YHEC
York Health Economics Consortium

Providing consultancy and research in health economics for the NHS, pharmaceutical and health care industries since 1986

WAKEFIELD CCG AND YORKSHIRE AND HUMBER AHSN

Evaluation of MyDiagnostick

Final Report

JO HANLON, Research Consultant
MICHELLE JENKS, Senior Research Consultant
NICK HEX, Associate Director

AUGUST 2016

UNIVERSITY of York

INVESTORS IN PEOPLE

Contents

All reasonable precautions have been taken by YHEC to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall YHEC be liable for damages arising from its use.

Executive Summary

1. INTRODUCTION

NHS Wakefield Clinical Commissioning Group (CCG) is keen to promote active diagnosis of atrial fibrillation (AF) in primary care. Public Health England estimates that the expected prevalence of AF in the Wakefield registered population in 2013/14 was 2.4%, compared to an observed prevalence of 1.7% [1]. Stroke prevention in patients with AF, by appropriate anticoagulation, is a priority work stream within Wakefield. To this end, the CCG, working together with Wakefield's Public Health team, is testing out the use of a new device for AF case finding in primary care in Wakefield.

The device being used is MyDiagnostick, distributed in the UK by the company Technomed Ltd and manufactured by MyDiagnostick Medical BV. MyDiagnostick is a two-lead ECG recorder in the shape of a stick, with metallic handles (electrodes) at both ends. If the device detects an irregular heartbeat, further investigations are necessary. In Wakefield, the patient will be referred for a 12 lead electrocardiogram (ECG) and if AF is confirmed, the patient follows the AF management pathway. The two-lead ECG record stored in MyDiagnostick can be documented in the clinical record.

York Health Economics Consortium has conducted an evaluation of the implementation of MyDiagnostick in Wakefield. The objectives of the evaluation were:

- To summarise the use of MyDiagnostick in Wakefield, based on the findings of the Wakefield CCG system analysis;
- To assess the acceptability of the MyDiagnostick device to healthcare professionals using the device;
- To assess the impact of MyDiagnostick on the running of clinics and clinician time;
- To assess the views of patients relating to acceptability of the device;
- To evaluate the cost implications and potential savings of using MyDiagnostick compared with the pulse check method as the first line method for diagnosing AF.

Initial intentions to use local data to assess the conversion rate from a positive MyDiagnostick result to a diagnosis of AF proved to be unachievable, due to the sample size required. The work undertaken to inform this feasibility analysis is described in Appendix A.

2. METHODS

A mixed methods approach was used to address the evaluation objectives. The use of MyDiagnostick by Wakefield practices was established through a pragmatic evaluation of GP databases (SystmOne and EMIS clinical systems). The number of positive screening tests with MyDiagnostick, and the subsequent number of positive diagnoses of AF, were determined using codes entered by clinicians during patient consultations.

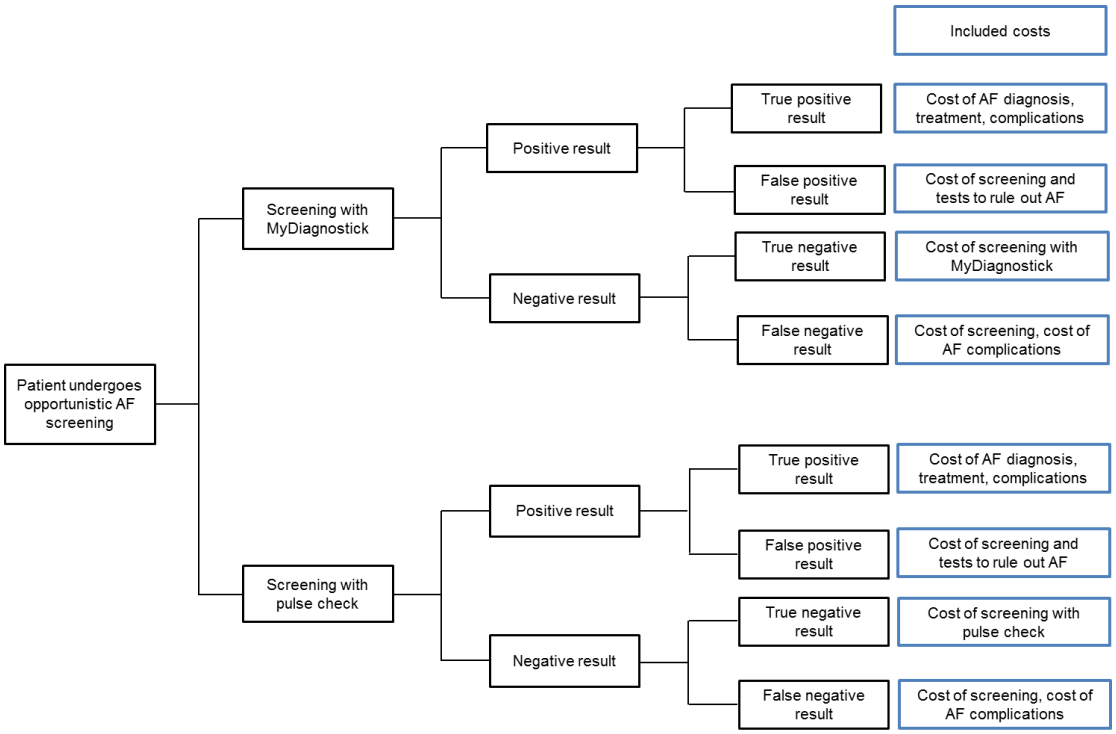
A qualitative survey of healthcare professionals was undertaken to assess the acceptability of the device in practice. An electronic survey was distributed via the commissioning lead in Wakefield. This asked about how the device was used in practice, the ease of use, advantages and disadvantages of using the device and the impact on consultation time. Respondents to the survey were asked if they were prepared to be involved in a short telephone interview to provide more detailed views on the device.

A paper based survey was developed to determine the views of patients on the ease of use of the device. The survey was designed to be quick and easy to answer in order to improve the response rate and all questions had multiple choice answers. The surveys were distributed to three practices with a large number of recorded uses of MyDiagnostick, to be returned to YHEC in a stamped addressed envelope.

A *de novo* cost analysis was performed by developing a user-friendly model in Microsoft *Excel*. The model structure took the form of a decision tree, whereby patients underwent screening for AF with either MyDiagnostick or the pulse check method. The head-to-head comparison of MyDiagnostick and the pulse check method is not fully in line with the feedback received from interviewees on how the device was being used within NHS Wakefield. However, this comparison was made based on the available clinical data which compared MyDiagnostick to the pulse check method. As such, patients fell into one of four categories:

- True negative: tested negative for AF and not in AF;
- False negative: test negative for AF, but in AF;
- True positive: tested positive for AF and in AF;
- False positive: tested positive for AF, but not in AF.

Costs were assigned to each of the four outcomes according to the following decision tree structure:



Information on the costs of each method, the number of patients, prevalence of AF, the effectiveness of the tests and the costs of complications, was derived from the interviews, the manufacturer and the available literature. These are described in detail in the full report. The model may be updated once further data become available.

The model presents both the cohort and per patient costs for those screened using MyDiagnostick and those screened using the pulse check method. The assumptions used to create the base case in the model were conservative. Sensitivity analysis was conducted around the model inputs that were uncertain, to test out which parameters had the most influence on the results.

3. RESULTS

The use of MyDiagnostick by general practices in Wakefield increased over time from July 2015. Up to 25th April 2016, 60% of the practices had used the device, with a mean of 53 recorded uses per practice. Of those patients with a positive reading recorded (n=42), 3.33% tested positive for AF with MyDiagnostick. Of these, 38.10% had since had AF confirmed (n=16). Unfortunately, due to the nature of the codes being used, it has not been possible to compare this conversion from positive test to AF diagnosis with the estimated rate of diagnoses in patients undergoing a pulse check.

The staff survey received 22 responses and three staff were interviewed. Those practices not taking up the offer of the device were asked for their views on why they were not interested in it, but no responses were received. The device was typically used by the nurses, around 10 times per week, in clinics for people with long term conditions, and also opportunistically in routine appointments where appropriate. Staff found it easy to use, adding very little time to the consultation. The use of MyDiagnostick within practices appears to be different to that anticipated by the project team, with interviewees suggesting that the device was actually used in addition to the pulse check. Those patients in whom an irregular pulse was suspected following a pulse check would then use MyDiagnostick to confirm or refute this finding, thereby 'ruling out' the need for a 12-lead ECG for those with a negative result.

A total of 23 patient surveys were returned to YHEC by the end of March 2016. The results of the survey showed that patients found the MyDiagnostick device easy to use and would be happy to use it again. All understood what the device was being used for and were aware of the test result.

The *de novo* cost analysis found the costs per patient screened with MyDiagnostick and the pulse check methods were £263.00 and £268.63 respectively. The base case results show that MyDiagnostick saves £5.63 per patient over a 10 year time horizon. These values take into account all of the associated costs for each method, not just the cost of performing the tests themselves. If a practice uses MyDiagnostick on average 10 times per week, the costs for the cohort of 520 patients per year would be £136,759 and £139,687 for MyDiagnostick and pulse check respectively, a difference of £4,013 over a 10 year time horizon. The model includes the option to generate outputs using evidence from a different study on the sensitivity and specificity of the MyDiagnostick device. This showed that the costs per patient screened with MyDiagnostick and the pulse check methods were £260.91 and £268.63 respectively. The costs for a cohort of 520 patients per year would be £135,674 and £139,687.

The sensitivity analysis showed that the key drivers of the cost analysis were the cost of a false negative result and the cost of a true positive result. The greatest costs in the analysis are those associated with further AF related costs i.e. the cost of treating diagnosed AF and/or the health consequences of unmanaged AF, such as a stroke. As the MyDiagnostick device is more sensitive than the pulse check method, it should reduce the number of false negative results. Therefore, the lower the cost of a false negative result to the health and social care system, the more unlikely the cost of the device would be offset by the savings. Similarly, as MyDiagnostick is more likely to detect true cases of AF, when the cost of treating true cases of AF rises, the savings from using the device would not offset the cost of subsequent treatment.

4. DISCUSSION

The combination of qualitative and quantitative evaluation methods has provided a comprehensive picture of the first stages of the implementation of MyDiagnostick in Wakefield. Furthermore, the staff survey and interviews provided valuable information to inform the cost analysis, with some parameters in the model (e.g. the time taken and number of patients using the device) being directly influenced by their experience of implementing MyDiagnostick.

The use of MyDiagnostick by general practices in Wakefield has increased over time from July 2015, with a mean of 53 recorded uses per practice up to 25th April 2016. The pattern of usage within individual practices appears to vary substantially, with some practices having fairly steady usage and some tailing off after an initial flurry of activity. It is, therefore, too early to tell if the use of MyDiagnostick will continue in the practices that currently appear to be enthusiastic about it. Furthermore, based on the uptake of MyDiagnostick over the nine month period studied, it appears likely that there will be some practices that choose not to use it at all. Local commissioners in Wakefield will continue to monitor the usage of the device over the coming months to inform the decision about whether to continue to promote it to practices.

The patient survey showed that the patient experience of MyDiagnostick was unanimously positive. The staff survey results suggest that the vast majority of respondents were also positive towards MyDiagnostick, finding that it is easy to use and it provides a greater degree of confidence in the result of a pulse check. A small number of respondents were less positive about the device and did not feel there was a need for it. As the respondents to the staff survey were self-selecting, it is possible that there could be further negative views about MyDiagnostick that have not come to light through this evaluation.

The device was used mainly by nurses and not by GPs. This is probably explained by the fact that the appointments when MyDiagnostick was most commonly used were the clinical contacts usually delivered by the nursing team e.g. long term conditions annual reviews. In these circumstances, the nurses were using MyDiagnostick as an additional tool to check for AF and not as a replacement for a pulse check.

While a prospective analysis of effectiveness was not possible using Wakefield's own data, the *de novo* cost analysis was able to evaluate the cost implications and potential savings of using MyDiagnostick, compared with the pulse check method. The cost analysis showed that the use of MyDiagnostick in general practice has the potential to achieve small cost savings from an NHS perspective. This is largely due to the differences in further AF related complication costs, some of which may be avoided by earlier diagnosis of AF by using a more sensitive and specific test.

The main limitation to the cost analysis is that it is theoretical modelling based on evidence from the literature and not actual diagnoses and prevalence of AF in Wakefield district. The results of the cost modelling appear quite robust to individual parameter uncertainty, as the sensitivity analysis showed that the changes required to take the parameter values over £0 are quite unlikely to occur. For example, the appointment time using MyDiagnostick would need to be an average of 7.5 minutes longer than when using the pulse check method.

5. CONCLUSIONS AND RECOMMENDATIONS

In conclusion, the evaluation has found that the MyDiagnostick device has been well received by those using it and has the potential to be moderately cost saving in the longer term. It is not possible to tell definitively whether the device is helping to detect undiagnosed AF more than with the pulse check method from this study; however, the inputs used within the cost model (from single arm studies) suggest that MyDiagnostick does detect more cases of AF than the pulse check method. Indeed, the device is not being used quite as anticipated and is not replacing the pulse check method in every case. The use of the device in Wakefield does not appear to have reached its optimum position in order to judge whether it is going to be embedded into routine practice in the longer term. It may be useful to repeat the local data analysis and clinician survey in 6 to 12 months' time to gain a fuller picture. The moderate cost savings may be insufficient to prompt take up of the device, particularly when the savings appear to be downstream from primary care. It does, however, have the potential to reduce the need for a 12-lead ECG to rule out a diagnosis of AF, which may be appealing.

The following recommendations are made based on the evaluation findings:

- The number of devices should be increased for those practices that report a positive experience of using MyDiagnostick and would find it helpful to have additional devices, for example, one in each of the practice nurse rooms;
- The usage of MyDiagnostick should continue to be monitored to observe whether the pattern of usage changes, whether late adopters choose to take up the device and whether the early adopters continue to use it;
- A follow-up staff evaluation later in 2016 may be useful to gain views once the device has been in place long enough to become embedded into routine practice;
- A head-to-head study comparing MyDiagnostick and pulse check method would provide information on the efficacy of the device and add more confidence to the results. This is not possible in the Wakefield CCG area, however.

Acknowledgements

York Health Economics Consortium would like to thank all participants for completing the survey, particularly those who took part in a follow-up interview. We are also grateful to those individuals who forwarded on the survey to people within their general practice. Thanks are also due to the project team from NHS Wakefield: Gillian Richardson, Paul Jaques, Nyasha Mareya and to David Philpott from Technomed Ltd.

Section 1: Introduction

1.1 BACKGROUND

NHS Wakefield Clinical Commissioning Group (CCG) is keen to promote active diagnosis of atrial fibrillation (AF) in primary care. Public Health England estimates show that the expected prevalence of AF in the Wakefield registered population in 2013/14 was 2.4%, compared to an observed prevalence of 1.7% [1]. Stroke prevention in patients with AF, by appropriate anticoagulation, is a priority work stream within Wakefield, and therefore the CCG is keen to ensure that all appropriate patients are being identified to reduce their risk of stroke.

To this end, the CCG, working together with Wakefield's Public Health team, is testing out the use of a new device for AF case finding in primary care in Wakefield. The device being used is MyDiagnostick, distributed in the UK by the company Technomed Ltd and manufactured by MyDiagnostick Medical BV. Funding was identified to purchase one MyDiagnostick for each of the 40[1] general practices in the NHS Wakefield CCG area. All general practices were asked if they would like a device and, if so, were trained in using the device.

MyDiagnostick is an easy to use, non-invasive medical device intended to discriminate AF from normal cardiac rhythm. It is a two-lead ECG recorder in the shape of a stick, with metallic handles (electrodes) at both ends. The patient holds the two ends of the stick for one minute, after which MyDiagnostick will display either a green or red indicator. A green indicator means that no AF was detected and the patient can be reassured. If the indicator turns red, the device has detected an irregular heart beat and further investigations are necessary. In Wakefield, the patient will be referred for a 12 lead electrocardiogram (ECG) and if AF is confirmed, the patient follows the AF management pathway.

MyDiagnostick is reportedly 98% sensitive and 96% specific for detecting AF in a population with a known 54% prevalence of AF [2]. The two-lead ECG record stored in MyDiagnostick can be retrieved through a USB connection by the physician, enabling clinical confirmation and documentation of AF in the clinical record. The device software can be used on the SystmOne and EMIS clinical systems.

[1] One of the 40 Wakefield practices is also the Walk-In Centre.

Prior to finalising the objectives of this evaluation, some preliminary work was undertaken. This work considered the feasibility of undertaking an evaluation to assess whether the conversion rate from a positive MyDiagnostick result to diagnosed AF is greater than the conversion rate from a positive pulse check to diagnosed AF (usual care). A secondary evaluation aim was whether MyDiagnostick increases the detection rate of AF above the current background rate in the patient cohort. Based on some initial analysis of the sample size required to demonstrate an observable effect, it was decided that meeting these objectives was unachievable. The work undertaken to inform this feasibility analysis is described in Appendix A.

1.2 OBJECTIVES OF THE EVALUATION

Following the feasibility study, the objectives of the evaluation were agreed as follows:

- To summarise the use of MyDiagnostick in Wakefield, based on the findings of the Wakefield CCG system analysis;
- To assess the acceptability of the MyDiagnostick device to healthcare professionals using the device;
- To assess the impact of MyDiagnostick on the running of clinics and clinician time;
- To assess the views of patients relating to acceptability of the device;
- To evaluate the cost implications and potential savings of using MyDiagnostick compared with the pulse check method as the first line method for diagnosing AF.

Each of these objectives is addressed in the remainder of this report. In Section 2, the evaluation within Wakefield CCG is summarised. In Section 3, a survey of healthcare professionals, to assess the acceptability of the device in practice and impact on clinician time, is reported. In Section 4, a survey of patients using the device is described. Finally, in Section 5, a *de novo* cost analysis of MyDiagnostick is reported.

Section 2: Evaluation within Wakefield CCG

2.1 INTRODUCTION

An evaluation is currently being conducted within NHS Wakefield CCG to determine the conversion rate from positive test with MyDiagnostick to AF diagnosis. The results of this evaluation, to 25th April 2016, are reported within this section.

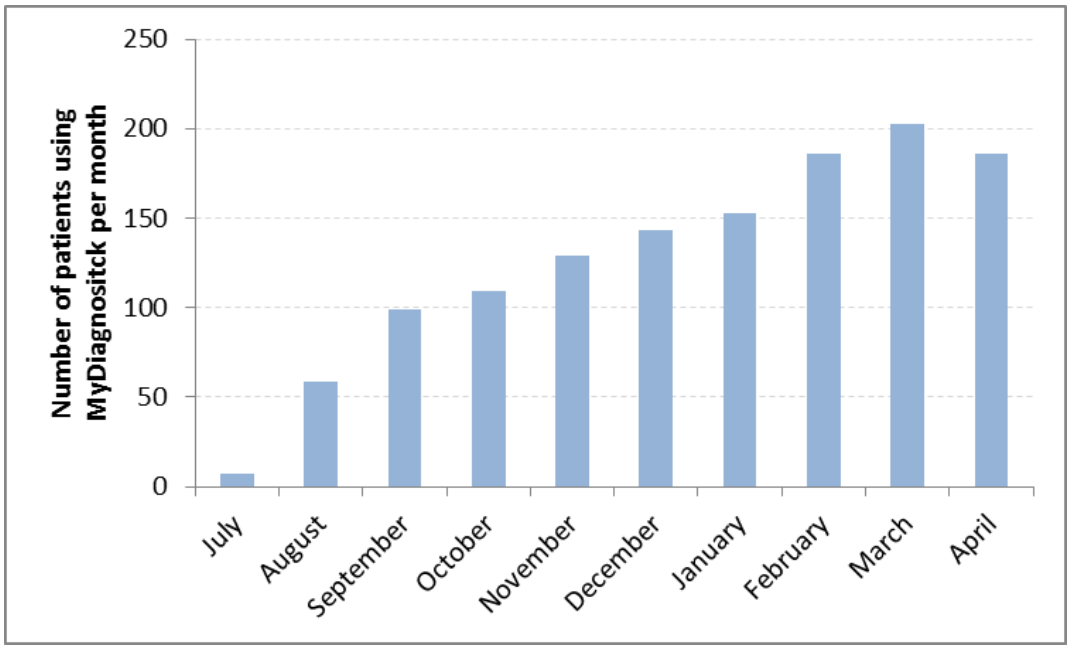
2.2 METHODS

The diagnosis rates with MyDiagnostick were established through a pragmatic evaluation of GP databases (SystmOne and EMIS clinical systems). The number of positive screening tests with MyDiagnostick, and the subsequent number of positive diagnoses of AF, were determined using codes entered by clinicians during patient consultations.

2.3 RESULTS

By 25th April 2016, 24 of the 40 (60%) general practices had registered the use of a MyDiagnostick device on their clinical system. A total of 1,274 checks with the device had been recorded, with this number increasing over time as more practices started using the device. The total number of patients using the device each month is shown in **Figure 2.1** (note: usage in April was only up to 25th of the month).

Figure 2.1: Use of MyDiagnostick per month in Wakefield

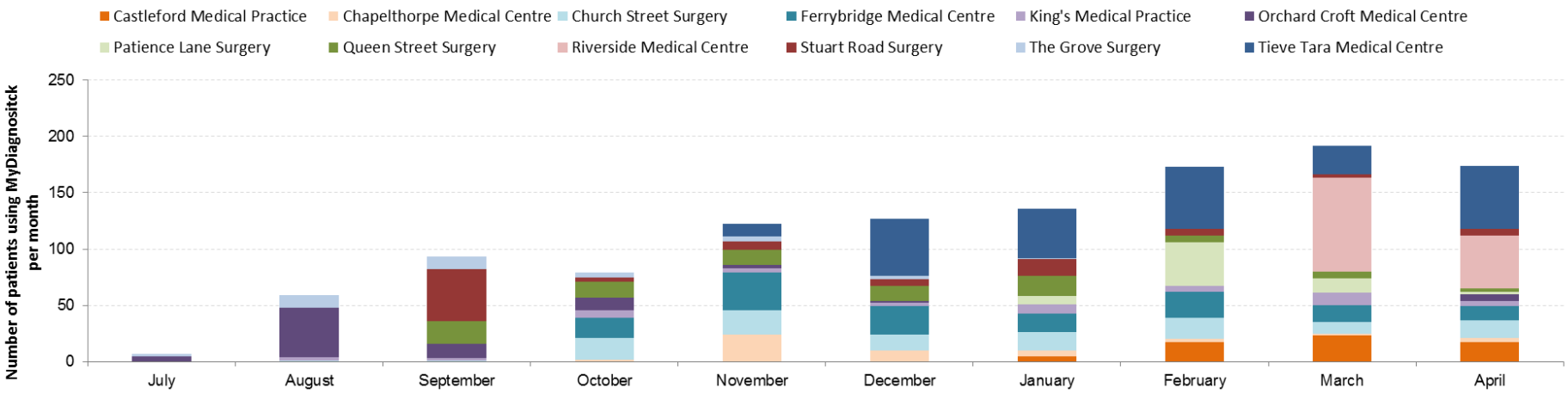


Further analysis was conducted based upon the number of times the device was being coded as being used by practices. The range of usage by practices varied widely (between 1 and 243 uses) with a mean of 53 uses per practice. The variation between practices is likely to be a result of: when practices started using their device; how large the practice is and how many clinical staff were enthusiastic about using the device. The number of uses per month by practice is shown in Figure 2.2. For clarity of interpretation, only the 12 most frequent users of the device have been included. The general practices not included in Figure 2.2 have used the device fewer than 30 times. It is apparent that some practices, for example, Orchard Croft, used the device a lot initially, but that this had tailed off in recent months. Conversely, other practices, such as Tieve Tara Medical Centre and Church Street Surgery, have maintained a consistently high usage. Riverside Medical Centre had only begun using the device (or coding the use of the device) in the last couple of months, but had very high usage.

The 1,274 uses of the device comprise 1,261 patients, suggesting that some patients had been checked on more than one occasion. Forty-two patients had a positive reading recorded, meaning 3.33% tested positive for AF with MyDiagnostick. Of these 42 patients, 16 (38.10%) had AF confirmed. Six of these 16 patients had a previous AF diagnosis on their record. Due to the nature of the coding it is impossible to tell whether the 26 remaining patients had a 12-lead ECG which confirmed they did not have AF; had a 12-lead ECG but the AF diagnosis has not yet been recorded or had not yet undergone a 12-lead ECG to confirm a diagnosis. Based on the data identified, 1.27% (16 of 1,261 patients) at high risk of AF had AF confirmed. Excluding those patients with a previous diagnosis, 0.80% (10 of 1,255 patients) of patients had AF confirmed.

Due to the nature of the codes being used, the rate of diagnoses of patients undergoing a pulse check cannot be estimated.

Figure 2.2: Use of MyDiagnostick per month by general practice



2.4 DISCUSSION

Whilst the coding data provide information on the number of people using MyDiagnostick, as well as the proportion of these that are positive, there are a number of limitations with these data.

The key limitation is that the analyses rely on the correct and complete coding of the use of the device. In order for accurate data to be generated, clinicians must have correctly coded each time they: used the device; had a positive result from the device and had a confirmation of this positive result. All of these elements are required simply to determine the true positive and false positive rates with MyDiagnostick. In addition, the device should be used in patients at risk of (but without an existing) AF diagnosis. It is apparent from the data identified that the device was being used in patients who were already known to have AF. The extent of missing coding is unknown.

A further limitation is that in order to make any comparison with the pulse check method, data on the use of pulse checks are required. Unfortunately, these data could not be retrieved for a number of reasons. Firstly, the clinical coding relating to the pulse check method does not always specify the site of the body at which the pulse check was taken. As such, those pulse checks of interest to the evaluation (wrist) could not confidently be disaggregated from all other pulse checks. Secondly, the reason for conducting a pulse check is not generally specified in the clinical record. Therefore, pulse checks being undertaken for other purposes than to screen for AF would have contributed to the number of patients with negative pulse check screening. Thirdly, feedback from clinicians demonstrated that the pulse check code may be used where pulse monitoring has been undertaken using blood pressure measurement or pulse oximetry rather than a manual pulse check. Thus, the comparator group is potentially not specific to pulse check alone. In addition, it was apparent from the data retrieved that clinicians were using *both* a pulse check and MyDiagnostick in the same appointment with the same patients. This was confirmed in interviews with clinicians (detailed in Section 3.3.8).

To conclude, whilst the coding data provides some information on the number of uses of the device and subsequent outcomes for patients using the device, there are serious limitations with the robustness of the data due to the reliance on complete and correct information being inputted by clinicians. Hence, the data should be interpreted with caution.

Section 3: Survey of Healthcare Professionals

3.1 INTRODUCTION

A survey of healthcare professionals with experience of using MyDiagnostick was conducted, to gain their views on acceptability of the device and the impact of using the device on their time and running of general practice.

3.2 METHODS

An electronic survey was developed by the authors and revised by the wider project team. This is shown in full in Figure 3.1. The survey was developed to find out whether the respondent was using MyDiagnostick and if so, their views on the device, including any advantages or disadvantages of the device. Respondents were also asked how using the device impacted upon clinic running times and whether the number of devices within their practice was adequate. The survey comprised a combination of closed and open-ended questions and was designed to be quick to answer in order to improve the response rate. To further improve the response, the link to the survey was distributed via email by Gillian Richardson, (Public Health Principal for Stroke & Cardiology) who had an existing relationship with the practices.

Respondents to the survey were asked if they were prepared to be involved in a short telephone interview to provide further, more detailed, views on the device.

The responses to the survey were analysed thematically in Microsoft Excel. For closed questions, summary statistics were generated, whilst for open-ended questions themes were determined and reported.

Figure 3.1: Survey of Healthcare Professionals

1. Have you used the MyDiagnostick device?
- **a.** Yes (go to question 3)
- **b.** No (go to question 2)

2. What is your reason for not using the MyDiagnostick device?
- **a.** I didn't know about it
- **b.** I knew about it, but have had no opportunity to use the device
- **c.** I knew about it, but have been unable to use it due to technical reasons (please specify reason in a free text box)
- **d.** I knew about it, but the device does not appeal to me (please specify reason in a free text box)
- **e.** Other (please specify in free text box)

Survey ends after this question for respondents not using the device.

3. How easy have you found the device to use:
- **a.** Very easy (go to question 5)
- **b.** Easy (go to question 5)
- **c.** Neither easy, nor difficult (go to question 5)
- **d.** Difficult (go to question 4)
- **e.** Very difficult (go to question 4)

4. Please explain what you found difficult about using the device.

5. Compared with using the pulse check method for detecting atrial fibrillation, does the MyDiagnostick device result in:
- **a.** A longer consultation time (go to question 6);
- **b.** A shorter consultation time (go to question 6);
- **c.** Consultation time of about the same length (go to question 7).

6. Please (a) provide an estimate of the difference in consultation time and (b) explain why this difference may occur.

7. Please describe any benefits of using the device.

8. Please describe any disadvantages of using the device.

9. Is the number of MyDiagnostick devices currently available within your practice adequate?
- **a.** Yes (go to question 11)
- **b.** No (go to question 10)
- **c.** I don't know (go to question 11)

10. How many devices would be required for optimal usage of the device?

11. Would you be prepared to be involved in a telephone interview to give further views on the use of the MyDiagnostick device in general practice?
- **a.** Yes (go to question 12)
- **b.** No (end survey)

3.3 RESULTS

The results of the survey are reported by theme in the following sub-sections.

3.3.1 Response to Survey

The survey was sent to 39 people at practices where training on MyDiagnostick had been provided and a device issued. These 39 people were a mixture of practice managers and clinical staff. In some cases more than one person per practice was sent the survey. Each contact was asked to forward the survey to those individuals within the practice who were using the device. Therefore, the total number of people who received the survey is unknown. By the initial closing date of 12th February 2016, 15 responses had been received. The time frame for responding to the survey was extended and a reminder email sent out on 18th February. At the final closing date of 26th February 2016, 22 responses had been received.

3.3.2 Use of MyDiagnostick

Of the 22 respondents, 19 (86%) had used the MyDiagnostick device. Of the remaining three respondents (14%), two had not used the device, as it did not appeal to them. These respondents felt that the device seemed to be "*a solution looking for a problem*" and "*did not see a place for it*". The third non-user had not used the device as they were the practice manager so were not qualified to do so, but noted, "*I wanted to check the survey before I sent it to the clinicians in my practice*". The 19 respondents who had used the device continued with the remainder of the survey. The findings from these 19 respondents are summarised below.

3.3.3 Ease of Use of the Device

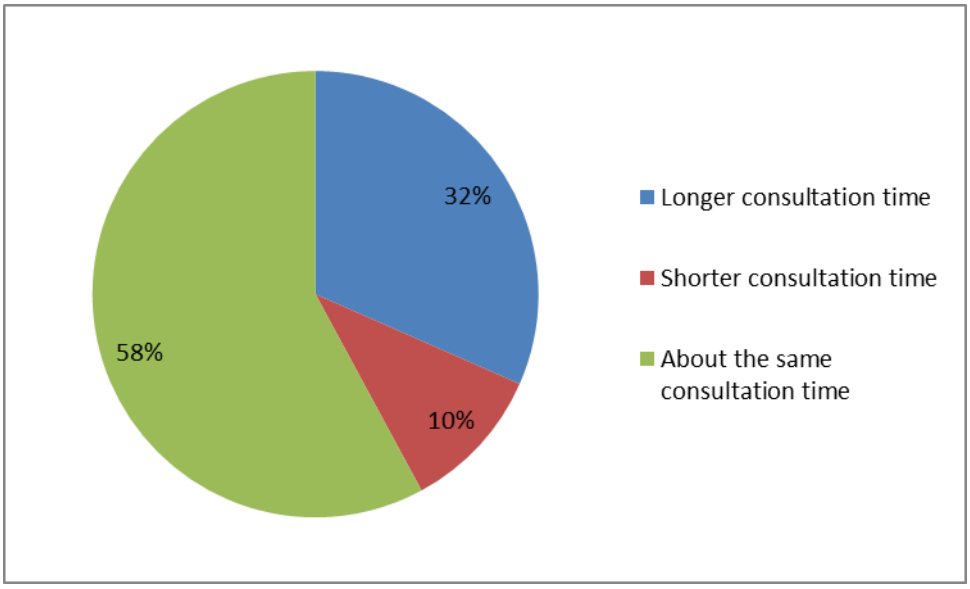
The survey asked respondents about how easy they found it to use the device. Sixteen respondents (84%) said the device was very easy to use, whilst the remaining three responders (16%) said the device was easy to use. As no respondents said that the device was difficult to use, no participants were asked to explain why this was the case.

3.3.4 Impact on Consultation Time

The 19 respondents reported how using MyDiagnostick, rather than the pulse check method, impacted on consultation time. The results are displayed in Figure 3.2, showing 58% (n=11) of respondents judged that the consultation time had not changed.

Of those (n=5) who reported that consultations took longer with MyDiagnostick, a range of 30 seconds to 3 minutes longer was stated. One respondent stated that consultations were now five minutes shorter. On average the six respondents reported that using MyDiagnostick resulted in a 0.375 minute (23 seconds) longer consultation time.

Figure 3.2: Implication of using MyDiagnostick rather than pulse check on consultation time



3.3.5 Benefits of Using MyDiagnostick

The 19 respondents were asked an open-ended question regarding the benefits of using MyDiagnostick. Eighteen of the 19 respondents provided positive responses, with the 19th stating that there was no benefit to using the device. Some respondents provided more than one benefit to using the device. Responses fell into the following themes:

- Seven respondents found the device easy to use and easy to explain to patients, describing the intervention as "fool-proof";
- Five respondents described how they find the device a useful and reliable diagnostic tool. One of these was pleased to no longer have to rely on a pulse check;
- Four respondents appreciated being able to get a download or print out of the two-lead ECG tracing;
- Two respondents found that the device provides reassurance to both the clinician and the patient;
- Two respondents stated that using MyDiagnostick was a quicker alternative to doing a 12-lead ECG, suggesting that in some cases, the use of the device had replaced 12-lead ECGs for the purposes of ruling out an AF diagnosis.

Two respondents provided answers that did not fall into the themes above. Firstly, that the device "*can be used by the patient whilst nurse doing other things or discussing issues with patient*" and secondly, that the device is "*convenient and well made*" and was used opportunistically more than expected.

3.3.6 Disadvantages of Using the Device

The 19 respondents were also asked an open-ended question around the disadvantages of using the device. Responses fell into the following themes:

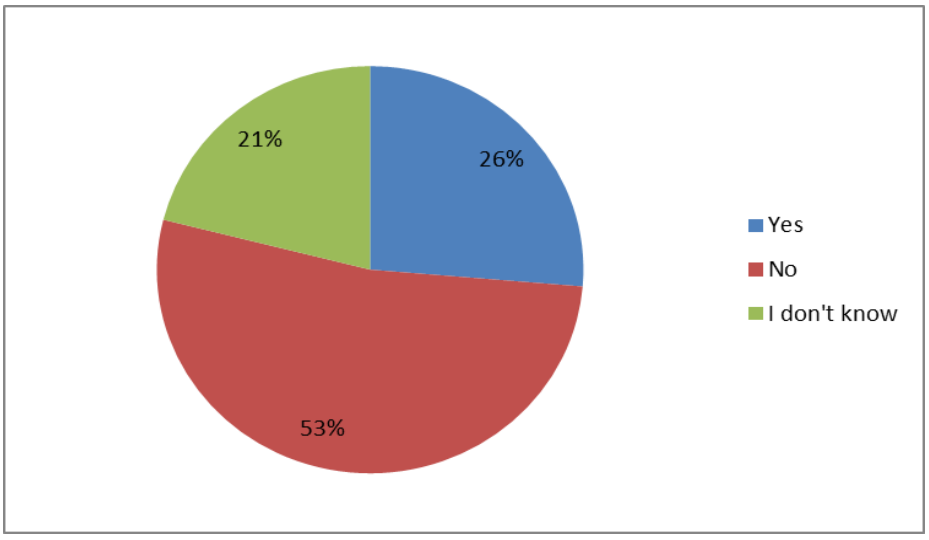
- Seven respondents felt that there were no disadvantages to using the device;
- Three respondents reported that the time taken to use the device is a disadvantage, with one expanding on this to say that they can't do anything else whilst the patient is holding the device;
- Three respondents mentioned the downloading process stating that a computer is required and that extra time is needed to download results (for those who test positive);
- Two respondents expressed concerns about the diagnostic accuracy of the device, with one mentioning false positive and the second mentioning false negative results. The second respondent expanded to say that they hoped in time as results corresponded with 12-lead ECG results, that confidence would improve;
- Two respondents stated that they still had to count the pulse as only the rhythm is provided by the device.

Two other responses did not fit in with the themes above. One stated that the battery life was a disadvantage of the device, but also that they "*only have one*". The second felt that the device was not required and that they can just examine the patient and send for a 12-lead ECG. This respondent was the person who stated that there was no benefit to using the device.

3.3.7 Number of Devices per Practice

Seventeen responses were provided regarding whether the number of devices per practice (one at time of survey) was adequate. The responses are displayed in Figure 3.3.

Figure 3.3: Responses to the question "is the number of MyDiagnostick devices currently available within your practice adequate?"



The ten respondents (53%) who felt the number wasn't adequate reported that a range of between two and ten devices would be ideal. The full results are displayed in Table 3.1.

Table 3.1: Number of devices ideally required by practice

Number	Frequency	Percentage
1	0	0%
2	2	20%
3	4	40%
4	2	0%
5	0	0%
6	0	0%
7	0	0%
8	1	10%
9	0	0%
10	1	10%
Total	**10**	**100%**

3.3.8 Follow-Up Telephone Interviews

Four respondents (18%) agreed to participate in a follow-up interview. Of these, three took part in the interview following an invitation to take part and a follow up email. One respondent no longer had time to be involved. The questions asked within the interview were agreed in advance by the project team. Following the telephone interview, responses were written up and shared with the respondent to ensure that nothing had been misrepresented. The questions asked and answers provided are displayed in Table 3.2.

Table 3.2: Responses in telephone interview

Question number	Question	Interviewee 1	Interviewee 2	Interviewee 3
1	When did you first start using MyDiagnostick in your practice?	Joined the practice in September and started using it then.	Had the device for about 5 months (at start of March).	Started using it immediately after the training session in December 2015.
2	How many times is MyDiagnostick used in the practice in the course of an average week?	Around 10 (or a few more) times per week.	Between about 18 and 30 times per week.	Approximately 10 times or more.
3	How are you using MyDiagnostick e.g. in clinic settings, in annual long term conditions reviews, opportunistically in routine consultations?	Mostly used at health check clinics, rather than opportunistically in routine consultations.	Mostly used at clinics for people with long term conditions and also at hypertension clinics. The device isn't really used opportunistically in consultations as there is only one device available.	Using it in appointments with patients with long term conditions and those with risk factors for AF. This can be opportunistically or as part of an annual review. They will print out the results and scan them into the patient's notes. If the result shows abnormal rhythm, a 12-lead ECG will be ordered and the result of this goes into the doctor's notes. Don't use it every time and will sometimes chose not to use it as it is stored in a separate room and cannot always justify taking the time to get it if AF seems unlikely.
4	Is the MyDiagnostick device passed between rooms/clinicians or does it tend to stay in one consulting room?	It is kept in one place and the clinician goes to collect it when they wish to use it.	It is passed around (about 5 nurses use the device).	It is stored in a utility room and each user will go to collect it and return it after use.
5	Do all clinicians make use of the MyDiagnostick device where possible or are some more inclined to use it than others?	It is used by most of nurses and some of the GP have requested using the device.	It is used by all of the nurses, but not by the GPs.	Two of the three practice nurses use it. The GPs don't use it.

Question number	Question	Interviewee 1	Interviewee 2	Interviewee 3
6	Would you need more MyDiagnostick devices in the practice in order for every clinician to have the opportunity to use it?	It would be useful to have more devices, but are managing with one.	Yes, it could be used more (including opportunistically in consultations) if more devices were available.	Yes. One in each room would be good (i.e. three). The health care assistant could use it as well.
7	What is a rough estimate of the proportion of usage of the MyDiagnostick device between GPs and nurses?	The device is used by nurses the majority of the time and only occasionally used by GPs.	The device is used by nurses only.	Only used by the nurses.
8	Has MyDiagnostick replaced the need for you to use the pulse check for AF, or do you use both methods?	The device is used in health checks to replace the need for a 12-lead ECG. Previously, within the health check an ECG was taken in additional to measures specified in guidelines. The pulse check is still used within the health check in addition to MyDiagnostick.	Both methods are currently being used. The pulse check is included on the template of things to check, so MyDiagnostick is used as an additional check to those included on the template and given to patients to hold as the nurses are speaking to them. In time, if more devices were available, the pulse check may be taken off the template.	Alison will tend to use MyDiagnostick anyway as she uses an automatic BP machine so is not able to hear the pulse when doing a BP.
9	How did the diagnosis of AF work in the practice before the availability of the MyDiagnostick device?	Diagnosed using a pulse check and then ECG.	Diagnosed using a pulse check and then an ECG.	If irregular rhythm was suspected from history or observation, patient was referred for a 12-lead ECG. There are several modes of hearing/observing pulse (wrist pulse, foot doppler, pulse oximetry).
10	How does it work now, with use of the MyDiagnostick device?	Still likely to be the same method of diagnosis, but may be discussed internally the potential for using MyDiagnostick when checking for AF to replace a pulse check.	Diagnosed using a pulse check, MyDiagnostick and then a 12-lead ECG. The use of MyDiagnostick allows for greater confidence/ reassurance in pulse check result.	Same as previously but with MyDiagnostick as an extra tool for diagnosis. Will only print out the MyDiagnostick result if it is abnormal.

Question number	Question	Interviewee 1	Interviewee 2	Interviewee 3
11	Do you think having the MyDiagnostick device will reduce your use of 12-lead ECGs to detect AF? If so, why?	Yes, the device has replaced the need for 12-lead ECGs within health checks (for those having a negative MyDiagnostick result).	No, not really as a 12-lead ECG is still required for diagnosis.	Possibly – if the result indicates AF a 12-lead ECG would definitely be ordered. However, if the MyDiagnostick result is OK then they would trust the result and not order an ECG which might have been ordered previously (when they weren't able to follow up a suspect pulse check with MyDiagnostick).
12	Do you have any other comments?	Will discuss internally potential for using MyDiagnostick rather than a pulse check for opportunistic diagnosis of AF.	An issue with the device is that downloading the data can sometimes take a while. This is only done when there is a positive result, so there may be a number of other test results that need deleting beforehand.	Two of the nurses use it. They like it and think it is a valuable tool.

The interviews conducted allowed further information on the use of MyDiagnostick to be elicited. In all three practices the device had been used since training on it was provided and it was used more than 10 times per week. The device was typically used in clinics for people with long term conditions. One of the three respondents also stated that the device was used opportunistically in routine appointments for people with long term conditions. The device tended to be used by most of the nurses within each practice. Hence, it was either stored in a central location (and taken when needed), or passed between the nurses. All three interviewees felt that it would be useful to have more devices, so there could be one in each room, avoiding the need for the device to be passed around. The device was rarely used by GPs, with only one practice reporting occasional use by a GP.

The use of MyDiagnostick within practices appears to be different to that anticipated by the project team. NHS Wakefield CCG expected that the device would be used to replace the need for a pulse check in people with long term conditions at raised risk of AF. Following the use of MyDiagnostick it was expected that a positive result would be confirmed with a 12-lead ECG. Input from the interviewees suggests that the device was actually used in addition to the pulse check. Those patients in whom an irregular pulse was suspected following a pulse check would then use MyDiagnostick to confirm or refute this finding. Patients who have a positive result from MyDiagnostick would then undergo a 12-lead ECG to confirm a diagnosis of AF. One interviewee reported that the use of MyDiagnostick had overridden the need for a 12-lead ECG for the purposes of ruling out AF. A second interviewee reported that the use of a 12-lead ECG was now potentially lower than previously given that a negative MyDiagnostick result in someone who previously would have had a positive pulse check result would no longer require a follow up 12-lead ECG.

3.3.9 Feedback from Practices not Utilising MyDiagnostick

All practices who had not taken up the opportunity to receive a free device were contacted by Gillian Richardson to elicit views on why they were not interested in the device. No responses were received.

3.4 DISCUSSION

The survey results suggest that the vast majority of respondents were positive towards MyDiagnostick, finding that it was easy to use and it provided a greater degree of confidence in the result of a pulse check. A small number of respondents were less positive about the device and did not feel there was a need for it. There was some variation in how the device was being implemented in practice in that some practices used both MyDiagnostick and the pulse check, whilst others had replaced the use of 12-lead ECGs with the device when wishing to rule out AF. It was also used mainly by nurses and not GPs. This is probably explained by the fact that the appointments when MyDiagnostick was most commonly used are the clinical contacts usually delivered by the nursing team e.g. long term conditions annual reviews. In these circumstances, the nurses were often using MyDiagnostick as an additional tool to check for AF and not as a replacement for a pulse check.

It was found not to add an unacceptable length of time to the consultations for most patients. For those with an abnormal cardiac rhythm, the additional time taken to upload the two-lead ECG results from MyDiagnostick appears to be outweighed by the benefit of having the test result in the clinical record

Whilst the results of the survey suggest that general practice in Wakefield welcomes the introduction of MyDiagnostick, the survey is subject to a number of limitations. Firstly, due to the way in which the survey was distributed, the number of people receiving the survey was unknown. Therefore, the response rate could not be determined. Consequently, it is impossible to know whether the views elicited throughout the survey were representative of a majority or minority of people using the device. Furthermore, the survey was sent to each practice, which was asked to distribute the survey to their staff, meaning that some practices may have been represented by multiple respondents whilst others not represented at all. Given that the survey was asking individuals for their views this should not have biased the findings substantially.

Secondly, no information was received from practices choosing not to use MyDiagnostick. A total of 16 practices had not used a MyDiagnostick device. Had these practices responded to the information request sent to them, more negative responses describing a lack of need for MyDiagnostick may have been received.

Thirdly, the respondents agreeing to be involved in a follow-up interview were self-selected. All three individuals were very positive about the device in both their survey response and their follow-up interview. Eliciting further information from people who were positive about the device may have biased the picture being presented around MyDiagnostick. Had follow-up interviews been conducted with those not using the device, a greater proportion of negative information would likely have been received. However, given these people were not using the device it may have been difficult to gain further insight from them.

Section 4: Survey of Patients

4.1 INTRODUCTION

A survey of patients on whom the MyDiagnostick device was used was conducted to evaluate their views of the usability of the device.

4.2 METHODS

A paper based survey was developed by the authors and revised by the wider project team. This is shown in full in Figure 4.1. The survey was developed to determine the view of patients on the ease of use of the device. The survey was designed to be quick and easy to answer in order to improve the response rate. To this end, all questions had multiple choice answers.

Figure 4.1: Survey of patients

1. How easy did you find the MyDiagnostick device to use?
- **a.** Very easy
- **b.** Easy
- **c.** Neither easy nor difficult
- **d.** Difficult
- **e.** Very difficult

2. How good was your understanding of why the MyDiagnostick device was being used?
- **a.** Very good
- **b.** Good
- **c.** Neither good nor poor
- **d.** Poor
- **e.** Very poor

3. What was the result of your test?
- **a.** Positive (a × was displayed)
- **b.** Negative (a ✓ was displayed)

4. Was the result of the test explained to you?
- **a.** Yes
- **b.** No
- **c.** I don't know

5. Would you be happy to use this device again if asked?
- **a.** Yes
- **b.** No
- **c.** I don't know

Those practices with a large number of recorded uses of MyDiagnostick were asked by Gillian Richardson if they would be willing to hand out the surveys to patients. The surveys were given out to patients by three practices: Ferrybridge Medical Centre; Tieve Tara Medical Centre and Orchard Croft Medical Centre. The practices were each sent 20 surveys, to give to patients during a consultation when the MyDiagnostick device was used. Patients were provided with a stamped addressed envelope and asked to post their response back to the authors.

The responses to the survey were analysed in Microsoft Excel and summary statistics were generated.

4.3 RESULTS

The patient surveys were distributed to practices in February 2016. A total of 23 surveys were returned to YHEC by the end of March 2016. It is not possible to know how many of the 60 surveys were handed out to patients. If all 60 were handed out this is a response rate of 38.3%. It is likely, however, that not all of the surveys were utilised, and the response rate from patients was in fact higher than this. One patient returned the survey without answering the questions and a comment "*Sorry – don't remember using this device*". All other 22 respondents completed the multiple choice questions.

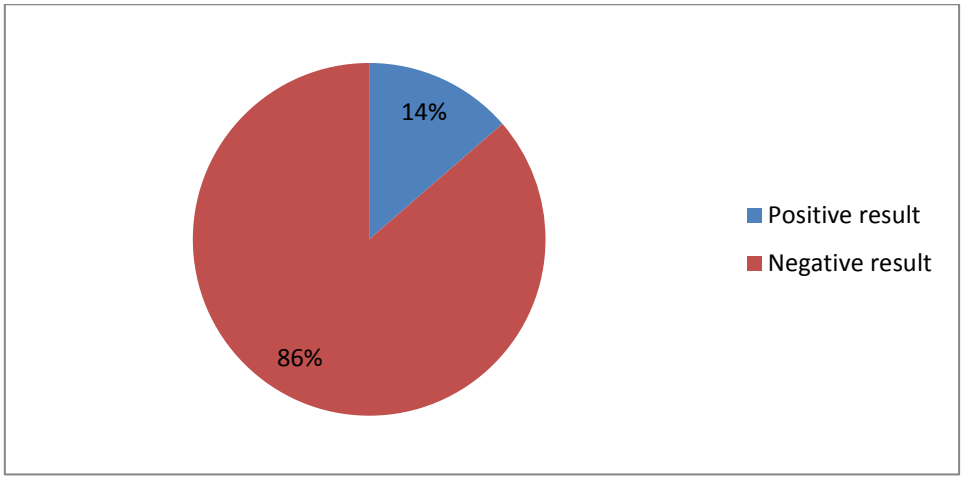
4.3.1 Ease of Use of the Device

The 22 respondents found the device very easy to use (21) or neither easy nor difficult (1). When asked how good their understanding of why the device was being used, all responded either very good (16) or good (6). All respondents would be happy to use the device again.

4.3.2 Test Results

Of the 22 responding, three reported that they had a positive result (a × was displayed) and 19 had a negative result (a ✓ was displayed), as shown in Figure 4.2. In all cases the patients reported that the result was explained to them. One respondent said that they were not shown the result but "*the nurse said it was fine*".

Figure 4.2: Proportion of patients with a positive and negative result



4.4 DISCUSSION

The results of the survey show that patients found the MyDiagnostick device easy to use and would be happy to use it again. All understood what the device was being used for and were aware of the test result. The respondents were not invited to make any additional comments over and above their responses to the multiple-choice questions. Of those that did add remarks however, there were no negative comments made.

The survey is subject to some limitations. Firstly, as it is not possible to know how many patients were asked to complete the survey, it is not known how many declined to do so. Consequently, it is possible that the responses are biased towards those who were more positive about the device. Secondly, the patients were drawn from only three of the 24 practices using MyDiagnostick. It is possible that they may not be representative of the wider patient cohort using the device.

Section 5: Cost Analysis

5.1 INTRODUCTION

A *de novo* cost analysis was developed in Microsoft *Excel* to evaluate the cost implications and potential savings of using MyDiagnostick compared with the pulse check method, as a first line method for detecting AF. The model was designed to be user friendly and suitable for update once further data become available.

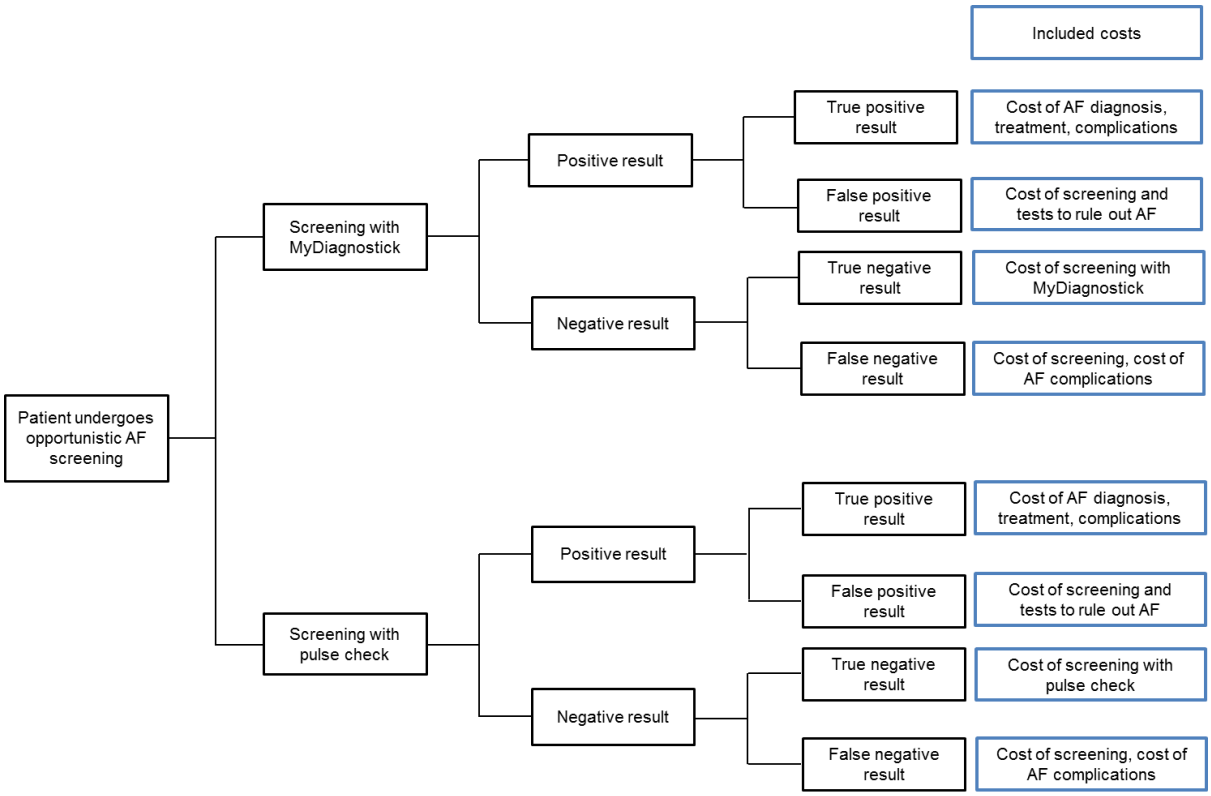
5.2 METHODS

The methods used to construct and populate the model are described below. The model's outputs and sensitivity analyses are described in Section 5.2.3.

5.2.1 Model Structure

The model structure took the form of a decision tree, whereby patients underwent screening for AF with either MyDiagnostick or the pulse check method. The head-to-head comparison of MyDiagnostick and the pulse check method is not fully in line with the feedback received from interviewees on how the device is being used within NHS Wakefield, reported in Section 3.3.8. However, this comparison was made based on the available clinical data which compared MyDiagnostick to the pulse check method. Those patients who had a positive result were split into true and false positives, based upon the accuracy of each test. Likewise, those patients who had a negative result were split into true and false negatives. As such, patients fell into one of four categories: true positive, false positive, true negative and false negative. Costs were assigned to each of the four outcomes (these are explained in more detail in Section 5.2.2). The full model structure is displayed in Figure 5.1. The model took a 10 year time horizon and an NHS and social care perspective.

Figure 5.1: Structure of the *de novo* economic model



5.2.2 Model Inputs

5.2.2.1 Set up

The number of patients using MyDiagnostick each year was used to determine the cost per patient of the device. Information on this was taken from staff interviews, whereby two of the three interviewees informed that that device is used around 10 times per week. The third interviewee reported that their device was used more often (18-30 times per week). Within the model, it was assumed that each device was used 10 times per week, or 520 times per year. Data taken from clinical coding reported in Section 2 suggests that the frequency of use is actually lower than 10 times per week in most practices, but that there was wide variability. However, as described in the previous section, this frequency depends upon complete and accurate coding, hence the true frequency of usage may be higher. Sensitivity analysis was conducted around the number of patients using the device each year using a minimum of 10 uses per month and maximum of 30 uses per week.

MyDiagnostick was reported by Technomed Ltd to have a lifespan of 8-10 years with heavy use. In the base case of the model it was conservatively assumed that the device lasted 8 years. This was varied in sensitivity analysis between 8 and 10 years.

In the base case, all uses of the device were deemed to occur during an appointment with a nurse. Similarly, all pulse checks were deemed to be carried out by nurses. This was based upon feedback from nurses during staff interviews who specified that they alone use the device. Sensitivity analysis was carried out whereby between 0 and 100% of appointments were nurse-led and the residual were GP-led.

5.2.2.2 Effectiveness

The sensitivity and specificity of MyDiagnostick and the pulse check were each combined with the prevalence of AF to determine the proportion of patients who fell into each of the following categories:

- True negative: tested negative for AF and not in AF;
- False negative: test negative for AF, but in AF;
- True positive: tested positive for AF and in AF;
- False positive: tested positive for AF, but not in AF.

The prevalence of AF describes the proportion of patients who truly have AF and thus combines those who test false negative and those who test true positive. The true positive patients can be determined by multiplying the prevalence by the sensitivity of the test; hence, a perfect test would be 100% sensitive and therefore identify all positive patients accurately.

The remaining proportion (100% minus the prevalence) comprises those who do not have AF, i.e. the false positives and the true negatives. The true negatives can be determined by multiplying the remaining proportion by the specificity of the test. Thus a test that is 100% specific would identify all negatives as true negatives.

Data for the sensitivity and specificity of both MyDiagnostick and the pulse check were taken from the published literature. The effectiveness of MyDiagnostick was derived from two sources:

- Tieleman *et al*. (2014):
  - Sensitivity = 100% (95% CI: 93%-100%);
  - Specificity = 96% (95% CI: 91.3%-98.1%) [3].
- Vaes *et al*. (2014):
  - Sensitivity = 94% (95% CI: 87%-98%);
  - Specificity = 93% (95% CI: 85%-97%) [2].

In the base case, values from Vaes *et al.* are used, given that these values are lower and will therefore generate more conservative results [2]. As both studies were single arm studies, effectiveness data describing the pulse check were taken from an alternative source published by Hobbs *et al*. (2005) who reported a sensitivity of 87.2% (95% CI: 82.1%-91.1%) and specificity of 81.3% (95% CI: 79.7%-82.8%) [4].

A prevalence of AF of 4.4% was used in the base case based upon data reported by Willits *et al.* (2012) [5]. This prevalence related to a population that would be screened for AF and therefore at higher risk than the general population. Alternative prevalence data were provided by Hobbs *et al*. who reported a prevalence of 7.7% in those at risk [4].

The sensitivity and specificity of both tests, as well as the prevalence of AF were varied during sensitivity analysis.

Table 5.1: Summary of effectiveness calculations

	Proportion in AF	**Proportion not in AF**	**Total**
Test positive	My Dx = 4.1 Pulse check = 3.8 (Prevalence of AF * sensitivity of test)	My Dx = 6.7 Pulse check = 17.9 ((1-Prevalence of AF)* (1-specificity of test))	**My Dx = 10.8** **Pulse check = 21.7** **(All those who test positive)**
Test negative	My Dx = 0.3 Pulse check = 0.6 ((1-Prevalence of AF)* sensitivity of test)	My Dx = 88.9 Pulse check = 77.7 (Prevalence of AF * specificity of test)	**My Dx = 89.2** **Pulse check = 78.3** **(All those who test negative)**
Total	**MyDx = 4.4** **Pulse check = 4.4** **(Prevalence of AF)**	**MyDx = 95.6** **Pulse check = 95.6** **(1 - Prevalence of AF)**	**My Dx = 100** **Pulse check = 100** **(All patients)**

5.2.2.3 Cost of screening for AF

The cost of screening for AF with either a pulse check or MyDiagnostick was determined. All patients entering each arm of the model incurred this cost. For the pulse check this comprised the cost of the appointment only (Table 5.2), and for MyDiagnostick the cost of an appointment, the cost of the device and the cost of training staff to use the device (Table 5.3).

Table 5.2: Cost of screening with pulse check

Component	**Value**	**Source**
Appointment time with GP (mins)	11.7	[6]
Unit cost per min (GP)	£3.80	[6]
Appointment time with nurse (mins)	15.5	[6]
Unit cost per min (nurse)	£0.783	[6]
Total appointment cost	**£12.14**	**Weighted by the proportion of patients seeing a GP/nurse (100% nurse in base case)**

Table 5.3: Cost of screening with MyDiagnostick

Component	Value	Source
Appointment time with GP (mins)	12.075	[6] plus additional time reported in staff survey (0.375 minutes)
Unit cost per min (GP)	£3.80	[6]
Appointment time with nurse (mins)	15.88	[6] plus additional time reported in staff survey (0.375 minutes)
Unit cost per min (nurse)	£0.783	[6]
Total appointment cost	£12.44	Weighted by the proportion of patients seeing a GP/nurse (100% nurse in base case)
Cost of device per patients	£0.10	Cost of device (£416) divided by the number of patients using the device each year over its 8 year life span
Cost of training on device per patient	£0.31	Training cost (£160 = 1 hour for 3 nurses and 1 advance nurse [6]) divided by the number of patients using the device in a year. The training cost is conservatively applied annually due to staff turnover and additional staff being trained
Grand total	**£12.84**	**Sum of appointment cost, device cost and training cost**

5.2.2.4 Further costs associated with AF

The further costs associated with AF include costs of treatment, costs of complications that may arise from treatment (e.g. bleeding) and costs of events that may arise as a result of either treated or untreated AF (e.g. stroke). Costs are applied within the model dependent on the patient's outcome from the model: true positive result, false positive result, true negative result and false negative result. The costs applied have been taken from a previous analysis conducted on behalf of NICE for a device used for AF screening [5]. The costs used within the current model are now described in turn by patient outcome.

True positive result

Patients who receive a true positive result with either MyDiagnostick or the pulse check method will undergo a 12-lead ECG test to confirm diagnosis and then receive treatment for their AF. This treatment comprises aspirin, warfarin or new oral anticoagulants (NOACs). Patients on treatment are at risk of bleeds and will also be at risk of stroke. The risk of stroke is lower for patients on treatment than those not being treated. The cost applied within the model for patients with a true positive result incorporates all AF related costs over a 10-year time horizon. This cost was derived from the work commissioned by NICE in which a longer term analysis was used to model the potential outcomes for patients with AF [5]. These long terms costs were discounted appropriately. In this study, the 10 year management costs of true positives were reported by CHADS2 score as outputs from the modelling exercise. These are displayed in Table 5.4.

Table 5.4: 10-year management costs for true positives

	Proportion of patients	10 year management cost
CHADS2 0	6.9%	£2,230
CHADS2 1	26.8%	£4,092
CHADS2 2	30.3%	£5,484
CHADS2 3	19.5%	£6,233
CHADS2 4	12.7%	£7,091
CHADS2 5	3.8%	£8,310
Weighted average		**£5,344**

The weighted average cost of £5,344 reported in Table 5.4 was inflated from 2011/12 prices to 2014/15 prices using the Hospital and Community Health (HCHS) Index [6]. A cost of £5,544 was derived. In addition to the 10-year management cost, these patients also require a 12-lead ECG to confirm the diagnosis of their AF, costing £50 [7]. As such, the total cost for true positives over 10-years was estimated as £5,594.

This cost assumes that no patients are treated with NOACs. This assumption was varied during sensitivity analysis. It is expected that patients being treated with NOACs will be more costly than the value used in the base case of the model, as previous economic evaluations comparing warfarin to the NOACs show the NOACs to be more costly when taking into consideration drug costs, bleed costs and stroke costs [8, 9]. However, the project team were aware that 93% of AF patients in Wakefield are treated with warfarin, meaning the impact of this assumption on the applicability of the model's results will be minimal.

False positive result

Patients who are deemed to be false positive, i.e. receive a positive result from MyDiagnostick or the pulse check, but do not have AF, incur the cost of a 12-lead ECG test only. This test will be conducted following the positive screening test, but will confirm that the patient does not have AF. Within the model, a cost of £50 for the 12-lead ECG test has been applied based upon a cost reported by South Devon Healthcare NHS Foundation Trust. This cost is based upon Payment by Results (PbR) tariff, plus premium to cover administration costs. A more relevant cost could not be identified by NHS Wakefield. However, the cost used in the base case has been varied widely in sensitivity analysis.

True negative result

Patients who test negative for AF during their screening test and do not actually have AF, incur no further costs within the model.

False negative result

Patients who test negative with MyDiagnostick or the pulse check, but do have AF, are deemed to be false negatives. These patients will not be sent for a confirmatory 12-lead ECG, nor will they be treated for their AF, hence no cost is attributed to the patients for either of these resources. However, these patients will be at risk of stroke, a risk that is higher than if they were receiving treatment for their AF. The cost applied within the model for patients with a false positive result incorporates the risk and cost of stroke over a 10-year time horizon. Again, this cost was derived from the longer term analysis commissioned by NICE [5]. In this study, the 10 year management costs of false negatives were reported by CHADS2 score as outputs from the modelling exercise. These are displayed in Table 5.5.

Table 5.5: 10-year management costs for false negatives

	Proportion of patients	10 year management cost
CHADS2 0	6.9%	£2,570
CHADS2 1	26.8%	£3,686
CHADS2 2	30.3%	£4,975
CHADS2 3	19.5%	£6,935
CHADS2 4	12.7%	£9,102
CHADS2 5	3.8%	£11,830
Weighted average		**£5,630**

The weighted average cost of £5,630 reported in Table 5.5 was inflated from 2011/12 prices to 2014/15 prices using the hospital and community health (HCHS) index [6]. A cost of £5,842 was derived.

5.2.3 Model Outputs and Sensitivity Analysis

The model presents both the cohort and per patient costs for those screened using MyDiagnostick and those screened using the pulse check method. The cohort of patients represents the number of patients using the device per year. The incremental cost has been calculated, that is the difference in costs using MyDiagnostick and the pulse check. The results have been presented graphically.

Sensitivity analysis has been conducted around model inputs that were uncertain. This univariate deterministic analysis has been presented in a tornado diagram. Within this analysis each included input parameter was varied whilst all other inputs remained at their base case value. The impact on the results of the model for each change is presented.

Scenario analysis has been conducted using the alternative MyDiagnostick effectiveness data. Within Section 5.3.3 the results and tornado diagram are presented using these data. Under this scenario, the sensitivity of MyDiagnostick was varied between 93% and 100% and the specificity between 91.3% and 98.1% [3].

5.2.4 Summary of Model Inputs

Table 5.6 provides a summary of all of the model's inputs and the values considered during sensitivity analysis.

Table 5.6: Summary of model inputs

Input name	Input value	Reference	Range considered
Number of patients using MyDiagnostick each year	520	Staff interviews based on 10 uses per week	120 – 1,560
Lifespan of device	8 years	Technomed Ltd	8 - 10 years
Proportion of appointments carried out by a nurse (rather than a GP)	100%	Staff interviews	0 - 100%
Prevalence of AF	4.4%	[5]	0 - 10%
Sensitivity of MyDiagnostick	94.0%	[2]	87 - 98%
Specificity of MyDiagnostick	93.0%	[2]	85 - 97%
Sensitivity of pulse check	87.2%	[4]	82.1 - 91.1%
Specificity of pulse check	81.3%	[4]	79.7 - 82.8%
Cost of screening with pulse check	£12.14	See Table 5.2	£10 - £20
Cost of screening with MyDiagnostick	£12.85	See Table 5.3	£10 - £20
Cost of true positive outcome	£5,594	[5, 7]	£2,000 - £8,500
Cost of false positive outcome	£50	[7]	£25 - £200
Cost of true negative outcome	£0	Assumption	Fixed as no cost will be incurred for people who do not have AF
Cost of false negative outcome	£5,842	[5]	£2,500 - £12,000

5.3 RESULTS

5.3.1 Base Case Results

Based on the input values noted in Table 5.6 above, the costs per patient screened with MyDiagnostick and the pulse check methods were £263.00 and £268.63 respectively. These values take into account all of the associated costs for each method, not just the cost of performing the tests themselves. The base case results showed that using MyDiagnostick saves £5.63 per patient over a 10 year time horizon. The breakdown of these costs can be seen in Table 5.7.

Table 5.7: Costs per patient screened for MyDiagnostick and pulse check method

Costs per patient	MyDiagnostick	Pulse check	Cost difference
Device cost	£0.41	£0.00	£0.41
Appointment cost	£12.44	£12.14	£0.29
Further AF related costs.	£250.16	£256.49	-£6.33
Total	**£263.00**	**£268.63**	**-£5.63**

It can be seen that the upfront cost resulting from the use of the device was £0.70 per patient (£0.41 plus £0.29), with £6.33 per patient being saved in the longer term. However, the point in time at which 'break-even' occurs is unknown.

If each test were to be used 10 times per week, the costs for the cohort of 520 patients per year would be £136,759 and £139,687 for MyDiagnostick and pulse check respectively. The breakdown of these costs can be seen in Table 5.8.

Table 5.8: Costs per cohort screened for MyDiagnostick and pulse check method

Costs for cohort	MyDiagnostick	Pulse check	Cost difference
Device cost	£212	£0	£212
Appointment cost	£6,466	£6,314	£153
Further AF related costs	£130,081	£133,374	-£3,293
Total	**£136,759**	**£139,687**	**-£2,928**

5.3.2 Sensitivity Analysis

Sensitivity analysis has been conducted around the model inputs that were uncertain, using the ranges described in Table 5.6. These are illustrated in the tornado diagram in Figure 5.2, which presents the results of varying the different parameters between two plausible values, referred to as a 'high value' and a 'low value'. The green and pink bar combined represents the range of results that were generated and allows the reviewer to assess which of the model input parameters have the greatest influence on the model's results.

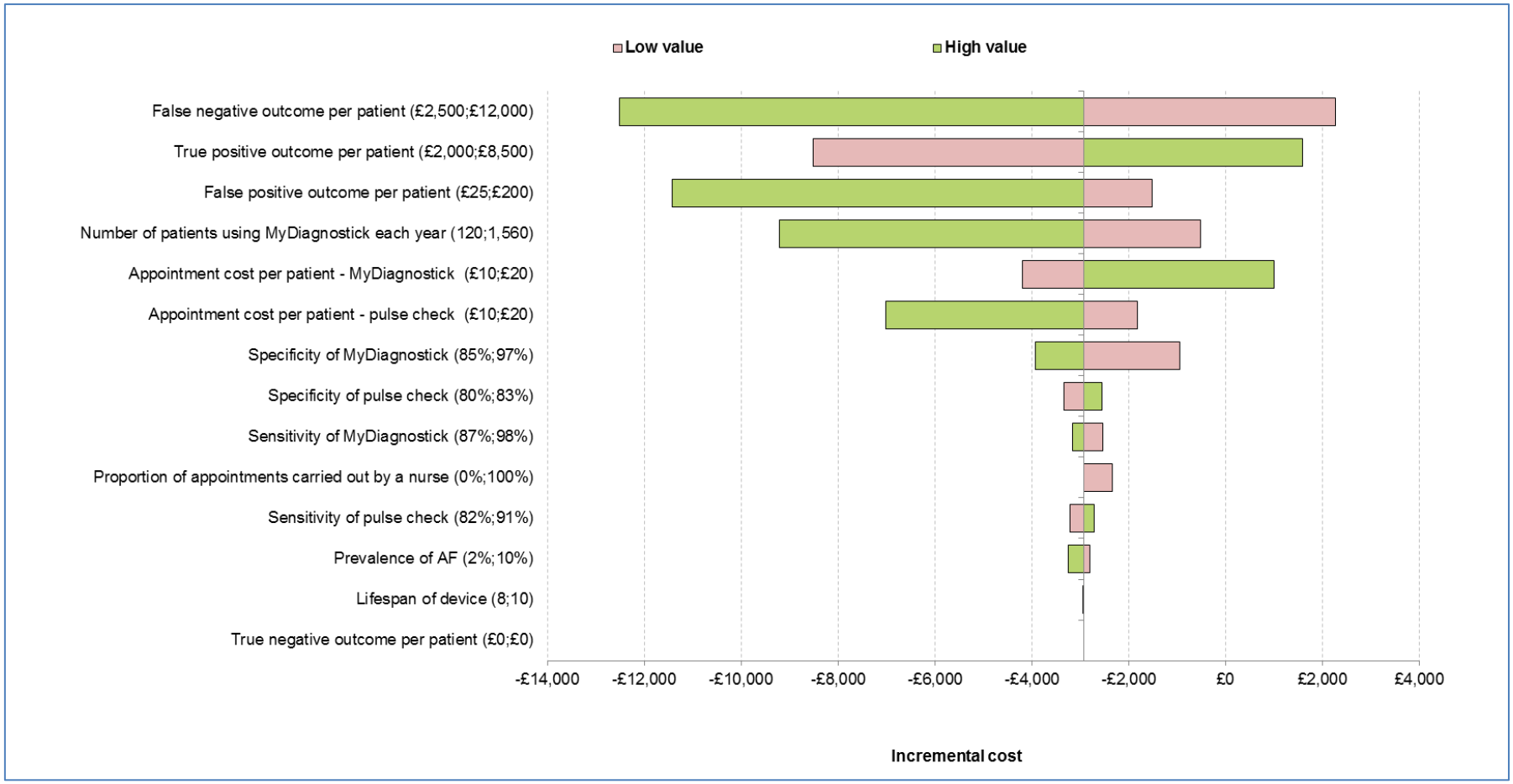
It can be seen from the tornado diagram that the key drivers of the cost analysis were:

- The cost of a false negative result;
- The cost of a true positive result.

The greatest costs in the analysis were those associated with further AF related costs i.e. the cost of treating diagnosed AF and/or the health consequences of unmanaged AF, such as a stroke. As the MyDiagnostick device is more sensitive than the pulse check method, it should reduce the number of false negative results. Therefore, the lower the cost of a false negative result to the health and social care system, the more unlikely the cost of the device would be offset by the savings.

Similarly, as MyDiagnostick is more likely to detect true cases of AF, when the cost of treating true cases of AF rises, the savings from using the device would not offset the cost of subsequent treatment.

Figure 5.2: Tornado diagram showing the results of sensitivity analyses around the bases case (Vaes *et al*, 2014)



A further parameter worthy of note is the 'Appointment cost per patient for MyDiagnostick'. The staff survey has shown that using the device in a consultation takes on average 23 seconds longer than using the pulse check method. When the appointment becomes much longer and more costly, the use of MyDiagnostick becomes cost incurring. From the model it can be seen that the appointment would need to be an average of 7.5 minutes longer than when using the pulse check method for the incremental cost to be zero (i.e. no savings made). Based on the information gained in this analysis, this is extremely unlikely to occur.

The varying of the proportion of appointments carried out by a nurse, rather than a GP, has very little influence on the incremental cost. This is because the type of clinician is likely to affect both methods equally so it is not a driver of the incremental cost.

5.3.3 Scenario Analysis Using Alternative MyDiagnostick Effectiveness Data

The cost analysis model has been based on the more conservative sensitivity and specificity data from the literature (Vaes *et al.*, 2014). For the purposes of comparison, the model includes the ability to use effectiveness data from another source (Tieleman *et al.,* 2014).

Using these values, the costs per patient screened with MyDiagnostick and the pulse check methods were £260.91 and £268.63 respectively. The base case results showed that MyDiagnostick saves £7.72 per patient over a 10 year time horizon. The breakdown of these costs can be seen in Table 5.9.

Table 5.9: Costs per patient screened for MyDiagnostick and pulse check method

Costs per patient	MyDiagnostick	Pulse check	Cost difference
Device cost	£0.41	£0.00	£0.41
Appointment cost	£12.44	£12.14	£0.29
Further AF related costs	£248.07	£256.49	-£8.42
Total	**£260.91**	**£268.63**	**-£7.72**

The costs for the cohort of 520 patients per year were £135,674 and £139,687 for MyDiagnostick and pulse check respectively, a difference of £4,013 over a 10 year time horizon. The breakdown of these costs can be seen in Table 5.10.

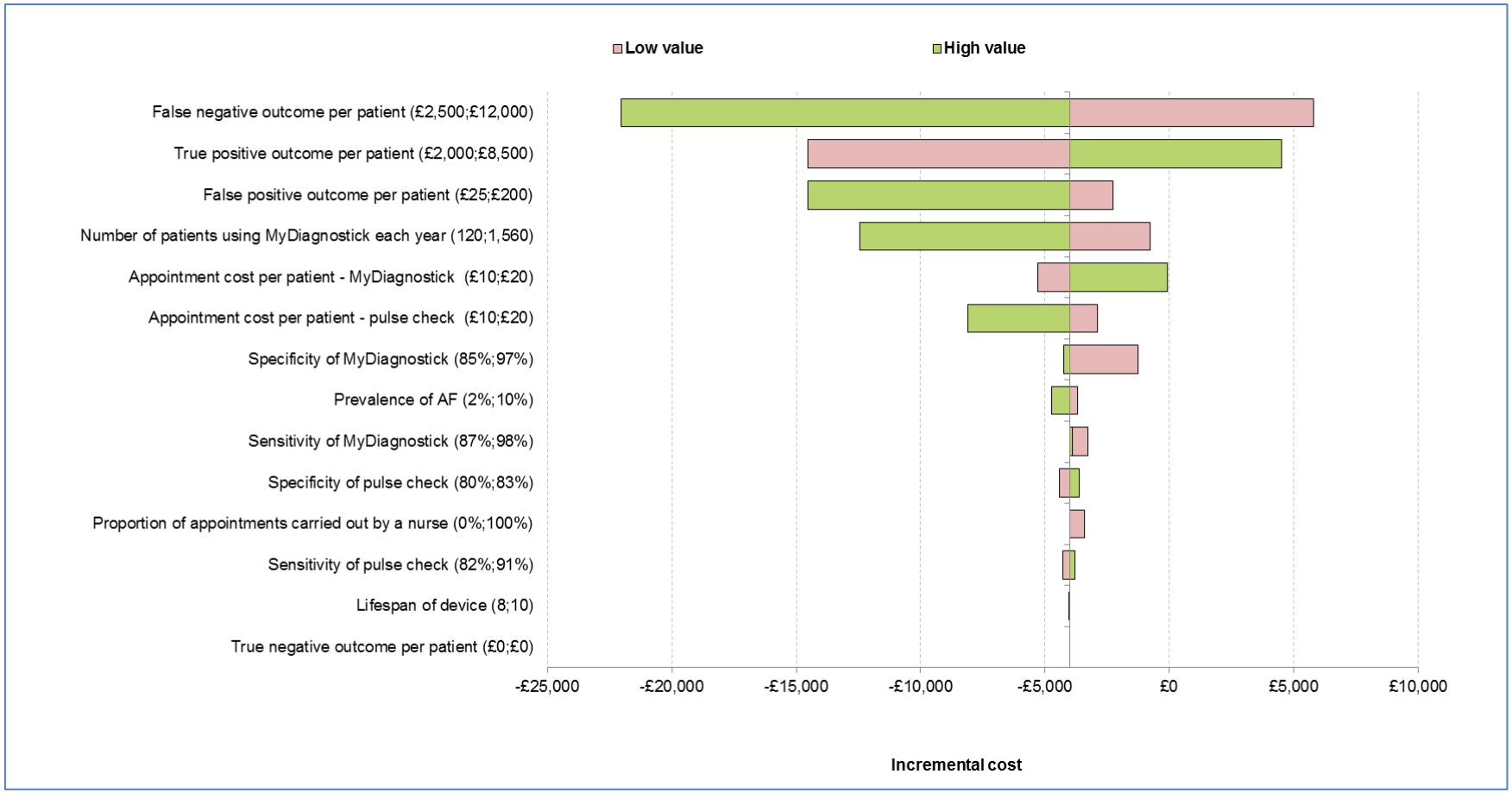
Table 5.10: Costs per cohort screened for MyDiagnostick and pulse check method

Costs for cohort	MyDiagnostick	Pulse check	Cost difference
Device cost	£212	£0	£212
Appointment cost	£6,466	£6,314	£153
Further AF related costs	£128,996	£133,374	-£4,378
Total	**£135,674**	**£139,687**	**-£4,013**

The results using the alternative effectiveness data show MyDiagnostick to be more cost effective than when using the less conservative figures from Tieleman *et al,* 2014. This is due to the increase in both sensitivity and specificity of the MyDiagnostick device.

The sensitivity analysis shows that the parameters having most influence on the incremental cost are the same as for the Vaes *et al* data i.e. the cost of a false negative result and the cost of a true positive result. While the influences on the incremental cost are the same as for the more conservative data, the bars in the Tornado diagram are wider, indicating that these parameters are less likely to affect the result.

Figure 5.3: Tornado diagram showing the results of sensitivity analyses around the bases case (Tieleman *et al*, 2014)



5.4 DISCUSSION

The cost analysis shows that the use of MyDiagnostick in general practice has a small cost saving from an NHS and social care perspective. This is largely due to the differences in further AF related complication costs, some of which may be avoided by earlier diagnosis of AF by using a more sensitive and specific test.

The sensitivity analysis shows that the extent of the savings is most influenced by changes in the cost of a false negative and a true positive result. This has implications for the population in which the device has most potential to achieve savings. In a population with a high proportion of patients with a low CHADS2 score the associated costs of undetected AF are likely to be lower, meaning that the potential for achieving savings is lower. However, when AF related complications are prevented in patients with a higher CHADS2 score, the potential for savings is higher.

When the cost of a true positive result is higher (due to expensive treatment) then the cost of finding more of these cases using MyDiagnostick is also higher. The cost of a true positive result is more likely to be higher in a population with a greater proportion of patients with higher CHADS2 scores. In reality, the patient population with undetected AF (who may be subject to AF detection methods such as MyDiagnostick) is likely to reflect a case mix of CHADS2 scores so the effect of changes in false negatives and true positives will be true across populations. Furthermore, in a population with relatively low CHADS2 scores the cost savings with MyDiagnostick decrease due to lower potential costs of undetected AF, but increase due to a lower cost of true positive patients meaning that the cost implications are likely to cancel out. The same is likely to be true in a higher risk population.

The results of the cost modelling are quite robust, as the changes required to take the parameter values over £0 are quite unlikely to occur. There are some limitations to the model, however. The modelled data was based on information from the literature, and was not specific to actual diagnoses and prevalence of AF in Wakefield district. It also doesn't reflect actual practice in all practices in Wakefield, whereby some clinicians are using both tests for AF, rather than replacing a pulse check with MyDiagnostick. The effectiveness data for the MyDiagnostick and pulse check methods were taken from three different, single arm studies, as no study directly comparing the two methods was found. It is possible therefore, that there may be confounding factors affecting the results for sensitivity and specificity that are not apparent. The long term costs related to AF were taken from the outputs of a previous modelling analysis of stroke related costs, conducted for NICE. While this is considered to be a reliable source, the detailed calculations used were not available and so this is not completely transparent.

Section 6: Conclusions

6.1 CONCLUSIONS

The purpose of introducing MyDiagnostick in NHS Wakefield CCG was to test out the use of a new device for AF case finding in primary care settings. The aim of this evaluation was to assess whether the device was acceptable to both clinicians and patients. It also aimed to understand the costs and potential savings of using MyDiagnostick, compared with the pulse check, as a method for detecting AF.

The evaluation used a combination of qualitative and quantitative methods to collect data and inform the *de novo* cost analysis. Although it was not possible to conduct a prospective evaluation using local data, the information gained by the other methods has provided a comprehensive picture of the first stages of the implementation of MyDiagnostick in Wakefield. Furthermore, the staff survey and interviews provided valuable information to inform the cost analysis, with some parameters in the model (e.g. the time taken and number of patients using the device) being directly influenced by their experience of implementing MyDiagnostick.

The use of MyDiagnostick by general practices in Wakefield increased over time from July 2015. Up to 25th April 2016, 60% of the practices had used the device, with a mean of 53 recorded uses per practice. 30% of practices had used the device 30 times or more. Of those patients with a positive reading recorded (n=42), 3.33% tested positive for AF with MyDiagnostick. Of these, 38.10% have since had AF confirmed (n=16). Unfortunately, due to the nature of the codes being used, it has not been possible to compare this conversion from positive test to AF diagnosis with the estimated rate of diagnoses in patients undergoing a pulse check. The key limitation of this analysis is the reliance on the correct and complete coding of the use of the MyDiagnostick device. The extent of missing coding is unknown and the data should, therefore, be interpreted with caution.

The pattern of usage within individual practices appears to vary substantially, with some practices having fairly steady usage and some tailing off after an initial flurry of activity. It is therefore too early to tell if the use of MyDiagnostick will continue in the practices that currently appear to be enthusiastic about it. Furthermore, based on the uptake of MyDiagnostick over the nine month period studied, it appears likely that there will be some practices that choose not to use it at all. Local commissioners in Wakefield will continue to monitor the usage of the device over the coming months to inform the decision about whether to continue to promote it to practices.

The staff opinion survey collected information from 22 survey respondents and three interviewees. The respondents were mostly positive about MyDiagnostick, finding that it is easy and quick to use. For those patients with an abnormal cardiac rhythm it was felt to be beneficial having the two-lead ECG results from MyDiagnostick in the clinical record.

The findings show that MyDiagnostick was being used mostly in appointments with the practice nurses and not the GPs. In general, the nurses were using MyDiagnostick as an additional tool to check for AF and not as a replacement for a pulse check. It was suggested that in some cases this could prevent the need for a follow-up 12-lead ECG for the purposes of ruling out an AF diagnosis, as staff have confidence in the negative result. In some instances it is apparent that the device was being used in patients who were already known to have AF. The use of MyDiagnostick to 'rule out' AF is somewhat different to what was anticipated by the project implementation team. A majority of respondents would welcome having more than one device in the practice, with three being the most commonly requested number.

A small number of staff respondents were less positive about the device and did not feel there was a need for it. As the respondents to the staff survey were self-selecting, it is possible that there could be further negative views about MyDiagnostick that have not come to light through this evaluation.

The patient survey collected information from 23 respondents. This showed that the patient experience of MyDiagnostick was unanimously positive. All understood what the device was being used for and were aware of their test result. As with the staff survey, the respondents were self-selecting and it is possible that they may not be representative of all patients using the device.

While a prospective analysis of effectiveness was not possible using Wakefield's own data, the *de novo* cost analysis was able to evaluate the cost implications and potential savings of using MyDiagnostick, compared with the pulse check method. Effectiveness data (sensitivity and specificity) from the literature was combined with evidence on the costs of the intervention, costs of treatment and of complications that may arise from treatment (e.g. bleeding) and the costs of events that may arise as a result of either treated or untreated AF (e.g. stroke).

The cost analysis showed that the use of MyDiagnostick in general practice has the potential to achieve small cost savings from an NHS and social care perspective. The costs per patient screened with MyDiagnostick and the pulse check methods were calculated to be £263.00 and £268.63 respectively, a saving of £5.63 per patient over a 10 year time horizon. The difference is largely due to the differences in further AF related complication costs, some of which may be avoided by earlier diagnosis of AF by using a more sensitive and specific test. The main limitation to the cost analysis is that it is theoretical modelling based on evidence from the literature and not actual diagnoses and prevalence of AF in Wakefield district. Also, the effectiveness data for the MyDiagnostick and pulse check methods were taken from three different, single arm studies.

The results of the cost modelling appear quite robust to individual parameter uncertainty, as the sensitivity analysis showed that the changes required to take the parameter values over £0 are quite unlikely to occur. For example, the appointment time using MyDiagnostick would need to be an average of 7.5 minutes longer than when using the pulse check method.

The sensitivity analysis showed that the extent of the savings are most influenced by changes in the cost of a false negative result (due to the consequent costs of untreated complications) and a true positive result (due to the costs of necessary AF treatment).

In conclusion, the evaluation has found that the MyDiagnostick device has been well received by those using it and has the potential to be moderately cost saving in the longer term. It is not possible to tell whether the device is helping to detect undiagnosed AF more than with the pulse check method from this study; however, the inputs used within the cost model (from single arm studies) suggest that MyDiagnostick does detect more cases of AF than the pulse check method. Indeed, the device is not being used quite as anticipated and is not replacing the pulse check method in every case. The use of the device in Wakefield does not appear to have reached its optimum position in order to judge whether it is going to be embedded into routine practice in the longer term. It may be useful to repeat the local data analysis and clinician survey in 6 to 12 months' time to gain a fuller picture. The moderate cost savings may be insufficient to prompt take up of the device, particularly when the savings appear to be downstream from primary care. It does, however, have the potential to reduce the need for a 12-lead ECG to rule out a diagnosis of AF, which may be appealing.

6.2 RECOMMENDATIONS

The following recommendations are made based on the evaluation findings:

- The number of devices should be increased for those practices that report a positive experience of using MyDiagnostick and would find it helpful to have additional devices, for example, one in each of the practice nurse rooms;
- The usage of MyDiagnostick should continue to be monitored to observe whether the pattern of usage changes, whether late adopters choose to take up the device and whether the early adopters continue to use it;
- A follow-up staff evaluation later in 2016 may be useful to gain views once the device has been in place long enough to become embedded into routine practice;
- A head-to-head study comparing MyDiagnostick and pulse check method would provide information on the efficacy of the device and add more confidence to the results. This is not possible in the Wakefield CCG area, however.

References

1. Public Health England. Atrial fibrillation prevalence estimates 2013/14. 2015. Available from: http://www.yhpho.org.uk/resource/view.aspx?RID=207902.
2. Vaes B, Stalpaert S, Tavernier K, Thaels B, Lapeire D, Mullens W*, et al.* The diagnostic accuracy of the MyDiagnostick to detect atrial fibrillation in primary care. BMC Family Practice. 2014;15(1):1-7.
3. Tieleman RG, Plantinga Y, Rinkes D, Bartels GL, Posma JL, Cator R*, et al.* Validation and clinical use of a novel diagnostic device for screening of atrial fibrillation. Europace. 2014;16(9):1291-95.
4. Hobbs FD, Fitzmaurice DA, Mant J, Murray ET, Jowett S, Bryan S*, et al.* A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study. Health Technol Assess. 2005;9(40):1-74.
5. Willits I, Keltie K, Craig J, Sims A. Cost impact of the WatchBP Home A used in a primary healthcare clinic environment 2012. [cited 25th April 2016]. Available from: https://www.nice.org.uk/guidance/MTG13/documents/watchbp-home-a-for-diagnosing-and-monitoring-hypertension-and-detecting-atrial-fibrillation-eac-additional-analysis-cost-impact-in-primary-care2.
6. Curtis L. Unit Costs of Health and Social Care 2015. 2015. [cited 25th April 2016]. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/index.php.
7. South Devon Healthcare NHS Foundation Trust. Private Patient and Overseas Visitor Price List 2015. [cited 26th April 2016]. Available from: http://www.torbayandsouthdevon.nhs.uk/uploads/23968.pdf.
8. National Institute for Health and Care Excellence. Rivaroxaban for the prevention of stroke and systemic embolism in people with atrial fibrillation (TA256). 2012. [cited 25th April 2016]. Available from: https://www.nice.org.uk/guidance/ta256/chapter/4-Consideration-of-the-evidence.
9. National Institute for Health and Care Excellence. Dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation (TA249). 2012. [cited 25th April 2016]. Available from: https://www.nice.org.uk/guidance/ta249/chapter/4-Consideration-of-the-evidence.

APPENDIX A

Previous Aims of the Evaluation

A.1 Previous Aim of Evaluation

The previous aim of the evaluation was to assess whether the conversion rate from a positive MyDiagnostick result to diagnosed AF is greater than the conversion rate from a positive pulse check to diagnosed AF. A secondary evaluation aim included whether MyDiagnostick increases the detection rate of AF above the current background rate in the patient cohort.

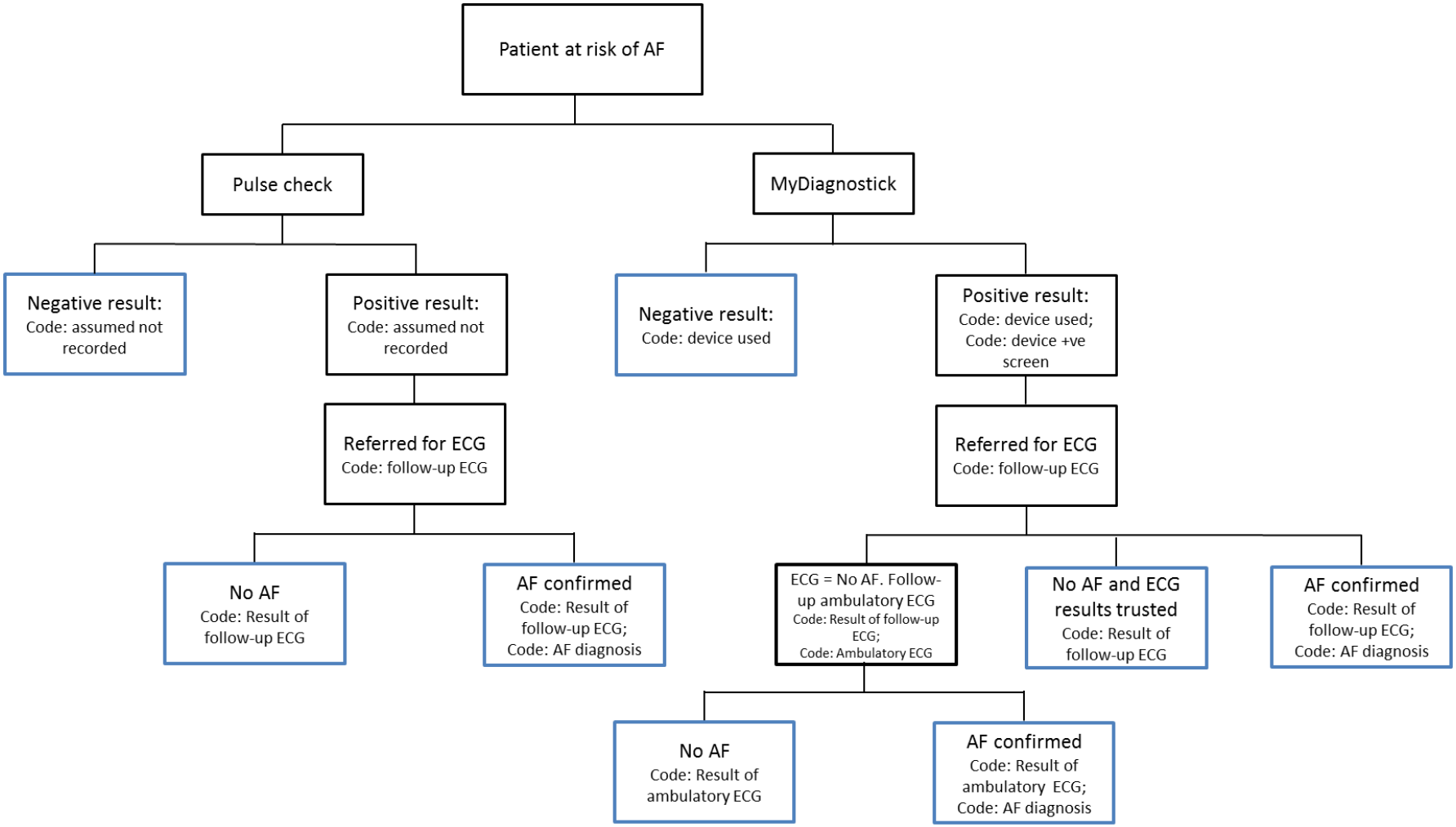
These questions were designed to be answered through a pragmatic evaluation of GP databases (SystmOne and EMIS clinical systems). The number of positive screening tests with the pulse check method or MyDiagnostick and, consequently, the number of positive diagnoses for AF were to be determined using codes entered by clinicians during patient consultations. Ideally, the sample size was to be sufficient to generate statistically significant results. Figure A.1 displays the patient flow algorithm showing the expected paths patients undergoing screening for AF may take.

The project team undertook some initial analysis of the patients undergoing screening with MyDiagnostick and the sample size required to achieve statistically significant results. Based on this, the following issues were identified:

- The clinical coding relating to the pulse check method does not always specify the site of the body at which the pulse check was taken. As such, those pulse checks of interest to the evaluation (wrist) cannot confidently be disaggregated from all other pulse checks;
- The reason for conducting a pulse check is not specified. Therefore, pulse checks being undertaken for other purposes than to screen for AF would have contributed to the number of patients with negative pulse check screening;
- Clinicians appear to have been using both the pulse check method and MyDiagnostick on around half of patients screened so far. Therefore, those patients using MyDiagnostick may have done so because they had an abnormal pulse check. As such, these patients may be at higher risk of AF than those undergoing pulse check screening;
- Feedback from clinicians demonstrated that the pulse check code may be used where pulse monitoring has been undertaken using blood pressure measurement or pulse oximetry rather than a manual pulse check. Thus, the comparator group is potentially not specific to pulse check alone;
- The sample size required to achieve statistically significant results is large and likely to be unfeasible (see Section A.2, A.3 and A.4).

As a result of the issues described above it was determined that any quantitative outcomes generated by the study were likely to be at high risk of bias, be highly uncertain and be non-significant. The project team therefore determined that the evaluation questions should focus on a primarily qualitative appraisal of MyDiagnostick. Such evaluation would be of most use to Wakefield CCG and any subsequent decision making regarding the purchase of additional MyDiagnostick devices.

Figure A.1: Patient flow algorithm



A.2 BACKGROUND

Sample size calculations have been undertaken to determine the sample size required to demonstrate a robust significant difference between the pulse check method and MyDiagnostick method for atrial fibrillation (AF) screening. The sample size required depends upon the outcome of interest. Therefore, within this document a number of options are presented for discussion with the project team. These calculations are based on the following:

- A 2:1 ratio of people in the pulse check and MyDiagnostick screening groups;
- 95% confidence level – the outcome of interest falls within the interval estimates 95% of the time. 95% confidence level is standard;
- 80% power – the likelihood that the study will detect an effect when there is an effect to be detected. 80% power is standard;
- Two-tailed test – the test will assess whether MyDiagnostick is significantly more or less effective than pulse check. Two-tailed tests are recommended if there is a possibility that MyDiagnostick could be less effective, even if this is highly unlikely.

The options presented provide a baseline figure of the sample size required. These can be updated and finalised if any of the bases in the calculations listed above are amended. The sample size calculator used can be found at:
http://epitools.ausvet.com.au/content.php?page=2Proportions

A.3 SAMPLE SIZE CALCULATION OPTIONS

Option A: Sample size for difference in true positive result for AF

Sample size calculations were undertaken to determine the sample size required to detect a significant robust difference in the true AF diagnoses in patients undergoing screening with either the pulse check method or MyDiagnostick. That is, the patients who have a positive screening test and go on to have confirmed AF.

This calculation was based on a prevalence of AF in the study population of 11.84%[2], of which 3.79% is likely to be undetected. This estimation is based on AF diagnosis data from NHS Wakefield CCG. Sensitivity (true positive detection rate) of the pulse check method of 94%[3] and sensitivity of MyDiagnostick of 100%[4] were used.

The total sample required is n = 237,351 with n = 79,117 being screened with MyDiagnostick and n = 158,234 being screened with the pulse check method. The numbers required are so large due to the relatively small difference in sensitivity of the two tests (100% versus 94%) once applied to the relatively low prevalence of undetected AF in the study population (3.79%).

[2] Davies *et al*. (2012): http://europace.oxfordjournals.org/content/14/11/1553.short
[3] Cooke *et al*. (2006): http://www.ncbi.nlm.nih.gov/pubmed/16451780
[4] Tieleman *et al*. (2014): http://www.ncbi.nlm.nih.gov/pubmed/24825766

Option B: Sample size for difference in false positive for AF

Similar sample size calculations were undertaken to determine the sample size required to detect a significant robust difference in the number of false positive AF diagnoses in patients undergoing screening with either the pulse check method or MyDiagnostick. That is, patients who have a positive screening test, but then do not have confirmed AF following an ECG or ambulatory ECG. Again, a prevalence of 3.79% for undetected AF in the study population was used. Specificities of 72% for the pulse check method[2] and 96% for MyDiagnostick were used[3].

The total sample required is n = 13,545 with n = 4,515 being screened with MyDiagnostick and n = 9,030 being screened with the pulse check method.

Option C: Sample size for 10 new cases of AF per month

Pragmatic calculations undertaken by NHS Wakefield CCG determined that an additional 10 AF diagnoses per month with MyDiagnostick would be required to show a clinically meaningful difference in diagnosis rates. The population meeting the study inclusion criteria in Wakefield CCG is expected to be 56,244. Based on a 3.79% prevalence of undetected AF in this population, 2,132 people would be expected to be diagnosed with AF. The mean current rate of AF detection with the pulse check method is 60 per month[5]. Therefore, sample size calculations have been conducted based on 70 AF detections per month with MyDiagnostick for 12 months.

The total sample required is n = 2,703 with n = 901 being screened with MyDiagnostick and n = 1,802 being screened with the pulse check method.

If MyDiagnostick was expected to detect fewer than 10 additional cases of AF per month compared with the pulse check method, the sample size would increase.

Option D: Sample size for non-significant results

Should the project team conclude that achieving significance in results it not essential, the sample size can be determined by other methods. For example, the timeframe of the project could be decided and sample size determined by the throughput of patients within said timeframe.

Calculations around diagnostic accuracy could still be undertaken and the study questions answered. However, any 'significant differences' found could be due to chance if the study is insufficiently powered (the usual level is 80%).

[5] Based on calculations by P Jacques using Wakefield data on new AF diagnoses, Nov 2013-May 2015.

A.4 CONCLUSION

The project team has identified the following as the principal question of interest:

> "Does MyDiagnostick improve the conversation rate from positive test to AF diagnosis compared with the pulse check method"?

The sample size calculations above show that in order to demonstrate a statistically significant answer to this question, a relatively large sample is required using Options A and B above. The likelihood of achieving these sample sizes will be determined by the take up of the device by GP practices and clinicians and also by the length of time the project continues. The sample size required for Option C is somewhat more achievable. A further option is to change the level of significance required or indeed, accept that any differences shown may not be statistically significant (Option D).

The chosen method will determine how calculations are undertaken in the final analysis but will not change the basic study design.

It appears that achieving an adequate sample size to demonstrate statistical significance with sufficient power in the study may be difficult. If the project team believe that based on the likely duration of the study and the throughput of patients that the sample size in Option C is achievable, then this will be a feasible option. However, it should be noted that should the required sample size be achieved, significant results that are suitably powered will only be gained if an additional 10 AF diagnosis per month for a one year duration are achieved.

Following discussion with the Project team regarding priorities for the evaluation, our recommendation is Option C if activity can be increased (noting the caveats listed above) or otherwise Option D. This option would allow for the study questions regarding diagnostic efficacy to be answered, with added value gained from the qualitative aspects of the evaluation.