect AF compared with pulse palpation, was within the Scope and relevant to the decision problem.

2.3.6 Subgroups

The sponsor did not provide a definition of subgroups in relation to the decision problem. However, within Section 4 of the submission (Regulatory information), the sponsor described the following as potential subgroups: patients who are elderly; hypertensive; diabetic; or those with underlying heart disease. These are risk factors that increase the risk of AF. Nevertheless, it is understood that any person can develop AF, including all those who receive screening for hypertension for whatever reason, although the risk increases exponentially with age.

One of the included studies concerned validation of the device to monitor BP in pregnant women with and without preeclampsia (11). However, this subgroup is not at increased risk of AF (12).

2.3.7 Special Considerations, Including Issues Related to Equality

The sponsor did not provide a definition of “Special considerations, including issues related to equality” in relation to the decision problem.

Within Section 4 (Regulatory information), the sponsor identified that patients who do not have access, or who cannot operate, an internet-connected PC at home would be unable to perform telemonitoring with the device. However, as the device automatically stores readings the patient could bring the device into the surgery or clinic, and the data could be downloaded there.

Some patients may be unable to use the device at home because of mental or physical incapacity. However, the EAC considered that this will be true of any HBPM device used and so is not an equality issue. The sponsor also claims that the device could reduce the need for 24-hour ECGs and therefore increase accessibility in some groups, but this has not been established.
3.1 CRITIQUE OF THE SPONSOR’S SEARCH STRATEGY

The sponsor provided details of seven published studies and three unpublished studies. These are described in the section in section 3.3.

3.1.1 Description of the Sponsor’s Search Strategy

There was no description of the search strategy used to identify published studies in Section 5.1.1 of the sponsor’s submission. The reader was referred directly to Appendix 10. There was a brief description of the identification of unpublished studies in section 5.1.2. The sponsor contacted experts, searched the ClinicalTrials.gov registry and also searched Google.

In Appendix 10 there was no clear description of the search strategy, its purpose, the methods used, or the rationale for the inclusion or exclusion of search terms. The sponsor reported that MEDLINE, Embase and the Cochrane Library were searched and this meets NICE’s database minimum requirement. However, it would appear that only PubMed (not MEDLINE specifically) was actually searched. This was not clearly reported in the search strategy; in fact PubMed was not mentioned at all. It was only through recognition of the search syntax and by rerunning the search that the EAC could deduce that the search was originally run in PubMed. There were no search strategies, database service providers, search dates, or results for either Embase or the Cochrane Library. The number of records retrieved was reported for PubMed, but neither the date of the search or the database date span was given.

The search strategy itself was rudimentary. There were five separate searches which could have been combined using the Boolean OR to have one set of results, leaving no need to remove duplicate records. There was no use of the medical subject headings (MeSH) available in PubMed. A number of possible MeSH terms might have been useful for this particular search strategy: Atrial Fibrillation/, Blood Pressure/, Hypertension/, Oscillometry/, Sphygmanometers/, Blood Pressure Monitors/, Blood Pressure Determination/, Blood Pressure Monitoring, Ambulatory/. PubMed does have the facility of searching for MeSH automatically, so the strategy will probably have searched for most, if not all of the suggested MeSH terms. However, it would have been better practice to have included the MeSH terms explicitly to ensure that they were being searched.

Search fields such as title and abstract were not used, and no limits were used, such as removing animal studies. Systematic searches would usually include such search commands and limits. Although PubMed does not allow proximity searching, it might have been better to search using the Boolean AND rather than specific phrase searches, e.g.
“atrial AND fibrilat*” rather than “atrial fibrilat*”. This would have produced more sensitive results.

The search could have been improved by searching for “blood pressure” alone rather using more specific search phrases for blood pressure monitoring: “blood pressure meas*”, “blood pressure monit*”, “blood pressure read*” and “automat* blood pressure”. This would have removed the need to search for these very precise phrases and resulted in a more sensitive search.

The fifth search for “microlife/Watch BP” made search four redundant, as “microlife” is searched for in both lines. It could be that line five would have been sufficient to retrieve all studies about “microlife/watch BP” if that was all the sponsor was looking to identify. However, if the sponsor wanted to identify comparators, then the search strategy was not sensitive enough.

More synonyms could also have been included, for example additional search terms for “atrial fibrillation” could have included the following: auricular fibrillation, AF, A-fib, and possibly atrial flutter. For “blood pressure monitors” the following search terms could have been included: sphygmomanometer, hand held ECG. Truncation was used effectively. The sponsor did search for unpublished studies by searching ClinicalTrials.gov and Google. However, the details of the search terms used were not reported, and so searches could not be reproduced. The sponsor contacted all research physicians known to them to be using the Watch BP device.

The sponsor’s search strategy was not systematic, was poorly described and inadequately reported, and did not appear to do all it reported, i.e. did not search Embase and the Cochrane Library.

There were no separate searches for adverse events.

3.1.2 Repeat of sponsor’s search

The EAC repeated the sponsor’s search as it was reported in Appendix 10 of the sponsor’s submission (that is using the search criteria on Medline, Embase, and the Cochrane library).

3.1.2.1 Database search

The sponsor stated that they searched three databases: Medline, Embase and the Cochrane library. The search terms used were:
1. Atrial fibrillat* AND (blood pressure meas* OR blood pressure monit* OR blood pressure read*)
2. Atrial fibrillat* AND oscildom*
3. Atrial fibrillat* AND (automat* blood pressure)
4. microlife AND atrial fibril*
5. Microlife OR WatchBP OR watch-BP [all] OR Watch BP
The sponsor did not state when they ran their search and they did not specify which search provider they used. The EAC attempted to replicate the sponsor’s search on 8th December 2011, and used Healthcare Databases Advanced Search (HDAS) in order to search the three databases specified. The sponsor did not supply a breakdown of retrieved articles from each of the three databases specified. The sponsor stated that they did not restrict their search by language, population or study design, but did not mention if they restricted by publication date—the EAC did not restrict its search by any of these parameters. The sponsor did not state if they used any synonyms or MeSH terms; therefore, the EAC did not use any in its search. The EAC searched title and abstract database fields apart from when “all fields” was specified. A comparison between the results of the sponsor’s search and the EAC’s is shown in Table 3.1.1.

Table 3.1.1. Database search results

<table>
<thead>
<tr>
<th>Search term</th>
<th>Sponsor search results</th>
<th>Medline</th>
<th>Embase</th>
<th>Cochrane</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>73</td>
<td>64</td>
<td>94</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>12</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>7</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>38</td>
<td>55</td>
<td>6</td>
</tr>
<tr>
<td>Combined</td>
<td>197</td>
<td>368</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duplicates removed</td>
<td>151</td>
<td></td>
<td></td>
<td>212</td>
</tr>
</tbody>
</table>

3.1.2.2 Inclusion and exclusion criteria

The EAC retrieved all 212 abstracts and applied the sponsor’s exclusion criteria. These were applied in the same three stage process as the sponsor performed and specified:

Stage 1

1. animal studies
2. articles not related to blood pressure

Stage 2

3. articles not related to atrial fibrillation

Stage 3

4. articles discussing the relationship between atrial fibrillation and blood pressure level
5. articles discussing the measurement of blood pressure in the presence of atrial fibrillation.
A comparison between the sponsor’s and EAC results of applying the sponsor’s inclusion/exclusion criteria is provided in Table 3.1.2.

Table 3.1.2. Exclusion criteria results

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
<th>Number of articles excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sponsor’s results</td>
</tr>
<tr>
<td>Stage 1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>83</td>
</tr>
<tr>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>Stage 2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Stage 3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Total excluded</td>
<td>144</td>
</tr>
<tr>
<td>Remaining articles</td>
<td>7</td>
</tr>
</tbody>
</table>

The seven articles that the sponsor identified from their search were:

1. Wiesel et al 2009 (13)
2. Chung et al 2009 (11)
3. Stergiou et al 2009 (14)
4. Stergiou et al 2007a (15)
5. Stergiou et al 2007b (16)
6. Wiesel et al 2007 (17)
7. Wiesel et al 2004 (18)

Three of the articles which the sponsor included in their submission did not meet their own inclusion criteria in that they were not related to AF (Stage 2 exclusion criteria) and should have been excluded. These articles were:

Chung et al 2009
Stergiou et al 2007a
Stergiou et al 2007b

After repeating the sponsor’s search and applying their exclusion criteria, the EAC identified all of the articles cited by the sponsor including the three articles which should have been excluded. These were:

Wiesel et al 2009
Chung et al 2009
Stergiou et al 2009
Stergiou et al 2007a
However, we also identified an additional relevant article which the sponsor did not cite. This was:

Wiesel et al 2010.

This was a poster presentation published in a journal, with the title “Detecting atrial fibrillation at home with a novel blood pressure monitor with atrial fibrillation detection functionality”. This study was later identified to describe the same data as that of Wiesel et al 2012, which is the Trial of Regular versus Irregular Pulse for the Prevention of Stroke (TRIPPS 2.0) trial (10). This is presently awaiting publication (journal title and anticipated date of publication unknown). A full manuscript was provided by the sponsors.

No additional papers were discovered from citations in the articles identified from the database search.

3.1.2.3 Grey literature

The sponsor stated that they performed further searches on the Internet using Google, but they did not define their search strategy. It was therefore not possible to replicate their search.

The sponsor stated that a search of clinical trial registration sites was made and they gave clinicaltrials.gov as an example, but they did not specify their search strategy, therefore we could not replicate the search. The sponsor indicated that they discovered one relevant unpublished paper from this search, which they obtained in confidence by contacting the lead author directly (Wiesel et al, 2012) (10).

The sponsor also said that they contacted research interested physicians from whom they knew to be using WatchBP devices. This led to the discovery of two additional studies. The first was Wessel et al 2010 (9), which we identified in our Embase search, but which should have been excluded from the submission because it was not related to AF (Stage 2 of the sponsor’s exclusion criteria). The second was an unpublished abstract (describing a trial protocol) and patient information sheet, which was again supplied in confidence (8).

3.2 CRITIQUE OF THE SPONSOR’S STUDY SELECTION

In their clinical evidence submission, the sponsor stated that “we included all studies related to the WatchBP Home device and/or the AF detection system. There would be no reason to exclude one of these studies”. The EAC considered that some of the included studies were not directly relevant to the decision problem. These were studies where the sole outcome was BP measurement (as the Watch BP Home A is a validated HBPM device and pulse
irregularity was the outcome under investigation) and studies in which 12 lead ECG was not used as the reference measurement.

3.3 INCLUDED AND EXCLUDED STUDIES

The sponsor identified seven studies they regarded as relevant to the decision problem in their literature search. These are listed in table 3.3.a. A further three unpublished studies were also included in the submission. These are listed in 3.3.b.

Table 3.3.a. List of published studies and relevance to decision problem.

<table>
<thead>
<tr>
<th>STUDY IDENTIFICATION</th>
<th>STUDY NUMBER</th>
<th>STUDY DESIGN AND BRIEF DESCRIPTION</th>
<th>RELEVANCE TO DECISION PROBLEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stergiou et al, 2007a</td>
<td>(15)</td>
<td>Validation study</td>
<td>Limited relevance. Validation and acceptability of home monitoring of BP only.</td>
</tr>
<tr>
<td>Stergiou et al, 2007b</td>
<td>(16)</td>
<td>Case series, questionnaire</td>
<td>Limited relevance. Validation and acceptability of BP measurement only.</td>
</tr>
<tr>
<td>Chung et al, 2009</td>
<td>(11)</td>
<td>Validation study</td>
<td>Limited relevance. Validation of BP in pregnant or preeclampsia women only.</td>
</tr>
</tbody>
</table>
Table 3.3.b. List of unpublished studies and relevance to decision problem.

<table>
<thead>
<tr>
<th>STUDY IDENTIFICATION</th>
<th>STUDY NUMBER</th>
<th>STUDY DESIGN AND BRIEF DESCRIPTION</th>
<th>RELEVANCE TO DECISION PROBLEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxford University study</td>
<td>(8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wessel et al, 2010</td>
<td>(9)</td>
<td>Randomised controlled trial (RCT).</td>
<td>Limited relevance. Comparison of BP monitoring regimens only. Poster abstract; EAC were supplied with full document.</td>
</tr>
<tr>
<td>Wiesel et al, 2012</td>
<td>(10)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The WatchBP Home A received the CE mark on the 14th June 2011. It is a validated BP monitoring device and is recommended as suitable for use in the clinic or home by the British Hypertension Society (www.bhsoc.org). The EAC judged that it was not appropriate for all the included studies to undergo further evaluation, and it was considered that studies that focussed on BP measurement only (i.e. did not include detection of AF) were not directly relevant to the decision problem. This is in accordance with the sponsor’s submitted literature search, which stated these studies should have been excluded at that stage because they bear “no relation with atrial fibrillation”. Furthermore, the sponsor did not critically appraise two of the excluded studies because they “were both validation papers that followed a strict protocol …… which is commonly accepted”.

The following studies were excluded from further analysis by the EAC:

- Stergiou et al, (2007a) (15). This study validated the WatchBP Home device with a mercury sphygmanometer comparator in 33 participants. The device was found to be compliant with International Protocol, but measurement of AF was not in the scope of the study.
- Stergiou et al, (2007b) (16). This was a descriptive article of the WatchBP Home monitor. Results of a questionnaire given to 20 participants indicated that there was good adherence with “diagnostic” mode. There was no discussion of the detection of AF in this study, although the issue of compliance may have some relevance in this context.

- Chung et al, (2009) (11). This study assessed the accuracy of the WatchBP Home in 45 pregnant women (15 of whom had recorded preeclampsia). Detection of AF was not in the scope of the study. Although the WatchBP Home A device has been validated to measure BP, it is not specifically indicated for this purpose, and AF is relatively uncommon in pregnancy (although other arrhythmias are more common) (19).

- Wessel et al, (2010) (9). This study was the only RCT that was submitted. However, it was designed to compare patient adherence of the “diagnostic mode” with the “usual mode” when measuring BP only (although it may be possible to extrapolate these results to AF detection). The study was presented as a poster abstract and is awaiting publication (i.e. it has not been peer reviewed). The EAC was provided with a full copy of the manuscript.

3.4 OVERVIEW OF METHODOLOGIES OF ALL INCLUDED STUDIES

The sponsor provided details of the methodology of nine out of ten included studies in tabulated form in section 5.4.1 on pages 27-35 of the clinical evidence submission document. The details of one study, Stergiou et al (2007b) were not reported (16).

After exclusion of the four studies that were deemed not directly relevant to the decision problem (9) (11) (15) (16), or did not provide usable data (8), five studies were included for further consideration. Of the included studies, four (one unpublished) were variations of cross-sectional diagnostic studies (10) (13) (14) (18), and one was a case series with diagnostic outcomes (17). The key characteristics of these studies are summarised in table 3.4.1. The critical appraisal of these studies, using a methodological checklist adapted from the Critical Appraisal Skills Programme (CASP), is summarised in table 3.4.2.

The four cross-sectional diagnostic studies were considered by the EAC to be of the appropriate study design to assess the diagnostic accuracy of the device (that is, they were suitable to identify how accurate the device identifies people with irregular pulse and possible AF, and how well the device excludes people who do not have an irregular pulse). Three of the studies used 12-lead ECG as the reference standard. This is widely regarded as the gold standard test for the diagnosis of AF, and because it uses an entirely different mechanism to detect AF, it is not subject to incorporation bias.

- The study by Wiesel et al (2004) (18) used a modified HBPM device from a different manufacturer (Omron 712C automatic sphygmomanometer). The sponsor confirmed that this device, as well as all the other devices described in the clinical submission, used the same algorithm for pulse irregularity detection as the WatchBP Home A, and is independent of the BP measurement functionality. However, it was noted by the EAC that the protocol to detect AF was two successive positive results for irregular pulse (as opposed to the “usual” mode of the WatchBP Home A which is
three successive positive readings). The study was conducted in two phases; the first recruited hospital inpatients, and the second recruited outpatients from a cardiac clinic. The AF status of the patients was known before recordings were performed, and a threshold irregularity index was selected such that all ECGs with AF exceeded this irregularity index value (thus the sensitivity of the device was fixed at 100%). Overall the study was considered to be of good methodological quality with a high degree of internal validity, though its application to clinical practice was limited.

- The study by Wiesel et al (2009) (13) used a Microlife HBPM device (BP3MQ1) with a different protocol to detect AF (two out of three positive readings were considered positive). This study used a boot strapping technique to estimate the sensitivity and specificity of a single measurement with the device, and retrospectively analysed patients who were classified as false positives. Overall the EAC considered that this study was of high quality with good internal validity. It is used as the basis for the clinical inputs used to populate the de novo economic model described in Section 4.

- The study by Stergiou et al (2009) (14) used a Microlife HBPM device (the BPA100), that in part used the same protocol to detect AF as the “usual” mode of the WatchBP Home A (i.e. three successive positive readings). Patients were simultaneously monitored with a 12-lead ECG, which may have reduced operator bias. The study appeared to have been of sound methodological quality, although the AF status of the patients at baseline was known, which may have implications for the external validity of the protocol.

The study by Wiesel et al (2012) (10) has not yet been published or peer-reviewed.

The other study that the EAC considered should be included in the assessment report was a prospective case series. Wiesel et al (2007) (17) recruited 19 patients with a past history of AF who were given the Omron 712C automatic sphygmomanometer to monitor their blood pressure and screen for AF over a period of five days to five months (three positive readings at various time points were regarded as positive for AF). If AF was “diagnosed”, the patients were asked to report to a clinic the next day to receive confirmation by 12-lead ECG. Case series studies lack internal validity because they have no controls, lack randomisation, often lack consecutive recruitment (selection bias), and have low patient numbers. Nevertheless this study has some merit because it more accurately represents one of the potential uses of the WatchBP Home A (i.e. it was a home study carried out over several weeks or months). However the patient group was not those asymptomatic to AF, and therefore not directly relevant to the decision problem.

Overall, the included studies were deemed by the EAC to have internal validity, as would be expected with this type of study design. The outcomes were generally appropriate for the questions the studies addressed. However, not all the studies adequately reported the uncertainty surrounding the point estimates of their sensitivity and specificity data (i.e. did
not report confidence intervals). It was noted that there was a lack of good prospective trials (e.g. RCTs) with clinical outcomes to assess the likely benefits of the device in a pragmatic setting, although intermediate outcomes were considered directly relevant to the decision problem. Of more concern, however, was the lack of generalisability, or external validity, that these studies had with respect to the decision problem. These issues are discussed in more detail in section 3.7.
### Table 3.4.1. Characteristics of included studies.

<table>
<thead>
<tr>
<th>STUDY</th>
<th>SETTING AND POPULATION*</th>
<th>INTERVENTION (I) AND COMPARATOR (C)</th>
<th>OUTCOMES</th>
<th>GENERALISIBILITY (TO UK SETTING)</th>
</tr>
</thead>
</table>
| Wiesel et al., 2004 (18)     | US urban cardiology practice.  
125 inpatients (53 documented with AF)  
446 outpatients (54 documented with AF) | I: Omron 712C automatic sphygmomanometer. 2 successive positive readings indicates AF.  
C: 12-lead ECG | Sensitivity, specificity and diagnostic accuracy of device to detect AF compared with gold standard (ECG) | Successful readings not equivalent to WatchBP Home A (either mode)  
US inpatient and outpatient setting not equivalent to UK community (e.g. greater prevalence of AF).  
Developed a threshold irregularity index to compare accuracies (i.e. sensitivity fixed at 100%). |
| Wiesel et al., 2007 (17)     | US patients with past episodes of AF recruited from inpatients or outpatients (n = 19)  
Patients monitored at home for between 5 days and 5 months | I: Home monitoring with Omron 712C automatic sphygmomanometer. Patients "diagnosed" with AF if 3 readings (initial, immediately after, 1 hour after) were positive.  
C: 12-lead ECG only in positive patients. Transtelephonic monitor. | Number of false positives generated by the device. | Protocol used does not match that of WatchBP Home A.  
All patients previously diagnosed with AF in secondary care – not equivalent to UK community practice. |
| Wiesel et al., 2009 (13)     | Two US cardiology offices  
Unselected outpatients (n = 405) | I: Microlife BP3MQ1 HBPM with AF detection. 2/3 positive readings indicates AF. (or 3 sequential readings)  
C: 12-lead ECG. | Sensitivity and specificity of device to detect AF compared with gold standard (ECG).  
Subgroup analysis of false positives | Protocol used does not match that of WatchBP Home A.  
US cardiology outpatient setting may not reflect UK general practice. |
| Stergiou et al., 2009        | Greek outpatient hypertension clinic  
Patients (n = 73) had known | I: Microlife BPA100 Plus HBPM with AF detection. 3 positive successive readings indicates AF.  
C: Simultaneous 12-lead ECG. | Sensitivity and Specificity of device compared with gold standard (ECG). | Protocol used matches "usual" mode only. Arrhythmia status |
Table 3.4.2. EAC quality assessment of included studies from adapted Critical Appraisal Skills Programme (CASP).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there a clear question for the study to address?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Was there a comparison with an appropriate reference standard?</td>
<td>Yes</td>
<td>Possibly</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Did all the patients get the diagnostic test and the reference standard?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Could the results of the test of interest have been influenced by the results of the reference standard?</td>
<td>Possibly</td>
<td>Yes</td>
<td>Possibly</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Is the disease status of the tested population clearly described? (is this at entry or post result?)</td>
<td>Possibly</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Were the methods for performing the test described in sufficient detail?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Can the results be applied to the relevant UK population?</td>
<td>No</td>
<td>No</td>
<td>Possibly</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Can the test results be applied to the relevant UK population?</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>Would the test lead to a change in clinical management in the relevant UK population?</strong></td>
<td>No</td>
<td>Possibly</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
3.5 OVERVIEW AND CRITIQUE OF SPONSOR’S CRITICAL APPRAISAL FOR EACH STUDY

3.5.1 Methodology

In section 5.4.3. of the submission, the sponsor acknowledged that most the provided studies were “comparative” in nature; that is the studies measured intra-patient detection of AF using two technologies (the intervention and the control, or index test). The EAC considered this an appropriate methodology for the measurement of the diagnostic accuracy of a device. The sponsor correctly stated that the participants were commonly outpatients receiving treatment in a specialist cardiovascular clinic, but did not elaborate on the implication of these settings in terms of external validity. The sponsor also commented that there were differences between the protocols used to detect to AF in the studies, but did not suggest how this may influence their interpretation.

In section 5.4.4., the sponsor indicated that post hoc analysis of false positives was undertaken in three studies. However, there was no further discussion of this subgroup analysis other than that most of the false positives “had some other form of arrhythmia”. Given that the WatchBP Home A detects pulse irregularity and not AF per se further clarification on this point would have been useful.

Section 5.4.6. is concerned with patient follow-up and withdrawal from the studies (e.g. whether intention to treat analysis was used). The sponsor correctly stated only one RCT was submitted (which was later excluded by the EAC), and although there were reports of withdrawal in other studies with home settings (10) (17), these are unlikely to be significant.

3.5.2 Critical appraisal

The sponsor has provided a qualitative critical appraisal of the included studies using an adapted CASP methodology, and the results are tabulated with comments (pages 38-44).

Although CASP checklists are suitable tools for critically appraising clinical studies, the sponsor has selected an inappropriate one for the studies in question. The CASP checklist for cohort studies was used, when in fact the CASP checklist for diagnostic test studies would have been more appropriate (see the table 3.4.2). This meant only some of the questions posed were relevant to the study type (for instance questions concerning confounding and follow up were not generally applicable for this study type). Nevertheless the sponsor captured important aspects of the critical appraisal including the assessment of recruitment, likelihood of bias, and inclusion of statistical analysis.

3.6 RESULTS

Where possible, the sponsor provided details of the results of the included studies in tabulated form in section 5.6.1. on pages 47-51 of the clinical evidence submission. The sponsor omitted the results of studies that were solely concerned with BP measurement as well as one ongoing study which had not yet reported data (8). The sponsor did not attempt to place the results of the reported studies in the clinical context of the decision problem.
A summary of the results of the five studies the EAC considered were suitable for inclusion are given in table 3.6. Three of the studies were cross-sectional diagnostic studies that provided sensitivity and specificity data of the device compared with 12-lead ECG in a clinical setting (13) (14) (18). One study was a case series in a home setting that evaluated the diagnostic accuracy of the device with next-day 12-lead ECG clinical confirmation of positive results only (17). The final unpublished study reported the diagnostic accuracy of the device in a home setting compared with an ECG event monitor (10). Some of the studies did not include statistical analysis or a measurement of uncertainty (i.e. confidence intervals) when reporting their results.

The studies that reported cross-sectional diagnostic accuracy derived from a clinical setting reported that the sensitivity of the device was 93-100% when one reading was taken, and specificity was 89-91% (13) (15) (16). Therefore the technology using one reading only was effective at ruling in people who did have AF, but less effective at ruling it out in people who did not have the condition. Increasing the number of consecutive readings (e.g. to two consecutive readings, or two out of three positive consecutive readings) improved the sensitivity to 98-100%, but had little effect on the specificity of the test (86-91%); thus there still was a tendency for people to be incorrectly diagnosed with the condition. However, it did not appear that repeated testing caused increased sensitivity and decreased specificity. One study reported the diagnostic accuracy of one positive reading out of three readings, and showed this approach caused unacceptably low specificity at 69% (95% CI 53% to 81%) (15).

The studies that investigated the use of an HBPM device with AF detection at home lacked the internal validity of the studies conducted in clinical settings, but provided interesting information about another use of the device (i.e. home monitoring). The case series (17) indicated that, when the device is used over a period of weeks or months, it is likely to result in some misdiagnoses (3 in 19 [16%] had false-positive irregular readings). However, using the methodology the study employed it is not possible to estimate how many cases of paroxysmal AF it may have missed, particularly asymptomatic AF (2).

Whilst the EAC considered that data from the studies set in home settings (9) (17) were potentially interesting, it considered these studies were not directly relevant to the decision problem because of the comparators used (see Section 3.7).
### Table 3.6. Results from included studies.

<table>
<thead>
<tr>
<th>STUDY</th>
<th>PATIENT NUMBER OR MEASUREMENTS</th>
<th>RESULTS</th>
<th>FOLLOW UP</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiesel et al, 2004 (18)</td>
<td>446 unselected cardiology outpatients</td>
<td>1 reading: Sensitivity 100%*</td>
<td>Single scheduled visit</td>
<td>*Sensitivity fixed at 100% using an irregularity index</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity 84%</td>
<td></td>
<td>Statistical significance and level of uncertainty not reported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 readings: Sensitivity 100%, Specificity 91%, Diagnostic accuracy 92%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiesel et al, 2007 (17)</td>
<td>19 consecutive patients discharged from hospital</td>
<td>3 false positive (16%)</td>
<td>5 days to 5 months</td>
<td>Small numbers (case series), no statistical significance reported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 true positive (37%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9 true negative (77%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stergiou et al, 2009 (14)</td>
<td>217 measurements (73 subjects). Outpatients and healthy volunteers.</td>
<td>1 reading: Sensitivity 93%, Specificity 89%.</td>
<td>Single scheduled visit</td>
<td>Level of uncertainty (confidence intervals) reported</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 readings: Sensitivity 100%, Specificity 76%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/3 readings: Sensitivity 100%, Specificity 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/3 readings: Sensitivity 100%, Specificity 69%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiesel et al, 2009 (13)</td>
<td>405 unselected cardiology outpatients</td>
<td>1 reading: Sensitivity 95%, Specificity 86%, PPV 68%, NPV 98%</td>
<td>Single scheduled visit</td>
<td>Subgroup analysis of false negatives.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 sequential readings: Sensitivity 97%, Specificity 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiesel et al, 2012 (10) (prepublication)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.7 DOES EACH RELEVANT STUDY INCLUDE THE PATIENT POPULATION(S), INTERVENTION(S), COMPARATOR(S) AND OUTCOMES AS DEFINED IN THE FINAL SCOPE?

The EAC considered that although the internal validity of the included studies was generally good, there was a high degree of uncertainty regarding their relevance to the decision problem. In particular, there were concerns with respect to the included populations, interventions, and comparators.

3.7.1 Population

The population described in the Scope essentially describes people without diagnosed AF who require screening or monitoring of BP in the community (in clinic or at home). In this context, the detection of AF can be considered a (valuable) incidental or opportunistic finding. Generally, this population is not reflected in the included studies.

The cross-sectional diagnostic studies were set in inpatient hospitals in part of one study (18); specialist cardiology or hypertension outpatient departments in three studies (13) (14) (18); and during consultation with a generalist in one study that is unpublished (8). The case series followed patients being monitored at home following discharge from hospital, where they had been diagnosed with AF that was resolved at the baseline of the study (17). In one study, the arrhythmia status of the participants was already known, with 27% of patients having AF when they were enrolled into the study, and a further 31% being healthy subjects in sinus rhythm (14). The AF status of the patients was also known in the seminal study by Wiesel (18).

The EAC considered that these patient groups exhibited potential referral bias (20), and were significantly different from those included in the final Scope. In all cases, the incidence and prevalence of AF would be expected to be significantly greater in the included studies than those outlined in the decision problem (i.e. they would have a greater pre-test probability of having AF). This is important because the prevalence of a condition impacts on the positive and negative predictive value (PPV and NPV) of a given test. For instance, if the studies had been repeated in the UK community population, the lower prevalence of AF in the study population would be expected to lead to a lower PPV; and consequently a greater number of false positives. This could have a significant negative impact to individual patients (i.e. increased anxiety and inconvenience), and on use of NHS resources.

An additional issue to consider is that the severity of AF, including the rate of ventricular rhythm and permanence of abnormality, may be greater in people who present in specialist settings. Consequently it is possible the WatchBP Home A detection algorithm could demonstrate greater diagnostic accuracy in these patients compared with those typically found in a community population (spectrum bias (20)). Another issue is that the device may exhibit better technical performance in a specialist setting rather than a community or home setting, due to improved adherence and practical issues (for example better fitting of the cuff).

None of the studies were set in the UK (four were in the US and one was in Greece). Different healthcare systems incorporate different incentives and trade-offs, as well as
distinct clinical care pathways, which might not reflect those of the English and Wales NHS, and lead to population bias (20).

In summary, the EAC believes the included studies were not conducted in populations comparable to the UK, and therefore the results could not be generalised as such.

3.7.2 Intervention

None of the studies specifically used the WatchBP Home A as part of the intervention. Two studies used a non-Microlife device (modified Omron device), whilst three used a different Microlife model (two used model BP3MQ1-2D, and one used model BPA100 plus). However, the sponsor has confirmed that the AF detection algorithm was independent of the BP measurement system used and that the devices should be considered equivalent.

There was considerable heterogeneity in the protocols used regarding the number and timing of measurements required to detect AF. These ranged from requiring two successive measurements (18); three successive positive measurements (14); one out of three successive measurements (14); two out of three successive positive measurements (10) (13) (14); and more complex protocols (17). Only one study had a protocol comparable to the “usual” mode of the WatchBP Home A as one of several protocols studied (14), and no studies used anything comparable to the “diagnostic” mode of the technology. This variation in protocols adds to the uncertainty of the assessment.

As well as variation in the BP measurement or AF detection protocols, there was also considerable variation in the settings the measurements were taken in. Most studies were in an inpatient or outpatient clinical setting, with only the small case series (17) and the unpublished trial (10) set in a home environment. The single study that is set in an environment relevant to the economic submission had not yet reported data (8). There could be significant differences in the results of a single measurement of BP in a clinical environment compared with several measurements in a home environment.

3.7.3 Comparator

The relevant comparator outlined in the final scope was defined as “Clinical examination including pulse palpation, followed by a confirmatory ECG to diagnose atrial fibrillation”. The included studies all used the second part of this pathway, ECG (mainly 12-lead), as the comparator. However, in the context of the decision problem this is not the most relevant comparator, as all patients who have AF detected using the WatchBP Home A monitor will also require confirmation with a 12-lead ECG. Therefore the EAC considered that the first part of the pathway, pulse palpation (for screening of AF), is the most pertinent comparator. In addition, it is because 12-lead ECG is regarded as a diagnostic procedure rather than a screening tool that it is widely regarded as the reference test or gold standard. In this context, the WatchBP Home A device can be regarded as an additional screening technique, which would “displace” pulse palpation from practice. In these circumstances, it is the incremental value or benefit of the new screening method compared to the existing method that needs to be addressed, not the comparison with the gold standard (21). The sponsor recognised that pulse palpation was the appropriate comparator in the Economic Evidence submission (see Section 4).
None of the included studies investigated the use of pulse palpation for screening of AF, so the decision problem has not been fundamentally answered in the Clinical Evidence submission. Neither the sponsor nor the EAC found any head-to-head trials in the literature that compare the WatchBP Home A (or similar devices) with pulse palpation. However, in the absence of these studies, it would seem reasonable to make an indirect comparison by comparing the diagnostic accuracy of pulse palpation with ECG, thus:

- WatchBP Home A versus 12-lead ECG
- Pulse palpation versus 12-lead ECG

By comparing both screening methods with a reference standard, it would be possible to make an indirect comparison between them, but the sponsor did not do this in their Clinical Evidence submission. Throughout the submission (for instance in sections 3.4, 3.5, 3.9, 3.10, and 5.93) the sponsor has made the claim that the use of the WatchBP Home A could reduce the need for “unnecessary” ECGs, which the EAC understands to mean that the device results in fewer false positives than manual palpation. However, in this claim has not been adequately explored. It is supported with reference to the SAFE study (5), but this trial did not investigate the diagnostic accuracy of the WatchBP Home A or similar device. Rather, it investigated the effectiveness of an opportunistic strategy to screen for AF with pulse palpation, compared with a systematic strategy to screen for AF with nurse or GP led ECG. The diagnostic accuracy of pulse palpation and a brief review of the SAFE study are discussed in section 3.10.2.

3.8 DESCRIPTION OF THE ADVERSE EVENTS REPORTED BY THE SPONSOR

The sponsor reported that there were no safety issues regarding the device, and they commented that the patients would be required to have their BP screened in any case. This seems reasonable as the WatchBP Home A is a non-invasive BP monitor. Possible adverse effects include haematoma caused by high inflation pressure of the cuff, but these are likely to be rare and not significantly problematic, and adverse effects caused by cuff inflation would be expected to occur with any HBPM device. Pulse palpation is not associated with any known adverse effect.

3.9 DESCRIPTION AND CRITIQUE OF EVIDENCE SYNTHESIS AND META-ANALYSIS CARRIED OUT BY THE SPONSOR

The sponsor did not carry out any meta-analyses or other forms of evidence synthesis, stating that there is a general lack of published studies on the WatchBP Home A device or other similar oscillometric devices. The EAC agreed that the included studies were too heterogeneous to synthesise, particularly in terms of the intervention and population used. For instance, the included studies encompassed several different HBPM devices, populations, and different measurement and AF detection regimens. This would have made the undertaking of a meta-analysis challenging and uninformative.
3.10 ADDITIONAL WORK CARRIED OUT BY THE EXTERNAL ASSESSMENT CENTRE IN RELATION TO CLINICAL EVIDENCE

3.10.1 Diagnostic accuracy of WatchBP Home A compared with pulse palpation

The EAC considered that the diagnostic accuracy of manual pulse palpation compared with WatchBP Home A was a relevant question that deserved further investigation. This was because pulse palpation and the WatchBP Home A are the appropriate methods for screening for AF, whereas 12-lead ECG is generally considered as a diagnostic technique. Although a brief search of the literature revealed no head-to-head studies that directly compared these screening techniques, it was considered that an indirect comparison of pulse palpation and WatchBP Home A should be undertaken to determine the diagnostic accuracy of the device compared with standard practice (opportunistic screening with pulse palpation).

A brief literature search using the search strategy “Atrial Fibrillation AND pulse palpation AND ECG” was performed which retrieved one systematic review and meta-analysis that was relevant to the decision problem (22). In addition, the NICE clinical guideline on Atrial Fibrillation (CG36) (2) was referenced for further information on the diagnostic accuracy of pulse palpation in the detection of AF, as was the SAFE study (a large pragmatic RCT relevant to the decision problem) (5).

The systematic review (22) identified three relevant studies comprising 2,385 patients that compared pulse palpation with consultant-led 12-lead ECG (the reference test). All the studies were carried out in a UK general practice setting. When the results from the studies were pooled, the sensitivity of the technique was found to be 94% (95% CI 84% to 97%), with a specificity of 72% (95% CI 69% to 75%). The authors noted that assuming a prevalence of 3% for undetected atrial fibrillation in patients older than 65 years, and given the test’s sensitivity and specificity, opportunistic pulse palpation in this age group would detect an irregular pulse in 30% of screened patients, requiring further testing with ECG. Among screened patients, 0.2% would have AF undetected with pulse palpation.

The three studies identified by this systematic review were also identified and cited by the NICE clinical guideline on Atrial Fibrillation (2). The authors of the guideline commented that “an irregular pulse was found to be sensitive to the presence of AF” and that “the negative predictive value of a regular pulse (>96%) was also emphasised”. This relatively high NPV is important because it reflects the low number of false negatives seen with pulse palpation and consequently the number of people who remain falsely undiagnosed and at increased risk of stroke. The authors concluded that “it would be prudent to consider any pulse irregularity as requiring further investigation to determine whether AF is present”.

The SAFE study was a large (n = 14,802), multicentre, pragmatic RCT commissioned by the NHS Research and Development HTA (Health Technology Assessment) Programme (5). It investigated the effectiveness and cost-effectiveness of an opportunistic strategy to screen for AF using pulse palpation (followed by confirmatory 12-lead ECG) with systematic
screening with nurse or GP led ECG (followed by confirmatory 12-lead ECG) in an elderly population. In addition, there was a control arm who received neither intervention. The opportunistic strategy proved to be most cost effective. Compared with the gold standard, pulse palpation was found to have a sensitivity of 87.2% (82.1% to 91.1%), a specificity of 81.3% (79.7% to 82.8%), a PPV of 30.1% (26.7% to 33.8%), and an NPV of 98.6% (97.9% to 99.0%).

Table 3.10.2. summarises the data from the three included cross-sectional studies (13) (15) (16) and the SAFE study (5) (the largest study with the highest methodological quality).

Table 3.10.2. Diagnostic accuracy of oscillometric BP device with AF detection indirectly compared with pulse palpation.

<table>
<thead>
<tr>
<th>SCREENING TECHNIQUE</th>
<th>SENSITIVITY</th>
<th>SPECIFICITY</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse palpation†</td>
<td>87.2% (82.1 to 91.1%)</td>
<td>81.3% (79.7 to 82.8%)</td>
<td>30.1% (26.7 to 33.8%)</td>
<td>98.6% (97.9% to 99.0%)</td>
</tr>
</tbody>
</table>

* Data range taken from all cross-sectional studies (13) (15) (16) and summary of all measurement protocols unless otherwise stated.

† Diagnostic parameters based on data from the SAFE study (5).

The sensitivity of the pulse palpation compared with 12-lead ECG appears to be less than that of an oscillometric BP device with AF detection; although, unlike the studies of the intervention, the SAFE study was conducted pragmatically in an environment directly relevant to the decision problem. The specificity of the new technology also appears to be superior to that of the traditional screening method. This could support the sponsor’s claims that the WatchBP Home A reduces the number of unnecessary diagnostic ECGs by reducing false positive results, but without statistical analysis or inclusion of further studies this claim cannot be made with any degree of certainty.

Although the PPV is superior for the WatchBP Home A than it is for pulse palpation, this may be due to the differing populations in the study settings, because PPV is dependent on disease prevalence (see Appendix A). The studies of the device were in high-risk populations (or in people whose AF status was already known) who were likely to have a high pre-test probability of having AF; whereas the studies of pulse palpation were carried out by nurses or GPs in primary care populations, and were consequently at much lower risk of AF.
In several places in the clinical evidence submission document the sponsors made reference to the Screening for Atrial Fibrillation in the Elderly (SAFE) study to support the claim that the WatchBP Home A could reduce the need for “unnecessary” ECGs (5). In this study, patients who were opportunistically screened had AF identified through pulse palpation before diagnosis through 12-lead ECG. Those who were systematically screened all received immediate 12-lead ECG, whereas the control group received current care, including a 12-lead ECG if an irregular pulse was detected. The authors found that significantly more cases of AF were detected in the intervention practices, although there was no significant difference between opportunistic and systematic screening (1.64% compared with 1.62%; difference 0.02%, 95% CI -0.5% to 0.5%). A within-trial economic evaluation showed the lowest incremental cost to be for the opportunistic arm, thereby supporting the use of opportunistic screening using pulse palpation from a healthcare system point of view. As the SAFE study did not investigate the use of the WatchBP Home A or an equivalent or similar device, it is not directly relevant to the decision question and could have provided a basis for the sponsor to make an indirect comparison, but this was not done.

3.11 CONCLUSIONS ON THE CLINICAL EVIDENCE

The WatchBP Home A oscillometric BP monitor is a technology that serves two functions. The EAC considered the main function of the device was for the measurement of BP for the screening, diagnosing, and monitoring of hypertension. The additional function of the device was for the incidental, or opportunistic, detection of AF, and possibly the ongoing home monitoring of this condition. The scope focussed on the latter aspect, as detection of AF was considered to be the innovative function of the device.

3.11.1 Measurement of blood pressure

The EAC confirmed that WatchBP Home A is fully validated for the measurement of BP in the absence of comorbidities, and considered further appraisal of this function was unnecessary. However, the reliability of the device to accurately measure BP in people with an irregular pulse (including those in AF) was not substantiated, which could have important implications for people with confirmed AF who also require ongoing monitoring of their BP. Conversely, it is understood that upper arm oscillometric BP devices generally perform favourably in these patients compared with wrist or auscultatory devices (12), and the sponsor of the WatchBP Home A verified that further research was being carried out in this area. Another claim the sponsor made concerning the device and BP measurement was that it could potentially reduce the need for clinic BP measurements. The EAC considered there was no evidence to support this assertion, and in most patients it would be contrary to national guidelines.

The EAC noted that one claimed advantage of the WatchBP Home A is the provision of a “diagnostic” mode. The EAC reflected that this mode was primarily designed to diagnose hypertension rather than detect AF. This view was supported by the inclusion of an unpublished RCT submitted by the sponsor that showed the WatchBP Home A was effective at increasing patient adherence for the measurement of BP, not the detection of AF (9). The
diagnostic mode measures BP twice in the morning and twice in the evening, which is consistent with guidelines from the European Society of Hypertension (www.eshonline.org), the British Hypertension Society (www.bhsoc.org) and NICE guidelines on Hypertension (1) (www.nice.org.uk). It is less clear whether the term diagnostic mode also applies to the measurement of AF (in which it is used as a screening device), although the algorithm would still detect AF in this mode.

3.11.2 AF detection

The EAC noted that HBPM monitors from other manufacturers are also equipped to detect “pulse irregularity”, but that the WatchBP Home A is the first to promote this function specifically for AF detection. It was considered that the simultaneous measurement of BP and detection of AF could potentially have significant benefits for the individual and healthcare system. The conditions are related, with hypertension being believed to be a causal risk factor of AF (2). Additionally, the incidence of both conditions increases with age, and both are frequently asymptomatic and suitable for screening or monitoring. Taken in this context, the dual functionality and simplicity of the device appear particularly appealing.

The EAC considered that to an extent the clinical evidence submission failed to focus on a specific, well defined, clinical question that in particular could address the main benefits of the device with a degree of certainty. Although the literature search was poorly implemented, it retrieved five studies that were considered directly relevant to the decision problem. There was no evidence of confirmation bias in the reported studies, and the EAC had no reason to believe that relevant studies had been omitted or remained undetected.

Three of the studies were cross-sectional diagnostic trials which had the principal objective of measuring the diagnostic accuracy of the AF detection algorithm used in the WatchBP home A device. The comparator used in these studies was 12-lead ECG (the gold standard), which is appropriate in this context, and the studies were regarded to be of good quality with a high degree of internal validity. The studies showed the device was highly sensitive in the detection of AF (thus nearly all people with AF are detected), and reasonably specific (thus a limited amount of people are incorrectly diagnosed by the device). Broadly speaking, the sensitivity of the algorithm was 93-100%, specificity 69-91%, PPV 68%, and NPV 98%. However, the differing measurement procedures used in each study (i.e. number of readings required to detect “AF”) introduced a level of uncertainty about the diagnostic accuracy of the device compared with the WatchBP Home A, which uses multiple measurements. Additionally, these studies were conducted in a clinical setting where the prevalence of AF was high or the AF status of the patients was already known (which has a large effect on the PPV and NPV). This meant the studies lacked generalisability to the NHS community setting in England and Wales, and therefore their interpretation should be treated with caution. However, overall the EAC concluded that the WatchBP Home A was likely to have a high degree of diagnostic accuracy in the detection of AF.

The EAC considered that although 12-lead ECG is an appropriate comparator for the validation of the device, it is not the most relevant comparator in NHS community practice, which would be pulse palpation. In the absence of direct head-to-head studies comparing
the device with pulse palpation, the EAC considered that, given the available literature (for example the HTA by Hobbs et al (5)), the sponsor could have made an indirect comparison between these screening methods. This perhaps demonstrated a lack of focus on behalf of the sponsor to adequately prove the benefits of the device in a clinical setting. For instance, one of the main claims in the clinical evidence submission by the sponsor, that the WatchBP Home A would reduce the amount of “unnecessary” confirmatory ECGs (compared with pulse palpation), was not adequately explored. It is, however, appreciated that this is mainly an economic concern, and these issues are discussed further in Section 4 (Economic evidence).

Two studies (one unpublished) investigated the use of the WatchBP Home A in a home setting. The published study was a small case series in a population that had had a prior diagnosis of AF, and as such had limited generalisability (17).

Home detection of paroxysmal and/or asymptomatic AF could be very important, as incidental AF detected in a home environment could lead to quicker treatment of the condition and the prevention of strokes, which would reduce mortality and morbidity, as well as free resources in the NHS. On the other hand, a greater number of false positive results compared with standard practice could also lead to an immediate increase in the number of ECGs performed and waste NHS resources (which would be contrary to the sponsor's expectations when the device is used solely in the clinic). More research with larger study populations are required to clarify the benefits and drawbacks of home monitoring of AF (see Section 6).

The EAC considered that, overall, the evidence base for the AF algorithm to detect irregular pulse was limited to the diagnostic accuracy of the device in specialised, controlled settings. In the setting of a GP’s or nurse’s office there are no studies that directly compare the WatchBP Home A with the present screening method (pulse palpation). In the setting of the home, there is a lack of data available to fully evaluate the benefits of the device, given the complexities of the conditions it is anticipated to detect (i.e. asymptomatic and/or paroxysmal AF) and potential difficulties with patient compliance. The sponsor provided qualitative statements to support the home use of the WatchBP Home A device in the clinical evidence submission, but there was a lack of data to support these. The use of the WatchBP Home A in the home environment was not explored in the de novo economic analysis supplied with the economic evidence submission (Section 4.3).

The sponsor did submit information about a trial protocol that is currently being conducted in the UK (8).
Section 4: Economic Evidence

4.1 CRITIQUE OF THE SPONSOR’S SEARCH STRATEGY

The sponsor reported that an additional literature was performed using the following terms:

- Atrial Fibrillat* AND screen*.
- pulse palp* OR Wrist AND Atrial Fibrill*.

Four studies were reported as being identified from the literature search, which were used for the clinical inputs for the *de novo* economic model. These were the studies from Wiesel (2009) (13), Stergiou (2009) (14), Morgan (2002) (23), and Hobbs (2005) (5). The sponsor also commented that their *de novo* analysis was largely based on the costing report derived from the NICE clinical guideline on *Atrial Fibrillation: the management of atrial fibrillation* (CG36) (2) (6). The sponsor indicated that this approach had been accepted by NICE prior to the economic submission.

The EAC noted that, given the relative paucity of directly relevant studies identified for the clinical evidence submission, it was unlikely that additional economic studies relevant to the decision problem would have been identified. However, it would have been appropriate for the sponsor to report details and results of a more systematic search, even if the results were negative.

4.2 CRITIQUE OF THE SPONSOR’S STUDY SELECTION

The sponsor’s literature search was crude and poorly reported. The sponsor did not report the results that were obtained from the search or report which databases or filters, if any, were used in the search. Assuming no limits or filters were used, the EAC repeated the searches using MEDLINE and found 460 studies from the first search and an additional 9 studies from the second.

4.2.1 Included and Excluded Studies

No new results from the literature search were reported so it would be assumed that the retrieved studies, with the exception of the four studies used to inform diagnostic accuracy, were excluded.

Studies included and excluded for the *de novo* economic analysis are discussed in section 4.3.4.2.

4.2.2 Overview of Methodologies of all Included Economic Studies

No economic studies were reported.
4.3 DE NOVO COST ANALYSIS

The sponsor provided a *de novo* cost analysis of the decision problem. This consisted of a 14 page narrative document that described the structure of the model, the base case analysis of results, and sensitivity analysis, including a graph concerning threshold sensitivity analysis.

The *de novo* analysis was accompanied with two appendices that described the various parameters and variables used. One appendix was an Excel spreadsheet that consisted of seven worksheets. The first worksheet contained the calculations for the base case analysis; the second to sixth worksheets contained the calculations for multivariate sensitivity analysis; and the final worksheet contained calculations for threshold analysis. The spreadsheet was fully executable; however it was not adequately annotated making it difficult to decipher. The other appendix was a PDF document that was essentially a non-executable copy of the spreadsheet.

4.3.1 Description of model

4.3.1.1 Perspective

The sponsor did not clearly state from whose perspective the economic analysis was performed. However, from the data inputs supplied and the output, the EAC assumed it to be from a third party payer perspective: that is the NHS of England and Wales and the Personal Social Services (PSS).

4.3.1.2 Model type

The sponsors adopted a deterministic cost consequence model for their economic analysis. The direct costs of two screening tests were established. These were the intervention; the WatchBP Home A; and the comparator, pulse palpation by a nurse or GP. The main consequences of each screening method were determined. These were referral for confirmation of AF with consultant-led 12-lead ECG; use of anticoagulant drugs and aspirin; adverse effects of anticoagulants and aspirin; and number of strokes prevented. Each consequence was assigned an associated monetary value which, combined with the direct costs associated with the interventions, were used to calculate the overall cost of each screening method. There was no attempt to assess utility or other measures of effectiveness, so that the incremental costs and benefits were not explicitly determined.

Although prevention of strokes is a long-term outcome, this was an uncomplicated model based on a one year time horizon (See Section 4.3.1.3). There was no attempt to extrapolate into the future (i.e. as in a life-time model) using, for instance, a state transition model (for example Markov model) or a decision tree. As the analysis was not based on a bespoke RCT, the parameters used in the model were estimated from the literature, mainly from four studies considered to be relevant by the authors (two of these were from the
Clinical Evidence submission) and the NICE costing study on *Atrial Fibrillation* (6). Sensitivity analysis was deterministic and largely restricted to the proportion of people with AF who are symptomatic, and the cost of stroke. In the case of the latter, threshold sensitivity analysis was calculated.

### 4.3.1.3 Model structure

This was a straightforward cost consequence model that investigated how the diagnostic accuracy of the WatchBP Home A (the intervention) to detect irregular pulse impacts on the costs associated with the current care pathway. The comparator chosen was pulse palpation which is appropriate, as it is the current screening procedure. The cost consequence analysis was informed by several inputs which were mainly estimated from the literature. The output of interest was the net yearly costs saved to the NHS in England and Wales and PSS.

**Inputs**

The following inputs were included in the analysis.

- The diagnostic accuracy (sensitivity and specificity) of the WatchBP Home A device (intervention) and pulse palpation (comparator) to detect pulse irregularity. These parameters were estimated from heterogeneous studies published in the literature (see Section 4.3.4.2).

- The absolute annual incidence rate of AF in England and Wales, taken from the NICE costing report on *Atrial Fibrillation* (6).

- Various parameters including the proportion of people with confirmed AF who receive anticoagulants or aspirin; the proportion of people who suffer an adverse effect on these drugs; and the number of strokes prevented (absolute risk reduction). These figures were largely taken from the NICE costing study (see Section 4.3.4.3).

- Costs associated for each of the parameters described above. Costs for pulse palpation were taken from Hobbs *et al* (5) and other costs were taken from the NICE costing study. All costs were adjusted for inflation (see section 4.3.4.4).

**Outputs**

The following outputs were used in the analysis (i.e. these values were calculated using formulae populated from the inputs described above).

- The proportion of the GP population presenting with AF (called prevalence in the model) was calculated using the sensitivity and specificity data for pulse palpation...
from the 2005 HTA authored by Hobbs et al (5) and the incidence data from the NICE costing report (6) (see Section 4.3.4.1).

- The calculated prevalence rate was used in conjunction with the sensitivity and specificity data of the WatchBP Home A device and pulse palpation to calculate the PPV and NPV. From this, the number of true positives, false positives, true negatives, and false negatives were determined.

- The overall number of ECGs required was calculated from the absolute number of true positives and false positives. This informed the total cost of the ECGs, as well as the drug costs associated with anti-coagulation and the costs associated with the adverse effects of these drugs.

- The number of strokes prevented was calculated from the number of true positives, the proportion of people given anticoagulants or aspirin, and the reduction in risk associated with anticoagulants or aspirin. The savings made from the reduction in strokes were then calculated based on the cost of having a stroke.

- The net cost saving using the WatchBP Home A device was calculated by the addition of all the associated costs for each consequence.

4.3.1.4 Time horizon

The time horizon was restricted to the course of one calendar year, with the assumption that a fixed number of new cases of AF, assumed to be the incidence rate of 87,000 estimated by the NICE costing report, will be screened each year at a constant rate. It also assumed that the incidence of stroke and the impact of anticoagulation treatment would occur at a constant rate each year, and did not take into account discounting of treatment effects or costs.

The EAC considered these assumptions to be adequate for an initial calculation of cost consequences based on a steady state model. The cost consequences of non steady state conditions were not considered. For example, if the WatchBP Home A device identified more cases of AF than current practice, it would be expected that in the early years of the introduction of the device that the rates of detection would be higher, to reflect the previously undiagnosed population pool (i.e. the calculation should be front-loaded). The best way to deal with these issues would have been for a state transition model to have been used, although the EAC was cognizant this would have involved a significant amount of resources and time to complete, and additionally the clinical inputs may not have been sufficiently robust to populate the model.

4.3.2 Patient population

The population specified in the Scope is described in Section 2.3.1. Briefly, it comprises of people who require measurement of BP in a GP or nurse’s office, for any reason, as the
main group; some of whom will have AF. The sponsor has described another group of people who are clinically suspected as having AF (for example people who have symptoms consistent with AF or who are at high risk of AF) as a secondary group. The sponsor has focussed on the latter group in their economic submission, by extrapolating the number of people who require screening for AF given the incidence of AF in the general population (see Section 4.3.4). However, this group of patients was not specified in the Scope.

The EAC considered that, in practice, the WatchBP Home A device would be used for the measurement of BP, which is potentially a very large population, as it would be expected to displace current HBPM devices. As an example of the potential size of this population, the NICE guideline on Lipid Modification (7) recommends that a systematic strategy to identify cardiovascular risk factors (including measurement of BP) should be carried out in all people aged 40-74 years, which would equate to a population of about 20 million in England (6). Although many people may be expected to decline an invitation for cardiovascular screening, many other people could require additional monitoring, and measurement of BP is performed routinely in GP practices as a point of good practice for many conditions.

Most of the population being screened for hypertension will not be at high risk of AF, as the risk of developing AF is highly dependent on age, with the prevalence roughly doubling with every decade of increasing age (2). For this reason, the number of positive tests that the device records in subjects who are not in AF (false positives) has important implications for the economic viability of the device, because each false positive would result in referral for an “unnecessary” ECG. However, the studies submitted by the sponsor that established the diagnostic accuracy of the AF detection algorithm generally lacked external validity to the setting defined in the Scope, as they were set in specialised outpatient settings in high-risk populations (see Section 3.7.1).

4.3.3 Technology and Comparators

The technology used in the de novo economic analysis was the WatchBP Home A to detect irregular pulse. It is assumed this would be carried out in a community setting by a physician or nurse using the device in “usual mode” (three successive positive results indicating likely AF). The WatchBP Home A also has a “diagnostic mode” but this is not applicable in an office setting.

The comparator used was nurse or GP-led pulse palpation to detect irregular pulse. The EAC considered that this was the appropriate comparator, but was not the comparator described in the Clinical Evidence section, which was 12-lead ECG (see Section 2.3.3).

4.3.4 Clinical Parameters and Variables

The sponsor estimated clinical parameters from the published literature. In the absence of evidence specific to the decision problem (for example from a bespoke RCT); this is an appropriate approach. However, due to the poor quality of the literature search and a somewhat arbitrary approach to study selection, the EAC had some concerns that there was the potential for confirmation bias (“cherry picking” of data, see section 4.3.2).
4.3.4.1 Estimate of incidence, prevalence, and sample size for screening

The annual incidence of AF, the population screened, and the numbers referred for ECG were essential parts of the sponsor's cost consequence model. In order to understand how the model was constructed, the EAC sought to establish how these parameters were derived.

**Incidence rate**

The sponsor reported that the incidence rate was taken from the NICE costing report on *Atrial Fibrillation* (6). The authors of this report estimated the value of 0.175% from the mid-range value from three epidemiological studies (0.05%, 0.17%, and 0.3%). By applying this value to the population of England (approximately 50,000,000) an absolute value of about 87,000 new cases of AF a year was calculated.

The EAC considered that this value from NICE was a broad estimate of the incidence of AF in England. It is known that the incidence of AF is not homogenously distributed among the population (being much more common in older individuals), and it was considered therefore that this absolute figure (achieved by multiplying the incidence by the entire population) was probably an overestimate of the true incidence, because AF is rare in people aged under 45 years. As the population of people over 45 in England and Wales is approximately 20 million (6), an annual number of incidences of 35,000 may have been more realistic.

**Prevalence**

The sponsor chose not to use a prevalence rate directly cited from the literature, but instead used a hybrid technique to calculate it using data from the study by Hobbs *et al* (5). This study stated that for every 5.7 patients detected as having irregular pulse, only one was confirmed as having AF by ECG. Using the inverse of this figure, the PPV can be calculated (0.175, or 17.5%). As the sensitivity (87.15%) and specificity (81.13%) of pulse palpation were known in this study, it is possible to calculate the prevalence of AF that would satisfy these values using equation 3.10 (Section 3.10.2). However, in order to replicate the prevalence value quoted by the sponsor, the EAC were required to perform an iterative calculation using mathematical simulation software (see Appendix A for details).

Using this technique, the sponsor estimated that the prevalence of AF in the population they anticipated would be screened for AF was 4.4%. The sponsor stated that the “incidence is higher than the general AF incidence but we estimate that this is representative for the population that is screened for AF (high risk, symptoms) as these are commonly elder people with certain risk factors”. However, the EAC considered that:

- The sponsor had not clearly defined the population the WatchBP Home A is intended to be used in (i.e. “high risk, symptoms”).

• The population is likely to be much broader than people who are at high risk of AF or who are symptomatic, and could potentially encompass all people who require measurement of BP (see Section 2.3.1).

• Using the sponsor’s methodology, the population requiring screening is calculated in reverse (i.e. the number of people successfully diagnosed with AF has been back-calculated to estimate the number of people who would have to have been screened to supply this result). The implication of this is that all the people screened must have been actively considered clinically at high risk of AF or as having had consistent symptoms with AF. The EAC deemed that this analysis was flawed (see below).

The EAC considered that a better approach would have been to work forward, taking a realistic estimate of the prevalence of AF in the population, and applying this to the sensitivity and specificity parameters reported in the published clinical studies. Sensitivity analysis could then have been performed to test these assumptions. Some examples of prevalence values in the literature are stated in table 4.3.4.1.

Table 4.3.4.1. Prevalence of AF reported in the literature.

<table>
<thead>
<tr>
<th>Prevalence (% of population with AF)</th>
<th>Population</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.28</td>
<td>General UK population</td>
<td>Majeed et al (2001) (24)</td>
</tr>
<tr>
<td>1.5</td>
<td>General UK population</td>
<td>Quality and Outcomes Framework (QoF) 2011 (4)</td>
</tr>
<tr>
<td>4.4</td>
<td>Unknown, poorly defined</td>
<td>Sponsor’s estimate</td>
</tr>
<tr>
<td>7.9</td>
<td>Patients over 65 years on GP register</td>
<td>Control arm of HTA by Hobbs et al (2005) (5)</td>
</tr>
</tbody>
</table>

The EAC performed its own sensitivity analysis on various parameters including the prevalence rate. See Section 4.6.

Sample size for screening

The consequence of the sponsor’s methodology and assumptions was that, for the base case analysis, a population of nearly two million (1,977,273) would have to be screened for irregular pulse because they were considered to be at high risk of AF or were exhibiting symptoms consistent with AF. Although the EAC were able to confirm the method of calculation used to reach this figure, it considered that alternative approaches may be valid in estimating the population likely to be screened using the WatchBP Home A.

Should the WatchBP be commissioned to all GP practices, the actual population that would be screened by the device (as an incidental finding) would be expected to match that of the population of people who have their BP measured in a calendar year (see Section 3.7.1). The EAC was unable to accurately estimate this figure, but it would be expected to be very high (probably in the tens of millions) considering:
- About 8.6 million people (8,586,974) were registered as hypertensive in the UK in 2011 (4).
- Most hypertensive patients (7,779,986) had had their BP measured in the preceding 9 months (4).
- Other patients are not known to be hypertensive but have their BP measured routinely during visits to their GP. Blood pressure is often measured opportunistically as a point of good clinical practice (e.g. see NICE guidance on Lipid Modification (7)).

The EAC considered that had the sponsor worked forward in their calculations instead of backwards, they could have performed sensitivity analysis on the amount of people being screened, and calculated how this would have impacted on their economic analysis. However, in order to do this they would have had to assume that the population being screened was at low risk of AF, which is not a population that was reflected in the included studies in the clinical submission (see Section 3.7.1).

### 4.3.4.2 Estimate of diagnostic accuracy

The sponsor estimated the diagnostic accuracy of the intervention, the WatchBP Home A, from two studies that were included in the clinical evidence submission. These were those of Wiesel et al (2009) (13) and Stergiou et al (2009) (14). The diagnostic accuracy of pulse palpation was estimated from two studies identified from the literature. These were Morgan et al (2002) (23) and Hobbs et al (2005) (5). These studies are listed in table 4.3.4.2.

**Table 4.3.4.2. Brief description of studies included in economic submission.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Study population</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiesel 2009</td>
<td>HBPM with AF detection (model BP3MQ1-2D). 2/3 positive readings for AF detection.</td>
<td>12-lead ECG</td>
<td>405 unselected patients from cardiology outpatients.</td>
<td>Sensitivity 96.8% Specificity 88.8%. Note: single reading Sensitivity 95.3% Specificity 86.4%.</td>
</tr>
<tr>
<td>Stergiou 2009</td>
<td>HBPM with AF detection (model BPA100 Plus). 2/3 positive readings for AF detection.</td>
<td>12-lead ECG</td>
<td>73 patients with AF status known: (37%) had AF, 23 (31%) non-AF arrhythmias and 23 (31%) sinus rhythm</td>
<td>Sensitivity 100% Specificity 89%</td>
</tr>
<tr>
<td>Morgan 2002</td>
<td>Nurse led pulse palpation</td>
<td>ECG rhythm strip</td>
<td>Aged over 65 years, community setting (n=3001)</td>
<td>Sensitivity 91% Specificity 74% (any irregularity)</td>
</tr>
<tr>
<td>Hobbs 2005</td>
<td>Nurse or GP-led pulse palpation</td>
<td>12-lead ECG</td>
<td>Aged over 65 years, community setting (n=14,802)</td>
<td>Sensitivity 87.2% Specificity 81.3%</td>
</tr>
</tbody>
</table>

**WatchBP Home A diagnostic studies**
Although not stated, the sponsor excluded three studies they had identified in the clinical submission.

- The seminal study from Wiesel et al (2004) was correctly excluded (18). This study reported the development of the threshold irregularity index (i.e. sensitivity was fixed at 100%).

- The case series by Wiesel et al (2007) was correctly excluded (17). This did not employ the appropriate methodology to investigate the diagnostic accuracy of the AF detection algorithm, and had a home setting. Although a home setting is within the Scope of the decision problem, the sponsor modelled the use of the WatchBP Home A in a clinic setting.

- The unpublished study by Wiesel et al (2012) was correctly excluded (10). This study did not use 12-lead ECG as the comparator.

The study from Stergiou et al (2009) was regarded as less suitable for inclusion (14) than that of Wiesel et al (2009) (13). The Stergiou study was performed in a controlled setting where the AF status of the patients was already known (37% had AF, 31% had non-AF arrhythmias, 31% had sinus rhythm), with the potential for operator or observer bias. There is also some evidence of confirmation bias on behalf of the sponsor, as the most favourable set of diagnostic results (those from two out of three positive readings indicating AF) were used for the economic analysis. This regimen is not equivalent to that of the WatchBP Home A (which requires three successive positive readings to indicate AF). Finally, this was a small study, with only 73 participants, but the sponsor made no effort to adjust the weighting to take this into account (instead opting for the mean of the results generated by the two studies). Taking all these factors into account, the EAC considered that the Stergiou study was probably not suitable for inclusion, and inputs from the Wiesel study only should inform the economic analysis.

**Pulse palpation diagnostic studies**

Regarding the reported studies that used pulse palpation as the intervention, the sponsor did not adequately describe how these studies were identified from the literature, raising concerns of confirmation bias. For instance, the sponsor failed to identify a systematic review that had compared pulse palpation with ECG (22). The sponsor did identify and exclude two studies that investigated the diagnostic accuracy of pulse palpation to identify AF.

- One study (25) was excluded because it was set in “only one GP’s practice and therefore may not be representative for overall healthcare”. This does not seem reasonable grounds for exclusion as it was a cross-sectional diagnostic study not intended to investigate overall healthcare processes. The study enrolled 84 patients, which was more than that of the included study by Stergiou et al (14) (73 patients), which was performed in a single specialist healthcare setting in Greece.
Excluded study demonstrated pulse palpation had a sensitivity of 96% and a specificity of 81%.

- One study by Sudlow *et al* (1999) (26) was excluded on the grounds that the comparator used was not 12-ECG (i.e. gold standard). The EAC agreed that this was grounds for exclusion. However, the study by Morgan *et al* (23) on pulse palpation used an ECG rhythm strip as its comparator and was included by the sponsor, leading the EAC to conclude that this study should also have been excluded on the grounds it did not include comparison with the gold standard. For this reason, only the study by Hobbs *et al* will be considered in the economic evaluation from here on. This study was of high methodological quality performed in a setting highly relevant to the decision problem. It was also very large (n = 14,802) and was carried out over the course of a full year, so in the opinion of the EAC it gave the best estimate of the diagnostic accuracy of pulse palpation in the population defined in the Scope.

The PPV and NPV, which inform the economic decision problem, depend on the sensitivity and specificity of the diagnostic test and on the population prevalence. Ideally, it would have been preferred if the studies of the WatchBP Home A had been performed in populations relevant to the decision problem (i.e. community practice in England and Wales) to establish the value of PPV and NPV directly (see Section 3.7.1). As this was not possible there is the possibility that there may be spectrum bias present (20); that is bias based on the underlying severity of AF, which may have been different to the WatchBP Home A study settings.

4.3.4.3 Other clinical variable estimates

The sponsor indicated that the clinical variable estimates in economic analysis were mostly derived from the NICE costing report from the *Atrial Fibrillation* clinical guidelines (6). These variables are listed in table 4.3.4.3.

*Table 4.3.4.3. Clinical variables used in economic analysis of WatchBP Home A (all parameters relate to patients who have been diagnosed with AF following detection with WatchBP Home A or pulse palpation).*

<table>
<thead>
<tr>
<th>Clinical Variable</th>
<th>Value used in model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability starting anticoagulant drugs (e.g. warfarin)</td>
<td>56</td>
</tr>
<tr>
<td>Probability starting antiplatelets (e.g. aspirin)</td>
<td>32 (Note: value from NICE costing report 34%)</td>
</tr>
<tr>
<td>Absolute risk reduction of having a stroke if on anticoagulants</td>
<td>4.3</td>
</tr>
<tr>
<td>Absolute risk reduction of having stroke if on antiplatelets</td>
<td>0.9</td>
</tr>
<tr>
<td>Probability of minor bleed (on treatment)</td>
<td>15.8</td>
</tr>
<tr>
<td>Probability of major bleed (on treatment)</td>
<td>2.4</td>
</tr>
</tbody>
</table>

The EAC considered that the clinical variables used in the economic submission had been derived from previous work by NICE and were appropriate for the model design.
The EAC noted that some possible clinical pathways in the model had been omitted. This included people who were diagnosed with AF and then received treatment with anti-arrhythmic drugs or received cardioversion, which could potentially reverse the condition. This would be expected to give clinical benefits as well as cost expenditure and savings. However, these care pathways were also omitted by Hobbs et al who had previously created a discrete event simulation to answer a similar decision problem (5). This was possibly because it would have made the model unwieldy, or because the drugs or procedures used were considered to be used infrequently for the demographics included (patients over 65 years).

4.3.4.4 Technology costs and usage

The sponsor derived the direct costs associated with the technology from the NICE costing report from the Atrial Fibrillation clinical guidelines (2) and the HTA by Hobbs et al (5). The costs directly associated with the WatchBP Home A and with pulse palpation were considered to be the cost of the time taken for the procedure itself, and the cost of confirmation of the diagnosis with ECG. These are shown in Table 4.3.4.4.

Table 4.3.4.4. Direct costs associated with WatchBP Home A and pulse palpation.

<table>
<thead>
<tr>
<th>Description of cost</th>
<th>WatchBP Home A</th>
<th>Pulse palpation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital cost</td>
<td>None</td>
<td>None</td>
<td>Sponsor highlighted most GP practices use HBPM systems of similar value</td>
</tr>
<tr>
<td>Time cost</td>
<td>None</td>
<td>£2.32</td>
<td>Cost is mean of 1 minute of GPs and nurses time (5).</td>
</tr>
<tr>
<td>ECG cost</td>
<td>£36.03</td>
<td>£36.03</td>
<td>(5) Updated by inflation.</td>
</tr>
</tbody>
</table>

The sponsor adjusted all the costs for inflation, and estimated this to be 5% per year (inflation rate is not the same as discount rate). This is equivalent to a multiplier of 1.27 from the years 2006 to 2011 inclusive. However, the EAC considered that this may be an overestimate considering that the actual rate of inflation in the UK has been significantly less than this over most this time period. Using the Hospital and Community Health Services (HCHS) index published by the Personal Social Services Research Unit, both price inflation and pay inflation were consistently below 5% in these years. In fact, price inflation was negative (-1.3%) in the year 2009/2010 (27). Given this, the EAC calculated an overall multiplier of about 1.12 might have been more realistic. The EAC also noted that the inflation multiplier was applied to the cost of ECG. The 2011 cost of a 12-lead ECG is £31 according to NHS reference costs (28), so the cost of ECG was also probably overestimated.

The sponsor did not include the capital cost of the WatchBP Home A in their analysis, stating that “the present cost-analysis the price of the monitor has not been taken into account as it is assumed that GP’s need to have a blood pressure monitor anyway”. The EAC considered that this might have been a significant omission from the sponsor. Whilst it is true that most GP and nurse offices will have an HBPM device, the WatchBP Home A would effectively
displace these devices and incur a significant replacement cost. If the device was supplied at the price stated by the sponsor in the submission (£75), this would amount to an initial cost of £2,700,000 presuming all registered GPs were allocated the device (36,000 GPs in England and Wales, see www.bma.org.uk). However, subsequently the incremental cost of replacing the device would depend on the price of alternative HBPM devices. To clarify this point, additional economic modelling could have been performed. The EAC was not provided with the relevant data inputs to perform this (e.g. bulk cost of the device to the NHS, length of warranty, servicing costs and schedule etc.).

The sponsors used inflation adjusted data from the Hobbs study to estimate the cost of pulse palpation (£2.32) (5). This was calculated by taking the mean average of one minute salary costs of a nurse and GP. However, it is unclear that stopping the practice of pulse palpation would release savings as assumed, or that there is a significant opportunity cost in this activity. Additionally, the WatchBP Home A requires three successive cuff inflations that are programmed to be spaced one minute apart (a minimum one minute gap is recommended by national guidelines (1)). Although the sponsor has clarified that there is an over-ride function that can significantly reduce the time between inflations (minimum 15 second gap), it is unlikely that the device could detect irregular pulse quicker than by manual pulse palpation, so there would be no time saving. The EAC therefore considered the sponsor’s baseline figure of £4,595,380 per year savings by eliminating the need for manual pulse palpation was unlikely to be achieved.

4.3.4.5 Indirect costs and savings

Indirect costs and potential resource savings of the WatchBP Home A include the cost of medication for management of confirmed AF (mainly warfarin or aspirin), the cost of treating adverse effects associated with these drugs, and savings from the number of strokes prevented. These costs and savings apply to both the WatchBP Home A and pulse palpation groups, but because the number of people successfully diagnosed in these groups will differ, so does the overall costs and savings. Costs associated with cardioversion or anti-arrhythmic drugs have not been modelled (5) (see Section 4.3.4.3.). The indirect costs and savings associated with the technology and pulse palpation were all derived from the NICE costing report from the Atrial Fibrillation clinical guidelines (2), and are listed in table 4.3.4.5.

Table 4.3.4.5. Indirect costs associated with WatchBP Home A and pulse palpation.

<table>
<thead>
<tr>
<th>Description of costs or savings</th>
<th>Cost (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of warfarin therapy</td>
<td>486</td>
</tr>
<tr>
<td>Cost of aspirin therapy</td>
<td>Not included</td>
</tr>
<tr>
<td>Cost of minor adverse effects</td>
<td>111</td>
</tr>
<tr>
<td>Cost of major adverse effects</td>
<td>1,998</td>
</tr>
<tr>
<td>Savings from strokes prevented</td>
<td>9,906</td>
</tr>
</tbody>
</table>

The EAC considered that the NICE costing report was a relevant and reliable source of data for the estimation of the indirect costs and savings associated with the WatchBP Home A.
However, all costs and savings were adjusted for inflation using a multiplier of 1.27, which as discussed in section 4.3.4.4 may have been an overestimate. This could be particularly significant when applying the multiplier to the cost of management of stroke.

The EAC noted that no costs were associated with the procurement of aspirin. However, costs associated with aspirin are low, and given the other uncertainties in the economic analysis is likely to be negligible at about £0.09 per day (29).

The EAC considered that the costs associated with stroke are associated with a high degree of uncertainty because of the heterogeneous nature of the condition, which can be fatal, debilitating, or cause minor disability. This was considered fully in the NICE costing report, which used the mean of two assessments of the annual direct costs of stroke of £3,900 and £11,700 to give a figure of £7,800. The sponsor recognised the uncertainty of this cost and performed univariate deterministic sensitivity analysis and threshold sensitivity analysis to investigate different costs of stroke.

4.3.5 Sensitivity Analysis

The sponsor performed univariate and limited multivariate deterministic sensitivity analysis on two inputs they considered to be of importance for the decision problem. These were the proportion of people with AF who are asymptomatic, and the cost of stroke. This type of sensitivity analysis is limited, because there are multiple sources of uncertainty that contribute to the economic model. However, the EAC accepted that more complex sensitivity analysis was unlikely to inform the decision model.

4.3.5.1 Asymptomatic patients

The sponsor ran the model using the assumption that 100%, 65%, or 50% of the patients were symptomatic. An assumption was made that asymptomatic patients would be screened for AF using the WatchBP Home A device, but not using pulse palpation. This meant that the numbers being screened using the WatchBP Home A device was kept constant (with the asymptomatic patients being incidental findings), but fewer people were screened using pulse palpation as the proportion of asymptomatic people increased. The sponsor noted that “Patients without symptoms are likely to be detected by the WatchBP device during routine blood pressure measurement, whereas they would remain undiagnosed by pulse palpation because there would be no indication for screening”.

The EAC did not agree that this was an appropriate approach to the decision problem for two main reasons:

- Firstly, by working backwards to calculate the number of people who would need to be screened for AF to arrive at the prevalence stated, it would mean that in the base case a very large proportion of the population (nearly two million people) would have to report to their GPs complaining of symptoms consistent with AF. This is not plausible and ignores the design of the studies that have been used to derive the clinical variables used in the model (for instance the study by Hobbs et al was carried
out in people over 65 years from the general population, and not targeted at people with symptoms (5)).

- Secondly, the sponsor has not taken into account the NICE clinical guideline on *Hypertension* (1), that states pulse palpation should be performed as a matter of course in all people who require measurement of their BP because automated devices may not be accurate if there is pulse irregularity (see Section 2.1.4).

The prevalence of asymptomatic AF in the general population is unknown (2). The sponsor’s estimate of 35% of cases of AF being asymptomatic was ultimately derived from a small crossover RCT (n = 43) (30), which could probably not be generalised to a national population. It was unclear to the EAC how the higher figure of 50% had been derived as it could not be identified in the text of the citation quoted (2). Therefore it was not clear how plausible these figures were, and, given that screening for irregular pulse should be carried out in most patients requiring BP measurement, what relevance they were. The EAC also considered the sponsor had deviated from the Scope in which the population is being opportunistically screened for irregular pulse as their BP is measured (and not screened on the basis of symptoms).

### 4.3.5.2 Cost of stroke

The baseline cost of stroke was calculated using data from the NICE costing report on *Atrial Fibrillation* (2) which had been modified for inflation. This equated to £9,906 per patient. Although there are uncertainties involved in the calculation of this figure (see Section 4.3.4.5), it could be considered as the best estimate available.

The sponsor identified another study that gave a much higher estimate of the cost of stroke. No information was given about how this study was identified or if any similar studies had been excluded, which gave rise to the possibility of confirmation bias. The study by Saka *et al* (2009) (31) estimated that the total cost of stroke to the UK economy was about £9,000,000,000 (nine billion pounds). From this, the sponsor estimated that the cost of stroke per patient was £44,000, although it was not clear to the EAC how this was calculated. This is nearly four and a half times the NICE estimate.

The EAC considered that this study was not relevant to the decision problem. The study took a societal perspective of the economic impact of stroke, whereas NICE adopt the perspective of the NHS and PSS in England and Wales. A societal perspective takes into account informal care costs (such as cost of carers) as well as indirect costs (such as loss of wages). This approach is associated with greater uncertainty than studies which adopt a third party payer perspective, and is not directly relevant to the use of NHS resources and therefore the decision problem at hand.

The sponsor presented a graph which illustrated the threshold analysis of cost of stroke compared with total costs of screening and treatment. The graph shows that both the WatchBP Home A and pulse palpation cost money compared to no screening, but that screening with pulse palpation is more expensive. To make the screening interventions cost
neutral, the cost associated with stroke would have to be £16,090 for the WatchBP Home A (data from Wiesel et al (13)), and £21,271 for opportunistic screening using pulse palpation (data from Hobbs et al (5)). Therefore, using the baseline cost of strokes, the modelled results report that both screening methods use more resources than they save (see Section 4.3.7), from a third party payer’s perspective.

### 4.3.5.3 Other inputs

Several other clinical parameters were subject to uncertainty but did not undergo sensitivity analysis by the sponsor. The EAC considered that the most important of these were:

- The measurement of sensitivity and specificity of the WatchBP Home A and pulse palpation. Although two studies were included for each screening method, the variation between these was not fully evaluated. Instead, the arithmetic mean (non-weighted) was used to calculate the definitive results provided in the narrative de novo submission. (Note: two studies were excluded by the EAC; see Section 4.3.4.2).

- The prevalence of AF in the population. There was a high degree of uncertainty regarding this value which the EAC suspected could have a large impact on all aspects of the analysis (section 4.3.4.1).

The EAC has considered these issues in *Additional Work* (see Section 4.6).

### 4.3.6 Results of Sponsor’s Base–Case Cost Analysis

The sponsor’s base case analysis assumed that all people with AF were symptomatic, and that the cost of stroke was £9,906. Using these figures, and excluding the studies the EAC deemed were unsuitable (see Section 4.3.4.2), the sponsor found that implementation of the WatchBP Home A device would lead to an annual saving of £9,165,000 for the NHS in England and Wales by displacing pulse palpation. The data for this calculation is shown in table 4.3.6.

*Table 4.3.6. Results of sponsor’s base-case cost analysis.*

<table>
<thead>
<tr>
<th>Cost</th>
<th>WatchBP Home A (£, thousands)</th>
<th>Pulse palpation (£, thousands)</th>
<th>Differential (£, thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anticoagulants</td>
<td>22,982</td>
<td>20,649</td>
<td>2,333</td>
</tr>
<tr>
<td>Major bleeds</td>
<td>2,266</td>
<td>2,036</td>
<td>230</td>
</tr>
<tr>
<td>Minor bleeds</td>
<td>825</td>
<td>741</td>
<td>84</td>
</tr>
<tr>
<td>Pulse palpation</td>
<td>0</td>
<td>4,595</td>
<td>-4,595</td>
</tr>
<tr>
<td>ECG</td>
<td>10,533</td>
<td>15,461</td>
<td>-4,928</td>
</tr>
<tr>
<td>Strokes prevented</td>
<td>-22,538</td>
<td>-20,249</td>
<td>-2,289</td>
</tr>
<tr>
<td>Total cost</td>
<td>14,068</td>
<td>23,233</td>
<td>9,165</td>
</tr>
</tbody>
</table>

Using the sponsor’s model, the WatchBP Home A was associated with increased expenditure in drugs used to prevent stroke and the adverse effects associated with these,
and a decreased expenditure in ECG costs due to a decrease in false positive results at the screening stage. There was also presumed to be no expenditure using the WatchBP Home A, whilst pulse palpation, which is applied to the entire screening population, was considered to use significant time resources, resulting in a large opportunity cost. In addition, the WatchBP Home A was associated with the prevention of 221 strokes, freeing NHS resources equivalent to £2,289,000 (less than half the time cost associated with pulse palpation). However, stroke has a large effect on a person’s quality of life as well as their life-expectancy, which would be an important factor should cost utility analysis be undertaken.

The EAC considered that there was significant uncertainty regarding many of the inputs to the model. For instance, there was no associated time cost with the Watch BP Home A, which in practice is likely to be unrealistic. If the time cost of pulse palpation is removed, the costs saved by the WatchBP Home A are reduced to £4,570,000. On the other hand, if it was assumed that the WatchBP Home A actually causes an excess time cost of one minute compared with pulse palpation (this is possible since it requires three successive inflations, see Section 4.3.4.4), then the device would actually be slightly more expensive than pulse palpation (by £25,000).

The EAC therefore noted that the model was highly sensitive to the assumptions used. This indicated the model was unlikely to be robust (see Section 4.5).

### 4.3.7 Results from Sensitivity Analysis

The sponsor identified the proportion of people with AF who were asymptomatic and the cost of stroke as parameters that should be investigated using sensitivity analysis. This was done using simple multivariate analysis, the results of which were discussed in the Results section under the Cost effectiveness IIa, IIb, IIIa, IIIb, and IIIc subheadings of the sponsor’s submission. The results of these analyses, with the EAC exclusion of two studies (see section 4.3.4.2.), are summarised in table 4.3.7

Table 4.3.7. Results of sensitivity analysis using data from Wiesel et al (13) and Hobbs et al (5).

<table>
<thead>
<tr>
<th></th>
<th>Cost WatchBP Home (£, thousands)</th>
<th>Cost pulse palpation (£, thousands)</th>
<th>Cost differential (£, thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case*</td>
<td>14,068</td>
<td>23,233</td>
<td>9,165</td>
</tr>
<tr>
<td>65% symptomatic</td>
<td>14,068</td>
<td>15,102</td>
<td>1,033</td>
</tr>
<tr>
<td>50% symptomatic</td>
<td>14,068</td>
<td>11,616</td>
<td>2,452</td>
</tr>
<tr>
<td>High cost stroke†, 100% symptomatic</td>
<td>-63,502</td>
<td>-46,460</td>
<td>17,042</td>
</tr>
<tr>
<td>High cost stroke, 65% symptomatic</td>
<td>-63,502</td>
<td>-30,199</td>
<td>33,303</td>
</tr>
<tr>
<td>High cost stroke, 50% symptomatic</td>
<td>-63,502</td>
<td>-23,230</td>
<td>40,272</td>
</tr>
</tbody>
</table>

* Base case stroke cost £9,906  
† High cost stroke cost £44,000  
- Denotes cost saving over no screening

The following can be seen from the sponsor’s economic model’s outputs:
• Assuming the base case cost of stroke from the NICE costing report on *Atrial Fibrillation* (2), an increasing proportion of asymptomatic patients leads to a reduced saving by the WatchBP Home A. In fact, if only 50% of patients are assumed to be symptomatic, the device costs £2,452,000 more than pulse palpation.

• In population which are 35% (or greater) asymptomatic, the WatchBP Home A is not cost effective if the cost of pulse palpation is excluded. Inclusion of capital investment costs (approximately £2.7 million) would increase the cost of the Watch BP Home A under this cost consequence analysis, thereby reducing reported savings.

• If the more expensive estimate of stroke is used (from Saka *et al* (31)), both pulse palpation and the WatchBP Home A device become cost effective and release resources for the NHS. However, the WatchBP Home A saves more resources than pulse palpation, and the differential increases with an increasing proportion of symptomatic patients.

However, the EAC considered that the results from this sensitivity analysis should be interpreted with caution because:

• The proportion of asymptomatic patients in the general population is not known (2), and it is uncertain how the sponsor selected the values of 35% and 50%. Furthermore, the assumption that AF is only screened for in symptomatic patients is fundamentally flawed and does not reflect the scope of the decision problem (see 4.3.5.1).

• The higher cost of stroke is not feasible and represents an estimate from a societal perspective, not that of the NHS and PPS in England and Wales (see Section 4.3.5.2).

### 4.4 MODEL VALIDATION

The EAC was able to fully reproduce the results reported by the sponsor. Calculation of the prevalence (4.4%) required iterative modelling to replicate (see Appendix A). The calculation of the number of true positives, false positives, true negatives, false negatives, and total population required to be screened were then calculated using the sensitivity and specificity data from the clinical studies, and the PPV and NPV from Hobbs *et al* (5) (see Appendix B). From this data the EAC were able to fully recreate the economic analysis and adjust the parameters to perform sensitivity analyses (see Section 4.5.2.).

Although the EAC was able to replicate the sponsor’s economic evaluation, it did not agree that their analysis was suitable for answering the decision question. In particular, the EAC considered that the backwards analysis the sponsor employed (calculating the factorial tables from the assumed incidence of AF and the calculated prevalence) was not appropriate and made too many assumptions. The EAC considered it would have been
better to calculate these parameters going “forward”. In addition the EAC did not agree that all the inputs were appropriate or correct. These issues are discussed in Sections 4.3.1, 4.3.2, and 4.3.3).

4.5 INTERPRETATION OF ECONOMIC EVIDENCE

4.5.1 Sponsor’s Interpretation

The sponsor provided a cost consequence model as evidence to support the economic benefits the WatchBP Home A might bring with regard to freeing resources within the NHS. For this model, the sponsor assumed that the device would only be used to screen for irregular pulse in populations who were at high risk of AF or had symptoms consistent with AF. The sponsor calculated that the prevalence of AF in this population would be 4.4% and used incidence data from the NICE costing report (6) (87,000 new cases of AF per year). By using this data and the PPV and NPV values reported in the HTA by Hobbs et al (5), the sponsor was able to calculate the proportion of patients who would be true positives, false positives, true negatives, and false negatives, as well as the total population required to be screened in order to achieve these results (nearly two million). From this, using data principally from the NICE costing report, the sponsor was able to calculate the net economic benefit of the WatchBP Home A compared with pulse palpation taking into account the cost of screening, the cost of anticoagulant and antiplatelet medication (and adverse effects associated with these), and the number of strokes prevented.

The sponsor assumed that the WatchBP Home A would have no associated capital investment or time costs. The sponsor provided multivariate sensitivity analysis by adjusting the number of people who were symptomatic of AF (100%, 65%, and 50%) as well as the cost of stroke (£9,906 and £44,000). The sponsor found that both screening techniques used net resources, unless the cost of stroke was considered to be about twice as high as stated in the NICE costing report, at which point screening begins to free resources. In most cases, the WatchBP Home A was found to be more economical than pulse palpation (the exception being 50% asymptomatic patients at the base case cost of stroke). The principal areas of savings of the device came from having no screening cost, reduced number of ECGs (i.e. fewer “unnecessary” ECGs), and more strokes prevented. Use of the WatchBP Home A incurred consequent costs for anticoagulation drugs and the adverse effects of these drugs, but these were outweighed by the aforementioned cost savings. To conclude, the sponsors asserted that the introduction of the WatchBP Home A device to community practices would lead to real resource savings for the NHS of England and Wales compared with the present method of pulse palpation.

The sponsor also discussed the potential economic benefits of using the WatchBP Home A at home to detect paroxysmal AF. Envisaged savings included a reduction in the need for ambulatory ECGs, and early detection of AF and consequent prevention of stroke. However, this use of the WatchBP Home A was not part of the de novo economic model, and no economic literature was provided to support this use.
4.5.2 EAC Critique of Sponsor’s Evaluation

The EAC considered that the results from the de novo economic model required caution in their interpretation. This was because some of the key assumptions made were questionable and there was a high degree of uncertainty concerning many of the parameters used to populate the model. These uncertainties are listed in table 4.5.2.a. (clinical parameters) and 4.5.2.b. (costing parameters), and discussed in more detail in Section 4.3.

**Table 4.5.2.a. Sources of uncertainty in clinical parameters.**

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>Description of uncertainty</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute incidence rate</td>
<td>87,000 new cases of AF per year.</td>
<td>Incidence rate of 1.75% was applied to entire population – likely to be an overestimate.</td>
</tr>
<tr>
<td>Prevalence</td>
<td>4.4% from hybrid calculation using data from SAFE study</td>
<td>Prevalence may not be representative of population being screened. Not included in sensitivity analysis.</td>
</tr>
<tr>
<td>Population number requiring screening</td>
<td>Value of 2 million obtained by back-extrapolation calculated using prevalence and data from SAFE study</td>
<td>The population estimated by this method is different from that identified in the Scope. The actual population screened should have reflected BP measurement (expected to be much higher).</td>
</tr>
<tr>
<td>Diagnostic accuracy of screening techniques</td>
<td>Doubts about external validity of WatchBP Home A studies.</td>
<td>Differences in populations and interventions used cause uncertainty. Indirect comparison with pulse palpation adds to this.</td>
</tr>
<tr>
<td>Number strokes prevented</td>
<td>Sensitive to the diagnostic accuracy of screening techniques</td>
<td>Doubts about the external validity of clinical studies for WatchBP Home A.</td>
</tr>
</tbody>
</table>

**Table 4.5.2.b. Sources of uncertainty in costing parameters.**

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>Description of uncertainty</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflationary multiplier</td>
<td>Inflation assumed to be 5% per annum to give multiplier of 1.27.</td>
<td>Recession in recent years makes this unlikely. Multiplier of 1.12 would more accurately reflect actual inflationary increases.</td>
</tr>
<tr>
<td>Cost of WatchBP Home A</td>
<td>Capital cost of Watch BP Home A (approx £2.8 M) not included in calculation.</td>
<td>Watch BP Home A would displace existing HBPM devices requiring upfront expenditure by incurring replacement costs.</td>
</tr>
<tr>
<td>Cost of pulse palpation</td>
<td>Time cost of one minute causes opportunity cost.</td>
<td>Elimination of need for pulse palpation may not achieve savings as calculated. No time</td>
</tr>
</tbody>
</table>
The EAC considered that economic model was flawed in that it derived important parameters in a reverse or backwards manner; that is it used the incidence data derived from different epidemiological studies together with diagnostic parameters derived from an RCT to inform about population prevalence. From this the number of people who would need to be screened to achieve this incidence was calculated. The EAC considered it would have been better to select and estimate the population consistent with the scope (i.e. screening, diagnosis or monitoring of hypertension), the likely prevalence of AF in this population, and report forwards from this point, using sensitivity analyses as appropriate.

In the opinion of the EAC, the unusual model structure combined with the uncertainty of some of the parameters that were used as inputs (see tables 4.5.2.a. and b.) caused the results of the economic analysis to be indeterminate. The sensitivity analysis that was performed did not usefully inform further about the decision question. Overall this meant that the EAC could not determine conclusively, from the sponsor’s submission, whether the WatchBP Home A was likely to free resources within the NHS.

### 4.6 ADDITIONAL WORK UNDERTAKEN BY THE EXTERNAL ASSESSMENT CENTRE IN RELATION TO ECONOMIC EVIDENCE

After deconstructing and replicating the sponsor’s economic model, the EAC conducted some additional multivariate sensitivity analyses to establish if alteration of key parameters would affect the outcome of the model. The following parameters were selected for sensitivity analysis:

- The prevalence of AF in the study population. Three values were chosen which were 1.28% (reflective of the prevalence of AF in the general population (6)); 4.4% (the base case value supplied by the sponsor); and 7.9% (reflective of the prevalence of AF in an over 65s population (5)).

- The sensitivity and specificity of pulse palpation and the WatchBP Home A. For pulse palpation, the point estimate and upper and lower 95% confidence intervals from the HTA by Hobbs et al (2005) were used (5). The study by Wiesel et al (2009) (13), used to estimate the diagnostic accuracy of the WatchBP Home A, did not report 95% confidence intervals, so these were estimated from the binomial distributions derived from the proportions described in the factorial tables.
The cost of pulse palpation. As the EAC considered the opportunity cost of pulse palpation was uncertain, values were calculated with and without this costing.

The population being screened was kept consistent with that of the sponsor (at almost two million). The costs per patient were also calculated; this allows potential cost savings to be scaled for any screening population size. All other parameters and assumptions were kept consistent with the sponsor’s base case analysis.

The results of the additional sensitivity analysis are shown in Appendix C.

From this sensitivity analysis it can be seen that:

- All the scenarios represent a cost to the NHS using this base case estimate of cost of stroke.
- Elimination of the time cost associated with pulse palpation did not make pulse palpation less costly than the WatchBP in any single scenario.
- The worst case sensitivity and specificity values increases costs in all cases; similarly the best case sensitivity and specificity values decrease costs in all cases compared with the base case sensitivity and specificity.
- Total cost is highly sensitive to prevalence of AF, with populations with a higher prevalence consuming more costs than those with a lower prevalence, both for the WatchBP and for pulse palpation. This is because of the greater usage of ECGs and the costs of anticoagulant and antiplatelet drugs (and their adverse effects), which are not offset by strokes prevented (using the base case cost of stroke).

Table 4.6. shows the ranges of costs per patient screened (one year time horizon) associated with the WatchBP Home A and its comparator (pulse palpation) based on the EAC’s additional analysis. For each value of AF prevalence considered, the range of per-patient costs associated with the use of the Watch BP is less than the range associated with manual pulse palpation, subject to the assumptions discussed previously.

Table 4.6. Ranges of cost per patient screened.

<table>
<thead>
<tr>
<th>Prevalence of AF (%)</th>
<th>Cost per patient screened</th>
<th>Pulse palpation (assuming no time cost)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WatchBP Home A</td>
<td></td>
</tr>
<tr>
<td>1.28</td>
<td>£3.98 to £6.11</td>
<td>£7.03 to £10.36</td>
</tr>
<tr>
<td>4.4</td>
<td>£6.29 to £8.18</td>
<td>£9.05 to £12.13</td>
</tr>
<tr>
<td>7.9</td>
<td>£8.89 to £10.50</td>
<td>£11.32 to £14.11</td>
</tr>
</tbody>
</table>

Appendix C shows that additional sensitivity analysis by the EAC, using the sponsor's model design, found that the savings per patient screened with the WatchBP compared with manual pulse palpation varied from £0.82 to £6.38, with a most likely saving of £2.58 per patient.
(1.28% prevalence of AF; no pulse palpation time saving; and pulse palpation accuracy based on the point estimate used in the SAFE study (5). In all cases examined, the WatchBP Home A delivered cost savings compared with pulse palpation.

4.7 CONCLUSIONS ON THE ECONOMIC EVIDENCE

The sponsor did not provide details of any additional economic studies retrieved from the literature, and it was assumed that no relevant studies were identified. The sponsor did supply a de novo economic model which could be described as a cost consequence model with consequences expressed in monetary units. This was a straightforward model that compared the WatchBP Home A with pulse palpation in a community setting (typically a GP’s surgery), with a time horizon of one year. Inputs into the model included the incidence and prevalence of AF in the population being screened and the diagnostic accuracy of the screening techniques for irregular pulse. Outputs included the proportion of people who received anticoagulant or antiplatelet drugs, and the number of strokes prevented. All consequences were assigned a monetary value and a final outcome of net cost to the NHS was calculated for each screening technique.

The sponsor found that in the base case analysis, the WatchBP Home A saved £11.6 million compared with pulse palpation by screening a population of nearly two million symptomatic patients. Additionally, the device would prevent 221 strokes. Sensitivity analysis showed that increasing the proportion of asymptomatic patients (those assumed not to be screened with pulse palpation) increased overall costs, because the cost of screening and treatment always outweighed resources saved by preventing strokes. However, limited multivariate and threshold sensitivity analysis suggested that the more expensive a stroke was estimated to be, the more cost saving the device was.

The EAC considered that the evidence provided from the economic model was subject to some uncertainties and several assumptions. These included:

- The model’s structure and fundamental assumptions. The model was constructed “backwards” using a hybrid technique and data inputs which were taken from the literature or calculated with the aid of diagnostic parameters taken from an RCT. The EAC considered that it would have been better to have created a “forward facing” model where cause and effect were clearly differentiated and data inputs were as relevant as possible.

- The population included was people who were at high risk of AF or who had symptoms consistent with AF, which was not consistent with the model described by the Scope. The absolute number of people screened was calculated from incidence data and estimated to be nearly two million people in the base case. This was considered to be an unfeasibly large number of symptomatic or high risk patients. In addition, this was not the population described in the Scope.
There was considerable doubt about the external validity of the studies that provided the sensitivity and specificity parameters for the WatchBP Home A, particularly regarding the setting, population, and intervention studied (number of readings taken by the AF detection algorithm).

There was uncertainty regarding the costing of the screening techniques used. In particular, the WatchBP Home A was assumed to have no associated capital cost for purchase, and pulse palpation was assumed to have a time cost which the device did not.

The EAC did not accept that the proportion of people who are symptomatic with AF, or the higher value of stroke used in the sensitivity analysis, were relevant to the decision problem.

The EAC performed further sensitivity analyses by varying the prevalence of AF in the screened population, the diagnostic accuracy of the screening techniques, and the assumption that pulse palpation incurs a time cost whereas the WatchBP does not. The EAC found that the difference in costs between the screening techniques was only moderately sensitive to changes in these inputs, and that the WatchBP Home A was associated with a lower range of costs than pulse palpation for each combination of parameters tested.

The sponsor discussed another potential use of the WatchBP Home A; that of a home setting to potentially detect or monitor paroxysmal AF. The EAC considered that this was an interesting and potentially useful application of the device. However, as no economic data for this indication was included, no conclusions were drawn.
Section 5: Conclusions

5.1 CONCLUSIONS ON THE CLINICAL EVIDENCE

The sponsor identified ten studies they considered were relevant to the decision problem (seven published and three unpublished). Of these, five completed studies (one unpublished) were directly relevant to the scope of the decision problem.

Three were cross-sectional diagnostic studies that were designed to measure the diagnostic accuracy (i.e. the sensitivity and specificity) of a pulse irregularity algorithm similar to that employed by the WatchBP Home A HBPM device in a clinical environment. The comparator used in these studies was 12-lead ECG, which in clinical practice is required to confirm AF and rule out differential diagnoses.

Two studies were set in a home setting. One of these was a case series with limited generalisability or usefulness in addressing the decision problem. The unpublished home study employed an ECG event monitor as its comparator, thus limiting its comparability to other studies of diagnostic accuracy.

5.1.1 Strengths of clinical evidence

The EAC considered that the WatchBP Home A’s BP measurement algorithm was validated in an outpatient population and that the AF detection algorithm had been well characterised compared with the gold standard (consultant-led 12-lead ECG). In this respect, the WatchBP Home A could be considered to have equal to or superior diagnostic accuracy than pulse palpation, particularly in terms of specificity. This means the device has the potential to improve the rule out of AF, and reduce expenditure on the number of ECGs required to confirm diagnosis compared with pulse palpation.

5.1.2 Weaknesses of clinical evidence and remaining uncertainties

The EAC considered that the included studies lacked generalisability to the setting of the NHS in England and Wales (community practice and home use). The populations studied were either at very high-risk of AF, or the AF status was known. In addition, the measurement regimens used to detect AF were different to that of the WatchBP Home A. The studies did not provide a head-to-head comparison with pulse palpation and the sponsor did not provide an indirect comparison (i.e. comparison of both with 12-lead ECG) in the clinical evidence section, which added a further degree of uncertainty to the diagnostic accuracy of the device in practice.

No evidence was presented about how accurately the device measures BP in people who are in irregular rhythm. Little evidence was presented in how effective the device is in detecting paroxysmal AF, which may be more relevant in the home setting.
5.2 CONCLUSIONS ON THE ECONOMIC EVIDENCE

The sponsor did not provide any economic evidence, relating specifically to their device, from the literature. Instead the sponsor presented a de novo cost consequence model with a time horizon of one year. In the model, the WatchBP Home A device to detect AF was compared with pulse palpation in a GP or nurse office setting. The inputs included the relative diagnostic accuracy of these screening techniques, and the use of anticoagulant or antiplatelet drugs to prevent stroke. The consequences were monetised, and the net economic benefit of each screening technique was calculated.

The sponsor calculated that, in the base case analysis, both techniques are a net cost to the NHS. However, use of the WatchBP Home A device costs less, would prevent 221 strokes per year, and would free resources for the NHS in England and Wales equivalent to £11.7 million per year, assuming that 1.98 million patients with previously undetected AF at a prevalence of 4.4% are screened per year (i.e. £5.91 saving per patient screened). Using sensitivity analysis, this saving was increased by assuming a higher cost of stroke. Additional sensitivity analysis by the EAC, using the sponsor’s model design, found that the savings per patient screened with the WatchBP compared with manual palpation varied from £0.82 to £6.38, with a most likely saving of £2.58 per patient (1.28% prevalence of AF; no palpation time saving; palpation accuracy based on HTA study). In all cases examined, the WatchBP Home A delivered cost savings compared with palpation.

5.2.1 Strengths of economic evidence

The cost consequence model adopted is accepted by NICE as being an appropriate method to calculate the economic benefit of the device. The model was populated using accepted reference sources and used an appropriate comparator. The EAC was able to replicate the sponsor’s calculations of potential cost savings.

5.2.2 Weaknesses of economic evidence and remaining uncertainties

The EAC considered that the model presented was subject to several uncertainties. These included:

- Model structure and fundamental assumptions. The population of the model was calculated “backwards” from an estimated incidence rate, which the EAC considered was not an optimal approach to the problem.

- Population. The population modelled (patients symptomatic of AF) was not that of the Scope (asymptomatic patients being screened or diagnosed for hypertension). In addition, the estimated prevalence of AF in the population was likely to be higher than the prevalence defined in the Scope.

- Time horizon. The time horizon was limited to one year which could have significantly underestimated the benefits and costs of the device. The sponsor did not consider costs and benefits over a longer period of time.
• Clinical inputs. The diagnostic sensitivity and specificity of the WatchBP Home A were derived from clinical studies the EAC considered lacked generalisability to the population specified in the final Scope.

• Time savings. In their model, the sponsor did not include time costs for using the WatchBP device itself, but did include costs associated with the time taken for pulse palpation. It remains uncertain whether or not eliminating the need for separate manual pulse palpation would release the full costs associated with the time saved in clinical practice, although additional analysis by the EAC which excluded this saving showed that the WatchBP remained cost saving overall compared with palpation, assuming other parameters were unchanged. It is also uncertain whether or not the "usual" mode operation of the WatchBP, which requires three successive inflations, may take longer than current practice.

• Device costs. The sponsor did not include the cost of purchasing the WatchBP in their submission. The EAC acknowledges that the device price is similar to alternative automated BP devices without an AF detection function, but in the absence of a more detailed lifetime cost analysis (including, for example relative reliability, replacement cuff prices, bulk purchase prices), there is residual uncertainty associated with the validity of the zero cost assumption.

• Sensitivity analysis. The EAC did not consider that all aspects of the sponsor's sensitivity analysis were relevant to the decision model (e.g. the consideration of different proportions of symptomatic AF). In addition, the model’s deterministic structure meant that the degree of uncertainty in the model was not accurately stated or visualised.

5.3 DIFFERENCES OF OPINION BETWEEN THE EAC AND SPONSOR

The EAC was not aware of any differences in opinion between the EAC and the sponsor.
The EAC considered that the WatchBP Home A is a promising device which, compared with screening by manual pulse palpation, could lead to improvements in the detection of AF, and consequent reduction in morbidity and mortality due to stroke and consequent cost savings. However, the current evidence base is poor and there remains uncertainty over whether these potential benefits will be realised in practice. The EAC has identified several areas for research to reduce these uncertainties.

- Assessment of the diagnostic accuracy of the WatchBP Home A in a pragmatic setting (community practice) using pulse palpation as the comparator. It is understood a relatively large clinical trial (n = 1000) with these criteria is currently in progress (8). Additional knowledge of the time taken, in practice, to use the WatchBP Home A in "usual" mode would help refine the economic model.

- Assessment of the accuracy of the WatchBP Home A device for the measurement of BP in people who have irregular heart rhythm.

- Assessment of the longer term use of the WatchBP Home A in a home setting, to opportunistically screen for paroxysmal AF, or monitor the effectiveness of anti-arrhythmic drugs.
Appendix A: Calculation of prevalence rate in population

The annual incidence of 0.175\% (6) was based on the proportion of entire population of the UK newly diagnosed with AF. If this annual incidence is applied to the population of England (49.9 million), we deduce that 87,000 new cases are diagnosed in England each year.

Out of every 5.7 patients with an irregular pulse, only 1 has the presence of AF confirmed by ECG. This translates to a PPV of $1/5.7=0.175$.

An iterative technique of numerical analysis has then been used to find the prevalence which gives a PPV value of 0.175 when the sensitivity and specificity stated by the SAFE-trial (Hobbs) of 87.15\% and 81.31\% respectively have then been placed into the equation:

$$PPV = \frac{(\text{sensitivity})(\text{prevalence})}{(\text{sensitivity})(\text{prevalence}) + (1 - \text{specificity})(1 - \text{prevalence})}$$

Matlab (Mathworks, Cambridge, UK) was used to simulate 10,000 values of prevalence between 0.0001 (0.01\%) and 1.0 (100\%). These were then combined with sensitivity and specificity in the above equation until the estimated value of PPV was $0.175 \pm 0.001$.

58 solutions were found which satisfied the above equation for an estimated PPV between 0.174 to 0.176. The median solution of prevalence was 4.4 (range 4.07 to 4.64) \%.

So for 1 patient to have AF detected out 5.7 with an irregular pulse, with a machine of 87.15\% sensitivity and 81.31\% specificity, the prevalence of AF in the population tested must be 4.4\%.
Appendix B: Replication of Sponsor’s *de novo* economic results

Inputs: Sensitivity 0.97\(^a\), specificity 0.89\(^a\), prevalence 0.044
Output: PPV = 0.28869047, R* = 3.46

<table>
<thead>
<tr>
<th>In AF</th>
<th>Not in AF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +ve</td>
<td>84,390</td>
<td>207,930</td>
</tr>
<tr>
<td>Test –ve</td>
<td>2,610</td>
<td>1,682,343</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>87,000</strong></td>
<td><strong>1,890,273</strong></td>
</tr>
</tbody>
</table>

Inputs: Sensitivity 1.00\(^b\), specificity 0.89\(^b\), prevalence 0.044.
Output: PPV=0.29498525, R* = 3.39

<table>
<thead>
<tr>
<th>In AF</th>
<th>Not in AF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +ve</td>
<td>87,000</td>
<td>207,930</td>
</tr>
<tr>
<td>Test –ve</td>
<td>0</td>
<td>1,682,343</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>87,000</strong></td>
<td><strong>1,890,273</strong></td>
</tr>
</tbody>
</table>

Inputs: Sensitivity 0.91\(^c\), specificity 0.74\(^c\), prevalence 0.044
Output: PPV=0.13873873, R* = 7.21

<table>
<thead>
<tr>
<th>In AF</th>
<th>Not in AF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +ve</td>
<td>39,585</td>
<td>245,735</td>
</tr>
<tr>
<td>Test –ve</td>
<td>3,915</td>
<td>699,400</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43,500</strong></td>
<td><strong>945,135</strong></td>
</tr>
</tbody>
</table>

Inputs: Sensitivity 0.8715\(^d\), specificity 0.8131\(^d\), prevalence 0.044
Output: PPV=0.17669144, R* = 5.66

<table>
<thead>
<tr>
<th>In AF</th>
<th>Not in AF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +ve</td>
<td>37,910</td>
<td>176,645</td>
</tr>
<tr>
<td>Test –ve</td>
<td>5,590</td>
<td>768,486</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43,500</strong></td>
<td><strong>945,131</strong></td>
</tr>
</tbody>
</table>

\(^a\) Sensitivity and specificity of Watch BP (Wiesel et al (13))
\(^b\) Sensitivity and specificity of Watch BP (Stergiou et al (15))
\(^c\) Sensitivity and specificity of pulse palpation (Morgan et al (23))
\(^d\) Sensitivity and specificity of pulse palpation (Hobbs et al (5))
Appendix C: Sensitivity Analysis of de novo economic model

<table>
<thead>
<tr>
<th></th>
<th>Input</th>
<th>Output</th>
<th>Cost Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion of population studied in AF (%)</td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
</tr>
<tr>
<td>Population size</td>
<td>1,977,273</td>
<td>4.4</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>4.4</td>
<td>96.7</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>4.4</td>
<td>91.9</td>
</tr>
<tr>
<td>Watch BP</td>
<td>1,977,273</td>
<td>7.9</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>7.9</td>
<td>96.7</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>7.9</td>
<td>91.9</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>96.7</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>91.9</td>
</tr>
<tr>
<td>Pulse palpation</td>
<td>1,977,273</td>
<td>4.4</td>
<td>91.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>4.4</td>
<td>87.2</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>4.4</td>
<td>82.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>7.9</td>
<td>91.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>7.9</td>
<td>87.2</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>7.9</td>
<td>82.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>91.1</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>87.2</td>
</tr>
<tr>
<td></td>
<td>1,977,273</td>
<td>1.28</td>
<td>82.1</td>
</tr>
</tbody>
</table>

a Total study population needed to give 87,000 AF cases assumed by sponsor.
b Upper 95% confidence limits of sensitivity and specificity calculated from binomial distribution
c Lower 95% confidence limits of sensitivity and specificity calculated using binomial distribution
d Prevalence rate in control arm of Hobbs et al (5)
e Prevalence rate of general UK population (Majeed et al, (24)).
f Upper 95% confidence limits of sensitivity and specificity (Hobbs et al (5))
g Lower 95% confidence limits of sensitivity and specificity (Hobbs et al (5))
Bibliography

8. Wessel S, Der NV, Den BV, Cammenga M, Montfrans GV. Microlife WatchBP for home blood pressure measurement more accurate in 'Diagnostic' mode compared to 'Usual' mode. 2010.
14. Stergiou GS, Giovas PP, Gkinos CP, Patouras JD. Validation of


