PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The organizational value of diagnostic strategies using high sensitivity troponin for patients with possible acute coronary syndromes: A trial-based cost-effectiveness analysis
AUTHORS	Juelicher, Paul; Greenslade, Jaimi; Parsonage, William; Cullen, Louise

VERSION 1 – REVIEW

REVIEWER	Arnoud van der Laarse Depts. of Cardiology and Clinical Chemistry and Laboratory Medicine Leiden University Medical Center Albinusdreef 2 2333ZA Leiden the Netherlands
REVIEW RETURNED	16-Aug-2016

GENERAL COMMENTS	The results of the present study, depicted in Figures 1 and 2, clearly show that after step 4, steps 5 and 6 offer hardly any added value. However, the authors do not mention that in the text. As to ACS patients, the "30-day hospital perspective" (see page 8) seems rather a long time. But on page 9, "the maximum length of stay was limited to 12 days in accordance with the assumptions made for the cost prediction model". Please explain this discrepancy in time. Criteria (page 12) is plural of criterion. In Table 2 the values tabulated for Low Risk group in the Cullen 2015 column and Study Cohort column are similar (2040 and 1530, resp.). Strange coincidence. Eleven of the 30 references (37%) are from the Cullen group, which seems rather abundant to this reviewer. This reviewer missed a good description of ADP (accelerated diagnostic protocol).

REVIEWER	Cheuk H. (Michael) Liu Yale-New Haven Hospital, New Haven, Connecticut, 06510. United States.
REVIEW RETURNED	29-Sep-2016

GENERAL COMMENTS	Feedback for the authors: (Line 18, page 2) should read "standard of care" rather than "usual care". The main outcomes (line 35, page 2) is the "measurement of total cost per patient" (we already know that you have stratify the group. In your conclusions (line 3 page 3) You may want to add that the algorithm presented decrease health care associated cost for patients seen in the ED with symptoms of

	ACS. In the Strengths and Limitations section please incorporate that if the final troponin level was measured after 6:30pm, the patient had to stay overnight, since this will likely skew your data. You may also want to differentiate what are the limitations and strengths of the study separately instead of lumping them in one paragraph. Introduction section: Statistical information about the US is unnecessary since the study was conducted overseas. Therefore, what happens in the US is irrelevant in the discussion. The ADAPT and Modified ADAPT studies were not referenced. You may want to include this since you talk about them several times throughout your paper. The aim of the study was not in-line with the study objective. May want to re-write that portion (line 44, page 5). Methods: Patient exclusion criteria you mention "non-ACS cause" but there are no examples of what is considered a non-ACS cause. May want to include some examples as well.(line 18,10 page 6) as well as if the staff considered that enrollment was inappropriate (there are no references of what guidelines the staff utilized to consider inappropriate enrollment). Line 44 page 6 speaks about a standardized case report form. You may want to include this form to the supplemental data, since it will show how the information and what information was collected. Paragragh 2 (start on line 14, page 7) needs to be re-written. It is unclear what endpoints were assigned or if all end-points were assigned to all patients. Cost prediction model section: Line 3 page 8 you mention that patients with a LOS > 12 days were excluded. Was this based on a previous study? May want to include an explanation why 12 days was the cut-off. Health Economic model section: This may be part of your limitations that the standard of care re-measures troponin levels after 6 hours of the first value and the high sensitive troponin level measures troponins after 2 hours. The test may appear beneficial but we do not know if it is the test itself or producing another level of troponins in 2 hours post symptom presentation. Line 7-8, page mentions low risk and high risk patients but there is no clear definition of what these patients are in the study and what characteristics were considered to make that difference. This is mention in the patient referral and management section. Perhaps you may want to include this classification when you first talk about the patients so the readers have a better understanding of what these categories are. Patient outcome and cost effectiveness: Line 12- page 12 states that the introduction of hsTnI did not alter overall diagnosis but increased the patients referred to ACS management. On the previous page you mention "the number of referrals to ACS management based on acute troponin findings did not differ". One statement contradicts the other. Please re-write. Line 30, page 32 mention that the "rule-in criteria was found to be superior". This is a cost evaluation, I think it would be inappropriate to utilize the word "superior" since you are not defining superiority and this is not a superiority trial. Discussion: Line 5, page 14 states that there is a substantial time reduction effect in troponin time. Please re-write, I am unsure of what the statement means and in comparison to what.

	"Usual care" should be substituted by "standard of care" since this is the way it appears in the references provided as well as the supplemental information. Paragraph 4 (line 38, page 14) needs to be re-written. I do not agree with the statement that you will be able to extrapolate short stay unit savings with an emergency department that does not have them. If so, please elaborate more on your explanation and provider objective measures of cost savings in these areas. CHEERS checklist: Please ensure that pages reported are accurate since the numbers do not match the sections of your paper.

REVIEWER	Maame Yaa A. B. Yiadom, MD, MPH Vanderbilt University, USA
REVIEW RETURNED	11-Oct-2016

GENERAL COMMENTS	This is a very thoughtful and interesting analysis presenting the potential health economic (and hospital resource utilization) impact of replacing current troponin assays with high sensitive troponins. However, the authors do not present a clear research question or scientific hypothesis grounding the study. In addition, the connection between the clinical context of ACS risk stratification, their study design, and structure of their analysis is not clear and very difficult to follow. The study design and analysis are not well positioned within the context of the emergency department decision-making process about how much testing (and time) is too much testing (or time). A stronger literature review about the clinical context of using protocols and the potential value added by replacing standard troponin with a high sensitive assay need be more strongly described and more clearly articulated. This review may help them refine their study question. This is needed to understand how they see their study impacting practice with the use of a high sensitive troponin. It is not clear if this was a single center, multi-centered, or multi-national study. They note that the patient sample is from Brisbane Australia, but their introduction cites United States statistics. The authors describe their primary outcome as 'total cost per correctly stratified patients.' However, the stratification scheme is not clear. They mention referral for acute coronary syndrome evaluation as part of stratification. However, there are multiple levels of stratification for potential ACS patients: upon arrival, upon emergency physician assessment, after the first troponin result is available, when a decision is being made about the likely disposition (home vs short stay vs hospital). Do they mean referral for non-infarction associated ischemia testing and further evaluation via prolonged ED stay vs. a short stay unit vs. hospital admission? The authors do a nice job of clarifying their assumptions about who is included on page 7, but it's not clear why these decision were made. Patients with cardiac diagnoses can have long hospital LOS, so why exclude those in-hospital for >12 days? Why did they exclude missing data or outliers rather than impute or perform sensitivity analyses with and without the outliers? It is not clear why they selected the independent or predictor variables included in the costing model. It can be inferred, but a clearer rationale should be

	provided. They should consider including language to the effect of the following: “We considered the major sources of hospital-based facility and provider costs associated with the acute coronary syndrome evaluation and included them in our costing model to estimate the cost of patient assigned to 3 treatment pathways.” Much of the language used is vague and makes the manuscript difficult to follow. To note a few: “Local cardiologist adjudicated the outcome (line 18).” How many? In what way? Why not an emergency physician if this is a study about informing emergency medicine practice. “End points were agreed on by consensus (line 27).” Consensus by whom? The 6 strategies compared using the model are not clear, nor is it clear how the costing model was applied to the strategies. Clarity is critical if the results are intended to change practice. I think they mean to say that they ‘used their model to estimate the cost of the recommended post ED care disposition for a patient per each of the strategies, then quantified the difference while testing for statistical significance.’ The table that lists the strategies includes many abbreviations that are difficult to connect with many of established protocols. The authors should consider including citations within the text or tables, fewer abbreviations, and more clearly outlining the 6 strategies. This a potentially impactful analysis that can inform a critical area of emergency medicine practice. They present multiple interesting findings. However, the manuscript is not structure such that a reader can understand the scientific question or see how the findings relate to the clinical problem at hand.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
1. The results of the present study, depicted in Figures 1 and 2, clearly show that after step 4, steps 5 and 6 offer hardly any added value. However, the authors do not mention that in the text.

**The key benefit while moving from Strategy 4 to Strategy 5 was a significant reduction in the admission rate to both the short-stay unit and ward ($p < 0.001$ for both indicators, eTable7). This is due to the direct rule-out of 6.1% of individuals previously assigned the intermediate-, or high-risk group. As a result, total costs appeared to be significantly reduced for Strategy 5 compared to Strategy 4 ($p=0.02$). Assuming an estimated number of 500,000 persons presenting with chest pain to Australian EDs, the cost reduction of $53 per patient would sum up to more than $26M, which was regarded as relevant by the authors.
It is correct to conclude that Strategy 6 did not significantly differ from Strategy 5 in terms of costs and diagnostic accuracy. Since all patients meeting the criteria of a highly elevated baseline hsTnI (≥52mg/L) were classified as high-risk by all other strategies, Strategy 6 did not result in a change in admission rates. However, the key value of Strategy 6 was the immediate referral to inpatient cardiology: 46.6% of patients finally diagnosed with ACS would receive earlier cardiac intervention. Given the fact that actually all patients in the underlying observational study were managed by standard care, data on potential outcomes effect of an earlier cardiac treatment were not available, and thus not captured in the health economic evaluation.

We have rephrased the paragraph in the discussion section to highlight these points.

2. As to ACS patients, the "30-day hospital perspective" (see page 8) seems rather a long time. But

on page 9, "the maximum length of stay was limited to 12 days in accordance with the assumptions made for the cost prediction model". Please explain this discrepancy in time.

**The cost prediction model was based on data that were limited to the assessment and stratification period (e.g. stress testing, ECG, troponin values, etc.). Treatment or inpatient management information other than inpatient time was not available.

The model compared different assessment strategies for individuals suggestive of having ACS. Since only 11% of patients finally had an ACS, we regarded it as important to not overestimate the early assessment costs by taking cost-outliers of non-cardiac patients into account. To mitigate the potential risk for bias, 2.5% of long-stay patients were excluded, and the costs prediction model was limited to an inpatient stay of 12 days. This threshold was selected, since it was the maximum length of stay threshold that did not affect quartiles, median, and the 95th percentile of the cost distribution of the original data, but also excluded effects of unknown inpatient activities from the prediction model.

The final adjudicated diagnosis was based on ACS endpoints occurring on presentation or within 30 days. This is a standard time frame within the ACS literature and allows the identification of patients who may have a missed event during their index presentation that is identified within 30 days. The model also considered follow-up costs for patients with a 30-day clinical outcome of ACS who were ruled-out by a strategy. Since these patients would have been rule-out inappropriately, it was assumed in the model that these patients would be readmitted within 30 days. Therefore, the time horizon of the model followed a 30-day perspective.

We added additional details to the eMethods section.

3. Criteria (page 12) is plural of criterion.

**Thank you. In response to additional reviewer comments, the paragraph was re-phrased (Pg 12, line 29) and the word criterion was removed. It now reads as:
“Strategy 5 (a protocol utitilizing hsTnI, ADP, and LOD) was found to be the dominant strategy in the study, providing better accuracy at lower costs (Figure 2).”

4. In Table 2 the values tabulated for Low Risk group in the Cullen 2015 column and Study Cohort column are similar (2040 and 1530, resp.). Strange coincidence.

**Both columns describe the actual costs of exactly the same patients. The “Study cohort” is an extract of the cohort described in Cullen 2015. Exclusion criteria (e.g. missing baseline troponin results) were described in eTable2 in the supplement. In sum, 219 patients were excluded from the underlying cohort reported in Cullen 2015. No patient classified as “Low Risk” (LR) was excluded for the health economic study. Thus, actual costs assigned to LR- patients remained exactly the same compared to the underlying cohort (Cullen 2015). The proportion of LR patients changed due to the reduced number of total patients (719 vs. 938).

5. Eleven of the 30 references (37%) are from the Cullen group, which seems rather abundant to this reviewer.

**Dr. Cullen contributed in a number of relevant papers in the area of risk assessment and stratification of patients arriving in the emergency department. In addition, the health economic

evaluation presented here was based on an observational study that was described previously with her contribution. Despite this, we have now removed several of the references from the Cullen group that we felt were not essential for this paper.

6. This reviewer missed a good description of ADP (accelerated diagnostic protocol).

**The ADP referred to the Modified ADAPT accelerated diagnostic protocol previously described in Cullen et al. JACC 2013. The described accelerated protocol was based on 0- and 2h hsTnI, TIMI risk scores, and ECG. It was demonstrated that patients with hsTnI below the diagnostic cutoff and TIMI score ≤1 could be safely discharged early.

We have now added a reference to the ADP in the introduction. We also added the following description of the ADP to the Methods section “Strategy 4 (hsTnI+ADP) utilised 0 and 2 hour hsTnI but enabled patients to be directly ruled out with no further work-up strategies using the modified ADAPT ADP. That is, patients could be ruled out if their TIMI risk score was <=1, their 0 and 2 hour troponin were below the diagnostic cutoff and their presentation ECG was non ischaemic. In addition, the ADP rule is described under Table 1, Figure 1 as “Accelerated rule-out applied to individuals with hsTnI values at 0 and 2h below the diagnostic cut-off and a TIMI risk score ≤1”, and in the eMethods.

Reviewer 2

1. (Line 18, page 2) should read "standard of care" rather than "usual care".

**We have consistently replaced usual care with standard care throughout the paper.

2. The main outcomes (line 35, page 2) is the "measurement of total cost per patient" (we already know that you have stratify the group.)

**We re-phrased this to: “The primary outcome was total costs”.

3. In your conclusions (line 3 page 3) You may want to add that the algorithm presented decrease health care associated cost for patients seen in the ED with symptoms of ACS.

**We have now rephrased to: “It would decrease costs, and would provide significant benefits for the hospital”. We did not include reference to patients seen in the ED as this would exceed our allowable word limit for the abstract.

4. In the Strengths and Limitations section please incorporate that if the final troponin level was measured after 6:30pm, the patient had to stay overnight, since this will likely skew your data.

**Thank you for your important comment. We ran a sensitivity analysis for the discharge threshold time (eFigure8). Since this threshold may well not be fixed in real life, we tested the impact with some flexibility. Data in eFigure 8 reveal no significant observable effect of a flexible threshold time on Strategy 2 (hsTnI), whereas Strategy 1 (cTnI, standard care) was strongly affected between 6 and 8pm. Although these findings depend on emergency department arrival pattern, results suggested that hsTnI enabled algorithms would be less affected by variation. Given the fact that the arrival pattern used in the model was derived from actual data, hsTnI supported protocols would likely lead to more stable and predictable emergency department processes.

We also added this aspect to the “Strength And Limitations: “ (…) In addition, we considered realistic management rules (e.g. patients not being discharged before 6.30pm had to stay overnight)”.

5. You may also want to differentiate what are the limitations and strengths of the study separately instead of lumping them in one paragraph.

**We had previously included these sections together as we thought this was a requirement of the journal. However, we are happy to have them separated and have now included separate headings to highlight strengths and limitations separately.

6. Introduction section: Statistical information about the US is unnecessary since the study was conducted overseas. Therefore, what happens in the US is irrelevant in the discussion.

**We have replaced US statistics with the following sentence “In Australia, over 500,000 persons per year present with chest pain, but fewer than 20% were diagnosed with ACS.2, 3”

7. The ADAPT and Modified ADAPT studies were not referenced. You may want to include this since you talk about them several times throughout your paper.

**We have now included the reference as requested. A description of the Modified ADAPT is also provided in the explanation for Table 1, and Figure 1 as “Accelerated rule-out applied to individuals with hsTnI values at 0 and 2h below the diagnostic cut-off and a TIMI risk score ≤1”

8. The aim of the study was not in-line with the study objective. May want to re-write that portion (line 44, page 5).

**We have rephrased the aim of the study to: “The aim of this study was to evaluate the hospital-specific health economic implications of different protocols utilizing hsTnI for assessment of emergency department patients with chest pain compared to standard care”.

In addition, we rephrased the title to clarify aim of the study, and to better distinguish this work from previous. It is now titled “The organizational value of diagnostic strategies using high sensitivity troponin for patients with possible acute coronary syndromes: A trial-based cost-effectiveness analysis”

9. Methods: Patient exclusion criteria you mention "non-ACS cause" but there are no examples of what is considered a non-ACS cause. May want to include some examples as well.(line 18,10 page 6) as well as if the staff considered that enrollment was inappropriate (there are no references of what guidelines the staff utilized to consider inappropriate enrollment).

**We have now changed the section as follows “Patients were excluded if there was a clear non–ACS cause for their symptoms (e.g., findings of pneumonia), they were unwilling or unable to provide informed consent, staff considered that recruitment was inappropriate (e.g., patients undergoing palliative treatment)”

10. Line 44 page 6 speaks about a standardized case report form. You may want to include this form to the supplemental data, since it will show how the information and what information was collected.

**The case report form is too lengthy to include as supplemental data. The standardized form includes all of the data outlined in the reference “Cullen L, Than M, Brown AF, et al. Comprehensive standardized data definitions for acute coronary syndrome research in emergency departments in Australasia. Emerg Med Australas 2010;22(1):35-55.”

11. Paragragh 2 (start on line 14, page 7) needs to be re-written. It is unclear what endpoints were assigned or if all end-points were assigned to all patients.

**We have now rewritten this section as follows “Each patient was assigned one or more endpoints to explain the reason for their index presentation, or any events occurring within 30 days of admission. There were fifteen possible endpoints, including both cardiovascular and non-cardiovascular endpoints. Patients were considered to meet the definition for ACS if they were assigned any of the following endpoints; cardiovascular death, cardiac arrest, revascularization procedure, cardiogenic shock, acute myocardial infarction, or unstable angina pectoris”.

12. Cost prediction model section: Line 3 page 8 you mention that patients with a LOS > 12 days were excluded. Was this based on a previous study? May want to include an explanation why 12 days was the cut-off.

**In alignment with the study focus, activities that were available by patient were limited to the risk assessment and stratification period: ECG, stress test, troponin testing, MPS, CTCA, angiography, etc.). Information about inpatient treatment and management other than inpatient time were not available. Thus, the prediction of total costs based on the available data was expected to be biased with increasing inpatient time. In fact, the average costs per inpatient day decreased with increasing stay until a slight increase appeared for patients staying more than 15 days. This was regarded as an indicator for costs accrued from activities not captured in the collected data. By further analyzing the data, we excluded 2.5% of patients with an inpatient stay of more than 12 days as this was the maximum length of stay threshold that did not alter quartiles, median, or the 95th percentile of the cost distribution of the original data, but also excluded effects of unknown inpatient activities from the prediction model.

We have rephrased the paragraph about the cost-prediction model to note that “Patients with a hospital length of stay (LOS) greater than 12 days were excluded to reduce bias from non-cardiac stays”. We also have added a “Cost prediction model” paragraph to the eMethods that is providing the all details mentioned above.

13. Health Economic model section: This may be part of your limitations that the standard of care re-measures troponin levels after 6 hours of the first value and the high sensitive troponin level measures troponins after 2 hours. The test may appear beneficial but we do not know if it is the test itself or producing another level of troponins in 2 hours post symptom presentation.

**The difference in protocol time of contemporary and high-sensitive troponin is an elementary factor considered in the strategies. As described in the methods section and in Table 1, Strategy-1 referred to a protocol considering troponin and a 6 hour protocol; all other strategies were defined by utilizing hsTnI and a 2h protocol. In addition to protocol time, the analytical performance of the assays (sensitivity , specificity) as well as effects arising from other management options (accelerated or direct rule-in and rule-out) were taken into account. We also acknowledge the point that it is unclear whether the use of a sensitive troponin taken at 2 hours would be equally effective and have now included the following sentence in the limitations. “The model compared a sensitive troponin assay at 6 hours to highly sensitive assay at 2 hours. For the models not utilizing the LoD, it is unclear whether a sensitive troponin taken at 2 hours would provide the same benefits outlined here with a highly sensitive assay”.

14. Line 7-8, page 9 mentions low risk and high risk patients but there is no clear definition of what these patients are in the study and what characteristics were considered to make that difference. This

is mention in the patient referral and management section. Perhaps you may want to include this classification when you first talk about the patients so the readers have a better understanding of what these categories are.

**We added the following description of the risk stratification in the methods:

“Patients were classified into risk groups according to the Heart Foundation of Australia/Cardiac Society of Australia and New Zealand guidelines (2011)”. In addition, we added a more detailed description and an illustration to the supplement (eFigure 1). We have also re-phrased the description of the pathway in the methods section (Page 9, line 3):

“The model structure and chest pain evaluation pathway is described eFigure 1. Individuals entering the model were stratified in the ED based on individual characteristics, first electrocardiogram, and troponin taken. Patients classified as high-risk were admitted to inpatient cardiology. Low-risk patients were kept in the ED to await final assessment. Intermediate-risk patients were referred to the short-stay unit (SSU) for further cardiac workup. Patients referred to SSU or inpatient ward were counted as admitted.”

15. Patient outcome and cost effectiveness: Line 12- page 12 states that the introduction of hsTnl did not alter overall diagnosis but increased the patients referred to ACS management. On the previous page you mention "the number of referrals to ACS management based on acute troponin findings did not differ". One statement contradicts the other. Please re-write.

**Both sections in the article refer to different aspects. The first discussed implications on the amount of patients referred to further clinical service. The latter on pg. 12 provided information about the accuracy and outcome of the referrals. The complete phrase on pg 12 is: “did not alter overall diagnosis but increased the patients referred to ACS management who had a final diagnosis of non-ACS”.
The number of total referrals comprised patients sent for ACS management with a final diagnosis of ACS, and those referred for ACS management without a final diagnosis of ACS. The total number of referrals did not differ significantly between contemporary and high-sensitivity troponin. However, a slight increase of patients referred for ACS management with non-ACS was found for the hsTnI strategy.

16. Line 30, page 32 mention that the "rule-in criteria was found to be superior". This is a cost evaluation, I think it would be inappropriate to utilize the word "superior" since you are not defining superiority and this is not a superiority trial.
Page 12, line 32.

**Superiority here refers to a strategy that would provide equal or better outcome at lower costs. In health economics, this is usually referred to as the dominant strategy. In order to avoid confusion with other definitions and be more consistent to the terminology in health economic studies, we appreciate the comment and replaced “superior” with “dominant”. In addition, we added a reference to this concept (Drummond, M., et al. (2005). Methods for the Economic Evaluation of Health Care Programmes. New York, Oxford University Press.)

17. Discussion: Line 5, page 14 states that there is a substantial time reduction effect in troponin time. Please re-write, I am unsure of what the statement means and in comparison to what.

**We rephrased the paragraph.
“The overall organizational benefits of the dominant strategy (Strategy-5) compared to standard care were caused by two effects: 1) a substantial time reduction in protocol time, and 2) significantly

improved stratification efficiency".

18. "Usual care" should be substituted by "standard of care" since this is the way it appears in the references provided as well as the supplemental information.

**Thank you, this has now been replaced.

19. Paragraph 4 (line 38, page 14) needs to be re-written. I do not agree with the statement that you will be able to extrapolate short stay unit savings with an emergency department that does not have them. If so, please elaborate more on your explanation and provider objective measures of cost savings in these areas.

**We have deleted this paragraph.

20. CHEERS checklist: Please ensure that pages reported are accurate since the numbers do not match the sections of your paper.

**We apologise. The pages in the checklist referred to the original manuscript that we submitted. The document provided to the reviewer was re-formatted by the journal. We have now updated the document to the best of our ability

Review 3

1. This is a very thoughtful and interesting analysis presenting the potential health economic (and hospital resource utilization) impact of replacing current troponin assays with high sensitive troponins.

2. The authors do not present a clear research question or scientific hypothesis grounding the study.

**We have now altered the aim within the introduction and we hope this clarifies the purpose of the study. The research question was to add a health economic evaluation thus complementing existing research that has been focusing on clinical outcomes. To improve clarity we rephrased the paragraph: "While research into novel accelerated strategies has usually reported clinical outcomes, few have assessed health economic implications of such protocols, or made comparisons to define optimum strategies. The incorporation of highly sensitive cardiac troponin I (hsTnI) assays into clinical practice may have additional health economic benefits on the hospital level; however, this aspect has not been explored to date. The aim of this study was to evaluate the hospital-specific health economic implications of different protocols utilizing hsTnI for assessment of emergency department patients with chest pain compared to standard care."

3. The connection between the clinical context of ACS risk stratification, their study design, and structure of their analysis is not clear and very difficult to follow. The study design and analysis are not well positioned within the context of the emergency department decision-making process about how much testing (and time) is too much testing (or time).

**We replaced Figure 1 with an illustration of the model structure and flow. We added an illustration of the risk stratification and process of care to the eMethods (Figure 1). We rephrased the description of the the risk stratification in the methods section.

We rephrased the title in order to clarify aim of the study: "The organizational value of diagnostic

strategies using high sensitivity troponin for patients with possible acute coronary syndromes: A trial-based cost-effectiveness analysis".

It remains unclear in the current literature about what can be considered as too much testing in the ED for ACS. What drives clinicians to investigate patients with possible ACS is perceived risk and guideline recommended care. An acceptable miss-rate of <1% for serious adverse events has been reported to be appropriate in this context. Many ADPs aim to achieve this level of safety.

In terms of time, many EDs are under constraints of time targets (the 4 hour rules) and this often places a focus on the length of assessment periods. Attempts to reduce unnecessarily long ED stays are appropriate.

4. A stronger literature review about the clinical context of using protocols and the potential value added by replacing standard troponin with a high sensitive assay need be more strongly described and more clearly articulated. This review may help them refine their study question. This is needed to understand how they see their study impacting practice with the use of a high sensitive troponin.

**We appreciate the advice of the reviewer that further detail in the literature review would be ideal. However, word limits make it difficult to add substantial further information into the paper. To the best of our knowledge, we reviewed and referenced all studies dealing with accelerated troponin protocols and high-sensitivity troponin. This review was actually also the baseline for defining the comparators or strategies used in the model. Table 1 summarized the strategies, references to clinical studies or guidelines. However, we have now reworded the second paragraph of the introduction to try and clarify the types of accelerated protocols used.

The aim of the study was not to discuss and compare different protocols regarding their clinical outcome, but to evaluate health economic implications resulting from different protocols. Since we used a very short-term hospital perspective, this evaluation provided relevant information about the likely impact of different protocols on ED management and costs. We appreciate the idea of a comprehensive review of existing, recommended chest pain protocols. This may well add value to the reader.

5. It is not clear if this was a single center, multi-centered, or multi-national study.

**The health economic evaluation was based on data from a single center.
We have now added: "The study utilized data from a prospective, single-center observational study in Brisbane, Australia."

6. They note that the patient sample is from Brisbane Australia, but their introduction cites United States statistics.

**We have now replaced US statistics with Australian data.

7. The authors describe their primary outcome as 'total cost per correctly stratified patients.' However, the stratification scheme is not clear. They mention referral for acute coronary syndrome evaluation as part of stratification. However, there are multiple levels of stratification for potential ACS patients: upon arrival, upon emergency physician assessment, after the first troponin result is available, when a decision is being made about the likely disposition (home vs short stay vs hospital). Do they mean referral for non-infarction associated ischemia testing and further evaluation via prolonged ED stay vs. a short stay unit vs. hospital admission?

**We apologise, the reference to stratification was unclear. The primary outcome has now been changed to total cost. We changed the description of the primary outcome to: “Differences between strategies were expressed in terms of total costs per patient and diagnostic accuracy. Diagnostic accuracy was defined as the percentage of correctly diagnosed patients compared to the final adjudicated diagnosis. In addition, LOS, referral rates, admission rates, and overnight stays were evaluated”.

8. The authors do a nice job of clarifying their assumptions about who is included on page 7, but it’s not clear why these decision were made. Patients with cardiac diagnoses can have long hospital LOS, so why exclude those in-hospital for >12 days?

**We agree that some cardiac patients can have long LOS. In fact, in our cohort, only 2.5% of all patients did stay more than 12 days. The purpose of the study, however, was to compare costs and effectiveness of different ACS evaluation strategies. The collected data used to develop a cost prediction model were restricted to activities for risk assessment and evaluation of ACS. Information about inpatient treatment other than LOS were not available. Thus, the cost prediction model was limited by the availability of data.

89% of patients analyzed in the study ended up with non-cardiac conditions. Actually, only 2.5% of patients were excluded from the baseline cohort with an LOS >12 days. We chose this threshold as this was the maximum length of stay threshold that did not affect quartiles, median, and the 95th percentile of the cost distribution of the original data, but also excluded effects of unknown inpatient activities from the prediction model. With that, we believed that this was a reasonable limitation of the model without

We added this information to the “Cost prediction model” paragraph in the eMethods section.
We added the limitation of the cost-prediction model to the limitations.

9. Why did they exclude missing data or outliers rather than impute or perform sensitivity analyses with and without the outliers? It is not clear why they selected the independent or predictor variables included in the costing model. It can be inferred, but a clearer rationale should be provided. They should consider including language to the effect of the following: “We considered the major sources of hospital-based facility and provider costs associated with the acute coronary syndrome evaluation and included them in our costing model to estimate the cost of patient assigned to 3 treatment pathways.”

**The actual costs data were based on administration records. Patients referred to CABG surgery were sent to an external hospital; cost data were collected from different sources, and had to be excluded for consistency reasons (n=14; 1.5%). In a few cases (n=6; 0.6%) costs were inconsistent or missing (Example: Inpatient costs of $105 were assigned to a STEMI patient staying 4 days in hospital). Exclusion criteria are given in eTable1.

Only 11% of all included patients did end up with a final diagnosis of ACS. In alignment with the study focus, the considered predictor variables were based on activities limited to the risk assessment and stratification period: ECG, stress test, troponin testing, MPS, CTCA, angiography, etc.). Information about inpatient treatment and management other than inpatient time were not available. It appeared that the average costs per inpatient day decreased with increasing stay until a slight increase appeared for patients staying more than 15 days. This was regarded as an indicator for costs accrued from activities not captured in the collected data. By further analyzing the data, we excluded 2.5% of patients with an inpatient stay of more than 12 days as this was the maximum length of stay threshold that did not affect quartiles, median, and the 95th percentile of the cost distribution of the original data, but also excluded effects of unknown inpatient activities from the prediction model.

The paragraph under METHODS / Cost prediction model was re-phrased to:

“The prediction model was developed in four steps. First, we analyzed the data and predefined exclusion criteria (eTable1). Patients who received coronary bypass surgery (CABG) were excluded because they were transferred to another hospital for surgery with no available outcome data and unknown accuracy of cost information. Cases with inconsistent or missing cost data were excluded. Patients with a hospital length of stay (LOS) greater than 12 days were excluded to reduce bias from non-cardiac stays. Second, we considered key activities for evaluating an acute coronary syndrome in a generalized Box-Cox transformed model. Third, we dropped non-significant variables (2nd troponin, p=0.9; stress echocardiography, p=0.6) from the predicting variables, checked for any relevant multicollinearity between variables, and excluded cases that showed extreme discrepancies to the predicted results (n=4; eTable1). Fourth, we run the final analyses that led to the cost prediction model and the 95% confidence interval of each predictor (eTable6). The final model was based on data of 891 individuals. The following predictors for ACS evaluation costs were used: emergency department time, inpatient time, performed activities (exercise stress test, myocardial perfusion scan, computed tomography coronary angiography, echocardiography, and angiography), admission to short-stay unit, or admission to an inpatient ward. More information is given in the supplement (eMethods). ”

Additional information was added to eMethods.

Uncertainty in the cost estimate was addressed by sampling each predictor from a uniform distribution of the 95% confidence interval of the regression coefficient for every individual patient. The very intensive sampling strategy resulted in variety of results as described in eFigure3.

In order to improve clarity, we re-phrased the respective paragraph in the methods section:

“A minimum required dataset was defined for patients included in the analysis (eTable 2), and 219 patients with missing troponin values were excluded.

Work-up, work-up duration, and length of stay were statistically analyzed from the model cohort and transformed into statistical distributions.

Patient attributes (age, sex, clinical characteristics, adjudicated diagnosis, electrocardiogram status, and troponin values) were individually sampled from the model cohort by bootstrapping. This created a hypothetical cohort of 40,000 patients who followed the model for each of the strategies. Work-up and times for each patient were randomly sampled from distributions. Costs were estimated by considering attributes, work-up activities, work-up duration, and length of stay in the cost prediction model with coefficients individually sampled from the 95% confidence interval of the respective predictor.

The model followed a 30-day hospital perspective.

Differences between strategies were expressed in terms of total hospital costs and diagnostic accuracy. Diagnostic accuracy was defined as the percentage of correctly diagnosed patients compared to the final adjudicated diagnosis. In addition, LOS, referral rates, admission rates, and overnight stays were evaluated.

We conducted one-way and probabilistic sensitivity analysis to test the robustness of the microsimulation. Model structure, parameters and assumptions are described in details in the supplement.”

10. Much of the language used is vague and makes the manuscript difficult to follow. To note a few: “Local cardiologist adjudicated the outcome (line 18).” How many? In what way? Why not an emergency physician if this is a study about informing emergency medicine practice. “End points were agreed on by consensus (line 27).” Consensus by whom?

**We have rephrased a number of paragraphs in the methods thus hoping to add clarity. “One cardiologist from a group of three potential cardiologists adjudicated the outcome independently. Cardiologists had knowledge of the clinical record, electrocardiogram and troponin results from standard care and used such information to determine whether the patient met the predefined criteria for the cardiovascular endpoints15. Patients not meeting such endpoints were classed as having a non-cardiovascular problem. A second cardiologist from the group conducted a blind review of all ACS cases and 10% of non–ACS cases. In cases of disagreement, endpoints were agreed on by consensus by the two cardiologists involved in endpoint adjudication and one emergency physician. This was achieved for all endpoints”.

.

While you correctly note that this is a study about emergency medicine practice, an emergency physician is not typically responsible for identifying the patient’s ultimate cardiac diagnosis. The use of a cardiologist to determine whether or not the patient ultimately had a cardiovascular endpoint reflects clinical practice and is the gold standard within ACS research

11. The 6 strategies compared using the model are not clear, nor is it clear how the costing model was applied to the strategies. Clarity is critical if the results are intended to change practice. I think they mean to say that they ‘used their model to estimate the cost of the recommended post ED care disposition for a patient per each of the strategies, then quantified the difference while testing for statistical significance.’

**The strategies basically differ by the troponin assay used. This had several potential consequences such as a change in protocol time, and additional management options.

We re-phrased the description of the strategies in the methods section:

“We compared six strategies (Table 1). The standard of care was based on a protocol using cardiac troponin I (cTnI) at baseline and 6 hours after arrival (Strategy 1). All other strategies utilized hsTnI at presentation and 2 hours. Strategy 2 (termed hsTnI) was the same as standard care except that a 2-hour highly sensitive troponin was used rather than a 6-hour sensitive troponin. Strategy 3 (hsTnI+LoD) also utilised a 2 hour hsTnI, but allowed a patient to be directly ruled out on admission with no further work-up strategies if their baseline hsTnI was below the assay’s limit of detection (LoD). Strategy 4 (hsTnI+ADP) utilised 0 and 2 hour hsTnI but enabled patients to be directly ruled out with no further work-up strategies using the modified ADAPT ADP. That is, if their TIMI risk score was <=1, with normal troponin and non ischaemic ECG. Strategy 5 (hsTnI+LoD+ADP) was a combination of Strategies 3 and 4 in that patients could be ruled out if their baseline hsTnI was below the LoD or if they met the criteria according to the modified ADAPT ADP. Finally, Strategy 6 (hsTnI+LoD+ADP+direct rule in) employed the same rule out criteria as Strategy 5, but also enable atients with hsTnI at presentation >52ng/L to be directly ruled-in and admitted for ACS management (strategy 6)..”

ADP was defined in the introduction as recommended by the reviewer.

A hypothetical cohort of 40,000 patients was created by bootstrapping attributes from the model cohort. Each patient was sent through the model following each strategy. The cost model was applied

each of these patients by considering attributes, work-up activities and time, and coefficients individually sampled from the 95% confidence interval of the respective predictor.

We tried to address this comment by re-phrasing parts in the methods section.

12. The table that lists the strategies includes many abbreviations that are difficult to connect with many of established protocols. The authors should consider including citations within the text or tables, fewer abbreviations, and more clearly outlining the 6 strategies.

**The strategies consisted of a combination of different assays (cTnI, hsTnI), and management options (Protocol time, early rule-in, early rule-out, accelerated rule-out) as suggested and described in literature. We tried to address the reviewer's comment by re-phrasing the description of the strategies in the methods section (see above).

We realize that a certain complexity exits, but we believe that describing the differences between strategies in a table provides the best clarity and details.

We re-phrased the description in the Methods section in order to improve clarity.

13. This a potentially impactful analysis that can inform a critical area of emergency medicine practice. They present multiple interesting findings. However, the manuscript is not structure such that a reader can understand the scientific question or see how the findings relate to the clinical problem at hand.

**We very much appreciate the thoughtful reviews, and the very good suggestions and comments. We tried to address every comment in details, thus hoping that the revised version provides better clarity.

**Additional changes / Erratum

1. Title
We rephrased the title in order to clarify aim of the study
The organizational value of diagnostic strategies using high sensitivity troponin for patients with possible acute coronary syndromes: A trial-based cost-effectiveness analysis

2. Table 2: p-values for LR and HR in Table 2
p-values for LR and HR groups were mixed up by mistake.
The values have been corrected in the revised manuscript: LR: p-value 0.95; HR: p-value 0.76.

VERSION 2 – REVIEW

REVIEWER	Arnoud van der Laarse Dept. of Cardiology and Clinical Chemistry Leiden University Medical Center Leiden the Netherlands
REVIEW RETURNED	28-Nov-2016

GENERAL COMMENTS	I have seen the answers to my comments and to the comments of other reviewers, and I recommend acceptance of the revised manuscript # bmjopen-2016-013653.R1 entitled "The organizational value of diagnostic strategies using high sensitivity troponin for

	patients with possible acute coronary syndromes: A trial-based cost-effectiveness analysis" for BMJ Open.

REVIEWER	Francis J Zamora Yale-New Haven Health System, US
REVIEW RETURNED	30-Nov-2016

GENERAL COMMENTS	The primary outcome is not "total cost". May want to re-write this statement as you are looking at reduction of total cost as well as reduction in false-positive and false negative. Please re-write statement. Page 5 line 49, when you speak about the aim of the study "health economic implications" may be too broad of a term. May consider substitute the sentence by "cost savings in the emergency department". More specific to what was actually measured. Page 6 line 10 you stated that patient presented to the ED with 5 minutes worth of chest pain. The statement is suggestive that this is part of the American Heart Association guidelines for the diagnosis of ACS and this is not the case. The ACS guidelines mention "Onset of symptoms 3-4.5 hours, and less than 10 minutes for initiation of treatment if the patient is suspicious of ACS once arrived in the ED. Please rewrite statement. Page 12 line 29 "A significant reduction in LOS was observed if hsTnI replaced cTnI, with a mean saving of 6.2 hours". Was there a p value to establish that 6.2 hours is "significant"? Please ensure that the objectives of this study are mention and capture in the conclusions. Once the primary outcome is establish "as cost reduction" not "total cost" please ensure that the conclusions corroborate this statement. Overall, great job at re-writing some of the issues presented in the first review. Very interesting study.

REVIEWER	Maame Yaa A. B. Yiadom, MD, MPH Vanderbilt University, USA
REVIEW RETURNED	19-Dec-2016

GENERAL COMMENTS	This is a much improved manuscript with revisions highly responsive to prior critiques. I recommended that this be accepted with minor revisions. I have a few additional comments for consideration 1. The authors have very clearly articulated the study questions (or their “aim”) in this revision, but still do not clearly state the scientific hypothesis (or null hypothesis) 2. The data presented in Table 2 is a reasonable means of presenting the model validation. However, it would be helpful to see a calibration curve of the health economics model (comparing the observed and expected costs for the derivation population) included in the supplement or main manuscript if figure limits permit. It will permit the reader to see where the model was strong and weak in a continuous fashion rather than with summary means and SDs and medians with IQRs, and aid the use of this model in real time clinical and hospital financing practice. 3. On page 9, line 44 they state that the max LOS was limited to 12 days in “accordance with the assumptions made for the cost prediction model.” They detail this is in the discussion, but would be helpful to explain here that this was done to avoid overestimating the effects from prolonged stays in patients with non-cardiac diagnoses.

VERSION 2 – AUTHOR RESPONSE

the abstract within the word limit.

The outcome of this study was total hospital costs but you correctly note that the aims and hypothesis were to identify whether there had been a reduction of total costs. We have now amended the aims and included a hypothesis to provide clarity around this endpoint.

2) Page 5 line 49, when you speak about the aim of the study "health economic implications" may be too broad of a term. May consider substitute the sentence by "cost savings in the emergency department". More specific to what was actually measured.

**We have changed the aim to the following: “The aim of this study was to evaluate the hospital-specific costs of different protocols utilizing hsTnI for assessment of emergency department patients with chest pain, compared to standard care. The hypothesis was that hsTnI enabled algorithms would streamline ED processes with equal or better assessment accuracy thus leading to savings in direct hospital costs when compared to standard care”.

3) Page 6 line 10 you stated that patient presented to the ED with 5 minutes worth of chest pain. The statement is suggestive that this is part of the American Heart Association guidelines for the diagnosis of ACS and this is not the case. The ACS guidelines mention "Onset of symptoms 3-4.5 hours, and less than 10 minutes for initiation of treatment if the patient is suspicious of ACS once arrived in the ED. Please rewrite statement.

**We have now changed the statement to read “…..presented to the emergency department with at least five minutes’ worth of chest pain suggestive of ACS, and were being evaluated for ACS. Pain suggestive of ACS was defined in accordance with American Heart Association case definitions.13”

4) Page 12 line 29 "A significant reduction in LOS was observed if hsTnI replaced cTnI, with a mean saving of 6.2 hours". Was there a p value to establish that 6.2 hours is "significant"?

**We have changed the sentence to read: A significant reduction in LOS was observed if hsTnI replaced cTnI, with a mean saving of 6.2 hours ($p<0.001$, Table 3).

5) Please ensure that the objectives of this study are mentioned and captured in the conclusions. Once the primary outcome is established "as cost reduction" not "total cost" please ensure that the conclusions corroborate this statement.

**We have changed the first section of the conclusions to state that: “This trial based economic modeling study sought to evaluate the impact of different hsTnI protocols on direct hospital costs and diagnostic accuracy compared to standard care. We found that emergency department assessment strategies utilizing hsTnI are very likely to be cost-effective and provide cost savings on a hospital level when compared to sensitive TnI protocols for patients presenting with symptoms consistent with ACS”.

Overall, great job at re-writing some of the issues presented in the first review. Very interesting study.

Reviewer: 3

This is a much improved manuscript with revisions highly responsive to prior critiques. I recommended that this be accepted with minor revisions. I have a few additional comments for consideration

1. The authors have very clearly articulated the study questions (or their "aim") in this revision, but still do not clearly state the scientific hypothesis (or null hypothesis).

**We have included the following hypothesis: The hypothesis was that hsTnI enabled algorithms would streamline ED processes with equal or better diagnostic accuracy, thus leading to savings in direct hospital costs when compared to standard care".

2. The data presented in Table 2 is a reasonable means of presenting the model validation. However, it would be helpful to see a calibration curve of the health economics model (comparing the observed and expected costs for the derivation population) included in the supplement or main manuscript if figure limits permit. It will permit the reader to see where the model was strong and weak in a continuous fashion rather than with summary means and SDs and medians with IQRs, and aid the use of this model in real time clinical and hospital financing practice.

**We have added a comparison of the actual and predicted costs for individuals with a final diagnosis of ACS and Non-ACS to the supplement (eFigure 9A and 9B).

3. On page 9, line 44 they state that the max LOS was limited to 12 days in "accordance with the assumptions made for the cost prediction model." They detail this is in the discussion, but would be helpful to explain here that this was done to avoid overestimating the effects from prolonged stays in patients with non-cardiac diagnoses.

** We have now changed the sentence to read: "The maximum LOS was limited to 12 days to avoid bias in the effects from prolonged stays in patients with non-cardiac diagnoses".

VERSION 3 – REVIEW

REVIEWER	Francis J Zamora PharmD Yale New Haven Hospital, US
REVIEW RETURNED	18-Jan-2017

GENERAL COMMENTS	In the study you provide a table (Table 4) with false positives and negatives but this information is not discussed in the limitations/strengths of the study. You may want to consider including the impact of the process in determining false positive and false negative (or if there is any impact at all) in this section. Otherwise I think you did a great job at stream lining the document and correcting some of the issues presented in the review. Besides this, I have no further revisions to the study. Good job.

VERSION 3 – AUTHOR RESPONSE

Thank you for your ongoing interest in our manuscript. The reviewer raised the following issue:

In the study you provide a table (Table 4) with false positives and negatives but this information is not discussed in the limitations/strengths of the study. You may want to consider including the impact of the process in determining false positive and false negative (or if there is any impact at all) in this section.

In response, we have changed the following:
We added to “Strengths”:
“Troponin results must be interpreted in concert with clinical presentation, ECG changes and other available information. Diagnostic accuracy used in this study refers to results of the complete pathway consisting of troponin results, ECG and cardiac workup. All hospital costs accrued from assessment, management and events during 30-days follow-up were considered in the analysis.”

We added to “Limitations”:
“Generalizability might be hindered by the variety of assessment processes. Exploiting the value of hsTnI relies on the appropriateness of testing and the implementation of adequate protocols.”

In the discussion, we changed (pg 13, line 14-15) to:
“(…) tend to increase the number of patients with a false positive diagnosis of ACS (p=0.056; Table 4)”

We added to the discussion (pg 15, line 46):
“A false positive troponin status can result in unnecessary referrals. In our study, 12.1% of individuals were categorized with an hsTnI status indicative for an acute event (eTable 8). 32% of individuals were referred for ACS management; this number was not different between Strategy 1 and 2. Most of the referrals were based on negative troponin findings followed by a positive cardiac work-up. Although we considered a conservative criterium with an absolute delta change between serial hsTnI tests of 2ng/L, an increase in total referrals for ACS management could not be found. However, a tendency for an increased number of patients with a false positive diagnosis of ACS was observed. It is however important to note that costs accrued from such interventions were considered in the analysis.”

We added to limitations (pg 16, line 17):
“Follow-up was limited to 30 days. Events happening after 30 days were not considered but may have an impact on the number of false-positives diagnosis. Troponin results must be interpreted in concert with clinical presentation, ECG changes and other available information. Diagnostic accuracy in this study refers to results of the complete assessment pathway consisting of troponin results, ECG and cardiac workup. In an approach to emphasize safety, we used a conservative dynamic cut-off between serial troponin tests. The impact of different absolute or relative changes was not evaluated. Age or gender specific troponin reference values may further improve the diagnostic accuracy but were not considered.”

We added to limitations (pg 17, line 5):
“It should be noted that exploiting the value of hsTnI fully relies on the appropriateness of testing and the implementation of adequate protocols.”