Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis


CRD summary
This well-conducted review evaluated the accuracy of brain natriuretic protein (BNP), N terminal-proBNP and electrocardiography in the diagnosis of left ventricular systolic dysfunction (LVSD). The authors reliably concluded that no one test can be recommended over another or combinations of tests. It is debatable whether the data support the conclusion that these tests can be useful in excluding a diagnosis of LVSD.

Authors' objectives
To evaluate the accuracy of electrocardiogram (ECG), brain natriuretic peptide (BNP), N terminal-pro brain natriuretic peptide (NT-proBNP) and combinations of tests for the diagnosis of left ventricular systolic dysfunction (LVSD) in the primary care setting.

Searching
MEDLINE, EMBASE and the Cochrane Library were searched from 1980 to March 2004; the search terms were reported. Relevant conference proceedings and two cardiology journals were handsearched. Citation searches of identified included studies and reviews were also conducted. The authors did not state whether any language restrictions were applied, although 11 studies were excluded as no translation was available.

Study selection
Study designs of evaluations included in the review
Diagnostic cohort studies were eligible for inclusion.

Specific interventions included in the review
Studies of ECG, BNP or NT-proBNP, either alone or in combination, were eligible for inclusion. Studies that assessed the diagnosis of acute decompensated heart failure or ventricular diastolic dysfunction alone were excluded. In the included studies, ECG reporting was performed by secondary care physicians, by primary and secondary care physicians, or was automated. The definition of an abnormal ECG varied but the majority of studies deigned six or more ECG abnormalities as abnormal. The definition of an abnormal BNP results ranged from 5 to 49 pmol/L. The definition of an abnormal NT-proBNP ranged from 5 to 250 pmol/L.

Reference standard test against which the new test was compared
Studies that included nuclear cardiology investigative techniques or 2-D echocardiography, and that defined LVSD using a quantitative or qualitative measure of ejection fraction, were eligible for inclusion. The reference standards used in the included studies were nuclear cardiology, nuclear cardiology or echocardiography, echocardiography and symptoms of chronic heart failure, and echocardiography alone. Studies that used echocardiography employed a range of measurement methods (both quantitative and qualitative). Where reported, the threshold for an abnormal ejection fraction ranged from 30 to 50%.

Participants included in the review
Studies of adults suspected of having LVSD with or without co-morbid conditions were eligible for inclusion. Studies in which the majority of the population had been on long-term treatment with angiotensin-converting enzyme inhibitors and/or diuretics for presumed heart failure were excluded. Where reported, the mean age of participants in the included studies ranged from 53 to 79 years. The studies were conducted in both primary and secondary care settings. LVSD prevalence ranged from 4.5 to 83% in primary care settings and from 23 to 55% in secondary care settings.

Outcomes assessed in the review
Studies had to provide sufficient data to construct 2x2 tables of test performance to be included. The primary outcome measures reported in the review were the sensitivity and specificity.

How were decisions on the relevance of primary studies made?
One reviewer screened titles and abstracts. Potentially relevant articles were obtained and reviewed by at least two independent reviewers, with any disagreements resolved by a third reviewer.

**Assessment of study quality**
Two reviewers independently assessed studies for methodological quality, with any disagreements resolved by a third reviewer. Studies were assessed for the following quality domains: selection bias, verification bias, measurement bias and treatment paradox and disease progression bias.

**Data extraction**
Two independent reviewers extracted the data, with uncertainty resolved by a third reviewer. Data about the disease spectrum (prevalence of LVSD), reference standard and 2x2 data were extracted into a spreadsheet. Where multiple thresholds were provided within a study, the worst and best estimates of the diagnostic odds ratio for the individual study were extracted.

**Methods of synthesis**
How were the studies combined?
Pooled sensitivities and specificities were calculated using the DerSimonian and Laird random-effects model in the absence of a threshold effect. Where there was evidence of a threshold effect but no evidence of heterogeneity, summary sensitivity and specificity were derived from summary receiver operating characteristic curves, calculated using the Moses-Littenberg method. If there was evidence of heterogeneity, ranges of sensitivities and specificities were reported and pooling was not undertaken. Worst and best estimates were used to derive pooled estimates.

How were differences between studies investigated?
Heterogeneity was assessed using the chi-squared and I-squared statistics. Where the p-value was less than 0.01 and I-squared was 50% or less, the data were considered sufficiently homogeneous to pool the results. A threshold effect was also investigated (method details not reported).

**Results of the review**
Thirty-two studies were included. Fourteen assessed ECG, sixteen assessed BNP, seven assessed NT-proBNP, two assessed BNP and ECG combined, and three directly compared ECG and BNP in the same population. Some studies investigated more than one test.

The quality of the included studies was variable. Studies were particularly subject to verification bias, which is likely to overestimate test accuracy. The majority of ECG studies appear to have been affected by selection bias, which is likely to overestimate sensitivity.

ECG (14 studies): there was no evidence of a threshold effect (p<0.639). The sensitivity ranged from 42% (corresponding specificity 87%) to 98% (corresponding specificity 66%). The majority of studies reported sensitivities above 80%; estimates of specificity were less good and more heterogeneous. Pooled estimates were not calculated because of the significant heterogeneity.

BNP (16 studies): there was evidence of a threshold effect (p<0.03). The sensitivity ranged from 20% (corresponding specificity 89%) to 100% (corresponding specificity 47%). The majority of studies reported sensitivities above 80%; estimates of specificity were less good and more heterogeneous. Pooled estimates were not calculated because of the significant heterogeneity.

NT-proBNP (7 studies): there was evidence of a threshold effect (p<0.001). The sensitivity ranged from 25% (corresponding specificity 95%) to 98% (corresponding specificity 23%). The majority of studies reported estimates of sensitivity above 80% with poorer and more heterogeneous specificity. Pooled estimates were not calculated because of the significant heterogeneity.

Direct comparison of ECG and BNP (3 studies): the sensitivity was similar for both tests in all 3 studies, whereas 2 studies showed greater specificity with BNP and one study showed no difference.
ECG combined with BNP compared with BNP alone and ECG alone (2 studies); both studies showed a greater specificity for the combination compared with ECG alone, but no improvement in sensitivity.

**Cost information**
The authors stated that the estimated cost of an ECG is £10 compared with approximately £20 for a natriuretic peptide test.

**Authors’ conclusions**
ECG, BNP and NT-proBNP are useful in excluding a diagnosis of LVSD. There is insufficient evidence to recommend the use of one test over another, or of using tests in combination.

**CRD commentary**
This was a well-conducted and clearly reported review. The objective was defined and supported by defined inclusion criteria. The literature search was reasonable although studies in some languages appear to have been excluded, thus the review may be subject to language bias. Details of the review process were reported and included appropriate steps to minimise errors and bias. Only limited study details were reported in the text; further details were published in supplementary material available on the journal website. The methods used to analyse the results were appropriate and included an investigation of heterogeneity. It is questionable whether the results showed that these tests have sufficient accuracy to justify the authors’ conclusions that these tests can be useful to exclude a diagnosis of LVSD. The conclusion that there is insufficient evidence to recommend one test over another, or the use of combinations of tests, is supported by the evidence presented.

**Implications of the review for practice and research**
Practice: The authors stated that ‘either BNP, NT-proBNP or the ECG should be used as part of the diagnostic work up of individuals with suspected chronic heart failure and that there is no evidence to justify the use of both tests’. They also stated that use of abnormal test results to select patients for echocardiography could overwhelm services.

Research: The authors stated that further research needs to directly compare the diagnostic performance of these tests in homogeneous representative primary care populations. They also stated that further work based on existing research would require the use of individual patient data, which would allow the investigation of variation in accuracy with changes in patient characteristics, application of the reference standard, and threshold used to define a positive result.

**Funding**
Not stated.

**Bibliographic details**

**PubMedID**
16438815

**Other publications of related interest**
These additional published commentaries may also be of interest.
Doust J. Review: electrocardiography, BNP, and N terminal-pro BNP are more sensitive than specific for chronic left ventricular systolic dysfunction. ACP J Club 2006;145:22.
Doust J. Review: ECG, BNP, and N terminal-pro BNP are more sensitive than specific for chronic left ventricular systolic dysfunction. Evid Based Med 2006;11:117.

**Indexing Status**
Subject indexing assigned by NLM
MeSH
Adult; Cohort Studies; Electrocardiography /standards; Humans; Natriuretic Peptide, Brain /blood; Quality Assurance, Health Care; Reproducibility of Results; Sensitivity and Specificity; Ventricular Dysfunction, Left /diagnosis

AccessionNumber
12006000631

Date bibliographic record published
21/02/2007

Date abstract record published
30/09/2008

Record Status
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.