PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Diagnostic Performance of an Automatic Blood Pressure Measurement device, Microlife Afib, for Atrial Fibrillation Screening in a Real World Primary Care Setting
AUTHORS	Chan, Pak Hei; Wong, Chun Ka; Pun, Louise; Wong, Yu Fai; Wong, Michelle Man-Ying; Chu, Daniel Wai-Sing; Wah Siu, Chung

VERSION 1 - REVIEW

REVIEWER	Natale Daniele Brunetti University of Foggia, Italy
REVIEW RETURNED	28-Aug-2016

GENERAL COMMENTS	We read with great interest this paper reporting on Diagnostic Performance of an Automatic Blood Pressure Measurement device, Microlife Afib, for Atrial Fibrillation Screening in a Real World Primary Care Setting. The study is well managed and resuts of interest. However some point should be addressed to improve the overall quality of the paper. Authors claim that the device has a high sensitivity (80%) for the detection of atrial fibrillation. However, the paper seems to ignore the much better performance in terms of accuracy (approximatively 98%) and negative predictive power. Please rewrite conclusion, discussion and strength of the study not focusing on sensitivity but on accuracy and negative predictive power. The optimal negative predictive power is particular useful for a screening tool. Figures are too many and partly redundant. Classifcation by age is not relevant given non significance of differences: please delete figure 3B,C,D, and Figure 4.

REVIEWER	Carlo Gandolfo Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics and Maternal Infantile Health, University of Genoa, Italy
REVIEW RETURNED	02-Nov-2016

GENERAL COMMENTS	Interesting and well conducted study,

REVIEWER	Martin Dawes University of British Columbia Canada
REVIEW RETURNED	17-Nov-2016

GENERAL COMMENTS	BMJ Review bmjopen-2016-013685 Diagnostic Performance of an Automatic Blood Pressure

	Measurement device, Microlife Afib, for Atrial Fibrillation Screening in a Real World Primary Care Setting Introduction – precise, well written, describes the problem of AF, (epidemic is a little strong), and the difficulty in detection, the trials that have been done, their small size and their specialty populations. At this stage I would expect a brief description of the trials in primary care, maybe the prevalence (given the low prevalence of your setting) and their weaknesses in using them to argue for the use of this technology Good pragmatic set of patients – but you should have described diabetes and age as risk factors for AF in the introduction. I would expect maybe the OR for >65 and Diabetes in the intro, along with hypertension. The fact that 14% of them are hypertensive is not the same as the risk of AF if you are hypertensive. The clinic looks as if it is just for hypertension These patients may be very different form family practice and so this is really an outpatient population In many countries, this would be regarded as secondary care. You need t look at those who did not consent and include them in your diagram - small numbers but should be done to make sure they are similar to those who entered trial, if possible The test had a sensitivity of just over 80% - this means it is not superb at ruling out AF if it is negative The test had a terrific specificity – so when positive it was very very likely the patient had AF Statistics are not my strong point but I think AUC is better used when you have varying thresholds of the test – and when you are comparing different methodologies (manual detection of the pulse). I was not sure you are using the data accurately in your description when you describe accuracy. I have requested a statistical opinion The data could have presented in a table – rather than text. I also think a lot of your tables /figures could be reformatted to display the data far more succinctly I was not sure why you compared it within the age groups – and did not look at DM presence or absence or level of BP, or drugs etc. The first paragraph of your discussion is very strong. The second paragraph is repeating your introduction. You then go onto describe the devices – but you have strong data. What I would like is your detection rate compared with studies that have looked at clinical judgement (the standard now) versus this tool. Certainly, compare the strengths of this tool – but use the sensitivity and specificity and discuss the populations. Overall this feels very secondary care to me and that is worrying. To balance this impression the primary care detection rates may well address that feeling. Finally I think you need to look at why the tool missed 1 in five patients. I may have missed this in the multiple figures but I would want to know the characteristics of the patients it missed.

REVIEWER	Martin Myers Sunnybrook Health Sciences Centre Canada
REVIEW RETURNED	28-Nov-2016

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer 1:

Comment 1: Authors claim that the device has a high sensitivity (80%) for the detection of atrial fibrillation. However, the paper seems to ignore the much better performance in terms of accuracy (approximatively 98%) and negative predictive power. Please rewrite conclusion, discussion and strength of the study not focusing on sensitivity but on accuracy and negative predictive power. The optimal negative predictive power is particular useful for a screening tool.

Answer: We thank the reviewer for the suggestion. The manuscript is revised accordingly to emphasise the importance of negative predictive power.

Comment 2: Figures are too many and partly redundant. Classifcation by age is not relevant given non significance of differences: please delete figure 3B,C,D, and Figure 4.

Answer: We thank the reviewer for the suggestion and have deleted the figures accordingly.

Response to Reviewer 3:

Comment 1: Introduction ..... At this stage I would expect a brief description of the trials in primary care, maybe the prevalence (given the low prevalence of your setting) and their weaknesses in using them to argue for the use of this technology.

Answer: We thank the reviewer for the suggestion. There is lack of study of AF screening in the real-world primary care setting and the epidemiology in other countries is not based on AF screening. We have modified the introduction to put more emphasis to explain the use of this technology.

Comment 2: Good pragmatic set of patients – but you should have described diabetes and age as risk factors for AF in the introduction. I would expect maybe the OR for >65 and Diabetes in the intro, along with hypertension.

Answer: We thank the reviewer for the suggestion and modified the introduction accordingly.

Comment 3: These patients may be very different form family practice and so this is really an outpatient population
In many countries, this would be regarded as secondary care. You need t look at those who did not consent and include them in your diagram - small numbers but should be done to make sure they are similar to those who entered trial, if possible.

Answer: We thank the reviewer for the opinion. The clinics were primary care clinic looking after patients with chronic diseases including hypertension, diabetes, and other problems such as

osteoarthritis. Overall, there are only less than 2% of patients who did not consent to the study. Unfortunately due to their refusal for study, their baseline demographics were not recorded but presumably similar to those who consented for study.

Comment: The data could have presented in a table – rather than text. I also think a lot of your tables /figures could be reformatted to display the data far more succinctly. I was not sure why you compared it within the age groups – and did not look at DM presence or absence or level of BP, or drugs etc.

Answer: We thank the reviewer for the suggestion. And also from suggestion of reviewer 1, we have deleted figure 3B, 3C, 3D and 4. We did look at the difference between presence or absence of DM, however, since there is no difference therefore we altogether skip presenting these sub-group data.

Comment: The first paragraph of your discussion is very strong. The second paragraph is repeating your introduction. You then go onto describe the devices – but you have strong data. What I would like is your detection rate compared with studies that have looked at clinical judgement (the standard now) versus this tool. Certainly, compare the strengths of this tool – but use the sensitivity and specificity and discuss the populations.

Overall this feels very secondary care to me and that is worrying. To balance this impression the primary care detection rates may well address that feeling.

Answer: We thank the reviewer for the comment and modified the discussion accordingly.

Comment: Finally I think you need to look at why the tool missed 1 in five patients. I may have missed this in the multiple figures but I would want to know the characteristics of the patients it missed.

Answer: We thank the reviewer for the inquiry. In fact, as a screening tool, a sensitivity of 1 in 5 is still regarded as reasonable. The sensitivity is largely due to the underlying algorithm that picks up beat-to-beat variation and had been mentioned in some previous studies. This has been emphasised in the discussion part in the revised manuscript.

VERSION 2 – REVIEW

REVIEWER	Natale Daniele Brunetti University of Foggia, Italy
REVIEW RETURNED	10-Jan-2017

GENERAL COMMENTS	Authors significantly improved the overall quality of the paper by addressing all reviewers observations.

REVIEWER	Martin Dawes University of British Columbia Canada
REVIEW RETURNED	04-Jan-2017

GENERAL COMMENTS	This is much clearer. great revision and I would accept with one proviso. I still think you need to be more explicit about the test missing 14 of the 72 patients with AF. We dont want clinicians thinking a negative result rules out AF. Your 95% CI had a lower rage of 69.5 so missing three out of ten possibly. You need to be very explicit about this

	finding. The tool is great, it really helps identify patients with AF in primary care if the result is positive, but if clinicians think a negative result implies no AF this will lead to harm. I think if you leave your conclusion as it is you will be having to respond to a lot of doctors identifying the sensitivity as not being "high" enough to be a rule out test.

VERSION 2 – AUTHOR RESPONSE

Thank you for considering our manuscript for publication in BMJ Open. We would also like to thank you for the time and efforts in reviewing our work.

We have addressed the reviewer's comments. Thank you the opportunity to revise our work. The comments have helped us to improve our manuscript further.

In the revised manuscript, a paragraph emphasising the sensitivity of the device is added, written as "Of note, the sensitivity of the device in this study is 80.6% (95% CI: 69.5-88.9), which means there is probability that 2 to 3 out of 10 patients with underlying AF could be missed by this screening tool. Physicians utilizing this device to screen for AF should be well aware of this potential drawback and should not solely rely on this device and hence producing a false sense of security. The possible ways to improve the sensitivity include repeated measurements with Microlife device at intervals and combining the use of other screening tools in AF detection." Such change should give the physicians caution about the use of Microlife device and warn against a false sense of security.

We hope that the change show that this paper makes significant new publishable contributions that will be of great interest to the readers of BMJ Open. Please do not hesitate to contact us if we can provide further clarifications or helpful edits.

VERSION 3 – REVIEW

REVIEWER	Martin Dawes University of British Columbia Canada
REVIEW RETURNED	21-Jan-2017

GENERAL COMMENTS	Thank you for entering that new paragraph. This is an important paper that will prompt debate about recognition of AF, as well as the appropriate tools for BP measurement. That two for one approach sounds very attractive.