PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Point-of-Care Assessment of Platelet Reactivity in the Emergency Department May Facilitate Rapid Rule-Out of Acute Coronary Syndromes: A Prospective Cohort Pilot Feasibility Study
AUTHORS	Przyklenk, Karin; Darling, Chad; Sala Mercado, Javier; Quiroga-Castro, Walter; Tecco, Gabriel; Zelaya, Felix; Conci, Eduardo; Sala, Jose; Smith, Craig; Michelson, Alan; Whittaker, Peter; Welch, Rob

VERSION 1 - REVIEW

REVIEWER	Issam Mikati Northwwestern university US
REVIEW RETURNED	23-Sep-2013

- The reviewer completed the checklist but made no further comments

REVIEWER	Sandrine Charpentier Emergency Department Rangueil University Hospital, Toulouse, France
REVIEW RETURNED	29-Sep-2013

GENERAL COMMENTS	The main objective should be the second objective: “assessment of closure times provides a specific test with robust positive predictive value for the rule out of ACS”. The second objective should be the others. Therefore, the main outcome measures are the ACS diagnosis. In the abstract, the aims mentioned are different from thus mentioned in the introduction section. Introduction: the introduction is clear Method: In the first paragraph, I understood that the period of the enrollment was not the same for the two sites. But for each site, were the patients included consecutively? The review chart used to establish final diagnosis is not clear. Were patients categorized with only one expert at the Institudo Modelo de Cardiologia Privadio ? In the other site, why patients were reviewed by emergency physician in case of ACS negative diagnosis but cardiologist in case of ACS positive diagnosis? What’s happening in case of disagreement between both experts? The final diagnosis was the main outcome and need to be robust. On the other hand, the criteria for NSTEMI and STEMI were inaccurate. The criteria should be criteria included in the last guideline for the management of STEMI or NSTEMI.

	For example, for STEMI diagnosis the authors defined STEMI as ST elevation on the ECG and/OR cardiac catheterization findings consistent with ACS. I don't understand the OR, I think it is AND. For UA diagnosis, the diagnosis could not be sure without cardiac diagnostic testing. It's very confusing to use specificity and VPP for ACS rule out. I understood the choice of the authors but it is not usual in the studies of diagnosis accuracy. I suggest to the authors modifying the presentation of the diagnosis accuracy and using sensibility and VPN for ACS rule out. To determine if closure time provide additional predictive value, I suggest to construct 2 ROC curves for model with and without the closure time and compare the AUC of these ROC curve. The authors could also use the Net reclassification improvement. Pepe MS, Janes H, Longton G, et al. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol. 2004;159:882-890 Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added predictive ability of a new marker: From area under the roc curve to reclassification and beyond. Stat Med. 2008;27:157-172; discussion 207-112 This article reported the result of a pilot study. The aim of the study is to determine whether 1/ point of care testing is logistically feasible in an emergency setting 2/assessment of closure times provides a specific test with robust positive predictive value for the rule out of ACS 3/closure time yields additional predictive value beyond that obtained via routine clinical assessment that may assist in the identification of ACS negative vs ACS positive patient. There are some major concerns. Firstly, the main objective should be the second objective: "assessment of closure times provides a specific test with robust positive predictive value for the rule out of ACS". The second objective should be the other. Therefore, the main outcome measures are the ACS diagnosis. In the abstract, the aims mentioned are different from thus mentioned in the introduction section. Introduction: the introduction is clear Method: In the first paragraph, I understood that the period of the enrollment was not the same for the two sites. But for each site, were the patients included consecutively? The review chart used to establish final diagnosis is not clear. Were

	patients categorized with only one expert at the Institudo Modelo de Cardiologia Privadio ? Why patients were reviewed by emergency physician in case of ACS negative diagnosis but cardiologist in case of ACS positive diagnosis? What's happening in case of disagreement between both experts? The final diagnosis was the main outcome and need to be robust. On the other hand, the criteria for NSTEMI and STEMI were inaccurate. The criteria should be criteria included in the last guideline for the management of STEMI or NSTEMI. For example, for STEMI diagnosis the authors defined STEMI as ST elevation on the ECG and/OR cardiac catheterization findings consistent with ACS. I don't understand the OR, I think it is AND. For UA diagnosis, the diagnosis could not be sure without cardiac diagnostic testing. Statistical analysis It's very confusing to use specificity and VPP for ACS rule out. I understood the choice of the authors but it is not usual in the studies of diagnosis accuracy. I suggest to the authors modifying the presentation of the diagnosis accuracy and using sensibility and VPN for ACS rule out. To determine if closure time provide additional predictive value, I suggest to construct 2 ROC curves for model with and without the closure time and compare the AUC of these ROC curve. The authors could also use the Net reclassification improvement. *Pepe MS, Janes H, Longton G, et al. Limitations of the odds ratio in gauging the* *performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol.* *2004;159:882-890* *Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added* *predictive ability of a new marker: From area under the roc curve to reclassification* *and beyond. Stat Med. 2008;27:157-172; discussion 207-112* The results were clearly reported.

REVIEWER	Deborah Diercks University of California, Davis Medical Center, Emergency Medicine
REVIEW RETURNED	22-Oct-2013

GENERAL COMMENTS	Thank you for allowing me to review this manuscript. This is a prospective study with the aim of to determine if point of care platelet testing is feasible and adds diagnostic value to rule-out ACS. This is an interesting topic and the manuscript is well written. The use of platelet activation as a marker of ACS is novel and provides the convenience of a biomarker for use. The results of this study are clinically interesting. I have provided some comments below. My major concerns of the manuscript are limited to the presentation of the data and include a preference for using a diagnosis test to rule-in a disease and suggest including net reclassification index as a mechanism to describe the added value of the test. Specific comments 1. Abstract: a. Please mention if use of an antiplatelet agent is an exclusion criteria. This will be a common question and can easily be addressed up front. b. Use of the 90th percentile raises concerns regarding the validity of the results. If this device FDA approved, what are the manufacturer thresholds? c. Likelihood ratios would be more helpful instead of PPV. As PPV is dependent on the prevalence of disease in the population studied. 2. Introduction a. The investigators do a nice job differentiating the difference between their proposed measures of platelet activation over existing studies. b. The use of a measure of a rule-out device is cumbersome and not intuitive to how clinicians frame diagnostic tests. 3. Methods a. Page 9, line 3: Was the ACS negative only based on testing and the initial visit or were patients followed? What happened if an outpatient test was ordered for follow-up? b. Was a kappa assessed on the diagnosis for ACS? c. Although the investigators present their rationale why they are looking for ACS negative, the double negative makes it more cumbersome to interpret the data. They could address one of their concerns by using net reclassification index or IDI to look for the additional diagnostic value of the test. In fact, this is probably the preferred way to assess the value of a new diagnostic test that is added to standard of care as it allows the evaluation of the direction of change. d. Page 19, para 2: please justify why the 90th percentile was selected as the threshold. 4. Results a. Page, 11: Why were those with thrombotic events excluded? If this is to be used at the time of presentation, data known at presentation will have to drive the test. b. Was a sensitivity analysis done excluding the diagnosis of UA as it is a relatively subjective diagnosis. My skepticism with this diagnosis could be modified with a table on how this diagnosis was made. c. Time from symptom onset or at least time from presentation to blood draw should be presented. 5. Discussion a. As the study appears adequately powered for the regression

	analysis, I would eliminate the word pilot from the opening sentence of the discussion. b. It would have nice to have some discussion why site was such a predictor of ASC negative. c. Limitation: if the diagnosis of ACS is based on data entirely from the enrolling visit, it is possible that misclassification occurred.

VERSION 1 – AUTHOR RESPONSE

I. Reviewer: Issam Mikati

No comments provided.
Response: n/a

II. Reviewer: Sandrine Charpentier

Comment 1: “The main objective should be the second objective: “assessment of closure times provides a specific test with robust positive predictive value for the rule out of ACS”. The second objective should be the others. Therefore, the main outcome measures are the ACS diagnosis.”

Response: Our primary objective upon designing the study was, in fact, to establish the feasibility of point-of-care testing at time of presentation to the emergency department. Nonetheless, we have modified the sequence of the study objectives as requested by the Reviewer (Introduction: page 7 par 1).

Comment 2: “In the abstract, the aims mentioned are different from thus mentioned in the introduction section.”

Response: The Aims are now presented in the same sequence in the Introduction (page 7 par 1), Abstract and Article Summary.

Comment 3: “Introduction: the introduction is clear.”

Response: n/a

Comment 4: “Method: In the first paragraph, I understood that the period of the enrollment was not the same for the two sites. But for each site, were the patients included consecutively?”

Response: The text has been revised to specify that, at each site, patients were enrolled on a convenience basis (page 7 par 2).

Comment 5: “Were patients categorized with only one expert at the Institudo Modelo de Cardiologia Privadio?”

Response: We now state that, at the Instituto Modelo de Cardiologia Privado, reviews were performed by a team of cardiologists (page 9 par 1).

Comment 6: “In the other site, why patients were reviewed by emergency physician in case of ACS

negative diagnosis but cardiologist in case of ACS positive diagnosis? What's happening in case of disagreement between both experts?"

Response: The procedural differences in the determination of final diagnosis at UMASS versus the Instituto Modelo de Cardiologia Privado reflect differences in the operational procedures at the two sites. At the Instituto Modelo de Cardiologia Privado, all patient care, including emergent care at the time of hospital presentation and diagnosis of ACS-negative patients, was provided by cardiologists, and chart reviews were performed by cardiologists. In contrast, at UMASS, patients presented to the Emergency Department and were first seen by Emergency Medicine physicians; subsequent diagnoses were determined using standard criteria in consultation with a cardiologist. In case of disagreement, the final diagnosis was adjudicated by the cardiologist. These issues are clarified on page 9 par 1 of the revised text.

Comment 7: "The final diagnosis was the main outcome and need to be robust. On the other hand, the criteria for NSTEMI and STEMI were inaccurate. The criteria should be criteria included in the last guideline for the management of STEMI or NSTEMI. For example, for STEMI diagnosis the authors defined STEMI as ST elevation on the ECG and/OR cardiac catheterization findings consistent with ACS. I don't understand the OR, I think it is AND. For UA diagnosis, the diagnosis could not be sure without cardiac diagnostic testing."

Response:
(i) We apologize for the typographical error (the criteria for diagnosis of STEMI were 'ST elevation on the ECG AND cardiac catheterization findings') and specifically state on page 9 par 2 of the revised and streamlined Methods that diagnoses were made in accordance with current AHA/ACC guidelines.

(ii) As noted in the Introduction and on page 9 par 2: to simplify the presentation and for consistency with current guidelines, we have re-formatted the text and Figure 2 such that data for NSTEMI and UA patients are presented in a single, combined group.

(iii) Finally, we emphasize that our primary analyses focuses on the comparison of patients in the two main outcome groups: non-cardiac chest pain and ACS-positive (encompassing STEMI, NSTEMI and UA).

Comment 8: "It's very confusing to use specificity and VPP for ACS rule out. I understood the choice of the authors but it is not usual in the studies of diagnosis accuracy. I suggest to the authors modifying the presentation of the diagnosis accuracy and using sensibility and VPN for ACS rule out."

Response:
(i) We acknowledge that our analysis, based on 'ACS-negative' as the outcome of interest, can be confusing and may be considered unusual. However, this choice is: (a) guided by the current emphasis on development of new strategies to aid in the timely discharge of ACS-negative patients; and, more importantly, (b) mandated by the PFA closure time data. As illustrated in Figure 2: although closure times are modestly shortened in patients with ACS, this is not helpful for the diagnosis of 'ACS-positive' because of the overlap in closure times between cohorts at the lower end of the range. Rather – and as highlighted on page 12 par 4 – the two cohorts are distinguished by differences in the proportion of patients with prolonged closure times. Accordingly, as stated on page 15 par 2, the strength (and potential value) of the measurement of closure time is the incremental value that this endpoint may provide in augmenting standard ED practices for the rule-out of ACS.

(ii) In an attempt to alleviate some of the confusion, we have modified the wording throughout the

manuscript to state that our analysis focused on a diagnosis of 'non-cardiac chest pain/symptoms', rather than 'ACS-negative'.

(iii) In response to the request to present sensitivity and negative predictive value: both are provided (together with confidence intervals) in Table 2.

Comment 9: "To determine if closure time provide additional predictive value, I suggest to construct 2 ROC curves for model with and without the closure time and compare the AUC of these ROC curve. The authors could also use the Net reclassification improvement."

Response:
(i) As requested by the reviewer, ROC curves, with and without closure time incorporated in the model, have been included as Figure 3 in the revised manuscript. The curves differ significantly (p=0.009) and, as expected from the analysis of sensitivity, specificity, positive and negative predictive values (Table 2), the greatest separation seen at the lower end of the sensitivity range.

(ii) We have considered augmenting the analysis by including the 'net reclassification index' (NRI) and the 'integrated discrimination improvement' (IDI) statistic to evaluate the incremental prognostic impact that measurement of PFA closure times might have if added to standard ED practices for the rule-out of ACS. However, there is growing skepticism among respected leaders in the field of epidemiology regarding the value provided by these indices: i.e.,
"Regardless of the intuitive appeal that the NRI statistic may garner, its potential for being inflated by miscalibrated or over-fit models is a very serious concern."
"In practice, one should not use the NRI statistic, or other prediction improvement performance measures such as delta-AUC, delta-ROC(f), delta-SNB(t), or delta-Brier for that matter, to determine if there is predictive information in a novel marker Y."
[from Pepe et al, "The net reclassification index (NRI): a misleading measure of prediction improvement with miscalibrated or overfit models" in the UW Biostatistics Working Paper Series: http://biostats.bepress.com/uwbiostat/. See also Hilden & Gerds, "A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index." Statistics in Medicine, 2013, DOI: 10.1002/sim.5804,.Kerr, Pepe et al, "Net reclassification indices for evaluating risk prediction instruments: a critical review." Epidemiology 2013; Epub 14 November 2013, and others.]
As these analyses have yet to be thoroughly vetted and are currently being questioned, we elected not to include the NRI or ID statistic in our already-complicated manuscript – a decision that is noted on page 11 par 2 of the revised text.

III. Reviewer: Deborah Diercks

"Thank you for allowing me to review this manuscript. This is a prospective study with the aim of to determine if point of care platelet testing is feasible and adds diagnostic value to rule-out ACS. This is an interesting topic and the manuscript is well written. The use of platelet activation as a marker of ACS is novel and provides the convenience of a biomarker for use. The results of this study are clinically interesting. I have provided some comments below. My major concerns of the manuscript are limited to the presentation of the data and include a preference for using a diagnosis test to rule-in a disease and suggest including net reclassification index as a mechanism to describe the added value of the test."

Comment 1(a): "Abstract: Please mention if use of an antiplatelet agent is an exclusion criteria. This will be a common question and can easily be addressed up front."

Response: Use of anti-platelets was not an exclusion criterion: this is now specified in the Abstract and on page 8 par 1.

Comment 1(b): “Use of the 90th percentile raises concerns regarding the validity of the results. If this device FDA approved, what are the manufacturer thresholds?”

Response: The ‘reference range’ reported by Siemens for collagen-ADP closure times in healthy subjects (n=176) is 71-118 seconds. There are, however, no ‘thresholds’ for either healthy subjects or patient cohorts: the manufacturer’s manual recommends that “each laboratory should establish its own reference range” and, for categorization, “each laboratory establish its own cut-off based upon site population and internal procedures”. This categorization refers to the intended use of the device (that is, detection of platelet dysfunction, including aspirin-induced platelet dysfunction, von Willebrand disease, Glanzmann thrombobasthenia, etc), not platelet reactivity in ACS. Indeed, as there are no established criteria or thresholds for shortened or prolonged closure times in the study population, this is the reason why we: (i) first analyzed the data using the 90th percentile of the distribution for all patients enrolled in the study (a choice that was prospective but, we acknowledge, arbitrary), and then examined the consequences of varying the threshold to the 95th and 80th percentile (page 16 par 1, Table 5); and (ii) emphasized that the current lack of an appropriate definition of ‘prolonged’ closure time (page 16 par 1).

Comment 1(c): “Likelihood ratios would be more helpful instead of PPV. As PPV is dependent on the prevalence of disease in the population studied.”

Response: Table 2B has been revised to include the likelihood ratio (6.52) together with its 95% confidence intervals (1.61 to 26.51). This outcome (i.e., a likelihood ratio >5) is consistent with a significant increase in the likelihood of non-cardiac symptoms. It is, however, important to note that reporting the likelihood ratio provides no advantage in terms of dependency on the prevalence of disease: as stated by Brenner & Gefeller (“Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence.” Stat Med 1997;16:981-991), “in contrast to the widespread belief among clinicians that sensitivity and specificity and the likelihood ratios of diagnostic tests are independent of disease prevalence”, likelihood ratios do vary with prevalence.

Comment 2(a): “Introduction: The investigators do a nice job differentiating the difference between their proposed measures of platelet activation over existing studies.”

Response: n/a

Comment 2(b): “The use of a measure of a rule-out device is cumbersome and not intuitive to how clinicians frame diagnostic tests.”

Response: A similar comment was made by Dr. Charpentier, and we responded as follows:
(i) We acknowledge that our analysis, based on ‘ACS-negative’ as the outcome of interest, can be cumbersome and may be considered unusual. However, this choice is: (a) guided by the current emphasis on development of new strategies to aid in the timely discharge of ACS-negative patients; and, more importantly, (b) mandated by the PFA closure time data. As illustrated in Figure 2: although closure times are modestly shortened in patients with ACS, this is not helpful for the diagnosis of ‘ACS-positive’ because of the overlap in closure times between cohorts at the lower end of the range.

Rather – and as highlighted on page 12 par 4 – the two cohorts are distinguished by differences in the proportion of patients with prolonged closure times. Accordingly, as stated on page 15 par 2, the strength (and potential value) of the measurement of closure time is the incremental value that this endpoint may provide in augmenting standard ED practices for the rule-out of ACS.

(ii) In an attempt to alleviate some of the confusion, we have modified the wording throughout the manuscript to state that our analysis focused on a diagnosis of 'non-cardiac chest pain/symptoms', rather than 'ACS-negative'.

Comment 3(a): "Methods, Page 9, line 3: Was the ACS negative only based on testing and the initial visit or were patients followed? What happened if an outpatient test was ordered for follow-up?"

Response: Patient follow-up was not incorporated into the design of this pilot study, a point that s acknowledged as a limitation on page 17 par 1. For patients who were discharged with a diagnosis of non-cardiac chest pain, this was the diagnosis used in the analysis.

Comment 3(b): "Was a kappa assessed on the diagnosis for ACS?"

Response: Our study design did not include the calculation of kappa values. As noted on page 9 par 1, at both sites, final decisions on diagnosis were adjudicated by a cardiologist.

Comment 3(c): "Although the investigators present their rationale why they are looking for ACS negative, the double negative makes it more cumbersome to interpret the data. They could address one of their concerns by using net reclassification index or IDI to look for the additional diagnostic value of the test. In fact, this is probably the preferred way to assess the value of a new diagnostic test that is added to standard of care as it allows the evaluation of the direction of change."

Response: This issue was also raised by Dr. Charpentier, and our response is:
We have considered augmenting the analysis by including the 'net reclassification index' (NRI) and the 'integrated discrimination improvement' (IDI) statistic to evaluate the incremental prognostic impact that measurement of PFA closure times might have if added to standard ED practices for the rule-out of ACS. However, there is growing skepticism among respected leaders in the field of epidemiology regarding the value provided by these indices: i.e.,
"Regardless of the intuitive appeal that the NRI statistic may garner, its potential for being inflated by miscalibrated or over-fit models is a very serious concern."
"In practice, one should not use the NRI statistic, or other prediction improvement performance measures such as delta-AUC, delta-ROC(f), delta-SNB(t), or delta-Brier for that matter, to determine if there is predictive information in a novel marker Y."
[from Pepe et al, "The net reclassification index (NRI): a misleading measure of prediction improvement with miscalibrated or overfit models" in the UW Biostatistics Working Paper Series: http://biostats.bepress.com/uwbiostat/. See also Hilden & Gerds, "A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index." Statistics in Medicine, 2013, DOI: 10.1002/sim.5804,.Kerr, Pepe et al, "Net reclassification indices for evaluating risk prediction instruments: a critical review." Epidemiology 2013; Epub 14 November 2013, and others.]
As these analyses have yet to be thoroughly vetted and are currently being questioned, we elected not to include the NRI or ID statistic in our already-complicated manuscript – a decision that is discussed on page 11 par 2 of the revised text.

Comment 3(d): “Page 19, para 2: please justify why the 90th percentile was selected as the threshold.”

Response: As detailed in our response to Comment 1(b), the prospective choice of the 90th percentile of the entire study population as the criterion for ‘prolonged’ closure time was arbitrary: this point is now specified on page 10 par 2 and at multiple points throughout the revised text, and in part provides the rationale for the discussion on page 16 par 1 regarding the consequences of changing the threshold to the 80th or 95th percentile.

Comment 4(a): “Results. Page 11: Why were those with thrombotic events excluded? If this is to be used at the time of presentation, data known at presentation will have to drive the test.”

Response: Given that this was a pilot study, we made the decision to focus exclusively on ACS rather than introduce an additional, small and heterogeneous group into the statistical analysis.

Comment 4(b): “Was a sensitivity analysis done excluding the diagnosis of UA as it is a relatively subjective diagnosis. My skepticism with this diagnosis could be modified with a table on how this diagnosis was made.”

Response:
(i) When the analysis is repeated focusing exclusively on patients with a diagnosis of STEMI or NSTEMI (UA excluded) and using the 90th percentile of the study population as the threshold for ‘prolonged closure time’ (143 sec, after exclusion of UA patients), the outcome was comparable to that reported in Table 2 in which all patients are included: specificity, sensitivity, positive and negative predictive values for a diagnosis of non-cardiac chest pain were 100% (95% CI: 94.7% to 100%), 11.5% (CI: 8.3% to 15.5%), 100% (CI: 90.8% to 100% and 18.9% (CI: 15.0% to 23.3%).

(ii) The criteria for diagnosis of unstable angina were symptoms and ECG changes consistent with ACS but without positive biomarkers at 24 hours, together with a requirement for coronary intervention or CABG or a previous diagnosis of UA in the chart. To alleviate some of the concern regarding uncertainties in the diagnosis of unstable angina, to simplify the presentation and for consistency with current guidelines, we have re-formatted the text and Figure 2 such that data for NSTEMI and UA patients are presented in a single, combined group.

Comment 4(c): “Time from symptom onset or at least time from presentation to blood draw should be presented.”

Response: We now specify on page 8 par 3 that all blood samples were drawn within <1 hour of presentation.

Comment 5(a): “Discussion: As the study appears adequately powered for the regression analysis, I would eliminate the word pilot from the opening sentence of the discussion.”

Response: Done.

Comment 5(b): “It would have nice to have some discussion why site was such a predictor of ASC

negative."

Response: We cannot provide a definitive explanation of why site (Cordoba versus UMASS) was a significant predictor of outcome. However, we speculate that the significant difference in aspirin use, together with differences in demographics and the prevalence of risk factors in the two populations, may play a role. These concepts are raised on page 15 par 3 of the revised Discussion.

Comment 5(c): "Limitation: if the diagnosis of ACS is based on data entirely from the enrolling visit, it is possible that misclassification occurred."

Response: Potential misclassification has been as a limitation on page 17 par 1.