Triage tests for identifying atrial fibrillation in primary care: a diagnostic accuracy study comparing single lead ECG and modified BP monitors

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Triage tests for identifying atrial fibrillation in primary care: a diagnostic accuracy study comparing single lead ECG and modified BP monitors

Karen Kearley[1], Mary Selwood[1], Ann Van den Bruel[1], Matthew Thompson[1], David Mant[1], Richard Hobbs[1], David Fitzmaurice[2], Carl Heneghan[1].

Karen Kearley Clinical Scientist, Mary Selwood Research Nurse, Ann Van den Bruel Clinical Lecturer, Matthew Thompson Clinical Reader, David Mant Professor, Richard Hobbs Head of Primary Care Health Sciences , Carl Heneghan Clinical Reader

[1]Department of Primary Care Health Sciences, University of Oxford, New Radcliffe House, 2nd floor, Walton Street, Jericho, OX2 6NW, UK

[2]Primary Care Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, Fitzmaurice D Professor,

Correspondence to: Karen.kearley@phc.ox.ac.uk

Prior presentations: *Society of Academic Primary Care Conference* November 2012 (Glasgow) and July 2013 (Nottingham), and *North American Primary Care Research Group* (Ottawa, 2013).

Word Count: 2789

Numbers of tables: 3

Numbers of figures: 3

1

Abstract

Objectives: New electronic devices offer an opportunity within routine primary care settings for improving the detection of atrial fibrillation (AF), which is a common cardiac arrhythmia and a modifiable risk factor for stroke. We aimed to assess the performance of a modified blood pressure monitor and two single-lead ECG devices, as diagnostic triage tests for the detection of AF.

Setting: Six General Practices in the United Kingdom

Participants: 1000 ambulatory patients aged 75 years and over

Primary and secondary outcome measures: Comparative diagnostic accuracy of nurse pulse palpation, modified blood pressure monitor and single-lead ECG devices, compared to reference standard of 12-lead ECG, independently interpreted by cardiologists.

Results: A total of 79 participants (7.9%) had AF diagnosed by 12-lead ECG. All three devices had a high sensitivity (93.9-98.7%) and are useful for ruling out AF. WatchBP is a better triage test than Omron autoanalysis because it is more specific - 89.7% (95% CI 87.5-91.6) compared to 78.3% (95% CI 73.0-82.9) respectively. This would translate into a lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message in the study population. The overall specificity of single-lead ECGs analysed by a cardiologist was 94.6% for Omron and 90.1% for Merlin.

Conclusions: WatchBP performs better as a triage test for identifying AF in primary care than the single-lead ECG monitors as it does not require expertise for interpretation and its diagnostic performance is comparable to single-lead ECG analysis by cardiologists. It could be used opportunistically to screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement.

Trial Registration – not required.

Strengths and Limitations of this study

Strengths

- Three devices for detecting atrial fibrillation were tested on an unselected elderly primary care population of 1000 individuals.
- The prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision.
- Our population is generalisable to similar primary care settings worldwide.
- A reference standard (12 lead ECG) was performed on all patients, and interpreted blindly by cardiologists.

Limitations:

- The specificity of one cardiologist was substantially lower than the other three.
- Only 12 new cases of atrial fibrillation were detected in this setting, which is as expected for a primary care population such as this

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Atrial fibrillation is the most common cardiac arrhythmia; present in more than 10% of patients aged 75 years or over [1,2], and it significantly increases morbidity and mortality.[3] The main significance of atrial fibrillation (AF) is as a major independent risk factor for stroke and thromboembolism,[4] particularly in older patients. A large evidence base supports the efficacy of oral anticoagulation in reducing AF stroke risk by two-thirds[5]. AF may be asymptomatic but can be identified by detecting a characteristic irregularity in pulse rhythm. Despite this, it was found to be undiagnosed in 3.8% of patients aged over 75 years in a large UK screening study[6]. Indeed, undiagnosed AF is associated with 3.8-6.1% of all strokes[7,8]. Given the importance of identifying stroke risk patients, the high prevalence of AF, and the effective prevention strategies, there is a strong case to consider screening for atrial fibrillation.

European guidelines for the management of AF recommend opportunistic pulse assessment by a primary-care practitioner with a follow-up electrocardiogram (ECG) for an irregular pulse, an approach which is more cost effective than systematic screening with a 12-lead ECG. [6,9] In primary care research studies the sensitivity of nurse pulse assessment is high, varying between studies from 87-97% although it has a lower specificity (70 to 81%) requiring one out of every three or four screened patients to have a follow-up ECG[6,10,11]. This is a significant additional workload and could be a barrier to widespread implementation of opportunistic screening. Although pulse examination is a simple screening technique for AF, its detection relies on a subjective assessment in a busy routine care setting, which may partly explain why AF remains undetected in many patients. Indeed the American Heart Association suggests there is a need to develop strategies to detect AF more effectively in individuals and populations.[12]

Several new electronic devices have the potential to be useful triage tests for AF and initial studies suggest these have higher specificity than pulse palpation, so could significantly reduce the need for confirmatory 12-lead ECGs - but none have been evaluated in a primary care settings [13,14,15]. The National Institute for Health and Clinical Excellence (NICE) has recently suggested that using a blood pressure monitor, WatchBP, which is modified to detect irregularity of pulse during BP measurement in the elderly, together with appropriate anticoagulation of those subsequently diagnosed with AF, has the potential to reduce the incidence of stroke[16]. As stroke is the third leading cause of global disease burden, this would have substantial economic and patient benefits worldwide [16,17].

Methods:

Study design

3

An observational study comparing the accuracy of a modified BP monitor (WatchBP), and two single-lead ECG devices, one with an auto-analysis function (Omron) and one without (Merlin), with the reference standard of blinded cardiologist diagnosis of atrial fibrillation from a 12-lead ECG.

Participants

We recruited participants aged 75 years or over, living at home, from six general practices in the UK between May 2011 and October 2012. Having excluded patients with implanted pacemakers or defibrillators, those unable to give informed consent, or patients in whom the GP considered participation was inappropriate (e.g. terminal illness), all other patients over 75 were invited to take part until the sample size had been achieved (i.e. 1000 participants had been recruited). A diagnosis of AF recorded in the patient GP record was not used as part of the selection criteria. Consenting study participants identified from GP records and meeting the eligibility criteria attended a 30-minute appointment with a practice nurse for the testing procedure.

Test procedure

All participants received the tests in the same order, which were delivered by nine registered nurses working within the practices.

Firstly, the nurse used WatchBP (Microlife, Switzerland) a modified oscillometric BP monitor which flashes when it detects an irregular pulse during automatic BP measurement.

Secondly, the nurse then applied an Omron monitor (model HCG-801, Omron Healthcare Europe, The Netherlands) which involved placing one electrode on the bare chest wall 5 cm below the nipple, whilst the patient held the other electrode with the right index finger. The monitor records a single-lead ECG tracing, and displays a message indicating the presence of possible AF. The device's analysis algorithm includes several cardiac rhythms which could potentially be AF, including fast and irregular, slow and irregular, irregular and those where analysis impossible. The single-lead recording and text message were recorded and saved for later downloading and analysis.

Thirdly, the nurse applied a Merlin ECG event recorder (Meditech Ltd, Hungary) which resembles a watch, on participants' left wrist. The participant covered the electrode on the face of the device with the palm of their right hand for 30 seconds. The recording, with no automated analysis, was saved to a computer for later downloading and analysis. Unlike the Omron, the Merlin monitor does not require removal of any clothing, making it possible for use in public settings, an advantage for participants experiencing an intermittent arrhythmia.

The nurse recorded the results of the WatchBP monitor and the Omron automated text message during the initial examination. Each single-lead ECG trace was sent for interpretation to two independent cardiologists after removing all clinical information and patient identification except for date of birth and the text message (Omron only).

Participants were asked if they had experienced palpitations, racing, pounding, fluttering or irregular heartbeat in the previous 4 weeks.

The reference standard

A 12-lead ECG was performed on all participants at the end of the visit, using standard procedures. ECGs were independently interpreted by a panel of two cardiologists, blind to all patient identifiers, clinical details, results of triage tests, but not blind to study objectives. They were asked to classify the ECG as atrial fibrillation or atrial flutter (yes/no/don't know). A third cardiologist, also blind to any of the other study results, interpreted the ECG readings for which there was no consensus or if

4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

both cardiologists were uncertain about the presence of AF. AF is defined as the absence of distinct 'p' waves, an absolutely irregular RR interval and an atrial cycle length <200ms (300bpm) on a 12-lead ECG[17].

To cope with the workload, the task was divided between two separate pairs of cardiologists. All had completed cardiology specialist training of between five and six years. The third cardiologist, whose role was to resolve uncertainty and moderate disagreement, was a specialist electro-physiologist.

Study power and analysis

We anticipated that the overall prevalence of AF, including both new and prior diagnoses, would be 10%. We therefore estimated that 1000 participants would be needed to provide sufficient precision to estimate the sensitivity of each device with a maximum 95% CI of +/-10%.

We compared the accuracy of the WatchBP, Omron (text and tracing) and Merlin monitors, with a 12-lead ECG for detection of AF. All inconclusive results were treated as positive in our analysis (because in the clinical situation these participants could not have AF ruled out and would need to have a 12-lead ECG). We calculated sensitivity, specificity, likelihood ratios and 95% confidence intervals for each test. Analyses for both single-lead ECGs were done per cardiologist and the results were subsequently combined using a random effects bivariate model calculating overall sensitivity and specificity[19]. We analysed the inter-observer agreement of whether the single lead ECG trace showed AF between the cardiologists within each pair together with the kappa value [20].

We performed subset analyses of the comparative accuracy in patients with and without existing diagnoses of AF (whether or not treated) recorded in GP records. In addition, we modelled a 2-stage screening process for AF by examining patients who were WatchBP flash positive and analysing the sensitivity and specificity of the Omron autoanalysis in this subgroup.

Data analysis was performed using Stata version 11S

5

Results

Of 6529 potentially eligible patients, we invited 2673 for a screening visit; no further invitations were sent after the required 1000 sample had been achieved (Figure 1). Participants had an average age of 79.7 years (range 75.1 to 99.8 years), 49.3% were male (Table 1). 110 had AF recorded in the medical summary (11.0%), of whom 67 (67/110, 61%) had evidence of AF on their 12-lead ECG at the study visit. AF was noted in a further 12 participants based on the 12-lead ECG (1.2%). The three cardiologists could not reach a consensus about whether AF was present in the 12-lead ECG in one participant, who was therefore excluded from further analyses which are therefore based on 999 participants.

Overall, 50 (5.0%) participants reported having experienced palpitations, racing, pounding, fluttering or irregular heartbeat during the previous four weeks. Of these, 16 were already known, and three were new cases of AF. The inter-observer agreement for the 12-lead ECGs was good for both pairs of cardiologists; the kappa was 0.69 and 0.96 respectively for pair 1 and pair 2.

<u>**Accuracy of WatchBP and Omron autoanalysis**</u>

Both WatchBP and Omron autoanalysis were highly sensitive (94.9% and 98.7%, respectively) (Table 2). The specificity of WatchBP was 89.7 % (95% CI 87.5-91.6) whereas the Omron text message was lower at 76.2% (95%CI 73.3-78.9).

The higher specificity of WatchBP would translate into a markedly lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message. The superior sensitivity of the Omron text message means the false negative rate is lower than with Watch BP (0.1% vs 0.4%).

The prevalence of AF was 1.4% (11) in the subset of 889 participants who did not have AF recorded on their medical summary. WatchBP and the Omron autoanalysis provided very similar sensitivity and specificity to the total study population and this translated to a follow-on ECG rate of 11% for WatchBP and 24% for the Omron text message. Of these follow on ECGs, 12% would be true positives for the WatchBP flash compared to 5% for the Omron text message.

We modelled a two stage screening process within our total sample, where patients who had a positive result with WatchBP went on to have the Omron autoanalysis. In participants who screened positive using the WatchBP, the sensitivity and specificity of the Omron autoanalysis was 100% (95%CI 95.2-100) and 35.8% (95% CI 26.2-46.3) respectively. This would slightly reduce the follow-on 12-lead ECG requirement from 17% to 13.6%.

<u>**Accuracy of Omron and Merlin single-lead ECGs interpreted by a cardiologist**</u>

The diagnostic accuracy of the Omron and Merlin single-lead ECGs is shown in figure 3; meta-analysis of the four cardiologist's interpretations gave Omron a summary sensitivity of 94.4% with a summary specificity of 94.6%; Merlin had a summary sensitivity of 93.9% with a summary specificity of 90.1%. Cardiologist variability in single-lead ECG analysis is shown in Table 3. There was fair interobserver agreement (kappa 0.36 and 0.28 for Omron and Merlin respectively) between cardiologist pair 1, and good interobserver agreement (kappa 0.76 for both devices) between the second pair. The cardiologists were unable to interpret whether the trace showed AF on 0-1.2% of patients using Omron and 0-7.2% of occasions using Merlin.

6

Discussion

Summary of main results

Our results show that all the methods that we tested have a high sensitivity for detecting AF in an elderly primary care population and thus are useful for ruling out AF. During the study we detected 12 new cases of AF and 7.9% of the study population overall were experiencing AF at the time of the study consultation. The total prevalence in the study sample was 11%, as some patients have paroxysmal AF and were not experiencing arrhythmias at the time of the study.

WatchBP outperformed Omron autoanalysis because it is more specific – 89.7% compared to 76.2% and 78.3% respectively. The specificity of the single-lead ECGs when interpreted by a cardiologist were comparable with WatchBP. Overall Omron produced fewer inconclusive results and achieved slightly better specificity than Merlin. However cardiologist interpretation of the ECG trace was a more powerful factor in determining specificity than the type of monitor used. In this study, operator variability in ECG analysis was partly due to differences in expertise, and partly due to trade-offs made by different cardiologists between sensitivity and specificity.

In our elderly population, only 5% of patients reported experiencing palpitations. Amongst patients who did report palpitations, prevalence of AF was 36%. Most of these patients already had a diagnosis of AF.

Strengths and weakness

The devices were tested on an unselected elderly primary care population. Although our sample did not include the housebound or patients with dementia, we recruited a large population with a prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision. We believe our population is generalisable to similar primary care settings worldwide.

The cardiologists analysed ECG recordings in line with recommended practice and all had a minimum of 5 years cardiology specialist training. The specificity of one cardiologist was substantially lower than the other three. Further research with a larger sample of cardiologists would be required to more fully interpret this finding. However, we do not believe this diminishes the reliability of our reference standard since the second pair of blinded cardiologists achieved extremely high concordance, and the moderating consultant cardiologist was a specialist electrophysiologist.

Comparison with existing literature

The operating characteristics of WatchBP in this elderly primary care population are comparable to previous studies undertaken in different populations. In a US study of 405 patients attending a cardiology outpatients where 39% of the sample had an abnormal ECG (AF prevalence 23%) it had a sensitivity of 95% (95% CI 93 -98) and specificity of 86% (95% CI 84-89) while in a Greek study of a diverse group of 73 individuals (healthy volunteers, outpatients, and inpatients) it had similar sensitivity of 93% (95% CI 74-99) and specificity of 89% (95% CI 79-96).

Omron has been evaluated in a German study of 505 patients in a hospital setting where 66% of the sample had an abnormal ECG (mean age 61 years, 66% male, AF prevalence 28%) with all analyses (including the reference standard) being undertaken by a single cardiologist. This study reported higher accuracy than we demonstrated, with a sensitivity of 99% (95%CI 96-100) and specificity of 96% (95% CI 94-98) for AF. We are unaware of any AF validation studies for Merlin.

7

Clinical implications

Our study supports the recent NICE recommendation that GPs should consider using WatchBP for elderly patients who have a BP measurement in order to opportunistically identify undiagnosed AF. Specificity is comparable with a single-lead ECG analysed by a cardiologist and the result does not require specialist interpretation, which is an advantage over the single-lead ECG monitors and keeps screening straightforward and costs low. It is easily portable for use in housebound patients.

Prior to wider implementation, clinicians need to consider the optimal screening frequency. Assuming a primary care population in which the majority of AF cases had already been identified, the WatchBP would indicate a 'positive flash' for 1 in 10 patients who would thus require a follow-up 12-lead ECG. As some patients have their blood pressure measured frequently, there is potential for recurrent false-positive results and a high demand for 12-lead ECGs. These issues should be addressed to ensure sustainable implementation in primary care.

Paroxysmal AF presents a particular diagnostic challenge because patients may not be experiencing an arrhythmia when they are being screened with one-off measures. The potential for patients to undertake self-monitoring and to capture an ECG tracing when an arrhythmia is occurring are desirable features of an AF screening method. In our study, only 5% of patients reported experiencing palpitations, which would limit the usefulness of asking patients to record a single-lead ECG at the time of symptoms. Amongst the patients who did report palpitations, the prevalence of AF was 36%, but most of these patients already had a diagnosis of AF which might have improved their awareness of an arrhythmia.

WatchBP will not identify atrial flutter, which is also a risk factor for stroke. The prevalence of this is much lower than AF; in our study it was 0.4%.

Future research

Screening for AF is not yet routine practice in primary care, therefore determining not only the optimal device, but the most appropriate and cost effective process is a priority. This could include an impact analysis using WatchBP as a triage test to evaluate its effectiveness (and cost effectiveness) for improving the detection of AF and prevention of stroke.

Conclusion

WatchBP performs better than the single-lead ECG monitors as a triage test for identifying AF in primary care as it does not require any expertise for interpretation and their diagnostic performance is comparable. It could be used to opportunistically screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement. It has advantages over pulse assessment because the result is objective and it also has a higher specificity, keeping follow-up 12-lead ECGs to a minimum.

Contribution of authors

KK, MS, DM, RH, DF, CH initiated the study and designed the protocol. KK, MS, AV, MT, CH were involved in the conduct and analysis of the trial. MS oversaw data management and monitoring of protocol adherence throughout the trial. KK, AV, MT co-wrote the manuscript. MS, DM, CH, RH, DF contributed to writing the manuscript.

Full ethical approval was granted of the study by NRES Committee South Central – Oxford A.

Competing interests

We received subsidised monitors from Microlife, Omron and Merlin for the submitted work; FDRH has received free or subsidised BP equipment from Omron and Microlife for other research and there are no other relationships or activities that could appear to have influenced the submitted work .

Copyright

Crown copyright is claimed for this publication, in accordance with NIHR guidelines.

References

1. NICE. Atrial Fibrillation: National clinical guideline for management in primary and secondary care. 2006.
2. Fitzmaurice DA, Hobbs FDR, Jowett S, Mant J, Murray E, Holder R et al. Screening versus routine practice for detection of atrial fibrillation in people aged 65 or over: cluster randomised controlled trial. *BMJ* 2007; 335:383-6.
3. Miyasaka Y, Barnes ME, Bailey KR, Cha SS, Gersh BJ, Seward JB, Tsang TS. Mortality trends in patients diagnosed with first atrial fibrillation – a 21-year community study. *J Am Coll Cardiol* 2007; 49:986-92.
4. Wolf PA, Abbott RD, Kannell WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. *Stroke* 1991 Aug; 22(8):983-8.
5. Aguilar MI, Hart R. Oral anticoagulants for preventing stroke in patients with non-valvular atrial fibrillation and no previous history of stroke or transient ischemic attacks. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Edited (no change to conclusions), published in Issue 1, 2009
6. Hobbs F, Fitzmaurice DA, Mant J, Murray E, 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 Oct; 9(40): 1-74.
7. Lin HJ, Wolf PA, Benjamin EJ, Belanger AJ, D'Agostino RB. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. *Stroke* 1995;26:152730.
8. Liao J, Khalid Z, Scallan C, Morillo C, O'Donnell M. Noninvasive cardiac monitoring for detecting paroxysmal atrial fibrillation or flutter after acute ischemic stroke: a systematic review. *Stroke* 2007;38:
9. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. An update of the 2010 ESC Guidelines for the management of atrial fibrillation. *European Heart Journal* (2012) 33, 2719–2747
10. Sudlow M, Rodgers H, Kenny RA Thomson R. Identification of patients with atrial fibrillation in general practice: a study of screening methods. *British Medical Journal* 1998;317:327
11. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. *British Journal of General Practice*. 2002;52:373–380
12. Estes, N, Sacco R, Al-Khatib S, Ellinor P, Bezanson J, Alonso A, et al. American Heart Association Atrial Fibrillation Research Summit. Circulation. 2011 July 19; 124(3): 363–372
13. Wiesel J, Fitzig L, Herschman Y, Messineo F. Detection of atrial fibrillation using a modified microlife blood pressure monitor. *Am J Hypertens 2009*; 22(8): 848-852.
14. Stergiou G, Karpettas N, Protogerou A Nasothimiou EG, Kyriakidis M. Diagnostic accuracy of a home blood pressure monitor to detect atrial fibrillation. *Journal of Human Hypertension* (2009) 23, 654–65
15. Kaleschke G, HoffmanB Drewitz I Steinbeck G, Naebauer M, Goette A, et al. Prospective, multicentre validation of a simple, patient-operated electrocardiographic system for the

9

detection of arrhythmias and electrocardiographic changes. *Europace* (2009) 11 (10): 1362-1368

16. NICE. WatchBP Home A for opportunistically detecting atrial fibrillation during diagnosis and monitoring of hypertension, 2013

17. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2095

18. Harris K, Edwards D, Mant J. How can we best detect atrial fibrillation? *J R Coll Physicians* 2012;42 Suppl 18:5-22.

19. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010

20. Altman D. Practical Statistics for Medical Research. Chapman and Hall/CRC, 1991.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

Table 1: Characteristics of participants included in diagnostic accuracy analyses

Total number of participants	999
Age (mean)	79.7 years (range 75.1-99.8) (95% CI 79.4-79.9)
Gender % male	49.3 (95%CI 46.2-52.5)
Co-morbidities (from medical summary) Atrial fibrillation Heart failure Hypertension Diabetes Stroke Transient ischaemic attack	110 (11%) 31 (3.1%) 533 (53.3%) 122 (12.2%) 31 (3.1%) 65 (6.5%)
Participants with AF on anti-arrhythmic medication	87(8.7%)

Table 2: Diagnostic performance of WatchBP and Omron autoanalysis for detecting atrial fibrillation

n=999* Prevalence 7.9%	WatchBP with AF indicator flash (95% CI)	Omron autoanalysis* (95%CI)
Sensitivity %	94.9 (87.5-98.6)	98.7(93.2-100)
Specificity %	89.7 (87.5-91.6)	76.2 (73.3-78.9)
Positive likelihood ratio	9.2 (7.6-11.2)	4.15 (3.69-4.67)
Negative likelihood ratio	0.057 (0.022-0.15)	0.017 (0.0024-0.12)
Positive predictive value	44.1 (36.5-51.9)	26.3 (21.3-31.7)
Negative predictive value	99.5 (98.8-99.9)	99.9 (99.2-100)

*text message 'irregular 'or 'analysis impossible' counted as positive test

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

11

Table 3: Diagnostic performance of Omron and Merlin single-lead ECGs for detecting atrial fibrillation

n=999* Prevalence 7.9%	Interpretation	Omron single-lead ECG (95% CI)	Merlin single-lead ECG (95% CI)
Sensitivity %	*Cardiologist: 1*	95 (82-99)	97 (85-100)
	2	95 (82-99)	91 (77-98)
	3	90 (77-97)	98 (87-100)
	4	98 (87-100)	88 (74-96)
Specificity %	*Cardiologist: 1*	75 (71-79)	76 (71-80)
	2	98 (96-99)	95 (93-97)
	3	97 (96-99)	92 (89-94)
	4	95 (92-97)	91 (88-94)
Positive likelihood ratio	*Cardiologist: 1*	3.8 (3.2-4.6)	3.99 (3.3-4.8)
	2	41.8 (22.5-77.5)	18.8 (12.2-28.8)
	3	35.8 (20.3-63.1)	11.5 (8.5-15.6)
	4	18.5 (12.6-27.2)	10.1 (7.4-13.9)
Negative likelihood ratio	*Cardiologist: 1*	0.07 (0.02-0.28)	0.04 (0.01-0.26)
	2	0.05 (0.01-0.21)	0.09 (0.03-0.27)
	3	0.098 (0.04-0.25)	0.03 (0.004-0.18)
	4	0.02 (0.004-0.17)	0.13 (0.06-0.3)
Positive predictive value	*Cardiologist: 1*	24.1 (17.4-31.9)	24.5 (17.6-32.5)
	2	77.8 (62.9-88.8)	60.4 (46.0-73.5)
	3	76.0 (61.8-86.9)	50.6 (39.3-61.9)
	4	62.1 (49.3-73.8)	47.4 (36.0-59.1)
Negative predictive value	*Cardiologist: 1*	99.4 (97.9-99.9)	99.7 (98.3-100)
	2	99.5 (98.3-99.9)	99.3 (97.9-99.8)
	3	99.1 (97.8-99.8)	99.8 (98.7-100)
	4	99.8 (98.8-100)	98.9 (97.3-99.6)

Figure 1: Patient recruitment flow chart

6529 Potentially eligible patients

Exclusions:
Those with pacemakers
Unable to give informed consent
House bound
Inappropriate

2673 invited until recruitment target reached

Non responders
Or
Practice discontinued involvement in study

1000 for inclusion in analysis

Excluded:
1 no consensus on AF on 12 lead ECG among cardiologist x3

999 included in final analysis against the gold standard of 12 Lead ECG; *where n < 999 indicates missing data*

983 Nurse Palpation vs 12 lead ECG

999 AF Flash indicator vs 12 lead ECG

998 Omron autoanalysis vs 12lead ECG

996 Omron ECG trace vs 12 lead ECG

980 Merlin ECG trace vs 12 lead ECG

Figure 2: Predictive value of WatchBP and Omron autoanalysis for detecting AF

All patients (999)		
	WatchBP	
	Omron message	
Patients without known AF (889)		
	WatchBP	
	Omron message	

Blue dots = prevalence; red dots: probability of AF after a positive test; green dots: probability of AF after a negative test (each with their 95% CI)

Figure 3: Forest plot of diagnostic accuracy per cardiologist of Omron and Merlin single-lead ECGs

OMRON

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cardiologist 1	35	110	2	332	0.95 [0.82, 0.99]	0.75 [0.71, 0.79]		
Cardiologist 2	35	10	2	432	0.95 [0.82, 0.99]	0.98 [0.96, 0.99]		
Cardiologist 3	38	12	4	463	0.90 [0.77, 0.97]	0.97 [0.96, 0.99]		
Cardiologist 4	41	25	1	450	0.98 [0.87, 1.00]	0.95 [0.92, 0.97]		

MERLIN

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cardiologist 1	34	105	1	326	0.97 [0.85, 1.00]	0.76 [0.71, 0.80]		
Cardiologist 2	32	21	3	410	0.91 [0.77, 0.98]	0.95 [0.93, 0.97]		
Cardiologist 3	41	40	1	432	0.98 [0.87, 1.00]	0.92 [0.89, 0.94]		
Cardiologist 4	37	41	5	431	0.88 [0.74, 0.96]	0.91 [0.88, 0.94]		

BMJ open

Triage tests for identifying atrial fibrillation in primary care:
a diagnostic accuracy study comparing single lead ECG and
modified BP monitors

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004565.R1
Article Type:	Research
Date Submitted by the Author:	10-Mar-2014
Complete List of Authors:	Kearley, Karen; Oxford University, Primary Health Care Selwood, Mary; University of Oxford, Primary Care Health Sciences van den Bruel, Ann; University of Oxford, Dept of Primary Health Care Thompson, Matthew; Oxford University, Department of Primary Care Health Sciences Mant, David; University of Oxford, Primary Health Care Hobbs, Richard; University of Oxford, Dept of Primary Health Care Fitzmaurice, DA; University of Birmingham, Primary Care and General Practice Heneghan, Carl; Oxford University, Primary Health Care
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Cardiovascular medicine, Diagnostics, General practice / Family practice
Keywords:	Adult cardiology < CARDIOLOGY, PRIMARY CARE, STROKE MEDICINE

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Triage tests for identifying atrial fibrillation in primary care: a diagnostic accuracy study comparing single lead ECG and modified BP monitors

Karen Kearley[1], Mary Selwood[1], Ann Van den Bruel[1], Matthew Thompson[1,2], David Mant[1], FD Richard Hobbs[1], David Fitzmaurice[3], Carl Heneghan[1].

Karen Kearley Clinical Scientist, Mary Selwood Research Nurse, Ann Van den Bruel Clinical Lecturer, Matthew Thompson Senior Clinical Researcher & Professor, David Mant Professor, Richard Hobbs Head of Primary Care Health Sciences , Carl Heneghan Professor

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, New Radcliffe House, 2nd floor, Walton Street, Jericho, OX2 6NW, UK

[2]Department of Family Medicine, University of Washington, Seattle, WA 98195-4696

[3]Primary Care Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, Fitzmaurice D Professor,

Correspondence to: Matthew.Thompson@phc.ox.ac.uk

Word Count: 2789

Numbers of tables: 3

Numbers of figures: 3

1

Abstract

Objectives: New electronic devices offer an opportunity within routine primary care settings for improving the detection of atrial fibrillation (AF), which is a common cardiac arrhythmia and a modifiable risk factor for stroke. We aimed to assess the performance of a modified blood pressure monitor and two single-lead ECG devices, as diagnostic triage tests for the detection of AF.

Setting: Six General Practices in the United Kingdom

Participants: 1000 ambulatory patients aged 75 years and over

Primary and secondary outcome measures: Comparative diagnostic accuracy of modified blood pressure monitor and single-lead ECG devices, compared to reference standard of 12-lead ECG, independently interpreted by cardiologists.

Results: A total of 79 participants (7.9%) had AF diagnosed by 12-lead ECG. All three devices had a high sensitivity (93.9-98.7%) and are useful for ruling out AF. WatchBP is a better triage test than Omron autoanalysis because it is more specific - 89.7% (95% CI 87.5-91.6) compared to 78.3% (95% CI 73.0-82.9) respectively. This would translate into a lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message in the study population. The overall specificity of single-lead ECGs analysed by a cardiologist was 94.6% for Omron and 90.1% for Merlin.

Conclusions: WatchBP performs better as a triage test for identifying AF in primary care than the single-lead ECG monitors as it does not require expertise for interpretation and its diagnostic performance is comparable to single-lead ECG analysis by cardiologists. It could be used opportunistically to screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement.

Trial Registration – not required.

Strengths and Limitations of this study

Strengths

- Three devices for detecting atrial fibrillation were tested on an unselected elderly primary care population of 1000 individuals.
- The prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision.
- Our population is generalisable to similar primary care settings worldwide.
- A reference standard (12 lead ECG) was performed on all patients, and interpreted blindly by cardiologists.

Limitations:

- The specificity of one cardiologist was substantially lower than the other three.
- Only 12 new cases of atrial fibrillation were detected in this setting, which is as expected for a primary care population such as this

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Atrial fibrillation is the most common cardiac arrhythmia; present in more than 10% of patients aged 75 years or over [1,2], and it significantly increases morbidity and mortality.[3] The main significance of atrial fibrillation (AF) is as a major independent risk factor for stroke and thromboembolism,[4] particularly in older patients. A large evidence base supports the efficacy of oral anticoagulation in reducing AF stroke risk by two-thirds[5]. AF may be asymptomatic but can be identified by detecting a characteristic irregularity in pulse rhythm. Despite this, it was found to be undiagnosed in 3.8% of patients aged over 75 years in a large UK screening study[6]. Indeed, undiagnosed AF is associated with 3.8-6.1% of all strokes[7,8]. Given the importance of identifying stroke risk patients, the high prevalence of AF, and the effective prevention strategies, there is a strong case to consider screening for atrial fibrillation.

European guidelines for the management of AF recommend opportunistic pulse assessment by a primary-care practitioner with a follow-up electrocardiogram (ECG) for an irregular pulse, an approach which is more cost effective than systematic screening with a 12-lead ECG. [6,9] In primary care research studies the sensitivity of nurse pulse assessment is high, varying between studies from 87-97% although it has a lower specificity (70 to 81%) requiring one out of every three or four screened patients to have a follow-up ECG[6,10,11]. This is a significant additional workload and could be a barrier to widespread implementation of opportunistic screening. Although pulse examination is a simple screening technique for AF, its detection relies on a subjective assessment in a busy routine care setting, which may partly explain why AF remains undetected in many patients. Indeed the American Heart Association suggests there is a need to develop strategies to detect AF more effectively in individuals and populations.[12]

Several new electronic devices have the potential to be useful triage tests for AF and initial studies suggest these have higher specificity than pulse palpation, so could significantly reduce the need for confirmatory 12-lead ECGs - but none have been evaluated in a primary care settings [13,14,15]. The National Institute for Health and Clinical Excellence (NICE) has recently suggested that using a blood pressure monitor, WatchBP, which is modified to detect irregularity of pulse during BP measurement in the elderly, together with appropriate anticoagulation of those subsequently diagnosed with AF, has the potential to reduce the incidence of stroke[16]. As stroke is the third leading cause of global disease burden, this would have substantial economic and patient benefits worldwide [16,17].

Methods:

Study design

An observational study comparing the accuracy of a modified BP monitor (WatchBP), and two single-lead ECG devices, one with an auto-analysis function (Omron) and one without (Merlin), with the reference standard of blinded cardiologist diagnosis of atrial fibrillation from a 12-lead ECG.

Participants

We recruited participants aged 75 years or over, living at home, from six general practices in the UK between May 2011 and October 2012. Having excluded patients with implanted pacemakers or defibrillators, those unable to give informed consent, or patients in whom the GP considered participation was inappropriate (e.g. terminal illness), all other patients over 75 were invited to take part until the sample size had been achieved (i.e. 1000 participants had been recruited). A diagnosis of AF recorded in the patient GP record was not used as part of the selection criteria. Consenting study participants identified from GP records and meeting the eligibility criteria attended a 30-minute appointment with a practice nurse for the testing procedure.

3

Test procedure

All participants received the tests in the same order, which were delivered by nine registered nurses working within the practices.

Firstly, the nurse used WatchBP (Microlife, Switzerland) a modified oscillometric BP monitor which flashes when it detects an irregular pulse during automatic BP measurement.

Secondly, the nurse then applied an Omron monitor (model HCG-801, Omron Healthcare Europe, The Netherlands) which involved placing one electrode on the bare chest wall 5 cm below the nipple, whilst the patient held the other electrode with the right index finger. The monitor records a single-lead ECG tracing, and displays a message indicating the presence of possible AF. The device's analysis algorithm includes several cardiac rhythms which could potentially be AF, including fast and irregular, slow and irregular, irregular and those where analysis impossible. The single-lead recording and text message were recorded and saved for later downloading and analysis.

Thirdly, the nurse applied a Merlin ECG event recorder (Meditech Ltd, Hungary) which resembles a watch, on participants' left wrist. The participant covered the electrode on the face of the device with the palm of their right hand for 30 seconds. The recording, with no automated analysis, was saved to a computer for later downloading and analysis. Unlike the Omron, the Merlin monitor does not require removal of any clothing, making it possible for use in public settings, an advantage for participants experiencing an intermittent arrhythmia.

The nurse recorded the results of the WatchBP monitor and the Omron automated text message during the initial examination. Each single-lead ECG trace was sent for interpretation to two independent cardiologists after removing all clinical information and patient identification except for date of birth and the text message (Omron only).

Participants were asked if they had experienced palpitations, racing, pounding, fluttering or irregular heartbeat in the previous 4 weeks.

The reference standard

A 12-lead ECG was performed on all participants at the end of the visit, using standard procedures. ECGs were independently interpreted by a panel of two cardiologists, blind to all patient identifiers, clinical details, results of triage tests, but not blind to study objectives. They were asked to classify the ECG as atrial fibrillation or atrial flutter (yes/no/don't know). A third cardiologist, also blind to any of the other study results, interpreted the ECG readings for which there was no consensus or if both cardiologists were uncertain about the presence of AF. AF is defined as the absence of distinct 'p' waves, an absolutely irregular RR interval and an atrial cycle length <200ms (300bpm) on a 12-lead ECG[17].

To cope with the workload, the task was divided between two separate pairs of cardiologists. All had completed cardiology specialist training of between five and six years. The third cardiologist, whose role was to resolve uncertainty and moderate disagreement, was a specialist electro-physiologist.

Study power and analysis

We anticipated that the overall prevalence of AF, including both new and prior diagnoses, would be 10%. We therefore estimated that 1000 participants would be needed to provide sufficient precision to estimate the sensitivity of each device with a maximum 95% CI of +/-10%.

We compared the accuracy of the WatchBP, Omron (text and tracing) and Merlin monitors, with a 12-lead ECG for detection of AF. All inconclusive results were treated as positive in our analysis

4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(because in the clinical situation these participants could not have AF ruled out and would need to have a 12-lead ECG). We calculated sensitivity, specificity, likelihood ratios and 95% confidence intervals for each test. Analyses for both single-lead ECGs were done per cardiologist and the results were subsequently combined using a random effects bivariate model calculating overall sensitivity and specificity[19]. We analysed the inter-observer agreement of whether the single lead ECG trace showed AF between the cardiologists within each pair together with the kappa value [20].

We performed subset analyses of the comparative accuracy in patients with and without existing diagnoses of AF (whether or not treated) recorded in GP records. In addition, we modelled a 2-stage screening process for AF by examining patients who were WatchBP flash positive and analysing the sensitivity and specificity of the Omron autoanalysis in this subgroup.

Data analysis was performed using Stata version 11S

5

Results

Of 6529 potentially eligible patients, we invited 2673 for a screening visit; no further invitations were sent after the required 1000 sample had been achieved (Figure 1). Participants had an average age of 79.7 years (range 75.1 to 99.8 years), 49.3% were male (Table 1). 110 had AF recorded in the medical summary (11.0%), of whom 67 (67/110, 61%) had evidence of AF on their 12-lead ECG at the study visit. AF was noted in a further 12 participants based on the 12-lead ECG (1.2%). The three cardiologists could not reach a consensus about whether AF was present in the 12-lead ECG in one participant, who was therefore excluded from further analyses which are therefore based on 999 participants.

Overall, 50 (5.0%) participants reported having experienced palpitations, racing, pounding, fluttering or irregular heartbeat during the previous four weeks. Of these, 16 were already known, and three were new cases of AF. The inter-observer agreement for the 12-lead ECGs was good for both pairs of cardiologists; the kappa was 0.69 and 0.96 respectively for pair 1 and pair 2.

Accuracy of WatchBP and Omron autoanalysis

Both WatchBP and Omron autoanalysis were highly sensitive (94.9% and 98.7%, respectively) (Table 2). The specificity of WatchBP was 89.7 % (95% CI 87.5-91.6) whereas the Omron text message was lower at 76.2% (95%CI 73.3-78.9).

The higher specificity of WatchBP would translate into a markedly lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message. The superior sensitivity of the Omron text message means the false negative rate is lower than with Watch BP (0.1% vs 0.4%).

The prevalence of AF was 1.4% (11) in the subset of 889 participants who did not have AF recorded on their medical summary. WatchBP and the Omron autoanalysis provided very similar sensitivity and specificity to the total study population and this translated to a follow-on ECG rate of 11% for WatchBP and 24% for the Omron text message. Of these follow on ECGs, 12% would be true positives for the WatchBP flash compared to 5% for the Omron text message.

We modelled a two stage screening process within our total sample, where patients who had a positive result with WatchBP went on to have the Omron autoanalysis. In participants who screened positive using the WatchBP, the sensitivity and specificity of the Omron autoanalysis was 100% (95%CI 95.2-100) and 35.8% (95% CI 26.2-46.3) respectively. This would slightly reduce the follow-on 12-lead ECG requirement from 17% to 13.6%.

Accuracy of Omron and Merlin single-lead ECGs interpreted by a cardiologist

The diagnostic accuracy of the Omron and Merlin single-lead ECGs is shown in figure 3; meta-analysis of the four cardiologist's interpretations gave Omron a summary sensitivity of 94.4% with a summary specificity of 94.6%; Merlin had a summary sensitivity of 93.9% with a summary specificity of 90.1%. Cardiologist variability in single-lead ECG analysis is shown in Table 3. There was fair interobserver agreement (kappa 0.36 and 0.28 for Omron and Merlin respectively) between cardiologist pair 1, and good interobserver agreement (kappa 0.76 for both devices) between the second pair. The cardiologists were unable to interpret whether the trace showed AF on 0-1.2% of patients using Omron and 0-7.2% of occasions using Merlin.

6

Discussion

Summary of main results

Our results show that all the methods that we tested have a high sensitivity for detecting AF in an elderly primary care population and thus are useful for ruling out AF. During the study we detected 12 new cases of AF and 7.9% of the study population overall were experiencing AF at the time of the study consultation. The total prevalence in the study sample was 11%, as some patients have paroxysmal AF and were not experiencing arrhythmias at the time of the study.

WatchBP outperformed Omron autoanalysis because it is more specific – 89.7% compared to 76.2% and 78.3% respectively. The specificity of the single-lead ECGs when interpreted by a cardiologist were comparable with WatchBP. Overall Omron produced fewer inconclusive results and achieved slightly better specificity than Merlin. However cardiologist interpretation of the ECG trace was a more powerful factor in determining specificity than the type of monitor used. In this study, operator variability in ECG analysis was partly due to differences in expertise, and partly due to trade-offs made by different cardiologists between sensitivity and specificity.

In our elderly population, only 5% of patients reported experiencing palpitations. Amongst patients who did report palpitations, prevalence of AF was 36%. Most of these patients already had a diagnosis of AF.

Strengths and weakness

The devices were tested on an unselected elderly primary care population. Although our sample did not include the housebound or patients with dementia, we recruited a large population with a prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision. We believe our population is generalisable to similar primary care settings worldwide.

The cardiologists analysed ECG recordings in line with recommended practice and all had a minimum of 5 years cardiology specialist training. The specificity of one cardiologist was substantially lower than the other three. Further research with a larger sample of cardiologists would be required to more fully interpret this finding. However, we do not believe this diminishes the reliability of our reference standard since the second pair of blinded cardiologists achieved extremely high concordance, and the moderating consultant cardiologist was a specialist electrophysiologist. Having cardiologists rather than general practitioners ascertain the presence or absence of AF on the Omron or Merlin single lead ECGs may have risked over estimating accuracy given the likelihood that cardiologists would have greater skill in ECG interpretation. Finally, given that the devices were used sequentially followed by the 12 lead ECG it is theoretically possible that AF may not have been consistently present or absent during the entire procedure, leading to an increase of false positives and false negatives, however we believe this is a small risk given the extremely short time frame of approximately 10 minutes.

Comparison with existing literature

The operating characteristics of WatchBP in this elderly primary care population are comparable to previous studies undertaken in different populations. In a US study of 405 patients attending a cardiology outpatients where 39% of the sample had an abnormal ECG (AF prevalence 23) it had a sensitivity of 95% (95% CI 93 -98) and specificity of 86% (95% CI 84-89) while in a Greek study of a diverse group of 73 individuals (healthy volunteers, outpatients, and inpatients) it had similar sensitivity of 93% (95% CI 74-99) and specificity of 89% (95% CI 79-96).

7

Omron has been evaluated in a German study of 505 patients in a hospital setting where 66% of the sample had an abnormal ECG (mean age 61 years, 66% male, AF prevalence 28%) with all analyses (including the reference standard) being undertaken by a single cardiologist. This study reported higher accuracy than we demonstrated, with a sensitivity of 99% (95%CI 96-100) and specificity of 96% (95% CI 94-98) for AF. We are unaware of any AF validation studies for Merlin.

Clinical implications

Our study supports the recent NICE recommendation that GPs should consider using WatchBP for elderly patients who have a BP measurement in order to opportunistically identify undiagnosed AF. Specificity is comparable with a single-lead ECG analysed by a cardiologist and the result does not require specialist interpretation, which is an advantage over the single-lead ECG monitors and keeps screening straightforward and costs low. It is easily portable for use in housebound patients.

Prior to wider implementation, clinicians need to consider the optimal screening frequency. Assuming a primary care population in which the majority of AF cases had already been identified, the WatchBP would indicate a 'positive flash' for 1 in 10 patients who would thus require a follow-up 12-lead ECG. As some patients have their blood pressure measured frequently, there is potential for recurrent false-positive results and a high demand for 12-lead ECGs. These issues should be addressed to ensure sustainable implementation in primary care.

Paroxysmal AF presents a particular diagnostic challenge because patients may not be experiencing an arrhythmia when they are being screened with one-off measures. The potential for patients to undertake self-monitoring and to capture an ECG tracing when an arrhythmia is occurring are desirable features of an AF screening method. In our study, only 5% of patients reported experiencing palpitations, which would limit the usefulness of asking patients to record a single-lead ECG at the time of symptoms. Amongst the patients who did report palpitations, the prevalence of AF was 36%, but most of these patients already had a diagnosis of AF which might have improved their awareness of an arrhythmia.

WatchBP will not identify atrial flutter, which is also a risk factor for stroke. The prevalence of this is much lower than AF; in our study it was 0.4%.

Future research

Screening for AF is not yet routine practice in primary care, therefore determining not only the optimal device, but the most appropriate and cost effective process is a priority. This could include an impact analysis using WatchBP as a triage test to evaluate its effectiveness (and cost effectiveness) for improving the detection of AF and prevention of stroke. Determining the cost effectiveness of these devices would also have to include doctor time for interpretation of ECGs

Conclusion

WatchBP performs better than the single-lead ECG monitors as a triage test for identifying AF in primary care as it does not require any expertise for interpretation and their diagnostic performance is comparable. It could be used to opportunistically screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement. It has advantages over pulse assessment because the result is objective and it also has a higher specificity, keeping follow-up 12-lead ECGs to a minimum.

8

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Contribution of authors

KK, MS, DM, RH, DF, CH initiated the study and designed the protocol. KK, MS, AV, MT, CH were involved in the conduct and analysis of the trial. MS oversaw data management and monitoring of protocol adherence throughout the trial. KK, AV, MT co-wrote the manuscript. MS, DM, CH, RH, DF contributed to writing the manuscript.

Full ethical approval

Was granted of the study by NRES Committee South Central – Oxford A.

Competing interests

We received subsidised monitors from Microlife, Omron and Merlin for the submitted work; FDRH has received free or subsidised BP equipment from Omron and Microlife for other research and there are no other relationships or activities that could appear to have influenced the submitted work .

Data Sharing Statement

Data on comparative accuracy are available from authors on request, sending email to the corresponding author.

Copyright

Crown copyright is claimed for this publication, in accordance with NIHR guidelines.

9

References

1. NICE. Atrial Fibrillation: National clinical guideline for management in primary and secondary care. 2006.

2. Fitzmaurice DA, Hobbs FDR, Jowett S, et al. Screening versus routine practice for detection of atrial fibrillation in people aged 65 or over: cluster randomised controlled trial. *BMJ* 2007; 335:383-6.

3. Miyasaka Y, Barnes ME, Bailey KR, et al. Mortality trends in patients diagnosed with first atrial fibrillation – a 21-year community study. *J Am Coll Cardiol* 2007; 49:986-92.

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

5. Aguilar MI, Hart R. Oral anticoagulants for preventing stroke in patients with non-valvular atrial fibrillation and no previous history of stroke or transient ischemic attacks. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Edited (no change to conclusions), published in Issue 1, 2009

6. Hobbs F, Fitzmaurice DA, Mant J, 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 Oct; 9(40): 1-74.

7. Lin HJ, Wolf PA, Benjamin EJ, et al. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. *Stroke* 1995;26:152730.

8. Liao J, Khalid Z, Scallan C, et al. Noninvasive cardiac monitoring for detecting paroxysmal atrial fibrillation or flutter after acute ischemic stroke: a systematic review. *Stroke* 2007;38:

9. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. An update of the 2010 ESC Guidelines for the management of atrial fibrillation. *European Heart Journal* (2012) 33, 2719–2747

10. Sudlow M, Rodgers H, Kenny RA et al. Identification of patients with atrial fibrillation in general practice: a study of screening methods. *British Medical Journal* 1998;317:327

11. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. *British Journal of General Practice*. 2002;52:373–380

12. Estes, N, Sacco R, Al-Khatib S, et al l. American Heart Association Atrial Fibrillation Research Summit. Circulation. 2011 July 19; 124(3): 363–372

13. Wiesel J, Fitzig L, Herschman Y, et al. Detection of atrial fibrillation using a modified microlife blood pressure monitor. *Am J Hypertens 2009*; 22(8): 848-852.

14. Stergiou G, Karpettas N, Protogerou A et al. Diagnostic accuracy of a home blood pressure monitor to detect atrial fibrillation. *Journal of Human Hypertension* (2009) 23, 654–65

15. Kaleschke G, HoffmanB Drewitz I Steinbeck G, et al. Prospective, multicentre validation of a simple, patient-operated electrocardiographic system for the detection of arrhythmias and electrocardiographic changes. *Europace* (2009) 11 (10): 1362-1368

16. NICE. WatchBP Home A for opportunistically detecting atrial fibrillation during diagnosis and monitoring of hypertension, 2013

17. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2095

18. Harris K, Edwards D, Mant J. How can we best detect atrial fibrillation? *J R Coll Physicians* 2012;42 Suppl 18:5-22.

19. Macaskill P, Gatsonis C, Deeks J, et al. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010

20. Altman D. Practical Statistics for Medical Research. Chapman and Hall/CRC, 1991.

10

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

11

Table 1: Characteristics of participants included in diagnostic accuracy analyses

Total number of participants	999
Age (mean)	79.7 years (range 75.1-99.8) (95% CI 79.4-79.9)
Gender % male	49.3 (95%CI 46.2-52.5)
Co-morbidities (from medical summary) Atrial fibrillation Heart failure Hypertension Diabetes Stroke Transient ischaemic attack	110 (11%) 31 (3.1%) 533 (53.3%) 122 (12.2%) 31 (3.1%) 65 (6.5%)
Participants with AF on anti-arrhythmic medication	87(8.7%)

Table 2: Diagnostic performance of WatchBP and Omron autoanalysis for detecting atrial fibrillation

n=999* Prevalence 7.9%	WatchBP with AF indicator flash (95% CI)	Omron autoanalysis* (95%CI)
Sensitivity %	94.9 (87.5-98.6)	98.7(93.2-100)
Specificity %	89.7 (87.5-91.6)	76.2 (73.3-78.9)
Positive likelihood ratio	9.2 (7.6-11.2)	4.15 (3.69-4.67)
Negative likelihood ratio	0.057 (0.022-0.15)	0.017 (0.0024-0.12)
Positive predictive value	44.1 (36.5-51.9)	26.3 (21.3-31.7)
Negative predictive value	99.5 (98.8-99.9)	99.9 (99.2-100)

*text message 'irregular 'or 'analysis impossible' counted as positive test

12

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 3: Diagnostic performance of Omron and Merlin single-lead ECGs for detecting atrial fibrillation

n=999* Prevalence 7.9%	Interpretation	Omron single-lead ECG (95% CI)	Merlin single-lead ECG (95% CI)
Sensitivity %	*Cardiologist: 1*	95 (82-99)	97 (85-100)
	2	95 (82-99)	91 (77-98)
	3	90 (77-97)	98 (87-100)
	4	98 (87-100)	88 (74-96)
Specificity %	*Cardiologist: 1*	75 (71-79)	76 (71-80)
	2	98 (96-99)	95 (93-97)
	3	97 (96-99)	92 (89-94)
	4	95 (92-97)	91 (88-94)
Positive likelihood ratio	*Cardiologist: 1*	3.8 (3.2-4.6)	3.99 (3.3-4.8)
	2	41.8 (22.5-77.5)	18.8 (12.2-28.8)
	3	35.8 (20.3-63.1)	11.5 (8.5-15.6)
	4	18.5 (12.6-27.2)	10.1 (7.4-13.9)
Negative likelihood ratio	*Cardiologist: 1*	0.07 (0.02-0.28)	0.04 (0.01-0.26)
	2	0.05 (0.01-0.21)	0.09 (0.03-0.27)
	3	0.098 (0.04-0.25)	0.03 (0.004-0.18)
	4	0.02 (0.004-0.17)	0.13 (0.06-0.3)
Positive predictive value	*Cardiologist: 1*	24.1 (17.4-31.9)	24.5 (17.6-32.5)
	2	77.8 (62.9-88.8)	60.4 (46.0-73.5)
	3	76.0 (61.8-86.9)	50.6 (39.3-61.9)
	4	62.1 (49.3-73.8)	47.4 (36.0-59.1)
Negative predictive value	*Cardiologist: 1*	99.4 (97.9-99.9)	99.7 (98.3-100)
	2	99.5 (98.3-99.9)	99.3 (97.9-99.8)
	3	99.1 (97.8-99.8)	99.8 (98.7-100)
	4	99.8 (98.8-100)	98.9 (97.3-99.6)

13

Figure legends

Figure 1: Patient recruitment flow chart

Figure 2: Predictive value of WatchBP and Omron autoanalysis for detecting AF

Figure 3: Forest plot of diagnostic accuracy per cardiologist of Omron and Merlin single-lead ECGs

Triage tests for identifying atrial fibrillation in primary care: a diagnostic accuracy study comparing single lead ECG and modified BP monitors

Karen Kearley[1], Mary Selwood[1], Ann Van den Bruel[1], Matthew Thompson[1,2], David Mant[1], FD Richard Hobbs[1], David Fitzmaurice[3], Carl Heneghan[1].

Karen Kearley Clinical Scientist, Mary Selwood Research Nurse, Ann Van den Bruel Clinical Lecturer, Matthew Thompson Senior Clinical Researcher & Professor, David Mant Professor, Richard Hobbs Head of Primary Care Health Sciences , Carl Heneghan Clinical ProfessorReader

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, New Radcliffe House, 2nd floor, Walton Street, Jericho, OX2 6NW, UK

[2]Department of Family Medicine, University of Washington, Seattle, WA 98195-4696

[3]Primary Care Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, Fitzmaurice D Professor,

Correspondence to: Matthew.Thompson@phc.ox.ac.uk

Word Count: 2789

Numbers of tables: 3

Numbers of figures: 3

Abstract

Objectives: New electronic devices offer an opportunity within routine primary care settings for improving the detection of atrial fibrillation (AF), which is a common cardiac arrhythmia and a modifiable risk factor for stroke. We aimed to assess the performance of a modified blood pressure monitor and two single-lead ECG devices, as diagnostic triage tests for the detection of AF.

Setting: Six General Practices in the United Kingdom

Participants: 1000 ambulatory patients aged 75 years and over

Primary and secondary outcome measures: Comparative diagnostic accuracy of modified blood pressure monitor and single-lead ECG devices, compared to reference standard of 12-lead ECG, independently interpreted by cardiologists.

Results: A total of 79 participants (7.9%) had AF diagnosed by 12-lead ECG. All three devices had a high sensitivity (93.9-98.7%) and are useful for ruling out AF. WatchBP is a better triage test than Omron autoanalysis because it is more specific - 89.7% (95% CI 87.5-91.6) compared to 78.3% (95% CI 73.0-82.9) respectively. This would translate into a lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message in the study population. The overall specificity of single-lead ECGs analysed by a cardiologist was 94.6% for Omron and 90.1% for Merlin.

Conclusions: WatchBP performs better as a triage test for identifying AF in primary care than the single-lead ECG monitors as it does not require expertise for interpretation and its diagnostic performance is comparable to single-lead ECG analysis by cardiologists. It could be used opportunistically to screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement.

Trial Registration – not required.

Strengths and Limitations of this study

Strengths

- Three devices for detecting atrial fibrillation were tested on an unselected elderly primary care population of 1000 individuals.
- The prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision.
- Our population is generalisable to similar primary care settings worldwide.
- A reference standard (12 lead ECG) was performed on all patients, and interpreted blindly by cardiologists.

Limitations:

- The specificity of one cardiologist was substantially lower than the other three.
- Only 12 new cases of atrial fibrillation were detected in this setting, which is as expected for a primary care population such as this

2

Introduction

Atrial fibrillation is the most common cardiac arrhythmia; present in more than 10% of patients aged 75 years or over [1,2], and it significantly increases morbidity and mortality.[3] The main significance of atrial fibrillation (AF) is as a major independent risk factor for stroke and thromboembolism,[4] particularly in older patients. A large evidence base supports the efficacy of oral anticoagulation in reducing AF stroke risk by two-thirds[5]. AF may be asymptomatic but can be identified by detecting a characteristic irregularity in pulse rhythm. Despite this, it was found to be undiagnosed in 3.8% of patients aged over 75 years in a large UK screening study[6]. Indeed, undiagnosed AF is associated with 3.8-6.1% of all strokes[7,8]. Given the importance of identifying stroke risk patients, the high prevalence of AF, and the effective prevention strategies, there is a strong case to consider screening for atrial fibrillation.

European guidelines for the management of AF recommend opportunistic pulse assessment by a primary-care practitioner with a follow-up electrocardiogram (ECG) for an irregular pulse, an approach which is more cost effective than systematic screening with a 12-lead ECG. [6,9] In primary care research studies the sensitivity of nurse pulse assessment is high, varying between studies from 87-97% although it has a lower specificity (70 to 81%) requiring one out of every three or four screened patients to have a follow-up ECG[6,10,11]. This is a significant additional workload and could be a barrier to widespread implementation of opportunistic screening. Although pulse examination is a simple screening technique for AF, its detection relies on a subjective assessment in a busy routine care setting, which may partly explain why AF remains undetected in many patients. Indeed the American Heart Association suggests there is a need to develop strategies to detect AF more effectively in individuals and populations.[12]

Several new electronic devices have the potential to be useful triage tests for AF and initial studies suggest these have higher specificity than pulse palpation, so could significantly reduce the need for confirmatory 12-lead ECGs - but none have been evaluated in a primary care settings [13,14,15]. The National Institute for Health and Clinical Excellence (NICE) has recently suggested that using a blood pressure monitor, WatchBP, which is modified to detect irregularity of pulse during BP measurement in the elderly, together with appropriate anticoagulation of those subsequently diagnosed with AF, has the potential to reduce the incidence of stroke[16]. As stroke is the third leading cause of global disease burden, this would have substantial economic and patient benefits worldwide [16,17].

Methods:

Study design

An observational study comparing the accuracy of a modified BP monitor (WatchBP), and two single-lead ECG devices, one with an auto-analysis function (Omron) and one without (Merlin), with the reference standard of blinded cardiologist diagnosis of atrial fibrillation from a 12-lead ECG.

Participants

We recruited participants aged 75 years or over, living at home, from six general practices in the UK between May 2011 and October 2012. Having excluded patients with implanted pacemakers or defibrillators, those unable to give informed consent, or patients in whom the GP considered participation was inappropriate (e.g. terminal illness), all other patients over 75 were invited to take part until the sample size had been achieved (i.e. 1000 participants had been recruited). A diagnosis of AF recorded in the patient GP record was not used as part of the selection criteria. Consenting study participants identified from GP records and meeting the eligibility criteria attended a 30-minute appointment with a practice nurse for the testing procedure.

3

Test procedure

All participants received the tests in the same order, which were delivered by nine registered nurses working within the practices.

Firstly, the nurse used WatchBP (Microlife, Switzerland) a modified oscillometric BP monitor which flashes when it detects an irregular pulse during automatic BP measurement.

Secondly, the nurse then applied an Omron monitor (model HCG-801, Omron Healthcare Europe, The Netherlands) which involved placing one electrode on the bare chest wall 5 cm below the nipple, whilst the patient held the other electrode with the right index finger. The monitor records a single-lead ECG tracing, and displays a message indicating the presence of possible AF. The device's analysis algorithm includes several cardiac rhythms which could potentially be AF, including fast and irregular, slow and irregular, irregular and those where analysis impossible. The single-lead recording and text message were recorded and saved for later downloading and analysis.

Thirdly, the nurse applied a Merlin ECG event recorder (Meditech Ltd, Hungary) which resembles a watch, on participants' left wrist. The participant covered the electrode on the face of the device with the palm of their right hand for 30 seconds. The recording, with no automated analysis, was saved to a computer for later downloading and analysis. Unlike the Omron, the Merlin monitor does not require removal of any clothing, making it possible for use in public settings, an advantage for participants experiencing an intermittent arrhythmia.

The nurse recorded the results of the WatchBP monitor and the Omron automated text message during the initial examination. Each single-lead ECG trace was sent for interpretation to two independent cardiologists after removing all clinical information and patient identification except for date of birth and the text message (Omron only).

Participants were asked if they had experienced palpitations, racing, pounding, fluttering or irregular heartbeat in the previous 4 weeks.

The reference standard

A 12-lead ECG was performed on all participants at the end of the visit, using standard procedures. ECGs were independently interpreted by a panel of two cardiologists, blind to all patient identifiers, clinical details, results of triage tests, but not blind to study objectives. They were asked to classify the ECG as atrial fibrillation or atrial flutter (yes/no/don't know). A third cardiologist, also blind to any of the other study results, interpreted the ECG readings for which there was no consensus or if both cardiologists were uncertain about the presence of AF. AF is defined as the absence of distinct 'p' waves, an absolutely irregular RR interval and an atrial cycle length <200ms (300bpm) on a 12-lead ECG[17].

To cope with the workload, the task was divided between two separate pairs of cardiologists. All had completed cardiology specialist training of between five and six years. The third cardiologist, whose role was to resolve uncertainty and moderate disagreement, was a specialist electro-physiologist.

Study power and analysis

We anticipated that the overall prevalence of AF, including both new and prior diagnoses, would be 10%. We therefore estimated that 1000 participants would be needed to provide sufficient precision to estimate the sensitivity of each device with a maximum 95% CI of +/-10%.

We compared the accuracy of the WatchBP, Omron (text and tracing) and Merlin monitors, with a 12-lead ECG for detection of AF. All inconclusive results were treated as positive in our analysis

4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(because in the clinical situation these participants could not have AF ruled out and would need to have a 12-lead ECG). We calculated sensitivity, specificity, likelihood ratios and 95% confidence intervals for each test. Analyses for both single-lead ECGs were done per cardiologist and the results were subsequently combined using a random effects bivariate model calculating overall sensitivity and specificity[19]. We analysed the inter-observer agreement of whether the single lead ECG trace showed AF between the cardiologists within each pair together with the kappa value [20].

We performed subset analyses of the comparative accuracy in patients with and without existing diagnoses of AF (whether or not treated) recorded in GP records. In addition, we modelled a 2-stage screening process for AF by examining patients who were WatchBP flash positive and analysing the sensitivity and specificity of the Omron autoanalysis in this subgroup.

Data analysis was performed using Stata version 11S

5

Results

Of 6529 potentially eligible patients, we invited 2673 for a screening visit; no further invitations were sent after the required 1000 sample had been achieved (Figure 1). Participants had an average age of 79.7 years (range 75.1 to 99.8 years), 49.3% were male (Table 1). 110 had AF recorded in the medical summary (11.0%), of whom 67 (67/110, 61%) had evidence of AF on their 12-lead ECG at the study visit. AF was noted in a further 12 participants based on the 12-lead ECG (1.2%). The three cardiologists could not reach a consensus about whether AF was present in the 12-lead ECG in one participant, who was therefore excluded from further analyses which are therefore based on 999 participants.

Overall, 50 (5.0%) participants reported having experienced palpitations, racing, pounding, fluttering or irregular heartbeat during the previous four weeks. Of these, 16 were already known, and three were new cases of AF. The inter-observer agreement for the 12-lead ECGs was good for both pairs of cardiologists; the kappa was 0.69 and 0.96 respectively for pair 1 and pair 2.

Accuracy of WatchBP and Omron autoanalysis

Both WatchBP and Omron autoanalysis were highly sensitive (94.9% and 98.7%, respectively) (Table 2). The specificity of WatchBP was 89.7 % (95% CI 87.5-91.6) whereas the Omron text message was lower at 76.2% (95%CI 73.3-78.9).

The higher specificity of WatchBP would translate into a markedly lower follow-on ECG rate of 17% to rule in/rule out AF compared to 29.7% with the Omron text message. The superior sensitivity of the Omron text message means the false negative rate is lower than with Watch BP (0.1% vs 0.4%).

The prevalence of AF was 1.4% (11) in the subset of 889 participants who did not have AF recorded on their medical summary. WatchBP and the Omron autoanalysis provided very similar sensitivity and specificity to the total study population and this translated to a follow-on ECG rate of 11% for WatchBP and 24% for the Omron text message. Of these follow on ECGs, 12% would be true positives for the WatchBP flash compared to 5% for the Omron text message.

We modelled a two stage screening process within our total sample, where patients who had a positive result with WatchBP went on to have the Omron autoanalysis. In participants who screened positive using the WatchBP, the sensitivity and specificity of the Omron autoanalysis was 100% (95%CI 95.2-100) and 35.8% (95% CI 26.2-46.3) respectively. This would slightly reduce the follow-on 12-lead ECG requirement from 17% to 13.6%.

Accuracy of Omron and Merlin single-lead ECGs interpreted by a cardiologist

The diagnostic accuracy of the Omron and Merlin single-lead ECGs is shown in figure 3; meta-analysis of the four cardiologist's interpretations gave Omron a summary sensitivity of 94.4% with a summary specificity of 94.6%; Merlin had a summary sensitivity of 93.9% with a summary specificity of 90.1%. Cardiologist variability in single-lead ECG analysis is shown in Table 3. There was fair interobserver agreement (kappa 0.36 and 0.28 for Omron and Merlin respectively) between cardiologist pair 1, and good interobserver agreement (kappa 0.76 for both devices) between the second pair. The cardiologists were unable to interpret whether the trace showed AF on 0-1.2% of patients using Omron and 0-7.2% of occasions using Merlin.

6

Discussion

Summary of main results

Our results show that all the methods that we tested have a high sensitivity for detecting AF in an elderly primary care population and thus are useful for ruling out AF. During the study we detected 12 new cases of AF and 7.9% of the study population overall were experiencing AF at the time of the study consultation. The total prevalence in the study sample was 11%, as some patients have paroxysmal AF and were not experiencing arrhythmias at the time of the study.

WatchBP outperformed Omron autoanalysis because it is more specific – 89.7% compared to 76.2% and 78.3% respectively. The specificity of the single-lead ECGs when interpreted by a cardiologist were comparable with WatchBP. Overall Omron produced fewer inconclusive results and achieved slightly better specificity than Merlin. However cardiologist interpretation of the ECG trace was a more powerful factor in determining specificity than the type of monitor used. In this study, operator variability in ECG analysis was partly due to differences in expertise, and partly due to trade-offs made by different cardiologists between sensitivity and specificity.

In our elderly population, only 5% of patients reported experiencing palpitations. Amongst patients who did report palpitations, prevalence of AF was 36%. Most of these patients already had a diagnosis of AF.

Strengths and weakness

The devices were tested on an unselected elderly primary care population. Although our sample did not include the housebound or patients with dementia, we recruited a large population with a prevalence of AF expected for this setting, which allowed us to determine the operating characteristics of each monitor with precision. We believe our population is generalisable to similar primary care settings worldwide.

The cardiologists analysed ECG recordings in line with recommended practice and all had a minimum of 5 years cardiology specialist training. The specificity of one cardiologist was substantially lower than the other three. Further research with a larger sample of cardiologists would be required to more fully interpret this finding. However, we do not believe this diminishes the reliability of our reference standard since the second pair of blinded cardiologists achieved extremely high concordance, and the moderating consultant cardiologist was a specialist electrophysiologist. Having cardiologists rather than general practitioners ascertain the presence or absence of AF on the Omron or Merlin single lead ECGs may have risked over estimating accuracy given the likelihood that cardiologists would have greater skill in ECG interpretation. Finally, given that the devices were used sequentially followed by the 12 lead ECG it is theoretically possible that AF may not have been consistently present or absent during the entire procedure, leading to an increase of false positives and false negatives, however we believe this is a small risk given the extremely short time frame of approximately 10 minutes.

Comparison with existing literature

The operating characteristics of WatchBP in this elderly primary care population are comparable to previous studies undertaken in different populations. In a US study of 405 patients attending a cardiology outpatients where 39% of the sample had an abnormal ECG (AF prevalence 23%) it had a sensitivity of 95% (95% CI 93 -98) and specificity of 86% (95% CI 84-89) while in a Greek study of a diverse group of 73 individuals (healthy volunteers, outpatients, and inpatients) it had similar sensitivity of 93% (95% CI 74-99) and specificity of 89% (95% CI 79-96).

7

Omron has been evaluated in a German study of 505 patients in a hospital setting where 66% of the sample had an abnormal ECG (mean age 61 years, 66% male, AF prevalence 28%) with all analyses (including the reference standard) being undertaken by a single cardiologist. This study reported higher accuracy than we demonstrated, with a sensitivity of 99% (95%CI 96-100) and specificity of 96% (95% CI 94-98) for AF. We are unaware of any AF validation studies for Merlin.

Clinical implications

Our study supports the recent NICE recommendation that GPs should consider using WatchBP for elderly patients who have a BP measurement in order to opportunistically identify undiagnosed AF. Specificity is comparable with a single-lead ECG analysed by a cardiologist and the result does not require specialist interpretation, which is an advantage over the single-lead ECG monitors and keeps screening straightforward and costs low. It is easily portable for use in housebound patients.

Prior to wider implementation, clinicians need to consider the optimal screening frequency. Assuming a primary care population in which the majority of AF cases had already been identified, the WatchBP would indicate a 'positive flash' for 1 in 10 patients who would thus require a follow-up 12-lead ECG. As some patients have their blood pressure measured frequently, there is potential for recurrent false-positive results and a high demand for 12-lead ECGs. These issues should be addressed to ensure sustainable implementation in primary care.

Paroxysmal AF presents a particular diagnostic challenge because patients may not be experiencing an arrhythmia when they are being screened with one-off measures. The potential for patients to undertake self-monitoring and to capture an ECG tracing when an arrhythmia is occurring are desirable features of an AF screening method. In our study, only 5% of patients reported experiencing palpitations, which would limit the usefulness of asking patients to record a single-lead ECG at the time of symptoms. Amongst the patients who did report palpitations, the prevalence of AF was 36%, but most of these patients already had a diagnosis of AF which might have improved their awareness of an arrhythmia.

WatchBP will not identify atrial flutter, which is also a risk factor for stroke. The prevalence of this is much lower than AF; in our study it was 0.4%.

Future research

Screening for AF is not yet routine practice in primary care, therefore determining not only the optimal device, but the most appropriate and cost effective process is a priority. This could include an impact analysis using WatchBP as a triage test to evaluate its effectiveness (and cost effectiveness) for improving the detection of AF and prevention of stroke. Determining the cost effectiveness of these devices would also have to include doctor time for interpretation of ECGs

Conclusion

WatchBP performs better than the single-lead ECG monitors as a triage test for identifying AF in primary care as it does not require any expertise for interpretation and their diagnostic performance is comparable. It could be used to opportunistically screen elderly patients for undiagnosed AF at regular intervals and/or during BP measurement. It has advantages over pulse assessment because the result is objective and it also has a higher specificity, keeping follow-up 12-lead ECGs to a minimum.

Contribution of authors

8

KK, MS, DM, RH, DF, CH initiated the study and designed the protocol. KK, MS, AV, MT, CH were involved in the conduct and analysis of the trial. MS oversaw data management and monitoring of protocol adherence throughout the trial. KK, AV, MT co-wrote the manuscript. MS, DM, CH, RH, DF contributed to writing the manuscript.

Full ethical approval was granted of the study by NRES Committee South Central – Oxford A.

Competing interests

We received subsidised monitors from Microlife, Omron and Merlin for the submitted work; FDRH has received free or subsidised BP equipment from Omron and Microlife for other research and there are no other relationships or activities that could appear to have influenced the submitted work .

Copyright

Crown copyright is claimed for this publication, in accordance with NIHR guidelines.

References

1. NICE. Atrial Fibrillation: National clinical guideline for management in primary and secondary care. 2006.
2. Fitzmaurice DA, Hobbs FDR, Jowett S, Mant J, Murray E, Holder R et al. Screening versus routine practice for detection of atrial fibrillation in people aged 65 or over: cluster randomised controlled trial. *BMJ* 2007; 335:383-6.
3. Miyasaka Y, Barnes ME, Bailey KR, Cha SS, Gersh BJ, Seward JB, Tsang TS. Mortality trends in patients diagnosed with first atrial fibrillation – a 21-year community study. *J Am Coll Cardiol* 2007; 49:986-92.
4. Wolf PA, Abbott RD, Kannell WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. *Stroke* 1991 Aug; 22(8):983-8.
5. Aguilar MI, Hart R. Oral anticoagulants for preventing stroke in patients with non-valvular atrial fibrillation and no previous history of stroke or transient ischemic attacks. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Edited (no change to conclusions), published in Issue 1, 2009
6. Hobbs F, Fitzmaurice DA, Mant J, Murray E, 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 Oct; 9(40): 1-74.
7. Lin HJ, Wolf PA, Benjamin EJ, Belanger AJ, D'Agostino RB. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. *Stroke* 1995;26:152730.
8. Liao J, Khalid Z, Scallan C, Morillo C, O'Donnell M. Noninvasive cardiac monitoring for detecting paroxysmal atrial fibrillation or flutter after acute ischemic stroke: a systematic review. *Stroke* 2007;38:
9. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. An update of the 2010 ESC Guidelines for the management of atrial fibrillation. *European Heart Journal* (2012) 33, 2719–2747
10. Sudlow M, Rodgers H, Kenny RA Thomson R. Identification of patients with atrial fibrillation in general practice: a study of screening methods. *British Medical Journal* 1998;317:327

9

11. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. *British Journal of General Practice*. 2002;52:373–380

12. Estes, N, Sacco R, Al-Khatib S, Ellinor P, Bezanson J, Alonso A, et al. American Heart Association Atrial Fibrillation Research Summit. Circulation. 2011 July 19; 124(3): 363–372

13. Wiesel J, Fitzig L, Herschman Y, Messineo F. Detection of atrial fibrillation using a modified microlife blood pressure monitor. *Am J Hypertens 2009*; 22(8): 848-852.

14. Stergiou G, Karpettas N, Protogerou A Nasothimiou EG, Kyriakidis M. Diagnostic accuracy of a home blood pressure monitor to detect atrial fibrillation. *Journal of Human Hypertension* (2009) 23, 654–65

15. Kaleschke G, HoffmanB Drewitz I Steinbeck G, Naebauer M, Goette A, et al. Prospective, multicentre validation of a simple, patient-operated electrocardiographic system for the detection of arrhythmias and electrocardiographic changes. *Europace* (2009) 11 (10): 1362-1368

16. NICE. WatchBP Home A for opportunistically detecting atrial fibrillation during diagnosis and monitoring of hypertension, 2013

17. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2095

18. Harris K, Edwards D, Mant J. How can we best detect atrial fibrillation? *J R Coll Physicians* 2012;42 Suppl 18:5-22.

19. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010

20. Altman D. Practical Statistics for Medical Research. Chapman and Hall/CRC, 1991.

10

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Characteristics of participants included in diagnostic accuracy analyses

Total number of participants	999
Age (mean)	79.7 years (range 75.1-99.8) (95% CI 79.4-79.9)
Gender % male	49.3 (95%CI 46.2-52.5)
Co-morbidities (from medical summary) Atrial fibrillation Heart failure Hypertension Diabetes Stroke Transient ischaemic attack	 110 (11%) 31 (3.1%) 533 (53.3%) 122 (12.2%) 31 (3.1%) 65 (6.5%)
Participants with AF on anti-arrhythmic medication	87(8.7%)

Table 2: Diagnostic performance of WatchBP and Omron autoanalysis for detecting atrial fibrillation

n=999* Prevalence 7.9%	WatchBP with AF indicator flash (95% CI)	Omron autoanalysis* (95%CI)
Sensitivity %	94.9 (87.5-98.6)	98.7(93.2-100)
Specificity %	89.7 (87.5-91.6)	76.2 (73.3-78.9)
Positive likelihood ratio	9.2 (7.6-11.2)	4.15 (3.69-4.67)
Negative likelihood ratio	0.057 (0.022-0.15)	0.017 (0.0024-0.12)
Positive predictive value	44.1 (36.5-51.9)	26.3 (21.3-31.7)
Negative predictive value	99.5 (98.8-99.9)	99.9 (99.2-100)

*text message 'irregular 'or 'analysis impossible' counted as positive test

11

Table 3: Diagnostic performance of Omron and Merlin single-lead ECGs for detecting atrial fibrillation

n=999* Prevalence 7.9%	Interpretation	Omron single-lead ECG (95% CI)	Merlin single-lead ECG (95% CI)
Sensitivity %	*Cardiologist: 1*	95 (82-99)	97 (85-100)
	2	95 (82-99)	91 (77-98)
	3	90 (77-97)	98 (87-100)
	4	98 (87-100)	88 (74-96)
Specificity %	*Cardiologist: 1*	75 (71-79)	76 (71-80)
	2	98 (96-99)	95 (93-97)
	3	97 (96-99)	92 (89-94)
	4	95 (92-97)	91 (88-94)
Positive likelihood ratio	*Cardiologist: 1*	3.8 (3.2-4.6)	3.99 (3.3-4.8)
	2	41.8 (22.5-77.5)	18.8 (12.2-28.8)
	3	35.8 (20.3-63.1)	11.5 (8.5-15.6)
	4	18.5 (12.6-27.2)	10.1 (7.4-13.9)
Negative likelihood ratio	*Cardiologist: 1*	0.07 (0.02-0.28)	0.04 (0.01-0.26)
	2	0.05 (0.01-0.21)	0.09 (0.03-0.27)
	3	0.098 (0.04-0.25)	0.03 (0.004-0.18)
	4	0.02 (0.004-0.17)	0.13 (0.06-0.3)
Positive predictive value	*Cardiologist: 1*	24.1 (17.4-31.9)	24.5 (17.6-32.5)
	2	77.8 (62.9-88.8)	60.4 (46.0-73.5)
	3	76.0 (61.8-86.9)	50.6 (39.3-61.9)
	4	62.1 (49.3-73.8)	47.4 (36.0-59.1)
Negative predictive value	*Cardiologist: 1*	99.4 (97.9-99.9)	99.7 (98.3-100)
	2	99.5 (98.3-99.9)	99.3 (97.9-99.8)
	3	99.1 (97.8-99.8)	99.8 (98.7-100)
	4	99.8 (98.8-100)	98.9 (97.3-99.6)

12

1
2
3   Figure legends
4
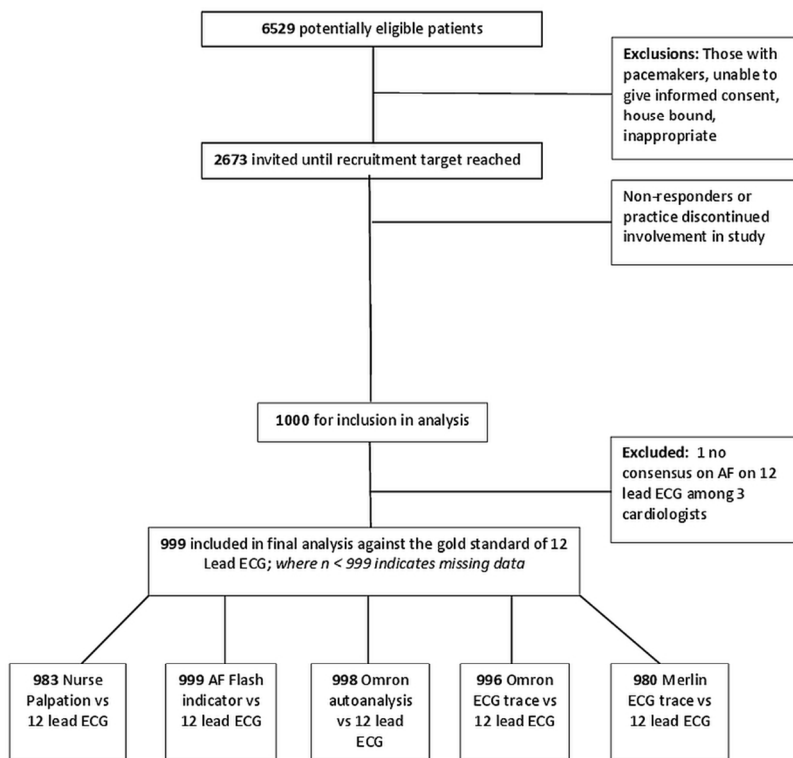5   Figure 1: Patient recruitment flow chart
6
7   Figure 2: Predictive value of WatchBP and Omron autoanalysis for detecting AF
8
9
10   Figure 3: Forest plot of diagnostic accuracy per cardiologist of Omron and Merlin single-lead
11   ECGs
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
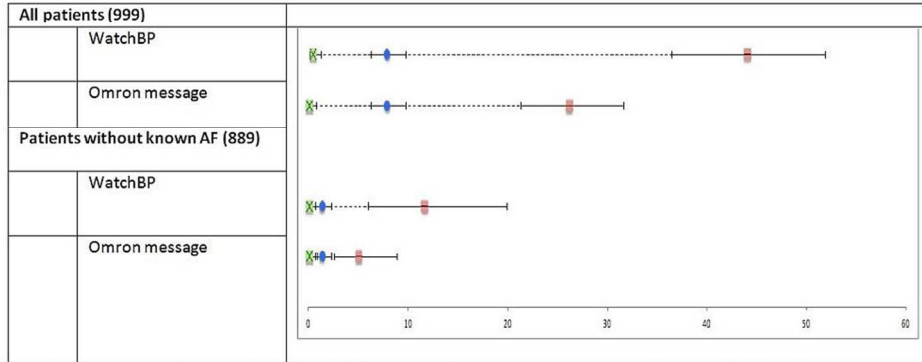60   13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1: Patient recruitment flow chart

90x116mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

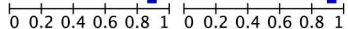Figure 2: Predictive value of WatchBP and Omron autoanalysis for detecting AF

Blue dots = prevalence; red dots: probability of AF after a positive test; green dots: probability of AF after a negative test (each with their 95% CI)

90x47mm (300 x 300 DPI)

OMRON

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cardiologist 1	35	110	2	332	0.95 [0.82, 0.99]	0.75 [0.71, 0.79]		
Cardiologist 2	35	10	2	432	0.95 [0.82, 0.99]	0.98 [0.96, 0.99]		
Cardiologist 3	38	12	4	463	0.90 [0.77, 0.97]	0.97 [0.96, 0.99]		
Cardiologist 4	41	25	1	450	0.98 [0.87, 1.00]	0.95 [0.92, 0.97]		

MERLIN

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cardiologist 1	34	105	1	326	0.97 [0.85, 1.00]	0.76 [0.71, 0.80]		
Cardiologist 2	32	21	3	410	0.91 [0.77, 0.98]	0.95 [0.93, 0.97]		
Cardiologist 3	41	40	1	432	0.98 [0.87, 1.00]	0.92 [0.89, 0.94]		
Cardiologist 4	37	41	5	431	0.88 [0.74, 0.96]	0.91 [0.88, 0.94]		

195x80mm (300 x 300 DPI)