PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prospective Register Of patients undergoing repeated OFfice and Ambulatory Blood Pressure Monitoring (PROOF-ABPM): protocol for an observational cohort study
AUTHORS	Sheppard, James; Martin, Una; Gill, Paramjit; Stevens, Richard; McManus, Richard

VERSION 1 - REVIEW

REVIEWER	Peter Lacy Institute of Cardiovascular Science, University College London, Gower Street, London. WC1E 6BT UK
REVIEW RETURNED	27-May-2016

GENERAL COMMENTS	This protocol describes a multi-centre observational cohort study, compiling a registry of blood pressure and other clinical data from patients undergoing cardiovascular evaluation and/or management in a primary and secondary care setting. Data from the registry will be used to test in a real-world setting, the sensitivity and specificity of a novel tool used to predict the direction of the clinic – out of office blood pressure difference. If proven, this tool has potential to assist in avoiding the misclassification of blood pressure status using office BP measurement alone, and to avoid the necessity for ambulatory blood pressure monitoring in many people with truly elevated blood pressure. The tool may also assist in the screening of apparently normotensive people with masked hypertension. Strengths: testing of a research tool in a real-world clinical setting, specialist nature of the registry, containing data on a wide range of patients with multiple BP readings. Weaknesses: Registry is likely to miss patients with masked hypertension and truly normal blood pressure (true negative). Comments. 1. Abstract: use of the term true "mean" pressure in describing the accuracy of blood pressure measurement may cause potential confusion as some may read this to suggest "mean arterial pressure" rather than the arithmetic mean of multiple blood pressure readings. 2. Whilst the tool has great potential for utility in primary care in avoiding misclassification in diagnosis of hypertension, it is not clear what advantage the tool will confer in secondary care where many patients are already diagnosed and undergoing treatment. Is there potential for using the tool to assist in attainment of target blood pressure following treatment titration i.e. in avoiding requirement for

	further out of office blood pressure measurement, or is more research required to elucidate this? Given that 500 of the estimated 1,000 participants required may be recruited from secondary care, consideration should be given to commenting on the utility of the tool in secondary care. 3. Will data from repeated clinic/ABPM measurements be collected? This data may be available from secondary care and used to assess the sensitivity of repeated measurement (potential for desensitisation of PROOF-BP response)? 4. Sample size accuracy is based on an assumption that 24% are true negatives (normotensive on both clinic and ambulatory BP). However, it is unlikely that many normotensive participants would be referred for ambulatory blood pressure monitoring and recruited into the study. Given that the study is well powered, is this assumption likely to influence accuracy? 5. Successful validation of the model is characterised by a 10% or greater improvement in patient classification or a reduction in ABPM use of 20% or greater. It is not clear how these threshold values were determined. 6. Anonymised data will be collected by default as is common in routine clinical audits and anonymised cohort studies. As participants will also be approached to consent to provision of attributed data, will those who do not wish to consent to providing attributed data also have the opportunity to opt out of providing anonymised data? 7. Does clinic blood pressure data collection depend on the use of oscillometric devices or can centres use devices based on auscultation? If so, are sensitivity analyses planned to compare use of the prediction tool with oscillometric and auscultatory devices?

REVIEWER	Ernest Vinyoles La Mina Primary Care Centre. University of Barcelona, Spain
REVIEW RETURNED	27-May-2016

GENERAL COMMENTS	Comments This is a very interesting protocol where authors will evaluate the predictive value of PROOF-ABPM algorithm for ambulatory blood pressure phenotype classification and cardiovascular outcomes in real clinical practice. However, there are some points that should be amended to accept the article for publication. 1. There is information about how automated office blood pressure will be collected. Nevertheless it would be interesting to add information about how 24 h-ABPM will be done. Validated devices, BP reading intervals (every 20 minutes?), how daytime and nighttime periods will be adjusted, which ABPM records will be excluded, for example. Please explain how the authors will guarantee the minimum ABPM quality (the gold standard). 2. What time range between office and ABPM will be considered acceptable to include a patient? 3. It is recommended to classify patients according the 24 h-BP mean (130/80 mmHg) in order to include the nighttime period and to improve the ambulatory BP prognostic value. If not, authors should justify very well why they don't include the nighttime period (and nighttime hypertension) 4. One limitation of the study will be the reproducibility of ABPM records, and consequently the ambulatory BP phenotype

	classification (J Hypertens 2007, 25:315–320). This issue should be commented in Discussion. 5. In table 1, authors will include some additional variables, if available. Please explain when these variables will be considered available. For ex, when a blood analysis will be acceptable to be included. What will be the acceptable time range between basal visit and blood analysis or waist circumference?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

1. Abstract: use of the term true "mean" pressure in describing the accuracy of blood pressure measurement may cause potential confusion as some may read this to suggest "mean arterial pressure" rather than the arithmetic mean of multiple blood pressure readings.
We have changed the term ‘mean’ to ‘underlying’ to avoid confusion

“Most BP measurements take place in a clinic setting, but it has long been recognised that readings taken out-of-office (via home or ambulatory monitoring) estimate true underlying blood pressure more accurately”

2. Whilst the tool has great potential for utility in primary care in avoiding misclassification in diagnosis of hypertension, it is not clear what advantage the tool will confer in secondary care where many patients are already diagnosed and undergoing treatment. Is there potential for using the tool to assist in attainment of target blood pressure following treatment titration i.e. in avoiding requirement for further out of office blood pressure measurement, or is more research required to elucidate this? Given that 500 of the estimated 1,000 participants required may be recruited from secondary care, consideration should be given to commenting on the utility of the tool in secondary care.
The reviewer is correct that the PROOF-BP tool has the potential to assist in the attainment of blood pressure targets in patients following treatment titration, which might be more relevant in a secondary care setting, and potentially help exclude white coat hypertension in patients being investigated for resistant hypertension. We have taken the opportunity to clarify this point:

“Utilised as a triaging tool for ambulatory blood pressure monitoring, the PROOF-BP prediction tool permits detection of those patients with a possible white coat or masked effect on the basis of data available in a routine Primary Care clinic. The tool could also be used to rule out the white coat effect in resistant hypertensives and assist in the monitoring of blood pressure target attainment following treatment titration which might be more relevant for secondary care populations.”

3. Will data from repeated clinic/ABPM measurements be collected? This data may be available from secondary care and used to assess the sensitivity of repeated measurement (potential for desensitisation of PROOF-BP response)?
Yes – where available this will be collected. We have added the following sentence:

“Where available, data from repeated clinic/ambulatory measurements in the same patient will be collected and used to assess the sensitivity the tool used with repeated measurements.”

4. Sample size accuracy is based on an assumption that 24% are true negatives (normotensive on both clinic and ambulatory BP). However, it is unlikely that many normotensive participants would be referred for ambulatory blood pressure monitoring and recruited into the study. Given that the study is

well powered, is this assumption likely to influence accuracy?
This is a valid point – it is possible that the cohort of patients enrolled will have a lower proportion of true negatives than were seen in our original validation study. However, the sample size is based on a predicted confidence interval and the projected sample size is sufficiently large that a change in the point estimate for each outcome, or indeed a slightly lower number of patients recruited in unlikely to have an effect on accuracy. We have added the following sentence to the methods section and included a new table clarifying the impact of our assumptions on the potential accuracy of the results:

"In a population of 1000 patients, it would be possible to estimate these rates with the following 95% confidence intervals: true positive 71% (68-74%), true negative 24% (21-27%), false positive 3% (2-4%) and false negative 2% (1-3%) (table 2). If these point estimates were to differ, or indeed a slightly lower number of patients were enrolled, the impact on the accuracy of the study results would be minimal."

The new table 2 demonstrates that these point estimates could be be established with sufficient accuracy with as few as 800 patients – our rationale for recruiting up to 1000 patients is to ensure sufficient numbers of patients are enrolled to answer both primary and secondary outcomes, assuming recruitment is appropriately distributed across clinic settings and patient characteristic sub-groups.

5. Successful validation of the model is characterised by a 10% or greater improvement in patient classification or a reduction in ABPM use of 20% or greater. It is not clear how these threshold values were determined.
These thresholds values were chosen as it was decided that such an improvement would represent a clinically significant improvement in diagnostic accuracy which might be expected to change clinical practice. We have added the following to the data analysis section:

"These thresholds were chosen because they were deemed to represent a clinically significant improvement in diagnostic accuracy which might be expected to change clinical practice.

6. Anonymised data will be collected by default as is common in routine clinical audits and anonymised cohort studies. As participants will also be approached to consent to provision of attributed data, will those who do not wish to consent to providing attributed data also have the opportunity to opt out of providing anonymised data?
Yes, all patients specifically requesting to opt out will be able to do so. We have added the following:

"Individual patient consent is not required for anonymised data collection and will not be sought, as is common in routine clinical audits and anonymised observational cohort studies, although those requesting to opt out will be able to if they wish."

7. Does clinic blood pressure data collection depend on the use of oscillometric devices or can centres use devices based on auscultation? If so, are sensitivity analyses planned to compare use of the prediction tool with oscillometric and auscultatory devices?
Sites are free to use whichever device they wish – in keeping with routine practice. If a large number use auscultatory devices they a sensitivity analysis may be carried out if there is sufficient power to do so, but this has not been pre-specified in the protocol.

Reviewer 2

1. There is information about how automated office blood pressure will be collected. Nevertheless it would be interesting to add information about how 24 h-ABPM will be done. Validated devices, BP reading intervals (every 20 minutes?), how daytime and nighttime periods will be adjusted, which ABPM records will be excluded, for example. Please explain how the authors will guarantee the minimum ABPM quality (the gold standard).
We will collect information with regard to how ABPMs are done (e.g. number of readings/intervals/device used). In our original paper describing the PROOF-BP tool, we validated our prediction model using carefully collected ABPM data from previous research studies, in which the ABPM quality could be guaranteed. The aim of this study is to collect routine data which reflects real life clinical practice, therefore no restrictions on devices or measurement protocols have been made. Were such restrictions put in place, we do not think the study would reflect true clinical practice. If the treating clinician deems the record to be of sufficient quality to use in their treatment decision, those data will be included in the study.

2. What time range between office and ABPM will be considered acceptable to include a patient?
We have added the following:
“Clinic readings taken at the time of referral for ABPM or at monitor fitting will be deemed acceptable for inclusion in the study, but no limit on the period of time between the two will be specified.”

3. It is recommended to classify patients according the 24 h-BP mean (130/80 mmHg) in order to include the nighttime period and to improve the ambulatory BP prognostic value. If not, authors should justify very well why they don’t include the nighttime period (and nighttime hypertension)
Nightime readings are not routinely collected in a Primary Care setting in the UK and therefore are not a requirement for inclusion in this study. We will include a sensitivity analysis examining the accuracy of the PROOF-BP prediction tool using 24 hr blood pressure to define the home-clinic blood pressure difference, were it is collected:

“Sensitivity analyses will exam the accuracy of the PROOF-BP prediction tool using 24 hr blood pressure to define the home-clinic blood pressure difference, where it is collected.”

4. One limitation of the study will be the reproducibility of ABPM records, and consequently the ambulatory BP phenotype classification (J Hypertens 2007, 25:315–320). This issue should be commented in Discussion.
The proposed tool is primarily used for deciding whether to refer for subsequent ABPM in routine practice (where repeated ABPMs are uncommon) so reproducibility is not of primary relevance to the study question (i.e. we are asking does the tool improve accuracy compared to classification of BP in routine practice, not is the classification of BP in routine practice accurate or not). Where repeat ABPM is conducted, records will be collected and reproducibility will be examined if sufficiently powered (see above).

5. In table 1, authors will include some additional variables, if available. Please explain when these variables will be considered available. For example, when a blood analysis will be acceptable to be included. What will be the acceptable time range between basal visit and blood analysis or waist circumference?
Blood analyses are only routinely conducted in secondary care – we have added the following footnote:

*Blood analyses are only routinely conducted in Secondary Care. Only measurements taken during the baseline visit will be deemed acceptable for inclusion.

VERSION 2 – REVIEW

REVIEWER	Peter Lacy Institute of Cardiovascular Science University College London Gower Street London WC1E 6BT
REVIEW RETURNED	28-Jul-2016

GENERAL COMMENTS	The authors have addressed the concerns I raised appropriately. The manuscript outlines the protocol for the Proof BP study which should yield intriguing results. I have just a couple of minor comments. 1. The authors have adequately addressed the comment regarding potential confusion in use of the term true "mean" pressure in the abstract of the manuscript. However this issue might also be addressed in the introduction where the term true "mean" pressure could be substituted by the term "underlying" blood pressure as in the abstract of the revised manuscript. 2. In response to questioning the accuracy of the assumption that 24% of the study population might fall into the group assigned as true negatives, the authors have demonstrated that serial 5-fold reductions in the point estimate for true negatives are unlikely to impact on accuracy. However, it is not obvious from the data in the new table 2 as to why the point estimates for false negatives in this exercise may both rise and fall when the point estimate for true negatives falls, although the number of people in the false negative category is very small. 3. Page 8 under heading "statistical analysis, outcomes" line 4 "Secondary outcomes will examined", missing word "be"; page 8 under heading "Data analysis" line 1 "agreed by steering committee prior conducting any analyses", missing word "to".

REVIEWER	Ernest Vinyoles CAP La Mina. Universitat de Barcelona.
REVIEW RETURNED	13-Jul-2016

GENERAL COMMENTS	Most of my concerns were addressed

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

1. The authors have adequately addressed the comment regarding potential confusion in use of the term true "mean" pressure in the abstract of the manuscript. However this issue might also be addressed in the introduction where the term true "mean" pressure could be substituted by the term

"underlying" blood pressure as in the abstract of the revised manuscript.

This has been revised accordingly.

2. In response to questioning the accuracy of the assumption that 24% of the study population might fall into the group assigned as true negatives, the authors have demonstrated that serial 5-fold reductions in the point estimate for true negatives are unlikely to impact on accuracy. However, it is not obvious from the data in the new table 2 as to why the point estimates for false negatives in this exercise may both rise and fall when the point estimate for true negatives falls, although the number of people in the false negative category is very small.

We agree that our arbitrary choice of point estimate adjustments may not have made complete sense to the reader. We have therefore revised table 2 to depict the changes in precision of the study results which would occur in four plausible scenarios as patients are accrued to the study: a) the prevalence of normotension is (50%) lower than expected, b) it is 75% lower than expected, c) the prevalence of white coat hypertension is (50%) higher than expected, d) it is 100% higher than expected.

3. Page 8 under heading "statistical analysis, outcomes" line 4 "Secondary outcomes will examined", missing word "be"; page 8 under heading "Data analysis" line 1 "agreed by steering committee prior conducting any analyses", missing word "to".

This has been revised accordingly.