Meta-analysis of B type natriuretic peptide and N-terminal pro B natriuretic peptide in the diagnosis of clinical heart failure and population screening for left ventricular systolic dysfunction

Ewald B, Ewald D, Thakkinstian A, Attia J

CRD summary
The authors concluded that B-type natriuretic peptide (BNP) was helpful in diagnosis and screening for heart failure and was superior to N-terminal pro-BNP. These conclusions should be interpreted with caution due to the possibility of publication bias and because data on comparative accuracy of the tests were based on a small number of inconsistent studies.

Authors' objectives
To determine the accuracy of tests for B-type natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP) for diagnosis of clinical heart failure in primary care and hospital settings and in population screening for left ventricular systolic dysfunction.

Searching
MEDLINE and EMBASE were searched from 2002 to mid 2005. The search strategy was available as a web appendix. Experts in the field were contacted to identify additional studies. No language restrictions were applied. Studies published before 2002 were identified through a previous meta-analysis (see Other Publications of Related Interest).

Study selection
Prospective diagnostic cohort studies that enrolled patients with unknown disease status and compared BNP or NT-proBNP with a reference standard were eligible for inclusion. Studies had to report sufficient data to allow construction of a 2x2 table of test performance. All patients had to receive both the index test and reference standard.

Studies were conducted either in symptomatic patients (cardiology referrals from general practice, general practice patients with suspected heart failure, breathless outpatients, emergency department patients, in-patients with suspected heart failure, intensive care patients) or in screening settings. In most studies the reference standard for a diagnosis of clinical heart failure was final diagnosis by a cardiologist or consensus of three cardiologists after reviewing all clinical data, tests and response to treatment. A variety of assays were used to assess BNP, NT-proBNP was assessed using the Roche NT-proBNP test or in-house assays. Thresholds varied across studies; some used age- and gender-specific thresholds.

The authors did not state how many reviewers performed the inclusion assessment.

Assessment of study quality
Study quality was assessed according to whether interpretation of the index test and reference standard was blinded, whether a detailed description of the method and criteria for both the index test and reference standard was reported and whether the index test and reference standard tests were performed on the same day.

It was unclear how many reviewers performed the quality assessment.

Data extraction
Two reviewers independently extracted data as 2x2 tables of test performance and used this to calculate diagnostic odds ratios (DOR), sensitivity and specificity. Where a 2x2 table had 0 cells, 0.5 was added to each cell. The ratio of DORs was calculated for studies that compared BNP with NT-proBNP directly. If age-specific thresholds were presented, data were extracted at these thresholds unless a single threshold gave better performance. If multiple thresholds were reported, the threshold that gave the greatest sum of sensitivity and specificity was used. When sex-specific thresholds were used, data were extracted separately for men and women and included as separate groups in the analysis.
multiple assays were used for each peptide, those of commercially available ones were selected. Discrepancies were resolved through discussion with a third reviewer.

**Methods of synthesis**
Summary DORs and ratios of DORs were estimated using random-effects models. Heterogeneity was assessed using the Breslow-Day method and quantified using the $I^2$ statistic. Meta-regression was used to investigate heterogeneity. The effect of study quality was assessed by removal of studies with a poor score on two of the three items assessed. Publication bias was assessed using the Egger test.

**Results of the review**
Forty-nine studies were included in the review (n=23,886). Study quality was generally described as adequate, although delays of up to a year between the index test and reference standard were reported in two screening studies.

Detection of heart failure in symptomatic patients (27 studies, n=7,062):
There was significant heterogeneity between the 15 studies ($p<0.001$, $I^2=79\%$) of BNP. Meta-regression showed that the DOR was decreased by a factor of 2.0 (95% CI 1.0 to 4.2) for each additional decade of age for studies of BNP. There was no significant effect of sex balance, in-patient and out-patient status, width of distribution of age and the use of a prespecified threshold. The results of nine studies of BNP in patients with a mean age of 80 years or less showed a summary DOR of 27.7 (95% CI 21.6 to 35.6). There was no evidence of heterogeneity ($p=0.34$, $I^2=5\%$).

There was significant heterogeneity between the nine studies of NT-proBNP ($p<0.001$, $I^2=68\%$). Meta-regression showed that the DOR was decreased by a factor of 2.5 (95% CI 1.7 to 3.7) for each additional decade of age for studies of BNP. After removal of one poor-quality study, the results studies of NT-proBNP in patients with a mean age of 80 years or less showed a summary DOR of 37 (95% CI 26.6 to 51.6). There was no evidence of heterogeneity ($p=0.32$, $I^2=12\%$).

Community screening studies (13 studies, n=14,590):
Studies that assessed BNP and used a severe definition of left ventricular systolic dysfunction (ejection fraction ≤40%) showed a summary DOR of 19.9 (95% CI 12.5 to 31.9). There was some evidence of heterogeneity ($p=0.14$). Meta-regression showed no significant effect of age, quality or age distribution on the DOR. Studies that assessed two levels of left ventricular systolic dysfunction severity consistently showed higher DOR for severe disease than for mild disease, which suggested that the assay was more accurate for more severe disease. There were only three studies of NT-proBNP and these showed a summary DOR of 9.3 (95% CI 4.7 to 19.0).

Mixed samples (eight studies, n=2,234):
Results of these studies varied widely and generalisability was unclear.

Comparison of assays (seven studies):
Accuracy of BNP was better than that of NT-proBNP in five studies; in two studies accuracy of NT-proBNP was greater. The pooled ratio of odds ratios showed that BNP was significantly more accurate than NT-proBNP (OR 1.77, 95% CI 1.06 to 2.95).

**Authors' conclusions**
Tests for BNP were helpful in the diagnosis of clinical heart failure or in screening for left ventricular systolic dysfunction and were superior to NT-proBNP. In the clinical setting, test performance declined with increasing patient age.

**CRD commentary**
The study objective was clearly defined and supported by relevant inclusion criteria. The literature search was adequate for published studies. It was unclear whether unpublished data were eligible and no specific attempts were made to locate such studies, so there was a possibility of publication bias. This was assessed in the review, but the methods used were not appropriate for diagnostic accuracy data. Appropriate steps were taken to minimise bias and error during data extraction; it was unclear whether such steps were taken during study selection and quality assessment. Study quality was assessed with appropriate criteria and considered in the review synthesis. Appropriate methods were used to pool data. Heterogeneity was investigated.

The authors' conclusions were supported by the data, but should be interpreted with some caution due to the possibility of publication bias. Conclusions related to the comparative accuracy of the two tests should be interpreted with care as these were based on a small number of inconsistent studies.

**Implications of the review for practice and research**
The authors did not state any implications for practice or research.

**Funding**
National Health and Medical Research Council of Australia.

**Bibliographic details**

**PubMedID**
18290826

**DOI**
10.1111/j.1445-5994.2007.01454.x

**Original Paper URL**
http://onlinelibrary.wiley.com/journal/119419339/abstract

**Other publications of related interest**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Aged; Aged, 80 and over; Female; Heart Failure, Systolic /blood /diagnosis; Humans; Male; Mass Screening; Middle Aged; Natriuretic Peptide, Brain /blood; Peptide Fragments /blood; Reproducibility of Results; Sensitivity and Specificity; Ventricular Dysfunction, Left /blood /diagnosis

**AccessionNumber**
12008005748

**Date bibliographic record published**
03/11/2008

**Date abstract record published**
11/08/2010

**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.