Diagnostic accuracy of exercise stress testing for coronary artery disease: a systematic review and meta-analysis of prospective studies
Banerjee A, Newman DR, Van den Bruel A, Heneghan C

CRD summary
This generally well-conducted review concluded that exercise testing was more useful at excluding coronary artery disease than confirming it. The reliability of the conclusions is compromised by the low predictive values of likelihood ratios of most tests, potential for missed studies, and methodological limitations of some of the included studies.

Authors' objectives
To determine the diagnostic accuracy of exercise stress testing for coronary artery disease on angiography.

Searching
Medion, MEDLINE, EMBASE, Central Register of Controlled Trials (CENTRAL) and The Cochrane Library were searched for English language articles indexed between 1966 and 2009; search terms were reported. Guidelines were identified using a search of the TRIP database. Bibliographies of included articles were screened for additional studies.

Study selection
Prospective studies were eligible if they assessed the presence or absence of coronary artery disease using exercise electrocardiogram (ECG) or echocardiogram compared to angiography, and post mortem or development of acute coronary syndrome during long-term follow-up. Studies had to recruit a representative adult (over 18 years) population, with either symptoms suggestive of, or risk factors for, coronary artery disease. Studies had to blind investigators to the results of coronary angiography and/or those performing angiography to the results of prior stress tests. Sufficient data had to be available to construct 2x2 tables of test performance. Studies of acute coronary syndrome, cardiomyopathies, valvular heart disease and cardiac failure, studies using pharmacological stressors, studies of imaging modalities without comparison, and studies of patients with previous myocardial infarction or known coronary artery disease were excluded.

Most studies were conducted in North American or Western European countries. The prevalence of coronary artery disease ranged from 0.07 to 0.83. Where reported, the mean age ranged from 45 to 68. Most of the studies used a threshold for ECG of 0.1mV or 1mm ST depression for 0.06 to 0.08 seconds. For the reference standard, the threshold of stenosis ranged from 50% to 75%.

Full papers were selected by two reviewers, with disagreements resolved by a third; it was unclear whether the titles and abstracts were also screened in this manner.

Assessment of study quality
Study quality was assessed using a 12-point QUADAS tool; representativeness of the patient spectrum (definition not given) and validity of the reference standard were used as inclusion criteria. If insufficient data were given to be confident that a criterion had been met, it was assessed as not being met. Studies were rated as: A if fulfilling all criteria; D if subject to partial verification bias or had unblinded interpretation of the index test; or C if subject to incorporation or progression bias, or had unblinded interpretation of the reference standard. All other studies were rated as B.

Two independent reviewers assessed study quality. Disagreements were resolved by discussion.

Data extraction
Two reviewers extracted the prevalence of coronary artery disease and data to construct 2x2 tables of test performance. Sensitivity, specificity and positive and negative likelihood ratios (LR+/-) were calculated.

Methods of synthesis
Hierarchical summary receiver operating characteristic (HSROC) curves were produced; these were available in online
appendices. Pre- and post-test probabilities of coronary artery disease were calculated for each study. Where a study evaluated multiple exercise tests in the same population, only one was included in the analysis; it was not reported how the choice of test was made. Subgroup analyses investigated the type of exercise test used, gender, age, prevalence of coronary artery disease and double-blinding.

**Results of the review**
Thirty-four studies were included in the review (3,352 participants; range 19 to 503). Of the 34 studies, 14 were classified on the quality assessment as A, 10 as B, three as C and seven as D.

Overall, echocardiography performed better than ECG for ruling in (echocardiography: LR+ 8.57; ECG: LR+ 3.22) and ruling out (echocardiography: LR- 0.19; ECG: LR- 0.38) coronary artery disease. This was true of treadmill tests (echocardiography: LR+ 7.94, LR- 0.24; ECG: LR+ 3.57, LR- 0.34) and bicycle tests (echocardiography: LR+ 11.34, LR- 0.17; ECG: LR+ 2.94, LR - 0.4). The bicycle echocardiography was the best when ruling coronary artery disease out (LR- 0.17).

Results from a range of other subgroup analyses are also given.

**Authors’ conclusions**
The diagnostic accuracy of exercise testing varied depending on the age, gender and clinical characteristics of the patient, prevalence of coronary artery disease and modality of test used. Exercise testing, by echocardiography or ECG, was more useful at excluding coronary artery disease than confirming it.

**CRD commentary**
The authors addressed a clear review question supported by appropriate inclusion criteria. Several relevant sources were searched. Only studies in English were included, and there was no specific search for unpublished studies, so language and publication bias could not be ruled out. Diagnostic terminology was used during the search which may have resulted in some studies being missed. It appeared that most of the review process was conducted in duplicate, which reduced the potential for error and bias, although this was unclear for the title and abstract stage of screening. Appropriate quality criteria were used, and although the results were used to give an overall grading, the importance of different biases were considered during this process. Appropriate methods of synthesis were used.

The conclusions drawn reflected the evidence presented, but only two of the likelihood ratios exceeded the commonly used thresholds for a strong prediction (LR+ greater than 10 and LR- less than 0.2). Overall, this was a well-conducted review, but the reliability of the conclusions is compromised by the low predictive values of most tests, potential for missed studies, and methodological limitations of the included studies.

**Implications of the review for practice and research**
**Practice:** Authors stated that the findings of the review were important for the choice of test modality in any particular clinical setting, and could help clinicians individualise coronary artery disease diagnosis.

**Research:** Authors stated that more prospective cohort studies of diagnostic tests in the setting of coronary artery disease were required, considering both the diagnostic and prognostic implications side by side.

**Funding**
National Institute for Health Research.

**Bibliographic details**

**PubMedID**
22512607

**DOI**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.