Author’s response to reviews

Title: Use of exercise tests in primary care: importance for referral decisions and possible bias in the decision process; a prospective observational study

Authors:

Gunnar Nilsson (gunnar.nilsson@jll.se)
Thomas Mooe (thomas.mooe@medicin.umu.se)
Lars Söderström (lars.soderstrom@jll.se)
Eva Samuelsson (eva.samuelsson@fammed.umu.se)

Version: 4 Date: 26 September 2014

Author's response to reviews: see over
Cover letter

Östersund, September 23, 2014

Dear Editor-in-Chief,

Thank you for your prompt response. We have carefully considered all of the reviewers’ comments and revised the manuscript accordingly. We hope that you find the revised version to be improved and acceptable for publication.

Below, we have included our responses to each of the reviewers’ comments. We have addressed all comments and suggestions individually.

Yours sincerely,

Gunnar Nilsson

MD, General Practitioner

Department of Public Health and Clinical Medicine

Umeå University, Sweden
Reviewer: Staffan Nilsson

1. Method; Recruitment and follow-up, line 159-160: How were myocardial infarction STEMI and NSTEMI defined in the study? I can’t find any description of the definitions in this manuscript nor in reference number 5. This is of importance if anyone else attempts to repeat the study in a different population.

   **Answer:** Myocardial infarctions were defined according to the universal definition by Thygesen K, Alpert JS, White HD: *Universal definition of myocardial infarction*. *European Heart Journal* 2007, 28(20):2525-2538. We have provided this reference in the revised manuscript. Since we did not describe patients with STEMI and NSTEMI separately, this division into two categories is unnecessary and was therefore changed to one diagnostic category, myocardial infarction. Please see the revised manuscript, section Methods, subsection Recruitment and Follow-up, line 142: "Registered events were: patients referred within six months to cardiologic evaluation, patients with hospitalisation for myocardial infarction, or hospitalisation for unstable angina and cases of cardiovascular death. Revascularisations (coronary artery bypass grafting or percutaneous coronary intervention) were recorded within 250 days from exercise testing, in some cases due to delays between referral and revascularisation. Myocardial infarctions were diagnosed in accordance with the universal definition [22]."

2. Section Results: Line 219: Since the mean age of the study population was 63.5 years, a substantial part of the study group must have been retired from their occupational work. The text and Tables does not demonstrate how many in the study group that were retired.
This matter may be of importance for the interpretation of results when it comes to socio-economic interactions and has to be accounted for by the authors.

**Answer:** The mean age of 63.5 years in our study indicates that many participants were retired. Our study file did not include data on retirement. Retirement is not always a clear-cut variable. Some patients may be on long-term sick leave before retirement, while others work part time after retirement, making this distinction difficult to use. In the SEI classification, a retired person is classified according to their previous main occupational background, and thus their socioeconomic status does not change upon retirement. Retirement from working life is still an important question, which we discuss in the revised manuscript in section Methods, subsection Measurement and Classification, line 170: "Retired participants were classified within the SEI system and not in a separate category; this classification also applied to participants on sick leave and to participants who were part-time retired." Please also refer to the Discussion, subsection Strengths and Limitations, line 307: "Since retirement was not recorded as a separate patient characteristic, we could not analyse the influence of retirement on referral to cardiologic evaluation, separated from age, or possible differences in retirement age in different socioeconomic groups."

**3. Line 316:** The authors state that there was no blinding of outcome data. I think this matter has to be discussed. Was there any risk of bias? Could it have been performed in an alternative way with blinding of data?

**Answer:** We designed an open observational study; the question of potential recall or classification bias should be addressed. The outcome of each exercise test was
classified at the end of each session by the physician responsible for the test at the Clinic of Clinical Physiology, Östersund Hospital. The outcome data, e.g. referral to the Heart Clinic, Östersund Hospital, were extracted from the medical record files at the heart clinic. All outcome data, such as referrals and cardiovascular events, were recorded as binary data (e.g. referral or no referral from a general practitioner (GP)). Patient characteristics and outcome data were determined independently from each other, an approach that was chosen to avoid recall or classification bias within the observational study design. Please refer to the revised Discussion, subsection Strengths and Limitations, line 298: "There was no blinding of outcome data, relative to patient characteristics, since all medical records were scrutinised in complete form. Patient characteristics and outcome data were determined independently from each other to avoid recall or classification bias within the observational study design."

4. Section: Measurements and classifications Line 188-192. As detailed descriptions of the classification of ECGs and the performance of exercise tests are given in reference number 5 also this reference should be given here.

Answer: We agree with the reviewer’s comment. This reference is now provided as reference [8] in the revised manuscript, section Methods, subsection Measurements and Classification, line 177.
5. **Table 3**: In the heading under “Women” “p” is missing. In the foot-note the number of women and men out of 99 patients referred to cardiologist should be stated in order to help the reader.

   **Answer**: We agree, the missing "p" has been corrected and the numbers of referred women (40) and men (59) are provided in the footnote to the table, which is provided as additional data in the revised manuscript (Additional file 3).

6. **Section: Recruitment and follow up Line 152 and following**. It would be interesting to know if there were any guidelines, to support GPs’ decision of referral for exercise testing and for referral to the cardiologist, for patients with suspected angina pectoris by the time of the study.

   **Answer**: This information is provided in the revised manuscript, section Methods, subsection Recruitment and Follow-up, lines 135-138: "At the time of the study there were no local guidelines to support GPs in the use of exercise tests or for referral to a cardiologist after testing. Enrolled study patients were referred due to suspected ischaemic heart disease at the discretion of the GPs who evaluated patients that accessed primary care."

7. In general the data appears to be sound. However the sub-groups are often small making the interpretation of findings less valid. This demands more studies in the area, which is also meritoriously mentioned by the authors (Line 336-337).
**Answer:** We agree with the reviewer’s comment. In a larger study sample, more robust conclusions could have been drawn. Please see the revised Discussion, subsection Strengths and Limitations, lines 320-322: "Some of our findings are close to the level of significance or are based on small numbers. Conclusions based on such results must be supported by other studies, preferably with larger samples, to allow for more robust conclusions."

8. How many patients returned with positive, inconclusive, negative or non-assessable exercise test results? This information is possible to find in Table 1. I think it should also be clearly stated in the text in order to make the reporting of results easier to follow.

**Answer:** We now provide this information in the revised manuscript. Please see the Results section, lines 217-219: "Patients with positive (n=55), inconclusive (n=142), negative (n=653), or non-assessable (n=15) exercise tests were referred to cardiologic evaluation in 67.3%, 26.1%, 3.5%, and 13.3% of cases, respectively."

9. In general the manuscript adhere to the relevant standards for reporting and data deposition. However, I think the authors should consider to reduce the number of tables e.g. Table 4 may be omitted.

**Answer:** We will carefully consider the total number of tables needed, in the light of comments from all reviewers. We suggest that Table 3 is provided as an additional file (Additional file 3).
10. Section: Interpretations of findings. Line: 358-362. Another interpretation of the finding that previous revascularization was associated with referral among women but not men, is that GPs may find it more difficult to evaluate symptoms, ECG-findings and results from exercise tests in women and therefore refer women more often when there is evidence of coronary heart disease in the history.

Answer: We agree that there are several possible interpretations of these data. Evaluation of symptoms and findings from various tests are somewhat subjective and are open to bias, to a certain extent. We addressed these issues in the revised Discussion, lines 344-346: "GPs may also find it more difficult to evaluate symptoms and physiologic test results in women, and therefore refer women more often when there is clear evidence of coronary disease in the patient’s history."
Reviewer: Stefan Bösner

Introduction: 1. In the introduction the authors describe both the utility of the exercise test for prognosis (e.g. predictive value of a negative exercise test with respect to cardiovascular events) and for diagnosis of CHD. It is difficult for the reader to distinguish these two aspects and it would be helpful to divide this more. The post test probability of a positive or negative exercise test in regard to CHD highly depends on the respective pre test probabilities which can be derived from different criteria of the patient history (Exercise test: LR + of 3.1 and LR – of 0.4 in patients with chest pain). This could be described better in the introduction and clearly be separated from the predictive value.

Answer: We agree that this issue could be described more clearly. In the revised manuscript, we more precisely describe the prognostic importance of negative exercise tests in age groups over and under 60 years. Please see the revised Background, lines 68-73: "The prognostic reliability of a negative exercise test was reported in a two-year follow-up conducted in Finnish primary care; 2% of patients aged less than 60 years and 3% of patients aged 60 years and older were diagnosed with coronary disease at the end of follow-up. As a diagnostic instrument for coronary disease, the usefulness of exercise testing depends on patient characteristics such as age, gender, and type of chest pain; best diagnostic yield is achieved when the pre-test probability of coronary disease is intermediate (15-65%)."

Methodology:
2. Line 138: Could you describe in one sentence the inclusion criteria (chest pain? Check-up in patients with known CHD?). Or was it left totally to the discretion of the referring GP whom to send for testing?

Answer: Please see the revised Method section, subsection Recruitment and Follow-up, lines 137-138, where this information is now provided: “Enrolled study patients were referred due to suspected ischaemic heart disease at the discretion of the GPs who evaluated patients that accessed primary care.”

3. Line 153: which patients were ‘potentially eligible’? (see comment above)

Answer: Potentially eligible patients were patients referred to exercise testing by GPs in the County of Jämtland, Sweden, from February 2010 until the end of February 2012. The final enrolled cohort consisted of consenting patients who were able to undergo exercise testing and were referred due to suspected ischaemic heart disease. Please refer to the revised manuscript, section Methods, subsection Recruitment and Follow-up, lines 134-135: "We invited potentially eligible patients referred to exercise testing from GPs from February 2010 until the end of February 2012 to participate in the current investigation."

4. Exercise test classification: Why did you count only a depression of the ST segment >0.1 mV and chest pain as a positive test? In other classifications the ST depression is sufficient. What about dyspnoea, delayed increase of the heart rate or blood pressure? Did these also count in your classification? Please also indicate the prerequisites for interpretation (e.g. the exercise threshold that had to be reached)
**Answer:** In this observational study, exercise tests were assessed according to the classification system that was in use during the study period, and which remains in use in 2014. In addition to chest pain and ECG changes, heart rate, blood pressure, and level of exertion were monitored during the test. The prerequisite for interpretation was not a fixed threshold of physical strain, yet we sought to conduct a maximal test in each case. The exercise test procedure and our classification of exercise tests were in accordance with national guidelines and were described in a previous paper from our group (revised manuscript, references 8, 30-32). Please see section Methods, subsection Exercise Test Classification, line 176: "... a complete description of the exercise test procedure and classification has been published elsewhere [8, 30-32].” See also the revised Discussion, subsection Strengths and Limitations, lines 293-297: "Compared to the complexity of coronary disease, any classification system based on exercise tests is a simplification. Classification into three categories yields a more complex result than classification into positive and negative tests. However, in the regression analysis of referrals, we implemented a bivariate approach because both positive and inconclusive test results need to be considered for patient management."

From the description of exercise test procedures and classification in reference (8): "**Exercise test procedure:** Exercise tests were conducted by nursing staff and physicians trained in clinical physiology. Exercise tests were conducted with a bicycle ergometer according to national guidelines [14, 15]. The initial workload was 30-50 W, considering the patient’s age, gender, and physical condition, with the aim that the patient should exercise for 6-10 min. The workload was increased by 10-20 W per min depending on the initial workload. We used a Cardiolex EC Sense for ECG recordings and a Rodby RE990 ergometer bicycle. A 12-lead ECG with computer-assisted reading of mean ST-segment depression was
registered during and one and four minutes after exercise. Visual assessment of the ECG recording was possible during the entire test. Systolic blood pressure, respiratory rate, symptoms of chest pain, and level of exertion were registered every two minutes (blood pressure) to every sixth minutes of exercise, until the end of the test. During the exercise test, chest-pain symptoms and perceived exertion were registered according to validated rating scales [32]. Patients were monitored until four minutes after exercise. The referring GP received a statement of the test results and held future responsibility for patient care.

**Classification of exercise tests:** The physician responsible for the test classified the exercise tests as positive, non-conclusive, negative, or non-assessable. A positive exercise test was defined as a horizontal or down-sloping depression of the ST segment >0.1 mV at 60 ms after the inflection point between the QRS and ST segments (J point) and chest pain suggestive of angina during the test. Non-conclusive tests were characterized by chest-pain symptoms or ST-segment depression during the test. Negative tests involved neither chest pain nor ST-segment depression during the test. Exercise tests in which the ST segment not could be assessed due to left bundle branch block, pacemaker, or digitalis medication were classified as non-assessable."

**5. Reference diagnosis:** How was this made in patients that did not undergo further cardiology work up. Did the researchers just take the respective entries out of the GPs documentation; or did a panel group discuss results of follow up examinations and then came to a conclusion? Please explain this better or insert a short paragraph named ‘reference diagnosis’
**Answer:** In this observational study, we did not use a panel group for reference diagnosis of ischaemic heart disease. Decisions on referral of patients to cardiology care were left to the discretion of the GPs, and the study group did not intervene in patient care. Please see the revised Method section, subsection Recruitment and Follow-up, line 151: “Our aim was to reflect normal care in which the exercise test results are signed and approved by a GP at the referring unit.” Please see also the Discussion, subsection Strengths and Limitations, lines 317-320: "We did not use a panel group for reference diagnosis in patients that did not undergo further cardiology work-up. Some cases of significant coronary disease probably remained undetected in patients who were not further examined after exercise testing. In an observational study such as this one, it is not possible to avoid that type of limitation."

**Discretionary Revisions 1.** You could formulate a clearer study question (“to assess the importance of exercise tests” is too vague).

**Answer:** Yes, as in the Abstract, we may use the term "utility" instead of "importance" to pose a clearer study question. Please see the last paragraph of the Background, lines 111-112, in the revised manuscript: "Our primary objective in this study was to assess the utility of exercise tests in selecting primary care patients for referral to further cardiology evaluation."

2. Results: Line 233 “Seventy-nine patients underwent coronary angiography”: it would be interesting to know the respective results (%age of patients where CHD was confirmed)

**Answer:** We have now added this information to the revised Results section,
In 63.3% of all patients with coronary angiography, coronary disease was confirmed.
Reviewer: Susan Smith

1. My one comment is that the presence of a positive exercise test is the strongest predictor of referral – this is unsurprising given that the test is used clinically to determine referral. The key result is that, in the adjusted analyses, there is no statistical difference in referrals based on age or gender.

   Answer: Please see the revised Discussion, line 272: "Overall, there were no significant differences in referral rates related to gender, socioeconomic status, or age." Our primary aim, which was to determine the utility of exercise tests for referral compared to clinical characteristics, could be of interest to readers even if our primary finding is not surprising.

2. I found it quite difficult to follow the denominators being presented. The aims do not clearly specify that the results presented refer only to those who have been referred. The methods state ‘all patients were followed up’ and the proportions presented in the results of the abstract actually only refer to the much smaller numbers who were referred (n=99). I am not sure that an analysis restricted to those referred addresses the primary objective as I would have thought this requires analysis of the full cohort? The authors do mention lack of follow up in those with negative exercise tests as a limitation but they say in other places that they were able to examine subsequent events in those not referred? – just needs to be more clearly presented

   Answer: All study participants (n=865) completed follow-up. We have expressed this more clearly in the revised Results section, lines 213-214: "Of 865 study
patients examined with clinical exercise testing upon referral from GPs, all completed follow-up. By six months, 99 patients were referred from primary care to evaluation at a heart clinic." Also see lines 217-219: "Patients with positive (n=55), inconclusive (n=142), negative (n=653), or non-assessable (n=15) exercise tests were referred to further cardiologic evaluation in 67.3%, 26.1%, 3.5%, and 13.3% of cases, respectively."

Please also see the revised Abstract, lines 37-38: "Of 865 participants, patients with positive, inconclusive, or negative exercise tests were referred to cardiologists in 67.3%, 26.1%, and 3.5% of cases, respectively."

Concerning follow-up, we clarify this in the revised Discussion, lines 317-329: "We did not use a panel group for reference diagnosis in patients that did not undergo further cardiology work-up. Some cases of significant coronary disease probably remained undetected in patients who were not further examined after exercise testing. " Even with a complete follow-up of the medical records of all study participants, some cases of significant coronary disease probably remained undetected in patients who were not further examined in primary care or at the hospital after exercise testing. In an observational study such as this one, it is not possible to avoid that type of limitation.

3. Please clarify patient selection and the definitions for a positive stress exercise test. It is also important for an international readership to know what the local referral guidelines are for exercise tests. In my clinical setting, GPs don’t have direct access to stress testing and I would have presumed that all those with a positive or even equivocal exercise test would be referred as these people usually have further investigations such as angiography or isotope scanning when they have a positive stress test here.
Answer: Patient selection is now clarified in the revised Method section, subsection Recruitment and Follow-up, lines 134-140: "We invited potentially eligible patients referred to exercise testing from GPs from February 2010 until the end of February 2012 to participate in the current investigation. At the time of the study there were no local guidelines to support GPs in the use of exercise tests or for referral to a cardiologist after testing. Enrolled study patients were referred due to suspected ischaemic heart disease at the discretion of the GPs who evaluated patients that accessed primary care. Of the 1191 potentially eligible patients, 265 declined to provide consent, eight were unable to carry out exercise testing, and 53 were referred for reasons other than ischaemic heart disease. The study group thus consisted of 865 patients, 438 men and 427 women."

Exercise test classification: Please see the revised Method section, lines 180-184: "Positive exercise tests were associated with a depression of the ST segment >0.1 mV, horizontal or down-sloping, and chest pain indicating angina during the test. Inconclusive tests were associated with either chest pain or ST-segment depression during the test. Patients with negative tests experienced neither chest pain nor ST-segment depression. ECG reactions impossible to assess due to left bundle branch block, pacemaker, or digitalis medication were classified as non-assessable." Also see line 176: "...a complete description of the exercise test procedure and classification has been published elsewhere [8, 30-32]."

Our previous report [8] from the same cohort includes a complete description of the exercise test classification: "A positive exercise test was defined as a horizontal or down-sloping depression of the ST segment >0.1 mV at 60 ms after the inflection point between the QRS and ST segments (J point) and chest pain suggestive of angina during the test. Non-conclusive tests were characterized by chest-pain symptoms or ST-segment
depression during the test. Negative tests involved neither chest pain nor ST-segment depression during the test. Exercise tests in which the ST segment could not be assessed due to left bundle branch block, pacemaker, or digitalis medication were classified as non-assessable.

4. The study design should be described in the abstract

**Answer:** In the revised Abstract, lines 30-31, we describe the study design: "We designed a prospective observational study of 438 men and 427 women from 28 Swedish primary-care clinics who were examined with exercise testing for suspected coronary disease."

5. I am concerned at the multiple testing with some very small numbers in some small subgroups. For example, much is made of the difference between self-employed women and other groups but there are only 32 self-employed women in the cohort

**Answer:** We agree with the reviewer that there are small numbers in some groups. We stress this as a limitation of our findings in the revised Discussion, subsection Strengths and Limitations, lines 320-322: "Some of our findings are close to the level of significance or are based on small numbers. Conclusions based on such results must be supported by other studies, preferably with larger samples, to allow for more robust conclusions." As one of our study aims, we sought to determine whether referral decisions were biased by gender or socioeconomic status. Our findings must be supported by future research.
6. There is a presumption in the discussion that lack of referral is problematic. It may be appropriate based on multi morbidity or patient preference as would be suggested by the age range but it does seem surprising that such a decision would not be recorded in the records.

**Answer:** Yes, we were surprised that the reason for not referring some patients with a positive exercise test was not given in some of the records. The records were scrutinised extensively, and all recorded data were evaluated in our study. Even though the question of continuity problems in primary care was beyond the scope of our study, our hypothesis remains as stated. As we mentioned in the revised Discussion, lines 347-350: "The reasons not to refer patients with a positive exercise test to cardiologic evaluation remain obscure; in 56% of such cases, the records did not provide any data reflecting the test result or actions taken. These cases are likely linked to continuity problems in primary care, with GPs working on short-term contracts. A more thorough exploration of this issue is beyond the scope of the present investigation."

7. Some of the Tables seem unnecessary. I don't think presenting the unadjusted ORs is helpful (Table 3).

**Answer:** Yes, we have decided to provide Table 3 as additional data (Additional file 3). In the revised Results section, lines 242-244 now read: "The unadjusted ORs and p values for referral of women and men, according to patient characteristics, are provided as additional data (Additional file 3)."
8. Need to be consistent in use of term sex or gender

   **Answer:** In the revised manuscript the term “gender” is now consistently used in tables and text.

9. I am not familiar with their term socioeconomy in English and suggest the term socioeconomic status is more usual

   **Answer:** In the revised manuscript we now use the term “socioeconomic status” consistently.