B-type natriuretic peptide assays in the diagnosis of heart failure

---

Part A – in the hospital emergency setting – November 2006

Part B – in the non-hospital setting – May 2007

MSAC Application 1087

Assessment report

ISBN 1-74186-201-9

Online ISBN 1-74186-202-7

ISSN (Print) 1443-7120

ISSN (Online) 1443-7139

First printed February 2008

Electronic copies of the report can be obtained from the Medical Service Advisory Committee's internet site at http://www.msac.gov.au/

Printed copies of the report can be obtained from:

The Secretary
Medical Services Advisory Committee
Department of Health and Ageing
Mail Drop 106
GPO Box 9848
Canberra ACT 2601

Enquiries about the content of the report should be directed to the above address.

The Medical Services Advisory Committee (MSAC) is an independent committee which has been established to provide advice to the Minister for Health and Ageing on the strength of evidence available on new and existing medical technologies and procedures in terms of their safety, effectiveness and cost-effectiveness. This advice will help to inform government decisions about which medical services should attract funding under Medicare.

MSAC recommendations do not necessarily reflect the views of all individuals who participated in the MSAC evaluation.

This report was prepared on behalf of the Medical Services Advisory Committee by Ms Tracy Merlin, Mr John Moss, Dr Anthony Brooks, Ms Skye Newton, Ms Hedyeh Hedayati and Professor Janet Hiller from Adelaide Health Technology Assessment (AHTA), Discipline of Public Health, The University of Adelaide. Statistical expertise was provided by Ms Kristyn Willson and Mr Tom Sullivan. The report was edited by Ms Jo Mason from MasonEdit, Adelaide. Part A of the report was endorsed by the Minister for Health and Ageing on 5 February 2007; with Part B of the report endorsed by the Minister for Health and Ageing on August 29, 2007.

Publication approval number: P3-1288

Contents

Part A – B-type natriuretic peptide assays in the diagnosis and monitoring of heart failure in the hospital emergency setting

Part B – B-type natriuretic peptide assays in the diagnosis of heart failure in the non-hospital setting

Tables

Figures

B-type natriuretic peptide assays in the diagnosis and monitoring of heart failure in the hospital emergency setting

---

November 2006

MSAC Application 1087

Assessment report (Part A)

Executive summary

Part A of this report assesses the use of two B-type natriuretic peptide assays (BNP and NT-proBNP) in three key areas (diagnosis, monitoring and prognosis) for suspected and diagnosed heart failure (HF) patients, with the diagnostic use occurring in the hospital emergency setting. Part B of this report assesses the diagnostic use of the two B-type natriuretic peptide assays to rule out HF in patients presenting in a non-hospital setting.

Although the prognosis section adds weight to the argument for the potential usefulness of B-type natriuretic peptide assays in assessing confirmed HF patients, these agents would not receive an MBS listing for this purpose; hence, the relevant methodology and assessment is presented in Part A, Appendix I.

The procedure

B-type natriuretic peptide testing involves a blood test to determine the level of cardiac neurohormone circulating in the blood of a patient suspected or diagnosed with HF. Levels of two types of cardiac neurohormone can be tested—brain natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP).

These B-type natriuretic peptides act as counter-regulatory hormones to stabilise circulatory function. In an attempt to maintain cardiac output from a failing heart, the renin-angiotensin-aldosterone system is activated to enhance blood volume retention, circulatory vasoconstriction and ventricular remodelling in order to maintain ventricular pre-load. This physiological response to the failing heart actually increases the workload of the heart because of an increase in vascular resistance and after-load. The circulatory volume overload stretches cardiac myocytes which then release the B-type natriuretic peptides to stabilise circulatory function.

Both peptides have been implicated as diagnostic biomarkers for suspected HF in clinical practice. In this context it is suggested that assays or tests of these peptides may complement conventional diagnostic strategies and thus assist with the identification of symptomatic patients with *suspected* HF. Patients with low levels of the cardiac neurohormones are '**ruled out**' for HF through these tests and are investigated for differential diagnoses; those **not excluded** from HF may go on to other confirmatory testing such as an echocardiogram.

B-type natriuretic peptides have also been suggested to assist the conventional clinical monitoring of patients *with* HF, as the concentration of these peptides in stabilised HF patients has been shown to correlate with HF prognosis (see Appendix I).

Medical Services Advisory Committee – role and approach

The Medical Services Advisory Committee (MSAC) was established by the Australian Government to strengthen the role of evidence in health financing decisions in Australia. The MSAC advises the Minister for Health and Ageing on the evidence relating to the safety, effectiveness and cost-effectiveness of new and existing medical technologies and procedures, and under what circumstances public funding should be supported.

A rigorous assessment of evidence is thus the basis of decision-making when funding is sought under Medicare. A team from Adelaide Health Technology Assessment (AHTA), Discipline of Public Health, School of Population Health and Clinical Practice, University of Adelaide was engaged to conduct a systematic review of the literature on B-type natriuretic peptide assays in the diagnosis, monitoring and prognosis of HF. An advisory panel with expertise in this area then evaluated the evidence and provided advice to the MSAC.

MSAC's assessment of B-type natriuretic peptide assays

Clinical need

Heart failure is commonly cited to afflict 300,000 Australians, with approximately 30,000 new cases occurring each year. However, these figures underestimate the number of patients *suspected* of having HF each year and thus who would potentially receive a B-type natriuretic peptide test. Patients presenting with symptoms like acute dyspnoea (breathlessness) may have HF or, alternatively, chronic obstructive pulmonary disease, pneumonia, emphysema or other lung diseases.

Using data from a key diagnostic randomised controlled trial and Australian Institute of Health and Welfare data on the annual hospitalisation rate for HF, it is estimated that approximately 98,000 patients each year will present to an emergency department (ED) with acute dyspnoea (suggestive of HF) that warrants investigation with a B-type natriuretic peptide test.

Safety

The likelihood of adverse events occurring during B-type natriuretic peptide testing is small and similar to that of other venepuncture blood tests. False positive or negative test results may theoretically cause harm through respective sequelae such as inappropriate or delayed treatment. However, there were no studies in the available evidence base that reported physical or psychological adverse events as a result of B-type natriuretic peptide testing.

Effectiveness

Diagnosis

BNP assays

The effectiveness of supplementing conventional diagnostic assessment with BNP testing was evaluated by a large volume of evidence, with the highest quality evidence obtained from one good quality level II direct intervention study as well as two good quality level II diagnostic accuracy studies. Results indicate that the BNP test has a strong ability to discriminate between the presence and absence of HF in symptomatic patients (sDOR = 46.81, 95%CI 21.5, 102.0). Variation in diagnostic accuracy between studies was possibly due to the different test thresholds (cut-off points) employed in the studies for ruling out HF. Overall, the body of evidence was relatively consistent in its findings that the BNP blood test is sensitive with a high negative predictive value. Its main role,

therefore, appears to be as a 'first line' test, as a negative result on the test 'rules out' the diagnosis of HF.

The impact of the BNP test on patient management was found to be mainly through the alteration of emergency physician diagnoses that were initially uncertain or were secondary diagnoses (level IV intervention evidence). In those situations where the clinical diagnosis was equivocal, the BNP assay added diagnostic value and resulted in a subsequent change in patient management (28% in one study). High level and good quality evidence (level II intervention evidence) assessed the impact on clinical management of supplementing the usual clinical diagnostic workup with BNP testing. In an intention-to-treat analysis, BNP-assisted diagnostic assessment significantly shortened hospital stay by a median of 3 days (p=0.001), except for patients with kidney disease. Time to treatment was significantly reduced by just under a median of half an hour in the group receiving BNP supplemented diagnostic assessment. More importantly, patients in this group were admitted to hospital [RR=0.88, 95%CI 0.81, 0.97] and intensive care [RR=0.62, 95%CI 0.42, 0.91] less often than those patients receiving conventional diagnostic assessment alone. Thus, only approximately 10 patients would need to be diagnosed with a BNP-supplemented diagnostic workup, as compared to conventional diagnostic strategies, to reduce one hospital or intensive care admission. This is presumably a consequence of patients being ruled out from HF earlier in the clinical pathway as a result of the test.

The direct impact of BNP testing on patient health outcomes was assessed by the same good quality randomised controlled trial (level II intervention evidence). These health outcomes were pre-specified as secondary outcomes, however, and so the trial was not necessarily powered to find a statistically significant difference in health outcomes between the trial arms. With respect to in-hospital mortality and 30-day mortality, patients receiving a BNP-supplemented diagnostic workup had a reduced rate of death compared to those receiving conventional diagnostic strategies, although the difference was not statistically significant (in-hospital mortality: RR=0.62, 95%CI 0.32, 1.22, p=0.21; 30-day mortality: RR=0.79, 95%CI 0.47, 1.34, p=0.45). However, for the elderly subgroup of patients followed in the trial, a pre-specified subgroup analysis found a particular benefit with BNP-supplemented diagnostic assessment, with a trend towards a reduction in in-hospital mortality (RR=0.46, 95%CI 0.21, 1.03, p=0.051) and a statistically significant and clinically important reduction in 30-day mortality (RR=0.51, 95%CI 0.26, 0.98, p=0.039). The latter indicates that 12 elderly patients would require diagnosis with a diagnostic strategy including BNP testing, compared to conventional diagnosis without BNP testing, to prevent one death within 30 days.

The populations studied in the included diagnostic studies are applicable to the target population in Australia, that is patients presenting to an ED with symptoms (eg acute dyspnoea) suggestive of HF. The results of the studies are largely generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in diagnosing and managing symptomatic suspected HF.

In conclusion, on the basis of the evidence presented, BNP testing appears to be a valuable 'first line' diagnostic test that, when added to conventional diagnostic assessment, assists the acute care physician to correctly 'rule out' HF in patients presenting with symptoms suggestive of HF, such as acute dyspnoea and oedema. It also appears to benefit the patient by reducing or preventing hospital stay and decreasing the time to treatment, and has the potential to reduce mortality rates in the short term in some patients.

NT-proBNP assays

The effectiveness of NT-proBNP testing added to conventional diagnostic assessment was evaluated by a reasonable volume of evidence, with three good quality level II and several average quality level III diagnostic accuracy studies being available for the hospital setting. Overall, the body of evidence was relatively consistent in its findings that NT-proBNP assays are sensitive with high negative predictive values (>90%), indicating the test effectively 'rules out' HF in patients with a negative result.

There were no studies available on the impact of NT-proBNP testing on clinical diagnoses formulated in the hospital setting. There were also no studies available to assess the direct impact of NT-proBNP testing on patient health outcomes. High level evidence of the effect of early treatment of HF (from linked evidence) on patient health outcomes indicates that early treatment is beneficial, although this was not investigated systematically in this report. The effect of early treatment could not be determined for those patients who would be 'ruled out' from HF and receive various alternative diagnoses, but potential benefits are likely if these patients present with acute or severe pathologies.

The populations studied in the available evidence base are applicable to the target population in Australia, that is patients presenting to a hospital ED with symptoms—primarily acute dyspnoea—suggestive of HF. The results of the studies are generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in diagnosing suspected HF.

In conclusion, on the basis of the evidence presented, NT-proBNP assays appear to be sensitive 'first line' diagnostic tests that, when added to conventional diagnostic assessment, may assist the acute care physician to correctly 'rule out' HF in symptomatic patients initially suspected of HF.

Monitoring

The overall body of evidence for monitoring of HF patients via B-type natriuretic peptides is limited in volume, with only two relevant studies contributing to answering this question. However, one of these studies is a well-designed randomised controlled trial and, as such, adds significant weight to a low volume evidence base. This small trial (level II intervention evidence) demonstrated that monitoring patients via NT-proBNP assays resulted in fewer cardiovascular deaths and total cardiovascular events than patients monitored via clinical criteria alone. The beneficial effect of the hormone-guided monitoring was presumably mediated through more predominant ACE-inhibitor and spironolactone use to achieve the target NT-proBNP concentrations. An abstract reported similar results in a randomised controlled trial assessing monitoring via BNP assays, but more detail is necessary to determine whether this study could be considered as supporting evidence.

The results of the published randomised controlled trial, taken together with the fact that B-type natriuretic peptide levels seem to provide important prognostic information over and above clinical criteria (see Appendix I), suggest that it is reasonable to hypothesise that adjusting pharmaceutical therapy to achieve lower hormone levels would improve the health outcomes of known HF patients compared to clinical monitoring alone. However, the potential benefit requires further evaluation. A larger randomised controlled trial currently being conducted in New Zealand should shed further light on this poorly researched area of monitoring of HF patients via B-type natriuretic peptides.

Economic implications

Unit cost of the B-type natriuretic peptide test

Current laboratory benchmarking data suggests that B-type natriuretic peptide tests would cost $50.59 per test. Because bulk-billing occurs in the vast majority of such cases, it is appropriate to regard these unit costs as representing the opportunity cost of the test. In contrast, the Medicare Benefits Schedule (MBS) fee for an echocardiogram is $231.

Emergency department setting

Relying on the results of the key randomised controlled trial of BNP-supplemented clinical diagnostic workup, the incremental direct costs and outcomes of the management of patients presenting to an emergency hospital department with acute dyspnoea as the primary symptom were examined from a societal perspective.

It was determined that the introduction of B-type natriuretic peptide testing into the EDs of Australian hospitals could lead to cost savings of $338 per patient presenting with acute dyspnoea (point estimate). The point estimate of incremental costs and incremental lives saved at both 30 days and 180 days suggests that the addition of B-type natriuretic peptide testing to the diagnostic workup *dominates* conventional diagnostic strategies alone. Thus, the two point estimates suggest that performing a B-type natriuretic peptide test in the ED setting leads to a superior health outcome at a lower cost. However, the 95% confidence interval of the joint probability distribution of incremental costs and the 30-day mortality rate indicate that the point estimates are subject to some uncertainty, but that 78.8 per cent of the joint probability distribution is in the dominant quadrant.

The drivers for cost-effectiveness of the B-type natriuretic peptide tests in the ED setting are: the time from presentation to the initiation of appropriate therapy is shorter; fewer patients are admitted to hospital; and fewer patients are admitted to intensive care. Avoidance of echocardiography in test negative patients might be expected but was not reported in the key trial.

With respect to financial outlays, the Australian Government will incur an additional expenditure of $352,000 under Medicare due to the introduction of B-type natriuretic peptide testing for private patients in private hospital EDs. This net amount incorporates the additional outlay of $1.78 million required for B-type natriuretic peptide testing as well as cost savings due to fewer echocardiograms and fewer private inpatient physician consultations. Offsetting this further will be a reduction in hospitalisation and length of stay, which will reduce private sector outlays. Although the majority of B-type natriuretic peptide tests will be performed in public hospital EDs, this is unlikely to lead to Australian health system expenditure savings because of capacity constraints, but may make additional public resources available for other patients in need.

Recommendations

'MSAC has considered the safety, effectiveness and cost-effectiveness of the use of assays of B-type natriuretic peptides (BNP) in the diagnosis of heart failure in patients presenting with dyspnoea in the hospital emergency setting and the use of the assays in monitoring the progress of patients with heart failure.

MSAC finds that there is sufficient evidence of the safety, effectiveness and cost-effectiveness of the use of these assays in the diagnosis of heart failure but insufficient evidence of effectiveness and cost-effectiveness for their use in monitoring the progress of patients with heart failure.

MSAC recommends that public funding be provided for the use of assays of BNP in the diagnosis of heart failure in the hospital emergency setting.'

The Minister for Health and Ageing accepted this recommendation on 5 February 2007.

Glossary

Absolute risk reduction	Absolute risk reduction (ARR) is the difference in the incidence of an outcome between the experimental group and the control group.
ACE	Angiotensin-converting enzyme
ACE-I	Angiotensin-converting enzyme inhibitor
ARB	Angiotensin II receptor blocker
Area under the curve	Calculated as the area under a receiver operator characteristic curve, the area under the curve (AUC) provides a numerical description of the accuracy of a diagnostic test. A test with no diagnostic value has an AUC of 0.5, while a perfect test has an AUC of 1.0.
BEACH	The BEACH Project is the 'Bettering the Evaluation and Care of Health' Project conducted by the Australian General Practice Statistics and Classification Centre (AGPSCC). AGPSCC is a Collaborating Unit of the Australian Institute of Health and Welfare (AIHW). http://www.fmrc.org.au/beach.htm
Begg funnel plot	The Begg funnel plot is a scatter plot of treatment effects against their associated standard errors. The plot is used to visually assess for the presence of publication bias.
BNP	Brain (or B-type) natriuretic peptide
Cochran's Q test	Cochran's Q test is used to assess the degree of heterogeneity between the statistical estimates of studies. A significant Q statistic indicates that an assumption of homogeneity is invalid.
COPD	Chronic obstructive pulmonary disease
Cox proportional hazards model	The Cox proportional hazards model is a regression model for survival time data. The model assesses the effects of explanatory variables on survival time under the assumption that hazard rates for different individuals (defined by their explanatory variables) remain proportional at all time points.
DerSimonian Laird random effects model	In meta-analysis a random effects model for integrating study results is often appropriate when there is evidence of heterogeneity between study estimates. Random effects models assume that treatment effects follow a distribution, thus allowing for between-study variation. The DerSimonian Laird method specifies the weighting scheme for combining study results in the model.
Diagnostic odds ratio	Diagnostic odds ratio (DOR) provides a measure of the diagnostic accuracy of a test. It is calculated as the odds of a positive test in those with the disease divided by the odds of a positive test in those without the disease.

Dyspnoea	A distressful sensation of uncomfortable breathing (also spelt dyspnea)
ECG	Electrocardiogram
ED	emergency department
Egger's test for publication bias	Egger's test for publication bias is a formal test for asymmetry in a Begg funnel plot.
Hazard rate	The hazard rate specifies the instantaneous risk of death or failure for a given point in time.
Heterogeneity	In meta-analysis heterogeneity refers to variability in the statistical estimates of studies.
HF	Heart failure
I^2 statistic	The I^2 statistic describes the percentage of variation in study estimates that can be attributed to heterogeneity rather than chance. Like the Cochran Q test, the I^2 statistic is used to assess statistical heterogeneity.
ITT	Intention-to-treat
Inter-quartile range	Inter-quartile range (IQR) is a measure of dispersion calculated as the difference between the 75th and 25th percentiles of a distribution.
LVEF	Left ventricular ejection fraction
LVSD	Left ventricular systolic dysfunction
MBS	Medicare Benefits Schedule
Meta-analysis	Meta-analysis refers to the statistical analysis of a number of individual study results for the purpose of integrating findings.
MSAC	Medical Services Advisory Committee
NT-proBNP	N-terminal proBNP (nucleotides 1–76)
Number needed To diagnose	The number needed to diagnose (NDD) specifies the number of patients who need to be diagnosed with the new test strategy, compared to the existing test strategy, in order to prevent one additional negative outcome. It is calculated as the inverse of the ARR.
NYHA	New York Heart Association
OECD	Organisation for Economic Co-operation and Development
Peto fixed effects model	In meta-analysis a fixed effects model for integrating study results is often used when there is little heterogeneity between study estimates. Fixed effects models assume that treatment effects are constant across different studies. The Peto method specifies the weighting scheme for combining study results in the model, and is commonly used when the estimate of interest is an odds ratio.

Poisson regression	Poisson regression refers to a linear regression model in which the outcome variable is distributed as a poisson random variable. The model is used to assess the effects of explanatory variables on either the occurrences of an event or the rate of event occurrence.
Power	Power refers to the ability of a statistical test to reject a false null hypothesis.
Publication bias	Publication bias occurs when studies reporting statistically significant effects are more likely to be published and cited.
QALY	Quality-adjusted life-year
Receiver operator characteristic curve	A receiver operator characteristic curve (ROC) is a plot of sensitivity against 1 minus specificity for different values of a diagnostic test. It highlights the trade-off between sensitivity and specificity, and gives an overall indication of the diagnostic accuracy of a test.
Relative risk	Relative risk (RR) is a measure of how much a particular risk factor influences the likelihood of an outcome. It is calculated as the incidence of an outcome in the experimental group divided by the incidence in the control group.
Restricted maximum likelihood	Restricted maximum likelihood (REML) is a method for estimating variance components and can be used in a number of statistical applications, including meta-analysis.
Sensitivity	Sensitivity refers to the proportion of people with a disease who report a positive test result.
Specificity	Specificity refers to the proportion of people without a disease who report a negative test result.
Univariate analysis	Univariate analysis involves the description of individual variables from a set of data. Estimates of dispersion and central tendency are usually the focus of univariate analysis.
Wald chi-square test	In the Cox proportional hazards model, the Wald chi-square test is used to assess the statistical significance of explanatory variables included in the model.

Introduction

The Medical Services Advisory Committee (MSAC) has reviewed the use of B-type natriuretic peptide assays for determining the diagnosis, prognosis and monitoring of patients with heart failure (HF). The MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Scheme in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. The MSAC adopts an evidence-based approach to its assessments, based on reviews of the scientific literature and other information sources, including clinical expertise.

The MSAC's terms of reference and membership are at Appendix A. The MSAC is a multidisciplinary expert body, comprising members drawn from such disciplines as diagnostic imaging, pathology, surgery, internal medicine and general practice, clinical epidemiology, health economics, consumer health and health administration.

Rationale for assessment

Roche Diagnostics Australia and Abbott Diagnostics Australia have made separate applications to the Medical Services Advisory Committee (MSAC) to have the Elecsys® proBNP and the AxSYM® BNP assays placed on the Medicare Benefits Schedule for the diagnosis, monitoring and prognosis of HF. These assays are performed on patient blood extracted through a simple blood test. They measure brain natriuretic peptide (BNP) or the by-product of the cleavage from the precursor of BNP to BNP (NT-proBNP) and would be performed by clinical laboratories, either in a public hospital or a private pathology laboratory. It is suggested that measurement of BNP and/or NT-proBNP will not replace traditional clinical investigations but may assist in the selection of patients who would benefit most from receiving further investigations. These assays are not designed to screen patients without risk factors but to act as a 'first line' test for individuals who are suspected of having HF due to various signs or symptoms (such as dyspnoea). This includes patients reporting to emergency departments (EDs), or where investigations are required in settings such as coronary care, respiratory or renal units, or by health practitioners such as cardiologists or general practitioners.

In addition, these assays may be used to determine the prognosis, and to monitor the treatment, of patients diagnosed with HF.

BNP and NT-proBNP testing are considered new medical services requiring a new Medicare item number.

Background

Suspected heart failure

There is no universal definition of heart failure (AIHW 2003). Heart failure (HF) occurs when the heart (typically the left ventricle) is unable to pump blood adequately to the rest of the body. A series of compensatory physiological adaptations ensues, with vasoconstriction and retention of fluid resulting in higher afterload and consequently accumulation of fluid in the lungs or legs (oedema). The heart muscles may fail to contract normally and expel sufficient blood (systolic HF) or they may fail to relax and fill normally (diastolic HF). Characterised by marked breathlessness (dyspnoea) with activity and while lying flat, the causes of HF include chronic hypertension, cardiomyopathy, valvular heart disease and myocardial infarction (Remme & Swedberg 2001; NHF & CSANZ 2002; AIHW 2003).

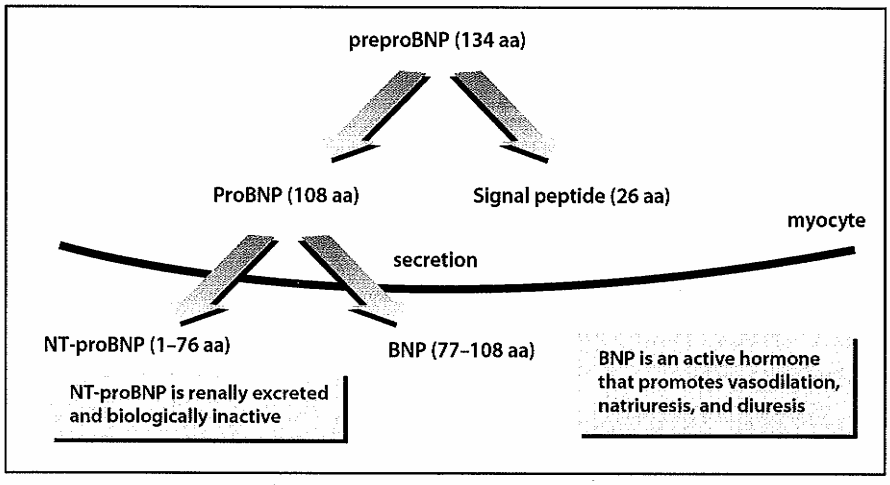
The diagnosis of HF is a clinical judgement based on the patient's history, physical examination and clinical investigations (Remme & Swedberg 2001). The clinical symptoms used to diagnose HF are neither sensitive nor specific (Craig et al 2004). They include dyspnoea, ankle oedema and fatigue at rest or during exertion, in addition to an objective measurement of cardiac dysfunction at rest (Remme & Swedberg 2001). Heart failure is difficult to diagnose correctly, and both underdiagnosis and overdiagnosis are common, which can lead to inadequate or inappropriate treatment (Doust et al 2004).

Ischaemic heart disease is the most common cause of HF in the industrialised world and is associated with left ventricular systolic dysfunction. However, it is increasingly recognised that HF may be present with preserved systolic function (Kirk et al 2004). Diastolic dysfunction may be diagnosed by exclusion, that is where clinical evidence suggests HF but where the left ventricular ejection fraction suggests preserved systolic function (Mottram & Marwick 2005). However, it is often difficult to obtain an accurate diagnosis of diastolic dysfunction (NHF & CSANZ 2002). It has been suggested that many patients who have preserved systolic function may be misdiagnosed as having diastolic dysfunction, when it is possible they have no cardiac abnormality (Banerjee et al 2004). It is therefore important that alternative diagnoses be excluded, such as pulmonary disease, obesity or myocardial ischaemia (Banerjee et al 2004).

B-type natriuretic peptides

Brain (or B-type) natriuretic peptide (BNP) is a cardiac neurohormone released as pre-proBNP, which is cleaved enzymatically to the active hormone BNP (77–108) and the inactive N-terminal proBNP (NT-proBNP, 1–76) fragment (McCullough & Sandberg 2003)—see Figure 1. The main stimulus for the constitutive release of pre-proBNP from cardiac myocytes is cardiac wall stretch resulting from volume or pressure overload in the heart. Under pathologic conditions, BNP is synthesised rapidly in the ventricles and/or atrium. The function of the natriuretic peptides is to protect the cardiovascular system from the effects of chronic volume overload by inducing vasodilation, sodium excretion and diuresis, thereby lowering blood volume and blood pressure (Azzazy & Christenson 2003).

Figure 1 Release of B-type natriuretic peptides



aa = amino acids

Source: McCullough et al (2003b)

Both BNP and NT-proBNP have been implicated as diagnostic biomarkers for HF in clinical practice (Maisel 2003). It has also been suggested that changes in BNP and NT-proBNP levels may be predictive of patient survival/death and would therefore provide useful prognostic and monitoring tools (Bettencourt 2004).

The BNP levels in a patient with decompensated HF may reflect the sum of their euvolemic 'dry' BNP level (their baseline level) plus that occurring from volume overload or acute pressure (their 'wet' BNP level) (Maisel & Zoorob 2005). Determining the baseline 'dry' level of BNP is valuable for monitoring and determining prognosis, as lower levels of 'dry' BNP weight may predict better survival (Maisel & McCullough 2003).

BNP levels are known to increase with advancing age, renal failure, myocardial infarction and acute coronary syndrome. In patients presenting with dyspnoea, HF is usually absent at BNP levels <100 pg/mL, possibly occurs between 100–500 pg/mL, and is probable at levels >500 pg/mL (McCullough & Sandberg 2003). Although elevated BNP levels (>100 pg/mL) may be present in other conditions, such as cor pulmonale, lung cancer and pulmonary embolism, it is not usually elevated to the same extent as in patients with HF (Maisel & McCullough 2003).

It is believed that NT-proBNP concentrations rise with age and are higher in women compared to men (Collinson et al 2004; Mcdonagh et al 2004). McCullough et al suggest that, due to age-dependent changes in NT-proBNP levels, two cut-off points for the detection of HF are necessary—125 pg/mL in patients under 75 years of age and 450 pg/mL in those over 75 years (McCullough & Sandberg 2003). NT-proBNP levels in the range 125–450 pg/mL in the elderly are considered non-diagnostic and would require more information to accurately determine a diagnosis. While it is unclear how NT-proBNP is metabolised, renal failure may impact more on its concentration, compared to BNP, because of differences in physiological clearance pathways between the molecules (Chenevier-Gobeaux et al 2005).

BNP and NT-proBNP neurohormones are dissimilar in their physiological stability, half-life, ease of measurement and relationship with patient age. No formal comparison of the two neurohormones will be made in this Assessment Report since, despite their dissimilarity in molecular structure, biological activity and physiological clearance pathways, clinical comparisons have found greater similarities than differences (Sikaris 2004). They are both affected by similar confounding variables independent of HF, such as increasing age, renal failure, myocardial infarction and acute coronary syndrome (expert opinion, MSAC Advisory Panel, 2005).

The procedure

B-type natriuretic peptides may be detected by using an assay for the active peptide BNP or the inactive hormone NT-proBNP. These assays are conducted on the patient's blood, collected through a simple blood test. While there are only two assays, they are available on several different platforms produced by different manufacturers (Sikaris 2004).

This Assessment Report is the result of two applications for funding under the Medicare Benefits Schedule (Elecsys® proBNP and AxSYM® BNP). The Elecsys® proBNP electrochemiluminescent assay uses purified synthetic NT-proBNP to measure the levels of circulating NT-proBNP in human serum and plasma. The assay takes approximately 20 minutes to complete (Roche Diagnostics 2002). The AxSYM® BNP, on the other hand, is an immunofluorescent assay which uses mouse monoclonal anti-BNP coated microparticles to measure circulating levels of BNP in human ethylenediaminetetraacetic acid (EDTA) plasma (Sallinen 2004).

Both the Elecsys® proBNP and the AxSYM® BNP assays have the advantage of being fully automated, which reduces the technologist's time and minimises human error.

The results of **all** available BNP and NT-proBNP assays were reviewed in this Assessment Report.

Intended purpose

In the terminology coined by Sackett and colleagues, the value of BNP and NT-proBNP assays is as **'**SnOut' tests. These tests are considered highly sensitive; therefore, a negative test result 'rules out' the diagnosis (Sackett et al 1991). It is proposed, therefore, that B-type natriuretic peptide tests would act as '**first line**' **diagnostic tools** to identify patients who should or should not be referred for echocardiography to confirm a clinical diagnosis of HF or to explore alternative diagnoses. As such they could be used *either* as supplemental *or* replacement tests in the diagnostic workup. For those patients 'ruled out' from HF they would *replace* the usual confirmatory HF tests. For those patients not excluded from HF, they would *supplement* the usual confirmatory HF tests. The role of the tests is not to act as a 'reference standard' as their specificity is generally not high.

BNP and NT-proBNP assays are also proposed as *supplements* to existing clinical strategies for **monitoring** clinical status or guiding the treatment of HF patients; and determining the prognosis of HF patients (see Appendix I).

Comparators

Diagnosis

Comparators

The most common method of diagnosing HF is by clinical examination. This involves a review of the patient's medical history and a physical examination, including observation, palpation and auscultation. The World Health Organization criteria (Table 1) may be used to diagnose HF. The subjective nature of a diagnosis made on clinical features alone is a weakness of this method. Clinical evaluation may be used in conjunction with objective tests, including electrocardiograms, chest X-rays and echocardiography when available (NHF & CSANZ 2002). Laboratory investigations (eg blood count, creatinine and urinalysis) are also part of the routine diagnostic evaluation for HF.

Table 1 Modified World Health Organization criteria for assessment of possible chronic heart failure, 1995 (Krum 2001)

Symptoms	Dyspnoea, chronic fatigue, oedema, exercise intolerance
Signs	Third or fourth heart sounds, heart murmur, cardiomegaly, pulmonary crackles, raised jugular venous pressure, dependent oedema
Causative factors	Angina, previous myocardial infarction, hypertension, valvular heart disease/rheumatic fever, cardiomyopathy
Possible HF is considered if patients have:	• 2 symptoms • ≥2 signs • ≥1 symptom and ≥1 sign • ≥1 symptom and ≥1 causative factor

HF = heart failure

Heart failure cannot be diagnosed or excluded reliably on the basis of clinical examination alone (see Table 2). Approximately 25–50 per cent of patients presenting to an ED with decompensation symptoms, such as dyspnoea and oedema, are misdiagnosed (Bayes-Genis et al 2004). Although incorrect treatment due to misdiagnosis may alleviate the patient's symptoms, it may obscure an underlying problem that worsens over time. This is particularly relevant in the elderly population, who are most at risk of HF and in whom multiple diseases are common (Remme & Swedberg 2001). Furthermore, when elderly patients experience symptoms upon exertion, they are likely to restrict their activity levels to reduce the symptoms, which can lead to deconditioning. Diagnosis within this population is difficult as onset may be slow, and HF may be asymptomatic at lower levels of exertion (Shamsham & Mitchell, 2000).

Table 2 European Society of Cardiology definition of heart failure (Swedberg et al 2005).

I	Symptoms of heart failure (at rest or during exercise)
	and
II	Objective evidence (preferably by echocardiography) of cardiac dysfunction (systolic and/or diastolic) (at rest)
	and (in cases where the diagnosis is in doubt)
III	Response to treatment directed towards heart failure

Despite the need for accurate diagnosis, many physicians, particularly in primary care, rely on clinical grounds alone to diagnose HF since the availability of echocardiograms is often limited. It also requires the services of an experienced cardiologist for interpretation, and patients with dyspnoea may find it difficult to lie down long enough for an echocardiogram (Hobbs 2002).

The reference standard

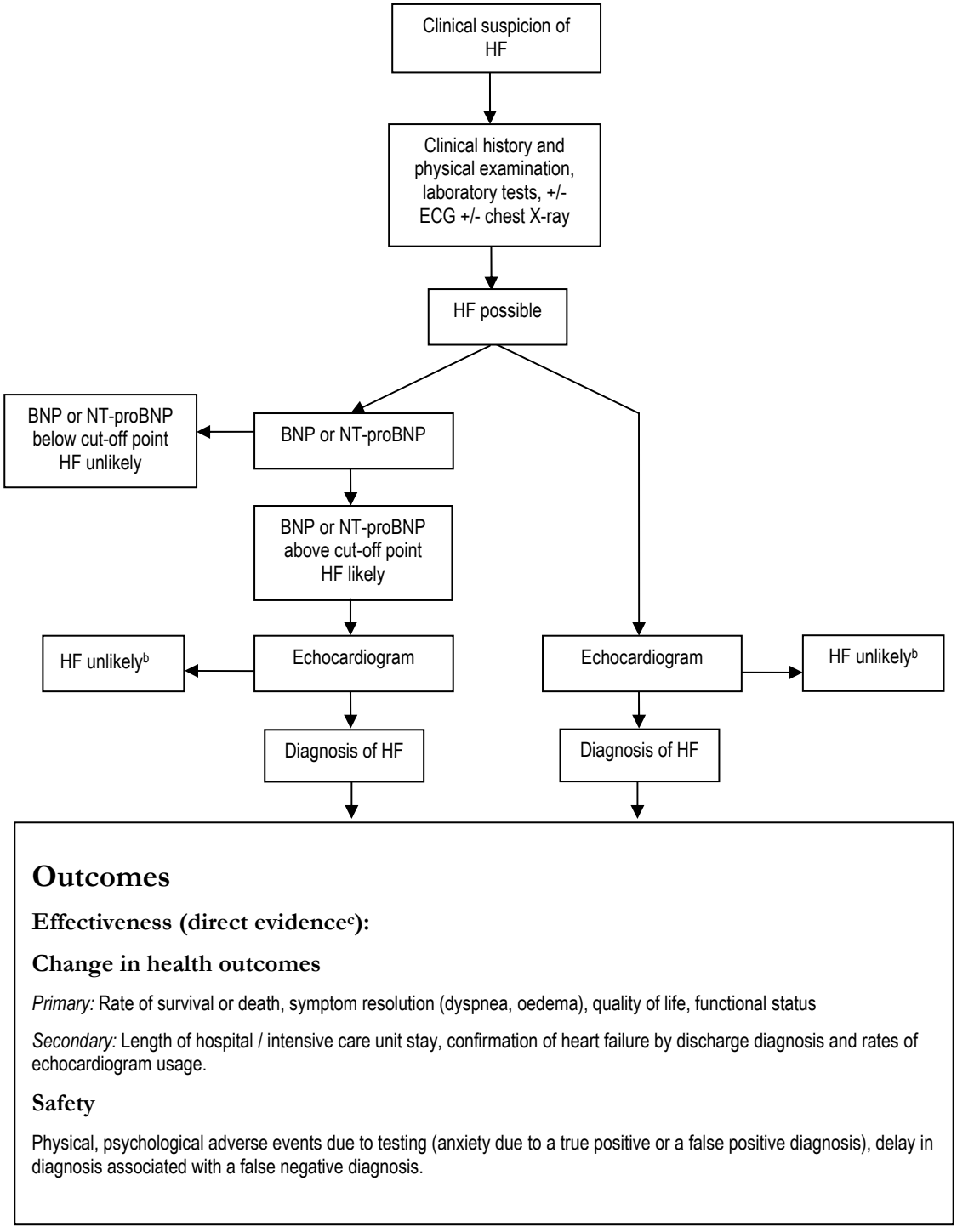
The diagnostic accuracy of the B-type natriuretic peptide assays can be assessed against the objective measure of ventricular function provided by the transthoracic echocardiogram (de Denus et al 2004). Echocardiography uses ultrasound to image the heart and surrounding tissues, providing structural and functional information. Left ventricular ejection fraction is the key parameter for distinguishing patients with cardiac systolic dysfunction from those with preserved systolic function (Remme & Swedberg 2001). Measurements of left ventricular relaxation (left ventricular end diastolic diameter) and filling pressures (via catheterisation or Doppler echocardiography) are considered the best objective measures to assess diastolic HF, but echocardiography may also be used for this purpose by assessing mitral inflow and pulmonary venous flow (Dhir et al 2004).

The transthoracic echocardiogram is a painless, non-invasive procedure that takes between 15 and 30 minutes, and involves the patient lying still on their back on the examination table with their chest exposed. The radiologist or technician applies gel onto the skin to allow the transducer to slide against the skin, emitting ultrasound waves that bounce back, or 'echo' off the structures of the heart (Penn State College of Medicine 2004).

However, echocardiography is an imperfect reference standard. The 'reference standard' should be the echocardiogram result taken in conjunction with a clinical diagnosis of HF (based on all information including signs, symptoms and other tests, eg chest X-ray). The 'gold standard' for diagnosing HF is usually *consensus* cardiologist opinion integrating clinical (signs and symptoms of HF) and objective tests, including echocardiography.

To assess the effectiveness of the BNP and NT-proBNP assays as 'first line' diagnostic tests, the effect on patient relevant outcomes of the addition of either of these assays to existing diagnostic strategies (ie clinical examination/diagnostic workup in conjunction with laboratory tests, chest X-ray, ECG and/or echocardiogram) would need to be compared to the effect on patient relevant outcomes of the existing diagnostic strategies alone (see Figure 2).

Figure 2 Clinical pathway for use of B-type natriuretic peptide assays in the diagnosis of heart failure[a] in the hospital setting



HF = heart failure; ECG = electrocardiogram; [a] Clinical pathway may differ slightly depending on the setting where tests are used; [b] Systolic heart failure unlikely, diastolic heart failure may still be a possibility; [c] Outcomes associated with the linked evidence approach for assessing diagnostic effectiveness are provided in Boxes 4 and 5.

Monitoring

Monitoring of the clinical status of HF patients enables the assessment of treatment efficacy. Consequently, alterations in treatment following changes in clinical variables, or lack thereof, may improve patient outcomes (Cardarelli & Lumicao Jr 2003). Commonly used monitoring tools include clinical evaluation, which may involve functional capacity assessment; fluid status; cardiac rhythm; and circulating levels of urea, electrolytes and creatinine (NICE 2003). Functional capacity may be measured by quality of life questionnaires, exercise tests such as the 6-minute walk or maximal exercise test, and peak VO_2 consumption. Fluid states are measured by physical examination such as assessing body weight, blood pressure and jugular venous distension; and cardiac rhythm is assessed by an electrocardiogram (NICE 2003). Patients such as those with valvular heart disease may benefit from serial assessment with echocardiography to assess changes in left ventricular size, even if asymptomatic (Vitarelli et al 2003).

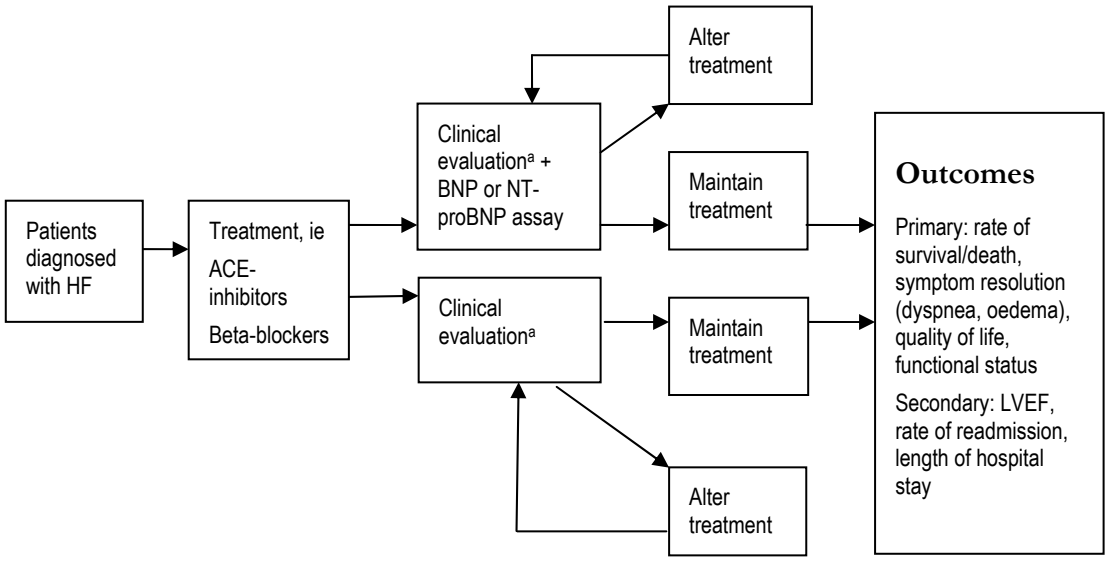
Once a diagnosis of HF has been established, its severity may be classified according to the New York Heart Association Classification (NYHA) system (see Table 3).

Table 3 Functional classification of patients for severity of heart failure

Class	New York Heart Association Classification (USA)
I	Normal daily activity does not initiate symptoms
II	Normal activities initiate symptoms, but subside with rest
IIA	*Slight limitation of physical activity = IIs (modified NYHA classification, Seino et al 2004)*
IIB	*Moderate limitation of physical activity = IIm (modified NYHA classification, Seino et al 2004)*
III	Minimal activity initiates symptoms; patient is usually symptom-free at rest
IV	Any type of activity initiates symptoms, and symptoms persist while at rest

For B-type natriuretic peptide assays to be deemed effective, adapting HF therapy according to the assayed hormone levels would need to result in superior patient health outcomes over and above that of monitoring HF patients via standard clinical management (see Figure 3).

Figure 3 Clinical pathway for use of B-type natriuretic peptide assays in the monitoring of heart failure



HF = heart failure; ACE = angiotensin-converting enzyme; LVEF= left ventricular ejection fraction; [a] including functional capacity, fluid states, cardiac rhythm, serum biochemistry

Clinical need and burden of disease

The Australian National Hospital Morbidity database of the Australian Institute of Health and Welfare (AIHW) registered 41,052 separations for HF in the financial year 2002–03 under the International Classification of Disease (ICD) code I50. The estimate is based on almost all Australian public hospitals and a majority of private hospitals. Seventy-eight per cent (n=31,879) of these HF separations were reported in public hospitals, with the remainder (22%; n=9,173) in private hospitals.

There were 2,612 deaths in Australia from HF during the period 2001–02 (AIHW 2004a). The population aged >75 years is the primary target for the diagnosis, monitoring and prognosis of HF because of their high mortality rate (Figure 4). In addition, prevalence of confirmed HF increases from 1 per cent in 50–59-year-olds to 50 per cent in people aged 85 years and older, and hospitalisation rates for HF are three times higher among people aged 75–84 years than in the 65–75 years age group (AIHW 2003; Krum 2001).

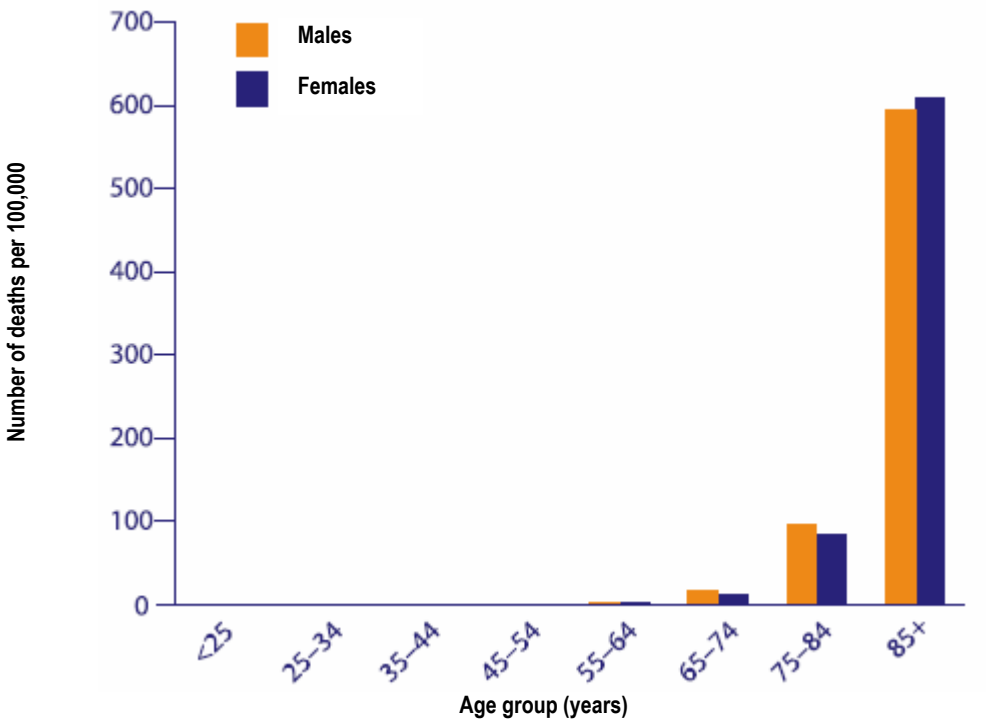


Figure 4 Number of deaths from heart failure in Australia (2002) by age group

Source: AIHW & NHF (2004)

The ageing population, improved survival after coronary events, and increasing incidence of diabetes and obesity may underlie the growing incidence and prevalence of HF in Australia (AIHW 2003; Kenchaiah et al 2002).

In the 2005 calendar year, 88,680,935 general practitioner (GP) attendances were claimed in Australia overall on the Medicare Benefits Schedule. The BEACH data on Australian General Practice Activity (AIHW 2005) reports that between April 2004 and March 2005 'shortness of breath, dyspnoea' was responsible for 779 of the total 94,386 GP encounters (0.83%; 95%CI 0.6%, 1.0%) assessed in the BEACH cohort. Applying this proportion to the total number of annual GP attendances, there would be 736,052 (95%CI 532,086; 886,809) presentations in Australian primary care in 2005 for dyspnoea. Thus, for the purpose of **monitoring** HF in order to titrate therapy (ie only for those patients who had a history of dyspnoea presentation), the number of tests would have ranged between 329,893 and 549,821 in 2005.

The clinical need for an additional tool to assist in the diagnosis of HF in the hospital setting is dependent on the number of patients suspected of having HF due to symptoms, signs and causative factors (primarily, acute dyspnoea). However, patients can present with similar symptoms to HF but receive alternative diagnoses. These can include chronic obstructive pulmonary disease, pneumonia, emphysema or other lung diseases, anaemia and asthma.

Mueller et al (2004b) reported that, of the 80 per cent of admitted patients with acute dyspnoea in their sample, 52.6 per cent were diagnosed with HF. This means that 42 per cent (0.8*0.525 = 0.42) of acute dyspnoea patients arriving at an ED were admitted to hospital with a primary diagnosis of HF. Applying this percentage to ICD I50 (Heart failure) separation data for 2002–03 (n=41,025 separations) above suggests an estimated

rate of 97,742 (41,025/0.42) ED presentations due to acute dyspnoea symptoms (suggestive of HF) per annum in Australia. These patients would be eligible to receive a B-type natriuretic peptide test.

Current treatments

Differential diagnoses

Given the wide variety of alternative diagnoses possible for patients suspected of HF who present primarily with acute dyspnoea, it is difficult to determine the effectiveness and availability of treatments for these pathologies without doing another systematic literature review on the subject.

It is, however, probable that correct and early identification of the alternative diagnosis to HF (eg chronic obstructive pulmonary disease) and prompt treatment would be beneficial for the patient, particularly in cases with a severe or acute presentation.

Heart failure

Early treatment of HF is important for preventing or retarding progression of the disease (Hammerer-Lercher et al 2001). Treatment strategies will depend on the cause and severity of the disease.

When a specific cause of HF is able to be identified, it should be addressed and, if possible, corrected (Leibovitch 2005). This could involve withdrawal of drugs which dampen cardiac function, or treating potentially reversible diseases. For instance, if HF is due to hypertension, thyroid dysfunction, sleep apnoea or renal failure, these causes should be addressed (Leibovitch 2005). Alternatively, if HF is due to an abnormal heart valve, the valve could be surgically replaced (American Heart Association 2005).

Most patients who experience HF will be advised to consider a number of non-pharmacological measures such as taking regular physical exercise, reducing their intake of dietary sodium to below 2,000 mg/day, and limiting fluid intake (1.5 L/day for mild to moderate HF and 1 L/day in severe HF). Smoking and alcohol intake are strongly discouraged. Patients' weight gain is monitored and they may be vaccinated against influenza and pneumococcal disease (NHF & CSANZ 2002).

A number of pharmacological agents are available for the treatment of systolic HF (left ventricular ejection fraction <40%), depending on the classification of the patient's presenting symptoms (eg New York Heart Association (NYHA) classes I–IV). Pharmacological agents include diuretics with or without angiotensin-converting enzyme (ACE) inhibitors. Depending on the progress of the patient and the reduction of symptoms, these agents may be supplemented by the use of beta-blockers. In cases of persistent oedema, spironolactone with or without digoxin may be prescribed. If these pharmacological measures are ineffective or cannot be tolerated by the patient, a heart transplant may be considered for patients <65 years of age with no other major comorbidity (NHF & CSANZ 2002).

Strong evidence of pharmacological effectiveness in treating HF is reported in two systematic literature reviews on beta-blockers (Shibata et al 2001) and ACE inhibitors (Garg & Yusuf 1995). The former review on beta-blockers analysed 22 single- or double-blinded randomised controlled trials that assessed five different beta-blocker agents. The pooled analysis included 10,480 patients (5,507 with active treatment; 4,973 as a placebo group) who were followed up for a mean of 11 months with a completeness of follow-up of 85 per cent. Most patients were categorised at baseline as NYHA functional class III (63.3%). The pooled effect (odds ratio, OR) measures due to beta-blocker therapy for all-cause mortality and hospitalisation were 0.65 (95%CI 0.57, 0.74; $p<0.00001$) and 0.63 (95%CI 0.56, 0.71; $p<0.00001$), respectively. Beta-blocker therapy therefore conferred a 35 per cent reduction in the chance of dying, and a 37 per cent reduction in the probability of hospitalisation, in HF patients relative to placebo treatment. A systematic review of ACE inhibitor effectiveness (Garg & Yusuf 1995), that included 32 randomised trials, resulted in a pooled analysis of 7,105 patients each randomised to a placebo or one of eight ACE inhibitors (predominantly Enalapril). The majority of HF patients included in the meta-analysis were classified as NYHA class II or III. A risk reduction of 23 per cent (OR = 0.77; 95%CI 0.67, 0.88) was reported for all-cause mortality and 35 per cent for hospitalisation (OR = 0.65; 95%CI 0.57, 0.74) relative to placebo treatment. Taken together, these systematic reviews suggest that beta-blockers and ACE inhibitors are clearly effective treatment options for HF.

Cardiac assist devices such as implantable cardioverter-defibrillators have been found to reduce mortality rates, but are associated with very high costs. Heart transplantation is also very effective, but a scarcity of resources (ie human hearts) limits availability of the technique (Leibovitch 2005).

In the Heart Outcomes Prevention Evaluation (HOPE) study, it was found that patients at risk of cardiovascular events—such as those with coronary artery disease, stroke, peripheral vascular disease or diabetes and one other risk factor such as hypertension, without any evidence of HF or left ventricular dysfunction—benefited from receiving treatment. In a randomised controlled trial it was found that patients who received ramipril (ACE inhibitor) were less likely to develop heart disease or experience cardiovascular events than patients who received a placebo (Aurbach et al 2004).

Potential impact of the test

The potential impact on the health system, should B-type natriuretic peptide assays be publicly funded for the diagnosis and the monitoring of HF, is likely to be extensive. In addition to the large population that would be eligible for testing on presentation of new HF-like symptoms in a hospital setting (diagnosis), the population potentially eligible for monitoring of existing HF is considerable. These patients currently represent the largest diseased population within Australia. Given the ageing of the Australian population, prevalence of risk factors (eg obesity, physical inactivity) and the fact that more individuals are surviving acute coronary events (ie myocardial infarction), HF incidence and prevalence will continue to rise.

The settings within which the B-type natriuretic peptide assays will be predominantly used for the diagnosis of HF are that of the non-hospital setting (see Part B of this Assessment Report) and the hospital ED. Cardiologists and/or GPs are also likely to use B-type natriuretic peptide assays to monitor HF patients.

The technology for testing BNP and NT-proBNP levels is already established. Pathology services currently offer testing via a fee-for-service arrangement. If B-type natriuretic peptide assays were to receive public funding, it is probable that the larger hospitals would conduct their own testing to expedite the measurement of B-type natriuretic peptides in patients who present to the ED. This arrangement would require the training of test operators and clinicians in addition to developing, implementing and monitoring relevant laboratory and clinical protocols. It is also likely that a large proportion of the cost impact of this testing would be borne by state health budgets as the majority of EDs are in public hospitals and the majority of patients in public hospitals are public patients. A substantial increase in use of current pathology services for B-type natriuretic peptides would be driven by primary care clinicians monitoring HF patients. This cost would be borne by the Commonwealth.

The unit cost of the B-type natriuretic peptide tests is estimated to be $50.59 per test on the basis of laboratory benchmarking data (see 'What are the economic considerations' section and Appendix G).

Marketing status of the technology

At the time of writing, all therapeutic products marketed in Australia require listing on the Australian Register of Therapeutic Goods (ARTG) unless they have an exemption.

According to the applications submitted to the MSAC, both the Abbott AxSYM BNP and Roche NT-proBNP analytic systems are exempt from the *Therapeutic Goods Administration Act 1998* because the proposed diagnostic tests are not used for blood screening, are not used by consumers, do not contain material of human/animal origin, are not listed on the Pharmaceutical Benefits Scheme, and are not used for human immunodeficiency virus or hepatitic C testing.

Current reimbursement arrangement

Although B-type natriuretic peptide assays are being used in Australia, there are currently no items on the Medicare Benefits Schedule that cover these products. This assessment is being conducted to determine whether these tests should receive public funding.

Approach to assessment

Objectives

The objective of this assessment is to determine whether there is sufficient evidence, in relation to clinical need, safety, effectiveness and cost-effectiveness, to use B-type natriuretic peptide assays (1) in the diagnosis of heart failure (HF) and (2) to monitor HF and guide treatment.

Research questions

Safety

- Is the use of the BNP assay as a 'first line' diagnostic test in the hospital setting, in conjunction with standard clinical assessment[1] ± echocardiography, as safe as, or safer than, standard clinical assessment ± echocardiography alone in the diagnosis of heart failure?

- Is the use of the NT-proBNP assay as a 'first line' diagnostic test in the hospital setting, in conjunction with standard clinical assessment ± echocardiography, as safe as, or safer than, standard clinical assessment ± echocardiography alone in the diagnosis of heart failure?

- Is the use of the BNP assay, in addition to clinical evaluation[2], as safe as, or safer than, clinical evaluation alone for the monitoring of heart failure?

- Is the use of the NT-proBNP assay, in addition to clinical evaluation, as safe as, or safer than, clinical evaluation alone for the monitoring of heart failure?

Diagnostic effectiveness

Direct evidence

- Is the use of the BNP assay as a 'first line' diagnostic test in the hospital setting, in conjunction with standard clinical assessment ± echocardiography, as, or more, effective at improving the health outcomes associated with suspected heart failure than standard clinical assessment ± echocardiography alone?

- Is the use of the NT-proBNP assay as a 'first line' diagnostic test in the hospital setting, in conjunction with standard clinical assessment ± echocardiography, as, or

[1] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs

[2] Functional capacity, fluid status, cardiac rhythm, laboratory assessment of serum biochemistry

more, effective at improving the health outcomes associated with suspected heart failure than standard clinical assessment ± echocardiography alone?

Linked evidence[3]

- What is the diagnostic accuracy of the BNP assay when used to diagnose heart failure in the hospital setting compared to clinical diagnosis and/or echocardiography?
- What is the diagnostic accuracy of the NT-proBNP assay when used to diagnose heart failure in the hospital setting compared to clinical diagnosis and/or echocardiography?
- Does the BNP assay affect the clinical management or treatment options available to patients suspected of heart failure in the hospital setting?
- Does the NT-proBNP assay affect the clinical management or treatment options available to patients suspected of heart failure in the hospital setting?
- Does the BNP assay and possible alterations in clinical management in the hospital setting impact on the health outcomes associated with suspected heart failure?
- Does the NT-proBNP assay and possible alterations in clinical management in the hospital setting impact on the health outcomes associated with suspected heart failure?

Diagnostic cost effectiveness[4]

- Is the BNP assay cost-effective as a 'first line' test in the hospital setting, in conjunction with standard clinical assessment ± echocardiography, in the diagnosis of heart failure compared to standard clinical assessment ± echocardiography alone?
- Is the NT-proBNP assay cost-effective as a 'first line' test in the hospital setting, in conjunction with standard clinical assessment ± echocardiography, in the diagnosis of heart failure compared to standard clinical assessment ± echocardiography alone?

Effectiveness of monitoring

- Is the use of the BNP assay, in addition to clinical evaluation[5], as or more effective than clinical evaluation alone for the monitoring of heart failure?
- Is the use of the NT-proBNP assay, in addition to clinical evaluation, as or more effective than clinical evaluation alone for the monitoring of heart failure?

---

[3] Used in situations where direct evidence of diagnostic effectiveness is not available or where it is limited

[4] Only investigated if evidence of clinical effectiveness was determined

[5] Functional capacity, fluid status, cardiac rhythm, laboratory assessment of serum biochemistry.

Cost-effectiveness of monitoring

- Is the use of the BNP assay, in addition to clinical evaluation[6], as or more cost-effective than clinical evaluation alone for the monitoring of heart failure?

- Is the use of the NT-proBNP assay, in addition to clinical evaluation, as or more cost-effective than clinical evaluation alone for the monitoring of heart failure?

Expert advice

An advisory panel with expertise in pathology, clinical biochemistry, general practice and consumer issues was established to evaluate the evidence from this Assessment Report and to provide advice to the MSAC from a clinical or consumer perspective. In selecting members for advisory panels, the MSAC's practice is to approach the appropriate medical colleges, specialist societies and associations and consumer bodies for nominees. Membership of the advisory panel associated with this MSAC assessment is provided at Appendix B.

Review of the literature

Literature sources and search strategies

The medical literature was searched to identify relevant studies concerning B-type natriuretic peptides for the period between 1988 and August 2005. B-type natriuretic peptide assays were first reported in 1988. Appendix C describes the electronic databases that were used for this search and other sources of evidence that were investigated. Grey literature was included in the search strategy. Unpublished literature, however, was not canvassed as it is difficult to search for this literature exhaustively and systematically, and trials that are difficult to locate are often smaller and of lower methodological quality (Egger et al 2003). It is, however, possible that these unpublished data could impact on the results of this assessment.

The search terms, presented in Appendix C, were used to identify literature in electronic bibliographic databases on the safety, effectiveness and cost-effectiveness of using B-type natriuretic peptide assays to diagnose or monitor HF.

Inclusion/exclusion criteria

In general, studies were excluded if they:

- did not address the research question;

- did not provide information on the pre-specified target population;

- did not include one of the pre-specified interventions;

[6] Functional capacity, fluid status, cardiac rhythm, laboratory assessment of serum biochemistry.

- did not compare results to the pre-specified comparator;
- did not address one of the pre-specified outcomes and/or provided inadequate data on these outcomes (in some instances a study was included to assess one or more outcomes but had to be excluded for other outcomes due to data inadequacies); or
- did not have the appropriate study design.

Where two (or more) papers reported on different aspects of the same study, such as the methodology in one and the findings in the other, they were treated as one study. Similarly, if the same data were duplicated in multiple articles, results from the most comprehensive or most recent article only were included.

The criteria for including studies relevant to determining the *safety* of the B-type natriuretic peptide assays can be found in Box 1.

Box 1 Study selection criteria for assessing safety

Selection criteria	Inclusion criteria
Population	Patients with (1) suspected HF in the hospital setting or (2) HF
Intervention	BNP or NT-proBNP assays in conjunction with (1) standard clinical assessment[a] ± echocardiography (diagnosis) or (2) clinical evaluation[b] (monitoring)
Comparator(s)	(1) Standard clinical assessment ± echocardiography (diagnosis) or (2) clinical evaluation (monitoring) alone
Outcomes	Adverse events—physical, psychological due to testing (anxiety due to a true positive or a false positive diagnosis), delay in management associated with a false negative diagnosis
Study design	Randomised or non-randomised controlled trials, cohort studies, registers, case series, case reports or systematic reviews of these study designs
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 04/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

HF = heart failure; [a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs; [b] Functional capacity, fluid status, cardiac rhythm, laboratory assessment of serum biochemistry.

Diagnostic assessment framework

This assessment of the diagnostic use of the BNP and NT-proBNP assays follows the methodology outlined in the MSAC *Guidelines for the assessment of diagnostic technologies* handbook (MSAC 2005).

In order to assess the effectiveness of the BNP and NT-proBNP tests in the diagnosis of HF in the hospital setting, there needed to be a consideration of their diagnostic accuracy (in comparison to a reference standard), their impact on the clinical management of the patient with suspected HF and their ultimate impact on patient health outcomes. The first goal of this assessment was to find *direct evidence* of the effectiveness of the BNP and NT-proBNP tests on health outcomes. Only limited direct evidence was available and so this was supplemented by a *linked evidence* approach. This is an approach where studies which assess, individually, the diagnostic accuracy of the tests, the impact on patient management and the impact on health outcomes are linked together through a narrative.

The criteria for including studies on *diagnostic effectiveness* are presented in Box 2 (for the direct evidence approach), and

Box 3 and Box 4 (for the linked evidence approach).

Box 2 Study selection criteria for assessing diagnostic effectiveness in the hospital setting (direct evidence approach)

Selection criteria	Inclusion criteria
Population	Patients suspected of heart failure due primarily to acute dyspnoea
Intervention	NT-proBNP or BNP diagnostic assays as 'first line' tests in conjunction with standard clinical assessment[a] ± echocardiography[b]
Comparator(s)	Standard clinical assessment ± echocardiography
Outcomes	Health outcomes: Primary: rate of survival/death, symptom resolution (dyspnoea, oedema), quality of life, functional status Secondary: hospital / intensive care unit length of stay, discharge diagnosis, rates of echocardiogram usage
Study design	Randomised or non-randomised controlled trials or cohort studies or case-control studies or systematic reviews of these study designs
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

[a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs; [b] Echocardiogram is likely to be used in the diagnostic pathway if, on the basis of the 'first line' tests (eg BNP, NT-proBNP, physical examination, chest X-ray, laboratory tests, ECG), the patient is still suspected of heart failure. Those patients 'ruled out' for heart failure on the basis of these 'first line' tests, however, will not receive an echocardiogram.

Box 3 Study selection criteria for assessing diagnostic accuracy in the hospital setting (linked evidence approach)

Selection criteria	Inclusion criteria
Population	Patients suspected of heart failure due primarily to acute dyspnoea
Intervention	NT-proBNP or BNP diagnostic assays
Comparator	Clinical diagnosis or echocardiography
Reference standard	Clinical diagnosis using all data, including echocardiogram
Outcomes	Sensitivity and specificity (and therefore rates of false positives and negatives), likelihood ratios and diagnostic odds ratios, negative predictive values, diagnostic yield
Study design	Cross-sectional studies where patients are cross-classified on the test and reference standard. Case-control diagnostic studies were only acceptable if cross-sectional studies were not available, or were limited. Systematic reviews of these study designs were also acceptable
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

Box 4 Study selection criteria for assessing effectiveness of diagnosis in the hospital setting (linked evidence approach)

	Change in management	Change in health outcomes
Selection criteria	**Inclusion criteria**	**Inclusion criteria**
Population	Patients suspected of HF due primarily to acute dyspnoea	Patients with HF or alternative diagnosis[c]
Intervention	NT-proBNP or BNP diagnostic assays as 'first line' tests in conjunction with standard clinical assessment[a] ± echocardiography[b]	Treatment for HF (eg ACE inhibitors, beta- blockers, surgery) or treatment for alternative diagnosis
Comparator(s)	Standard clinical assessment ± echocardiography	No (or delayed) treatment for HF or the alternative diagnosis
Outcomes	Primary: treatment rates, method of treatment, time to diagnosis, rate of referral to specialist	Primary: rate of survival/death, symptom resolution (dyspnoea, oedema), quality of life, functional status
	Secondary: rates of echocardiogram/ supportive diagnostic testing	Secondary: confirmation of HF by left ventricular ejection fraction (LVEF) <50%[d], hospital length of stay, rate of readmission
Study design	Randomised or non-randomised controlled trials or cohort studies, uncontrolled before-and-after case series (with 20 or more participants) or systematic reviews of these study designs	
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005	
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base	

HF = heart failure; [a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs; [b] Echocardiogram is likely to be used in the diagnostic pathway if, on the basis of the 'first line' tests (eg BNP, NT-proBNP, physical examination, chest X-ray, laboratory tests, ECG), the patient is still suspected of HF. Those patients 'ruled out' for HF on the basis of these 'first line' tests, however, will not receive an echocardiogram; [c] Given the multitude of alternative diagnoses for patients presenting with HF-like symptoms, it was not possible to assess treatment effectiveness systematically in this patient group; [d] It is acknowledged that LVEF is only effective at detecting systolic dysfunction, and may not detect diastolic dysfunction.

The criteria for including studies relevant to determining the *effectiveness of monitoring* by B-type natriuretic peptide assays are provided in Box 5.

Box 5 Study selection criteria for effectiveness for monitoring

Selection criteria	Inclusion criteria
Population	Patients with heart failure
Intervention	Clinical evaluation and NT-proBNP or BNP assays
Comparator(s)	Clinical evaluation—functional capacity, fluid status, cardiac rhythm, laboratory assessment of serum biochemistry
Outcomes	Primary: rate of survival/death, symptom resolution (dyspnoea, oedema), quality of life, functional status Secondary: left ventricular ejection fraction (LVEF) <50%[a], rate of readmission, length of hospital stay
Study design	Randomised or non-randomised controlled trials or cohort studies, uncontrolled before-and-after case series (with 20 or more participants) or systematic reviews of these study designs
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

[a] It is acknowledged that LVEF is only effective at detecting systolic dysfunction, and may not detect diastolic dysfunction.

Search results

The process of study selection for this report went through six phases:

1. All reference citations from all literature sources were collated into an Endnote 8.0 database;

2. Duplicate references were removed;

3. Studies were excluded, on the basis of the complete citation information, if it was obvious that they did not meet the inclusion criteria. All other studies were retrieved for full-text assessment;

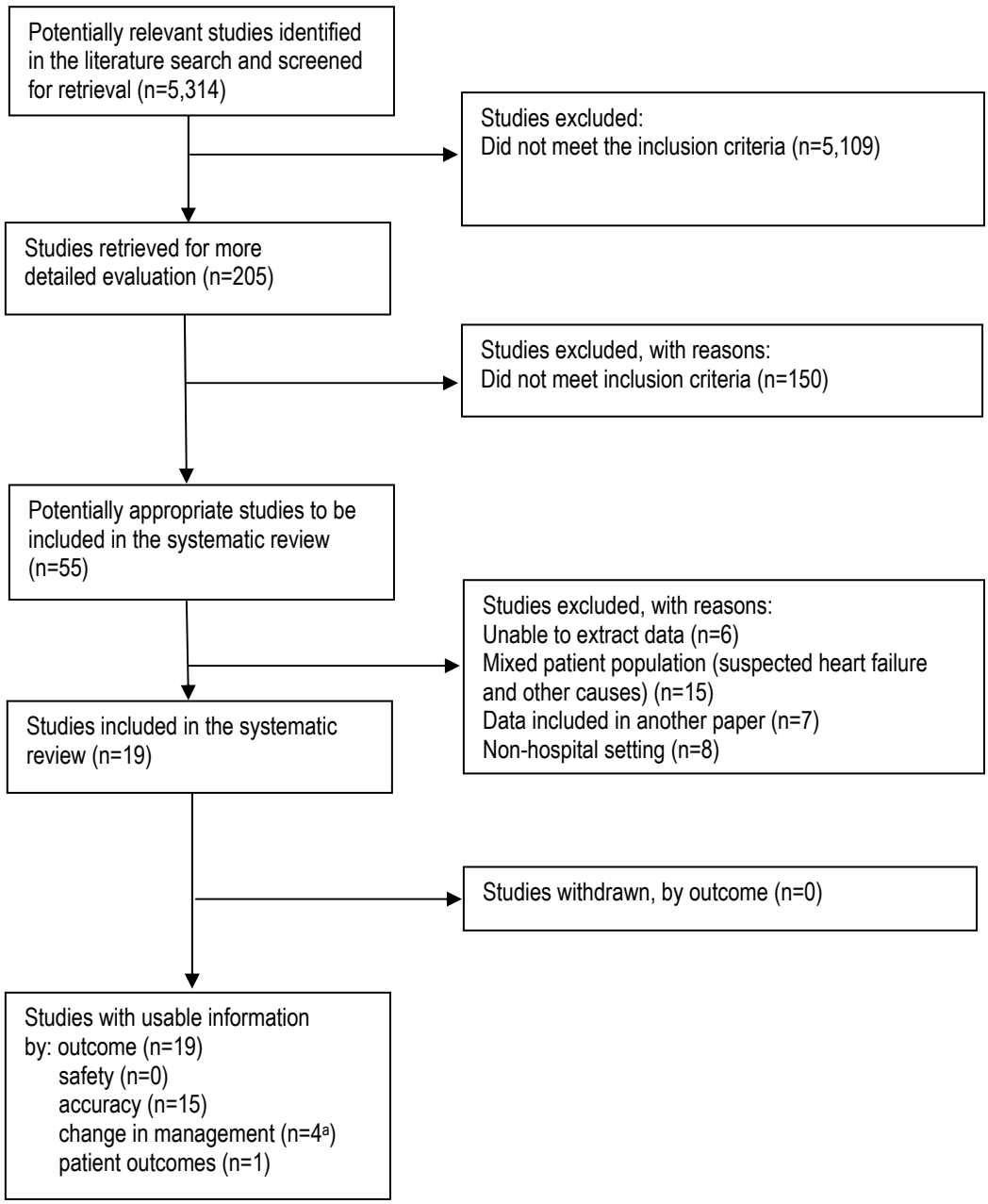
4. Inclusion criteria were independently applied to the full-text articles by two or more researchers. Those articles meeting the criteria formed part of the evidence base. The remainder provided background information;

5. The reference lists of the included articles were pearled for additional relevant studies. These were retrieved and assessed according to phase 4; and

6. The evidence base consisted of articles from phases 4 and 5 that met the inclusion criteria.

Any doubt concerning inclusions at phase 4 was resolved by group consensus. The results of the process of study selection—to collate the evidence base for assessing *diagnostic effectiveness* in the hospital setting—are provided in Figure 5. Figure 6 outlines the process that was used to select studies to assess the effectiveness of B-type natriuretic peptide assays at *monitoring* HF.

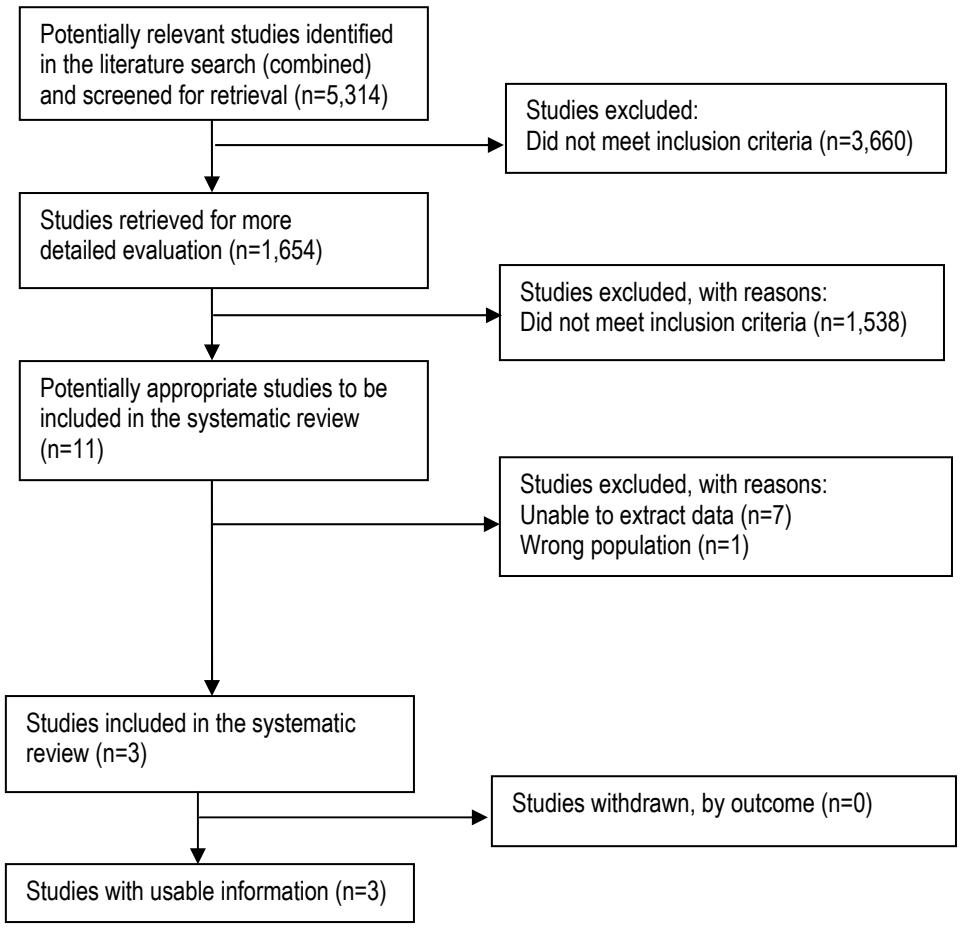
Figure 5 Summary of the process used to identify and select studies for the assessment of diagnostic effectiveness



[a] One of these studies also reported on patient outcomes

Adapted from Moher et al (1999)

Figure 6 Summary of the process used to identify and select studies for the assessment of monitoring effectiveness



Adapted from Moher et al (1999)

Data extraction and analysis

A profile of key characteristics was developed for each included diagnostic and monitoring study (Appendix E). Studies that were unable to be retrieved or that met the inclusion criteria but contained insufficient or inadequate data for inclusion are provided in Appendix F. Definitions of all technical terms and abbreviations are provided in the Glossary.

Diagnostic studies

The appropriate population for diagnostic accuracy studies (in linked evidence) included in this assessment consisted of patients suspected of HF due primarily to symptoms of acute dyspnoea. Studies were excluded that recruited patients based on referral for echocardiography without indicating whether the referral was for clinically suspected HF. In studies reporting diagnostic accuracy for HF as well as diagnostic accuracy for left ventricular dysfunction, only data referring to the former were extracted. However, data were extracted on diagnostic accuracy for left ventricular dysfunction when that was all that was presented. For *direct* evidence of diagnostic effectiveness and *linked* evidence **intervention studies** (eg assessing change in management/treatment through use of the test and change in health outcomes from treatment), descriptive statistics (eg means, standard deviations) were extracted or calculated from the individual studies for all safety and effectiveness pre-specified outcomes. A statistically significant difference in outcomes was assumed at $p<0.05$. When assessing the impact of B-type natriuretic peptide assays on outcomes such as reduction in mortality rate or symptom resolution (in the form of count data), the data were presented as relative risks and the number needed to diagnose, and their 95% confidence intervals.

Relative risks (RR) were calculated as the incidence[7] (risk) of an outcome in the experimental group divided by the incidence in the control group. The number needed to diagnose (NND) was calculated as the inverse of the absolute risk reduction[8]. It was defined as the number of patients who need to be diagnosed with the new test strategy, compared to the existing test strategy, to prevent one outcome.

The evidence base was limited for determining the *direct* effectiveness of B-type natriuretic peptide assays in the diagnosis of HF. A qualitative or narrative synthesis of the available data was therefore presented.

[7] Cumulative incidence = number of patients experiencing the outcome over a certain time period divided by the number of patients at risk of the outcome over the same time period.

[8] ARR = difference in the incidence of an outcome between the experimental group and the control group.

Data from the supportive, *linked* evidence on the diagnostic accuracy of the B-type natriuretic peptide tests was extracted using the classic 2 x 2 table, whereby the results of the diagnostic test were cross-classified against the results of the reference standard (Armitage et al 2002; Deeks 2001), and Bayes' Theorem was applied:

		Cardiac status (based on reference standard—clinical diagnosis and/or echocardiogram)		
		Heart failure (HF)	*Normal*	
Index test *(BNP or NT-proBNP)*	*Test +*	True positive	False positive	Total positive
	Test -	False negative	True negative	Total negative
		Total with HF	Total without HF	

Primary measures

The sensitivity of the index test (BNP or NT-proBNP) was therefore calculated as the proportion of people with HF who have positive diagnostic test results:

Sensitivity (true positive rate) = Number of true positives / total with HF * 100

The specificity of the index test (BNP or NT-proBNP) was calculated as the proportion of people without HF who have normal diagnostic test results:

Specificity (true negative rate) = Number of true negatives / total without HF * 100

When a 95% confidence interval was not provided in the relevant study, it was calculated using exact binomial methods.

Summary measures

The diagnostic odds ratio (DOR) was calculated as the ratio of the odds of a positive test in those people with HF compared to those without HF.

DOR = (true positives * true negatives) / (false positives * false negatives)

The summary receiver operator characteristic curve (SROC) plots the estimated sensitivity versus 1–specificity from different studies to produce a global measure of test accuracy.

Meta-analysis

Individual results of test sensitivity and specificity with 95% confidence intervals were plotted for the included studies. A chi-square test and I^2 statistic was used to conservatively test for heterogeneity in the sensitivity and specificity results reported across studies.

The method of pooling (meta-analysing) the diagnostic accuracy studies in this assessment report depended upon whether heterogeneity existed between the studies.

In the presence of heterogeneity it was not appropriate to present pooled sensitivity and specificity rates. A pooled DOR was therefore calculated using a random effects model (DerSimonian & Laird 1986) to provide an estimate of diagnostic accuracy. This could only be done when the study presented results so that a 2 x 2 table could be constructed. Results were then tested for publication bias using Egger's test for publication bias and the Begg funnel plot.

Finally, the results were plotted in the ROC plane to enable the investigation of likely threshold effects using the method of Littenberg and Moses (1993). If the SROC curve was symmetrical, then the pooled or summary DOR was constant and any variability between studies was probably due to differences in test threshold (cut-off points). The pooled DOR gave an overall estimate of test accuracy. Values larger than 1 indicated the strength of the test to discriminate between the presence or absence of HF; a value equal to 1 indicated that the test did not provide any useful diagnostic information; and values below 1 indicated that the test identified more positives among those without HF than with HF.

All statistical calculations and testing were undertaken using the statistical computer package Stata version 8.2 (Stata Corporation 2004). All data regarding B-type natriuretic peptide assay levels were presented as pg/mL. Data presented as pmol/L in the original studies were converted to pg/mL by multiplying by 8.457 and 3.456 for NT-proBNP and BNP, respectively (molecular weights of NT-proBNP and BNP are 8,457 and 3,456 respectively) (Januzzi & Maisel 2004). Further, to convert these data to Standard International Units, as used in Australia, the pg/mL should be converted to their original pmol/L values (by dividing by 8.457 and 3.456 for NT-proBNP and BNP, respectively) and then converting the pmol/L to mmol/L[9].

Monitoring studies

Descriptive statistics were extracted or calculated for all the pre-specified safety and effectiveness outcomes in the individual monitoring studies.

When assessing the impact of monitoring with B-type natriuretic peptide assays on outcomes such as reduction in mortality rate (in the form of count data), the data were presented as relative risks and number needed to treat (NNT) to benefit, along with 95% confidence intervals. When assessing the impact on continuous outcomes (ie improvement in LVEF from baseline), the data were presented in terms of the mean change from baseline (and standard deviation), and statistical analyses that adjusted for baseline differences (eg poisson regression) were reported.

A small evidence base for determining the effectiveness of B-type natriuretic peptide assays in monitoring HF patients warranted only a qualitative or narrative synthesis of the available data.

Appraisal of the evidence

The evidence presented in the selected studies was assessed and classified using the dimensions of evidence defined by the National Health and Medical Research Council (NHMRC 2000).

These dimensions (Table 4) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from the

[9] 1 mmol is 10^{-3} of a mole and 1 pmol is 10^{-12} of a mole; therefore, the conversion factor from pmol/L to mmol/L can be achieved by dividing by 10^{8}.

literature identified as informing a particular intervention. The last two require expert clinical input as part of their determination.

Table 4 Evidence dimensions

Type of evidence	Definition
Strength of the evidence	
Level	The study design used, as an indicator of the degree to which bias has been eliminated by design.[a]
Quality	The methods used by investigators to minimise bias within a study design.
Statistical precision	The *p*-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect.
Size of effect	The distance of the study estimate from the 'null' value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

[a] See Table 5.

Strength of the evidence in individual studies

The three subdomains (level, quality and statistical precision) are collectively a measure of the strength of the evidence.

Level

The 'level of evidence' reflects the effectiveness of a study design to answer a particular research question. Effectiveness is based on the probability that the design of the study has reduced or eliminated the impact of bias on the results.

The new version of the NHMRC evidence hierarchy provides a ranking of various study designs ('levels of evidence') by the type of research question being addressed (NHMRC 2005). Table 5 is an abbreviated version of this evidence hierarchy and includes the research questions relevant to an assessment of diagnosis and monitoring. To assess *direct evidence* of diagnostic effectiveness, the *Intervention* column in this evidence hierarchy was used. To assess *linked evidence* relating to diagnostic effectiveness, the *Diagnosis* column in this evidence hierarchy was used for diagnostic accuracy studies and the *Intervention* column was used for the studies on change in management due to the test and change in health outcomes due to treatment. To assess the evidence provided by the monitoring studies included in this report, the *Intervention* column in the evidence hierarchy was used.

Quality

Study quality was presented in this Assessment Report both in terms of the components of quality (eg selection bias, misclassification bias, reviewer bias) and as an overall quality score.

The quality of the included studies for assessing the safety and effectiveness, with respect to clinical (patient-relevant) outcomes, of adding NT-proBNP or BNP assays to current *diagnostic* testing strategies (*direct evidence* approach) was determined using a checklist developed by Downs and Black and modified for this assessment (Downs & Black 1998).

With respect to a *linked evidence* approach, the appraisal of studies pertaining to the diagnostic accuracy of the NT-proBNP and BNP assays was conducted using the QUADAS tool, a checklist developed by the Centre for Reviews and Dissemination,

York, United Kingdom (Whiting et al 2003). Studies assessing change in management and change in patient health outcomes were critically appraised using the Downs and Black instrument (Downs & Black 1998).

Table 5 Designation of intervention and diagnostic levels of evidence

Level	Intervention §	Diagnosis **
I *	A systematic review of level II studies	A systematic review of level II studies
II	A randomised controlled trial	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, §§ among consecutive patients with a defined clinical presentation ††
III-1	A pseudorandomised controlled trial (i.e. alternate allocation or some other method)	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, §§ among non-consecutive patients with a defined clinical presentation††
III-2	A comparative study with concurrent controls: • Non-randomised, experimental trial † • Cohort study • Case-control study • Interrupted time series with a control group	A comparison with reference standard that does not meet the criteria required for Level II and III-1 evidence
III-3	A comparative study without concurrent controls: • Historical control study • Two or more single arm study ‡ • Interrupted time series without a parallel control group	Diagnostic case-control study ††
IV	Case series with either post-test or pre-test/post-test outcomes	Study of diagnostic yield (no reference standard) ‡‡

* A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence; § Definitions of these study designs are provided on pages 7–8 in *How to use the evidence: assessment and application of scientific evidence* (NHMRC 2000); c This also includes controlled before-and-after (pre-test/post-test) studies, as well as indirect comparisons (ie using A vs B and B vs C, to determine A vs C); ‡ Comparing single arm studies, that is case series from two studies; ** The dimensions of evidence apply only to studies of diagnostic accuracy. To assess the effectiveness of a diagnostic test there also needs to be a consideration of the impact of the test on patient management and health outcomes. See MSAC (2004) *Guidelines for the assessment of diagnostic technologies.* Available at: <www.msac.gov.au>; §§ The validity of the reference standard should be determined in the context of the disease under review. Criteria for determining the validity of the reference standard should be pre-specified. This can include the choice of reference standard(s) and its/their timing in relation to the index test. The validity of the reference standard can be determined through quality appraisal of the study. See Whiting et al 2003; †† Well-designed population-based case-control studies (eg population-based screening studies where test accuracy is assessed on all cases, with a random sample of controls) do capture a population with a representative spectrum of disease and, thus, fulfil the requirements for a valid assembly of patients. However, in some cases the population assembled is not representative of the use of the test in practice. In diagnostic case-control studies a selected sample of patients already known to have the disease are compared with a separate group of normal/healthy people known to be free of the disease. In this situation patients with borderline or mild expressions of the disease, and conditions mimicking the disease, are excluded, which can lead to exaggeration of both sensitivity and specificity. This is called spectrum bias because the spectrum of study participants will not be representative of patients seen in practice; ‡‡ Studies of diagnostic yield provide the yield of diagnosed patients, as determined by an index test, without confirmation of the accuracy of this diagnosis by a reference standard. These may be the only alternatives when there is no reliable reference standard.

Note 1: Assessment of comparative harms/safety should occur according to the hierarchy presented for each of the research questions, with the proviso that this assessment occurs within the context of the topic being assessed. Some harms are rare and cannot feasibly be captured within randomised controlled trials; physical harms and psychological harms may need to be addressed by different study designs; harms from diagnostic testing include the likelihood of false positive and false negative results; harms from screening include the likelihood of false alarm and false reassurance results.

Note 2: When a level of evidence is attributed in the text of a document, it should also be framed according to its corresponding research question, eg level II intervention evidence, level IV diagnostic evidence, level III-2 prognostic evidence.

Source: NHMRC (2005)

Summary appraisal of strength of the diagnostic evidence

Individual studies assessing diagnostic effectiveness were graded according to the pre-specified quality and applicability criteria (MSAC 2005), as shown in Table 6.

Table 6 Grading system used to rank included diagnostic studies

Validity criteria	Description	Grading system
Appropriate comparison	Did the study evaluate a direct comparison of the index test strategy versus the comparator test strategy?	C1 direct comparison CX other comparison
Applicable population	Did the study evaluate the index test in a population that is representative of the subject characteristics (age and sex) and clinical setting (disease prevalence, disease severity, referral filter and sequence of tests) for the clinical indication of interest?	P1 applicable P2 limited P3 different population
Quality of study	Was the study designed to avoid bias?	Study design: NHMRC level of evidence
	High quality = no potential for bias based on pre-defined key quality criteria	Study quality: Q1 high quality
	Medium quality = some potential for bias in areas other than those pre-specified as key criteria	Q2 medium
	Poor quality = poor reference standard and/or potential for bias based on key pre-specified criteria	Q3 poor reference standard poor quality or insufficient information

Statistical precision

Statistical precision was determined using standard statistical principles. Small confidence intervals and p-values give an indication as to the probability that the reported effect is real (NHMRC 2000).

Size of effect in individual studies

It is important to establish whether statistically significant differences are also clinically important. The size of the effect needs to be determined, as well as whether the 95% confidence interval includes only clinically important effects. Where appropriate, rank scoring methods were used to assess the clinically important benefit of the effect size in the studies available (NHMRC 2000).

Relevance of evidence in individual studies

Similarly, the outcome being measured in the studies should be appropriate and clinically relevant. Inadequately validated (predictive) surrogate measures of a clinically relevant outcome should be avoided (NHMRC 2000). Where appropriate, rank scoring methods were used to determine the clinical relevance of the safety or effectiveness outcome being assessed in the controlled studies (NHMRC 2000).

The body of evidence

Appraisal of the body of evidence was conducted along the lines suggested by the NHMRC in their guidance on clinical practice guideline development (NHMRC 2005). Five components are considered essential by the NHMRC when judging the body of evidence:

- the volume of evidence—which includes the number of studies sorted by their methodological quality and relevance to patients;
- the consistency of the study results—whether the better quality studies had results of a similar magnitude and in the same direction, that is homogenous or heterogenous findings;
- the potential clinical impact—appraisal of the precision, size and clinical importance or relevance of the primary outcomes used to determine the safety and effectiveness of the test;
- the generalisability of the evidence to the target population; and
- the applicability of the evidence—integration of this evidence for conclusions about the net clinical benefit of the index test in the context of Australian clinical practice.

A matrix for assessing the body of evidence for each research question, according to the components above, was adapted for this assessment (see Table 7) (NHMRC 2005).

Table 7 Body of evidence assessment matrix

Component	A Excellent	B Good	C Satisfactory	D Poor
Volume of evidence	several level I or II studies with low risk of bias	One or two level II studies with low risk of bias or a SR/multiple level III studies with low risk of bias	Level III studies with low risk of bias, or level I or II studies with moderate risk of bias	Level IV studies, or level I to III studies with high risk of bias
Consistency	All studies consistent	Most studies consistent and inconsistency may be explained	Some inconsistency reflecting genuine uncertainty around clinical question	Evidence is inconsistent
Clinical impact	Very large	Substantial	Moderate	Slight or restricted
Generalisability	Population/s studied in body of evidence are the same as the target population	Population/s studied in the body of evidence are similar to the target population	Population/s studied in body of evidence different to target population but it is clinically sensible to apply this evidence to target population	Population/s studied in body of evidence different to target population and hard to judge whether it is sensible to generalise to target population
Applicability	Directly applicable to Australian healthcare context	Applicable to Australian healthcare context with few caveats	Probably applicable to Australian healthcare context with some caveats	Not applicable to Australian healthcare context

Results of assessment

Are B-type natriuretic peptide assays safe?

Summary – Safety of B-type natriuretic peptides

None of the studies that met the inclusion criteria for this assessment reported physical harms ensuing from the B-type natriuretic peptide testing procedure. Similarly, none of the diagnostic studies available for this assessment of B-type natriuretic peptide testing investigated the impact of the diagnosis on the patients' psychological wellbeing.

B-type natriuretic peptide testing involves a simple blood test. Blood is extracted using a standard venepuncture technique and is collected in tubes containing ethylenediaminetetraacetic acid. Common after-care is to apply manual pressure and/or a dressing to the wound to assist with haemostasis. B-type natriuretic peptide testing is therefore a minimally invasive procedure and in most cases would be one of several blood tests that the patient would undergo during the diagnostic process, particularly in situations involving an acute presentation of symptoms. In such instances, one needle would be inserted and several tubes of blood drawn for several different analyses. Like all blood tests, harms can occur if the venepuncture procedure is done incorrectly by the health practitioner. Similarly, patients with blood clotting disorders or receiving blood thinners require careful observation to ensure bleeding from the wound is controlled.

None of the studies included in this Assessment Report mentioned physical harms occurring as a result of B-type natriuretic peptide testing.

Psychological harms are a theoretical risk for patients undergoing B-type natriuretic peptide testing. False positive test results could mean that the patient undergoes the stress of receiving an initial diagnosis of heart failure (HF), along with a battery of generally more invasive diagnostic tests and, in some cases, treatment or medications that prove, eventually, to be completely unnecessary. False negative test results provide false reassurance to the patient that he/she is well, potentially resulting in poor health outcomes due to inappropriately delayed treatment.

The impact of B-type natriuretic peptide testing on patients' psychological wellbeing was not evaluated in any of the diagnostic studies that met the criteria for inclusion in this Assessment Report.

Are B-type natriuretic peptide assays effective in the diagnosis of heart failure in the hospital setting?

BNP assays (direct evidence of effectiveness)

Summary – Direct evidence of diagnostic effectiveness of BNP

High-level evidence (level II) suggests that the time to discharge and treatment and the number of hospital and intensive care unit admissions were reduced in most symptomatic patients receiving BNP-assisted diagnostic assessment to rule out HF, compared to conventional assessment. The impact on patient health outcomes was in the right direction (ie a reduction in mortality) but the trial was limited by a lack of statistical power, and so the result was only statistically significant in the pre-specified subgroup of elderly patients.

One good quality, single-blind (outcome assessment) randomised controlled trial assessed the direct impact of BNP testing on patient management and patient health outcomes (Mueller et al 2004b). This trial, known as the 'B-type natriuretic peptide for acute shortness of breath evaluation (BASEL) study', was conducted in Basel, Switzerland, on patients presenting to the ED of a university hospital with acute dyspnoea (or breathlessness). Dyspnoea is a primary symptom associated with HF and certain other pulmonary conditions.

The patients in this trial had an average age of around 70 years, with slightly higher male representation. Prevalence of HF in the trial participants was approximately 50 per cent, with slightly higher rates in those with comorbid kidney disease and, to a lesser extent, the elderly. A large proportion of the trial participants were receiving medications for existing conditions, with approximately one-quarter receiving treatment with beta-blockers, one-half on diuretics, and 40 per cent receiving ACE[10] inhibitors. The trial population and results are therefore applicable to the Australian healthcare context.

Patients consenting to participate in the trial were randomised to receive (1) conventional diagnostic assessment or (2) conventional assessment supplemented by BNP testing. A decision algorithm was followed by physicians in the BNP diagnostic strategy trial arm, which indicated that: patients with BNP levels <100 pg/mL were unlikely to have HF and so differential diagnoses should be investigated; BNP levels >500 pg/mL most likely indicated HF, so rapid therapy was recommended; intermediate (100–500 pg/mL) levels suggested that clinical judgement should be used in conjunction with further diagnostic testing.

Three pre-specified subgroup analyses were undertaken to investigate if there were differential effects of a BNP-supplemented diagnostic strategy on:

- women (Mueller et al 2004a)

[10] Angiotensin-converting enzyme inhibitor or angiotensin-receptor blocker

- the elderly (≥70 years) (Mueller et al 2005a)
- patients with kidney disease (Mueller et al 2005b).

Details of this randomised controlled trial and the subgroup analyses can be seen in Table 8 and Appendix E.

Do BNP assays improve patient management?

The BASEL study was powered for the primary outcome, namely time to hospital discharge. Other outcomes that were indicators of the impact of a BNP-supplemented diagnostic strategy on patient management included time to treatment, and admission rates to hospital and to intensive care.

Results for the whole trial population (see Table 9) suggest that BNP-assisted diagnostic assessment significantly shortened the hospital stay of patients presenting to the ED by a median of 3 days (p=0.001). A similar result was also reported in the subgroups of women and of elderly patients but not for the patients with kidney disease. The difference in time to treatment between all patients receiving conventional diagnostic assessment and all those receiving BNP-assisted diagnostic assessment was statistically significant and just under a median of half an hour. In the elderly and kidney disease patient subgroups, the difference before receiving treatment was also just under half an hour but the difference was not statistically significant.

Admission rates to hospital and the intensive care unit (ICU) were approximately 10 per cent lower, a statistically significant difference, in those patients receiving BNP-assisted diagnostic assessment compared to conventional diagnostic assessment. Approximately 10 patients would need to be diagnosed with the BNP-supplemented strategy, compared to the conventional diagnostic strategy, to reduce one hospital or ICU admission. The difference in hospital admission rates was slightly more marked, and in the same direction, in the subgroup of female patients. In the subgroup of elderly patients the admission rates were also lower in the BNP-assisted diagnostic group, although the difference was only greater than chance for the ICU admissions. In the more comorbid subgroup with kidney disease, there was no difference in admission rates between the groups receiving the different diagnostic strategies (see Table 9).

Do BNP assays improve health outcomes?

The three health outcomes measured in the BASEL study were secondary outcomes and therefore the study was not powered to assess statistically significant differences between the two diagnostic strategies for these outcomes.

Results are provided in Table 9 and indicate that in-hospital mortality was reduced by 38 per cent in patients receiving BNP-assisted diagnostic assessment, relative to conventional diagnostic assessment. The result, however, was not statistically significant. This relative reduction in mortality rate was more pronounced in the subgroups of women (although again not above chance) and of elderly patients, where there was a distinct trend for a statistically significant effect. Thirty-day mortality rates were again all in the same direction, with reduced risk in the BNP-assisted diagnostic assessment group. The effect was most marked in the subgroup of elderly patients, where the risk of dying within 30 days of presentation to the ED was reduced by 49 per cent in the BNP-supplemented diagnostic group compared to the conventional diagnostic group. The result was statistically significant. Twelve elderly patients would need to be diagnosed with the BNP strategy,

compared to conventional diagnosis, to reduce one death within 30 days. However, as this was a secondary outcome of the trial, it would need to be tested further.

Hospital readmission rates after 30 days were no different for patients receiving either diagnostic strategy in the whole trial population and in the subgroups.

Table 8 Summary of included BNP studies (direct evidence)—characteristics and quality appraisal

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests	Outcomes assessed	Study quality	Applicability
			N	Selection criteria				
(Mueller et al 2004b)	Randomised, single-blind controlled trial	University hospital – ED Basel, Switzerland 'B-type natriuretic peptide for acute shortness of breath evaluation (BASEL) study'	452 *BNP arm:* n=225 *Control arm:* n=227	Consecutive patients presenting to ED with acute dyspnoea as primary symptom Kidney disease considered present if GFR <60 mL/min /1.73m^2 at presentation	Medical hx Clinical exam Blood tests Pulse oximetry ECG CXR [echocardiography and pulmonary-function tests recommended]	Change in management • Admission - hospital - ICU • Time to treatment • Time to discharge Change in health outcome • In-hospital mortality • 30-day mortality • Readmission	Level II intervention evidence Q1	P1
(Mueller et al 2004a)	Pre-specified subgroup analysis based on: • Gender		*BNP arm:* Women – n=93 *Control arm:* Women – n=97					
(Mueller et al 2005b)	Pre-specified subgroup analysis based on: • Renal disease status		*BNP arm:* Kidney disease – n=113 *Control arm:* Kidney disease – n=127					
(Mueller et al 2005a)	Pre-specified subgroup analysis based on: • Age		*BNP arm:* ≥70 yrs – n=136 *Control arm:* ≥70 yrs – n=133					

ED = emergency department; GFR = glomerular filtration rate; hx = history; ECG = electrocardiogram; CXR = chest X-ray; ICU = intensive care unit

Table 9 Summary of included BNP studies (direct evidence)—results and precision estimates

Study Author(s) (Year)	Study quality	Study population			Reference standard	Index test specifications	Comparator specifications	Results [95%CI]			
		N	Characteristics	Disease prevalence				Outcome	BNP	Control	Statistic
(Mueller et al 2004b)	Level II intervention evidence Q1	452 *BNP:* n=225 *Control:* n=227	*BNP:* Age: 70 yrs [95%CI 68,72] M/F: 132/93 *Control:* Age: 71 yrs [95%CI 69,73] M/F: 130/97	*BNP:* HF: (45%) *Control:* HF: (51%)	Clinical diagnosis and echo-cardiogram	Clinical diagnosis + BNP ± echocardiogram [BNP - Fluorescence immunoassay, Biosite Diagnostics]		Time to treatment (mins) – median, IQR	63 (16–153)	90 (20–205)	p=0.03[b]
								Time to discharge[a] (days) – median, IQR	8.0 (1–16)	11.0 (5–18)	p=0.001[a]
								Hospital admission, no. (%)	169 (75)	193 (85)	RR=0.88 [0.81,0.97] NND=10 [6, 38] p=0.008[c]
								ICU admission, no. (%)	33 (15)	54 (24)	RR=0.62 [0.42,0.91] NND=11 [6, 52] p=0.01[c]
								In-hospital mortality, no. (%)	13 (6)	21 (9)	RR=0.62 [0.32, 1.22] p=0.21[d]
								30-day mortality, no. (%)	22 (10)	28 (12)	RR=0.79 [0.47, 1.34] p=0.45[d]
								30-day readmission rate, no. (%)	26 (12)	23 (10)	RR=1.1 [0.67, 1.94] p=0.63[d]

Study Author(s) (Year)	Study quality	Study population			Reference standard	Index test specifications	Comparator specifications	Results [95%CI]			
		N	Characteristics	Disease prevalence				Outcome	BNP	Control	Statistic
(Mueller et al 2004a)		*BNP:* Women: n=93 *Control:* Women: n=97	*BNP:* Age: 72±18 yrs *Control:* Age: 72±17 yrs	*BNP:* HF: 40/93 (43%) *Control:* HF: 53/97 (55%)				Time to discharge[a] (days) – median, IQR	6.0 (1–17)	10.0 (4–21)	p=0.023[a]
								Hospital admission, no. (%)	68 (73)	83 (86)	RR=0.85 [0.74, 0.99] NND=8 [4, 96] p=0.034[c]
								ICU admission, no. (%)	11 (12)	22 (23)	RR=0.52 [0.27, 1.01] NND=9 [5, 409] p=0.048[c]
								In-hospital mortality, no. (%)	4 (4)	10 (10)	RR=0.42 [0.14, 1.28] p=0.165[d]
								30-day mortality, no. (%)	8 (9)	13 (13)	RR=0.64 [0.28, 1.48] p=0.292[d]
								30-day readmission rate, no. (%)	5 (5)	8 (8)	RR=0.65 [0.22, 1.92] p=0.433[d]

Study Author(s) (Year)	Study quality	Study population			Reference standard	Index test specifications	Comparator specifications	Results [95%CI]			
		N	Characteristics	Disease prevalence				Outcome	BNP	Control	Statistic
(Mueller et al 2005b)		*BNP:* Kidney disease: n=113 *Control:* Kidney disease: n=127	*BNP:* Age: 76±12 yrs M/F: 58/55 *Control:* Age: 75±11 yrs M/F: 68/59	*BNP:* HF: 69/113 (61%) *Control:* HF: 88/127 (69%)				Time to treatment (mins) – median, IQR	60 (0–149)	87 (20–196)	p=0.135[b]
								Time to discharge[a] (days) – median, IQR	11.0 (4–19)	13.0 (8–19)	p=0.291[a]
								Hospital admission, no. (%)	99 (88)	119 (94)	RR=0.94 [0.86, 1.02] p=0.103[c]
								ICU admission, no. (%)	20 (18)	33 (26)	RR=0.68 [0.42, 1.12] p=0.122[c]
								In-hospital mortality, no. (%)	9 (8)	16 (13)	RR=0.63 [0.29, 1.37] p=0.241[d]
								30-day mortality, no. (%)	15 (13)	20 (16)	RR=0.84 [0.45, 1.57] p=0.588[d]
								30-day readmission rate, no. (%)	18 (16)	12 (9)	RR=1.69 [0.85, 3.34] p=0.130[d]

Study Author(s) (Year)	Study quality	Study population			Reference standard	Index test specifications	Comparator specifications	Results [95%CI]			
		N	Characteristics	Disease prevalence				Outcome	BNP	Control	Statistic
(Mueller et al 2005a)		*BNP:* ≥70 yrs: n=136 *Control:* ≥70 yrs: n=133	*BNP:* M/F: 132/93 *Control:* M/F: 130/97	*BNP:* HF: 68/136 (50%) *Control:* HF: 76/133 (57%)				Time to treatment (mins) – median, IQR	70 (15–161)	94 (20–203)	p=0.234[b]
								Time to discharge[a] (days) – median, IQR	9.0 (2–17)	11.0 (7–21)	p=0.029[a]
								Hospital admission, no. (%)	117 (86)	119 (89)	RR=0.96 [0.88, 1.05] p=0.389[c]
								ICU admission, no. (%)	16 (12)	29 (22)	RR=0.54 [0.31, 0.95] NND=10 [5, 85] p=0.027[c]
								In-hospital mortality, no. (%)	8 (6)	17 (13)	RR=0.46 [0.21, 1.03] NND=14 [7, 5935] p=0.051[d]
								30-day mortality, no. (%)	12 (9)	23 (17)	RR=0.51 [0.26, 0.98] NND=12 [6, 214] p=0.039[d]
								30-day readmission rate, no. (%)	18 (13)	11 (8)	RR=1.6 [0.79, 3.26] p=0.189[d]

HF = heart failure; IQR = inter-quartile range; ICU = intensive care unit; RR = relative risk or rate ratio; NND = number needed to diagnose; [a] Primary outcome of trial; [b] Mann-Whitney U test; [c] Chi-square test; [d] Fisher exact test

BNP assays (linked evidence of effectiveness)

Summary – Linked evidence of diagnostic effectiveness of BNP

The linked diagnostic studies were relatively consistent in their findings that BNP tests are sensitive with a high negative predictive value, meaning a negative test result effectively 'rules out' HF in a patient. A strong pooled diagnostic odds ratio also indicates that BNP tests discriminate between the presence or absence of HF in symptomatic patients. Variation in diagnostic accuracy between studies was possibly due to the different test thresholds employed in the studies for ruling out HF. Low level evidence suggests that the impact of the test was most noticeable in terms of how the physician managed a suspected HF patient. The test appeared to have most effect in cases where the clinical diagnosis was initially uncertain.

Fourteen studies met the inclusion criteria for providing linked evidence of the diagnostic effectiveness of BNP testing in a hospital or acute care setting (see Table 10 and Table 11). Twelve studies assessed the diagnostic accuracy of several BNP assays, while two studies provided evidence of the impact of BNP tests on patient management. In general, the patient populations were adult with symptoms of HF, primarily acute dyspnoea (breathlessness). Only one study specifically looked at a paediatric population (Koulouri et al 2004).

The prevalence of HF in the studies conducted in an acute care setting (see Table 11) was in the range 33%–71%. The latter (highest) rate was in a patient population with acute *severe* dyspnoea who presented to an ED (Logeart et al 2002).

Are BNP assays accurate?

Twelve studies provided diagnostic accuracy outcomes in terms of the ability of a BNP test to correctly identify or rule out HF in a hospital setting.

There was marked heterogeneity in the sensitivity and specificity outcomes reported in these studies. This was probably due to the various assay types that were assessed in the studies, along with the different cut-off points to rule out HF. Pooled sensitivity and specificity results are inappropriate when there is between-study heterogeneity. However, by stratifying results by the type of test and setting, it is possible to see the range of diagnostic accuracy outcomes associated with BNP testing. They are as follows:

The Triage immunofluorescence assay (Biosite Diagnostics) was by far the most commonly used BNP assay in these studies (11 studies). At the manufacturer's recommended threshold for ruling out HF (100 pg/mL), the sensitivity of the tests was in the range 78%–98% and specificity in the range 31%–94%. Negative predictive values at an optimised cut-off point were in the range 91%–98% (see Table 11).

The largest study to date where the Triage kit was used was the multicentre 'Breathing Not Properly' study (Maisel et al 2002). In this average quality study the negative predictive value was 96 per cent at 50 pg/mL for ruling out HF; while in a subgroup analysis of the same data, Knudsen et al (2004) reported a negative predictive value of 88 per cent at a cut-off point of 100 pg/mL. McCullough et al (2002) re-analysed the data from this study and placed it in the context of the effect of BNP testing alone or in

combination with clinical judgement in the diagnosis of HF. They determined that a BNP level of 100 pg/mL would add to clinical judgement and increase diagnostic accuracy from 74 per cent to 81.5 per cent ($p<0.0001$). With the use of ROC curves it was determined that clinical judgment in the ED had an area under the curve of 0.86 [95%CI 0.84, 0.88] at correctly discriminating HF from other conditions; for BNP testing it was 0.90 [95%CI 0.88, 0.91], and using both diagnostic modalities it was 0.93 [95%CI 0.92, 0.94].

One study used the AxSYM enzyme immunoassay (Abbott Laboratories) to test for BNP levels (Mueller et al 2005c). At four different cut-off points the highest sensitivity achieved was 96 per cent, the highest specificity was 86 per cent and the optimal negative predictive value achieved was 93 per cent (see Table 11).

Finally, one study using an in-house assay (Fleischer et al 1997) reported a sensitivity of 84 per cent, specificity of 95 per cent and an optimal negative predictive value of 92 per cent.

In general, therefore, BNP tests are usually characterised by high sensitivity but lower and variable specificity. The negative predictive values are, however, uniformly high (>90%).

Meta-analysis

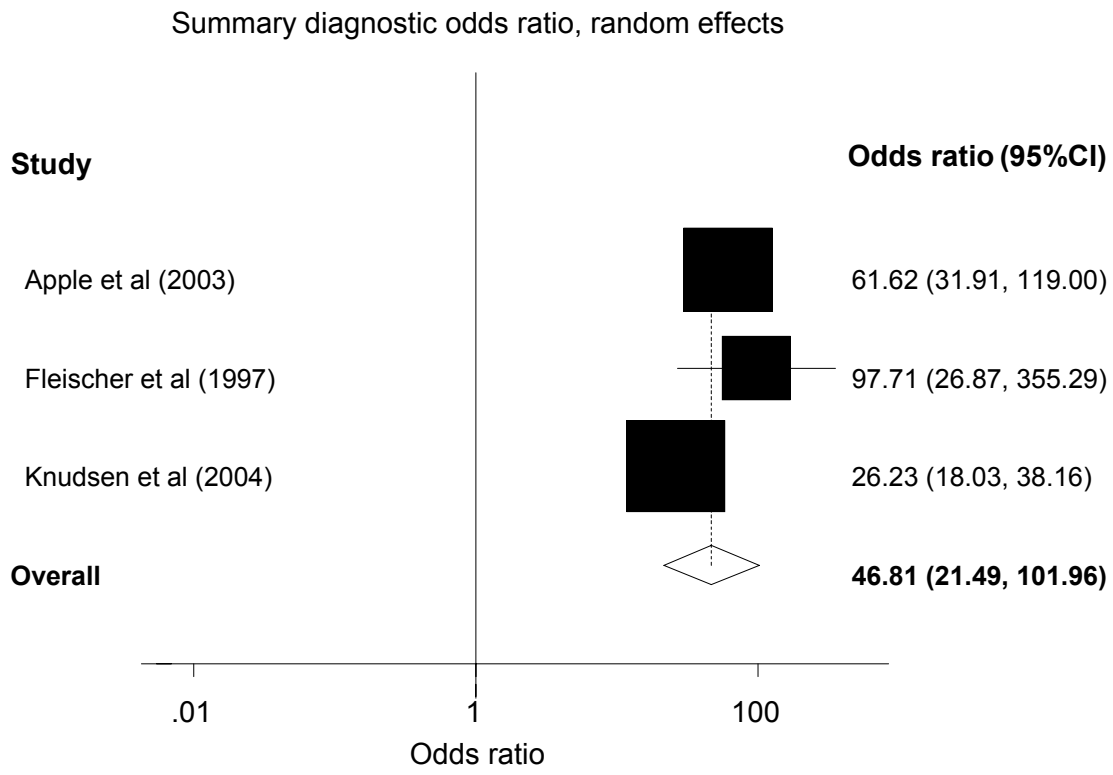
A summary diagnostic odds ratio (sDOR) was calculated based on the three BNP diagnostic studies in a hospital setting where raw data could be extracted. One of these studies used radioimmunoassays to test for BNP levels (Fleischer et al 1997), while two used the commercial immunofluorescence assay Triage kit (Apple et al 2003; Knudsen et al 2004). The comparator (or reference standard, where applicable) varied a little across the studies—although all used some form of clinical, or consensus clinical, diagnosis. A random effects model was used as there was statistically significant heterogeneity between the studies ($\chi^2 = 7.54$, d.f. = 2, $p=0.023$; I^2 variation in odds ratio attributable to heterogeneity = 73.5%). The pooled odds of a positive BNP test in those patients with HF, compared to those without HF, was nearly 47 times (sDOR = 46.81, 95%CI 21.5, 102.0; test of OR = 1 : z = 9.68 $p<0.0001$) (see Figure 7).

Both the Begg and Egger tests for publication bias were inconclusive (Egger coef = 5.9, 95%CI -42.5, 54.3, $p=0.365$), as with only three studies in the meta-analysis the analysis of publication bias was underpowered.

In order to determine whether the heterogeneity across these studies was due to differences in test threshold (Knudsen et al (2004) and Apple et al (2003) used cut-off points of 100 pg/mL for the Triage assay, whereas Fleischer et al (1997) used a cut-off point of 173 pg/mL for their in-house assay), a summary ROC curve was constructed. This was done after transforming the true positive rate (sensitivity) and false positive rate (1–specificity) for each study through the logarithm of their odds and then conducting a meta-regression analysis on the transformed data. The estimated regression coefficient (slope) of -0.3 was not statistically significant ($p=0.570$, 95%CI -4.44, 3.92), indicating that the heterogeneity observed could possibly be explained by a threshold effect, and therefore the accuracy of the test may be explained by a symmetrical ROC curve and a single diagnostic odds ratio (sDOR = 46.81, 95%CI 21.5, 102.0). It should be cautioned, however, that although the confidence intervals of the individual studies appear to overlap, as there were only three studies in this meta-regression analysis it is possible that the null hypothesis (of homogeneity in diagnostic accuracy across the studies) was not

rejected due to a lack of statistical power rather than any heterogeneity being explained by a threshold effect.

Figure 7 Diagnostic meta-analysis of the odds of a BNP test to accurately identify heart failure



Do BNP assays (linked evidence) change patient management?

Kosowsky et al (2003) in a research forum abstract presented in 2003 give a brief description of the impact of BNP testing on medical decision-making concerning older patients presenting to an ED with dyspnoea (level IV intervention evidence). Patients were recruited consecutively and physicians were blinded to BNP results obtained from a commercial assay. It was found that on disclosure of BNP levels, physicians rarely changed a primary clinical diagnosis of HF. High BNP levels only confirmed the diagnosis. However, for patients with a secondary or differential diagnosis of HF, the physician changed the diagnosis to a *primary* diagnosis of HF in 22 per cent of patients and changed management in 28 per cent of cases. In patients without a diagnosis of HF, disclosure of BNP levels changed the physician's decision to a primary diagnosis of HF in one case (2%) and a secondary diagnosis in four cases (9%); and changed medical management in five cases (11%).

Teboul et al (2004) reported the impact of BNP testing on the clinical diagnosis initially provided by emergency physicians travelling with mobile intensive care units (MICU) as part of the Emergency Medical Service in Paris, France. Emergency calls for 52 patients with dyspnoea were addressed, with 43 patients exhibiting respiratory distress upon the MICU's arrival. Physicians made an initial diagnosis, conducted a BNP test using the Triage assay, and then either retained the initial diagnosis or changed it on the basis of the BNP level result. This pre-test/post-test case series design (level IV intervention evidence) has the potential for bias and, due to the lack of a control group, it is difficult

to determine whether the results are associated with the BNP intervention being tested or another factor. Nevertheless, this study found that physicians changed their diagnoses for 18 of 52 patients based on the BNP result (although one of the 52 patients had a missing BNP value due to technical difficulties). All diagnoses where the physician had attributed a cardiac cause were confirmed by the BNP result. The majority of the 18 corrected diagnoses occurred where the physician had indicated that he/she was uncertain of the diagnosis and the BNP level subsequently provided additional information as to whether the dyspnoea could be ruled out as a cardiogenic or, alternatively, a pneumologic cause. The actual (reference standard) diagnosis is unknown, as the patients were not followed to see whether they had confirmed HF.

Does treatment on the basis of a BNP assay change health outcomes?

Linked evidence was not systematically assessed to address the impact of treatment on patient health outcomes, as there was direct evidence available (Mueller et al 2004b) for the effect of BNP testing, as well as subsequent earlier or more accurate treatment, on patient health. The section in the Background of this document also discusses—in a non-systematic manner—the high-level evidence supporting the well-known beneficial effects of the current treatments and medications for HF.

Table 10 Summary of included BNP diagnostic accuracy studies in the hospital setting—characteristics and quality appraisal

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests[a]	Outcomes assessed	Study quality[b]	Applicability[b]
			N	Selection criteria				
(Alibay et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Boulogne Billancourt, France	160	Referred to ED for dyspnoea	Medical hx Clinical exam	Sensitivity Specificity Negative predictive value	Level II diagnostic evidence Q1	P1
(Apple et al 2003)	Retrospecti ve cohort – cross-classified	Two hospitals – hospital-wide usage Minneapolis, MN and Hartford, CT. USA	334	Chart reviews of patients ruled in/out for HF or being monitored for HF therapy decisions	Not stated	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P2
(Dao et al 2001)	Prospective cohort – cross-classified	Hospital – ED San Diego Veteran's Healthcare System, USA	250	Convenience sample presenting to ED with symptoms of dyspnoea	Medical hx Clinical exam Blood tests CXR	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P2
(Dokainish et al 2004)	Prospective cohort – cross-classified	Hospital Houston, Texas, USA	122 [ITT: 145]	Hospital inpatients referred to cardiology consult service for suspected HF	Medical hx Clinical exam Laboratory tests Radiographic tests	Sensitivity Specificity	Level III-1 diagnostic evidence Q2	P2
(Fleischer et al 1997)	Prospective cohort – cross-classified	Hospital Christchurch, New Zealand	123	Patients requiring urgent admission to hospital for acute dyspnoea	Medical hx Clinical exam Blood tests Spirometry ECG CXR	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P1
(Knudsen et al 2004) *Subgroup analysis of Maisel 2002*	Prospective cohort – cross-classified	Multicentre study of five US and two European teaching hospitals – EDs	880 [ITT: 1586]	Patients presenting to ED with sudden onset of dyspnoea or worsening of chronic dyspnoea	Medical hx Clinical exam Blood tests ECG CXR	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests[a]	Outcomes assessed	Study quality[b]	Applicability[b]
			N	Selection criteria				
(Koulouri et al 2004)	Prospective cohort – cross-classified	Hospital – ED, paediatric ICU and wards, cardiothoracic ICU	49 [ITT: 51]	Infants and children presenting with objective findings of respiratory distress	Medical hx Clinical exam CXR	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P2
(Krishnaswamy et al 2001)	Prospective cohort – cross-classified	Hospital San Diego Veteran's Healthcare System, USA	400	Inpatients and outpatients referred for echo-cardiography to evaluate left ventricular function	Medical hx Clinical exam	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P2
(Lainchbury et al 2003)	Prospective cohort – cross-classified	Hospital – ED Christchurch, New Zealand	205	Patients presenting to ED with dyspnoea	Medical hx Clinical exam Blood tests CXR Other diagnostic tests	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Logeart et al 2002)	Prospective cohort – cross-classified	Hospital Clichy, France	163 [ITT: 235]	Patients presenting to ED with acute severe dyspnoea	Medical hx Clinical exam ECG CXR	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Maisel et al 2002)	Prospective cohort – cross-classified	Multicentre study of five US and two European teaching hospitals – EDs	1586	Patients presenting to ED with predominant symptom of dyspnoea	Medical hx Clinical exam Blood tests CXR Other diagnostic tests	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Mueller et al 2005c)	Prospective cohort – cross-classified	Hospital – ED St John of God Hospital, Linz, Austria	251 [ITT: 276]	Patients presenting to ED with predominant symptom of dyspnoea	Medical hx Clinical exam ECG Blood tests CXR Liver sonography	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P2

HF = heart failure; hx = history; ECG = electrocardiogram; CXR = chest X-ray; ITT = intention-to-treat; ED = emergency department; ICU = intensive care unit; [a] Only tests that were mentioned or inferred from the study are included – there may be other prior tests that were not specifically reported; [b] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

Table 11 Summary of included BNP diagnostic accuracy studies in the hospital setting—results and precision estimates

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
(Alibay et al 2005)	Level II diagnostic evidence Q1	160	Age: 80±14 yrs M/F: 76/84 Unclear if HF hx	HF: 60/160 (38%)	Consensus clinical diagnosis – all data (cardiologists)	BNP Single-use fluorescence immunoassay Triage kit, Biosite Diagnostics		50 100 150 200	99 [95,100] 98 [93,100] 94 [87,98] 87 [79,93]	31 [22,41] 47 [37,57] 61 [51,71] 64 [54,73]
(Apple et al 2003)	Level III-2 diagnostic evidence Q3	334[b]	Age: HF – mean of 67 yrs; non-HF – mean of 61 yrs M/F: 52%/48% Unclear if HF hx	HF: 172/334 (52%)		BNP Single-use fluorescence immunoassay Triage kit, Biosite Diagnostics	Clinical diagnosis Physician discharge dictations based on NYHA clinical criteria; and ICD-9 codings	100	95 [91,97]	77 [70,83]
(Dao et al 2001)	Level III-1 diagnostic evidence Q2	250	Age: 63±0.9[c] yrs M/F: 94/6 Patients with HF hx included	HF: 97/250 (39%)	Consensus clinical diagnosis – all data (cardiologists)	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		80 100 115 120 150	98 [93,100] 94 [89,97] 90 [83,95] 90 [82,95] 87 [78,92]	92 [86,96] 94 [89,97] 96 [91,98] 96 [92,99] 97 [93,99]
(Dokainish et al 2004)	Level III-1 diagnostic evidence Q2	122 [ITT: 145]	Age: 56±13 yrs M/F: 62/60 Patients with HF hx included	HF: 70/122 (57%)	Clinical diagnosis – all data (cardiologist)	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		250	86 [78,92]	77 [68,85]
(Fleischer et al 1997)	Level III-2 diagnostic evidence Q3	123	Age: 68 (23–90) yrs M/F: 69/54 Patients with HF hx included	HF: 43/123 (35%)		BNP In-house assay	Clinical diagnosis – based on intent to treat HF with diuretic therapy within 24 hours of admission	173	84 [69,93]	95 [88,99]
(Knudsen et al 2004)	Level III-1 diagnostic	880 [ITT: 1586]	Age: 64±16 yrs M/F: 482/398	HF: 447/880	Consensus clinical diagnosis – all	BNP Fluorescence		100 200	90 [87,92] 80 [71,87]	75 [71,79] 87 [79,93]

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
Subgroup analysis of Maisel 2002	evidence Q2		Patients with HF hx included	(51%)	data (independent cardiologists)	immunoassay Triage kit, Biosite Diagnostics		300 400	71 [61,80] 64 [54,73]	90 [82,95] 92 [85,96]
(Koulouri et al 2004)	Level III-2 diagnostic evidence Q3	49 [ITT: 51]	Age: n/a for overall group M/F: n/a Unclear if HF hx	HF: 23/49 (47%)	Clinical diagnosis – New York University Pediatric Heart Failure criteria and echocardiography	BNP Fluorescence immunoassay Triage kit, Biosite		40 60 80 100	91 [72,99] 83 [61,95] 78 [56,93] 78 [56,93]	77 [56,91] 77 [56,91] 81 [61,93] 85 [65,96]
(Krishnaswamy et al 2001)	Level III-1 diagnostic evidence Q2	400	Age: n/a for whole group M/F: 385/15 Patients with HF hx included	HF: 132/400 (33%)	Clinical diagnosis – Framingham criteria and echo-cardiography (cardiologists)	BNP Fluorescence immunoassay Triage kit, Biosite Diagnostics		107	86 [78,92]	70 [60,79]
(Lainchbury et al 2003)	Level III-1 diagnostic evidence Q2	205	Age: 70±14 yrs M/F: 100/105 Patients with HF hx included	HF: 70/205 (34%)	Consensus clinical diagnosis – all data (independent cardiologists)	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		69 104 208 277 346	97 [91,99] 97 [91,99] 94 [87,98] 83 [74,90] 77 [68,85]	44 [34,54] 49 [39,59] 70 [60,79] 78 [69,86] 84 [75,91]
(Logeart et al 2002)	Level III-1 diagnostic evidence Q2	163 [ITT: 235]	Age: n/a for whole group M/F: 109/54 Patients with HF hx included	HF: 115/163 (71%)	Consensus clinical diagnosis – all data (two cardiologists and a pneumologist)	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		80[d] 100 200 300 400	97 [91,99] 96 [90,99] 93 [86,97] 88 [80,94] 79 [70,87]	27 [19,37] 31 [22,41] 56 [46,66] 87 [79,93] 93 [86,97]
(Maisel et al 2002)	Level III-1 diagnostic evidence Q2	1586	Age: 64±17 yrs M/F: 883/703 Patients with HF hx included	HF: 744/1586 (47%)	Consensus clinical diagnosis – all data (independent cardiologists)	BNP Immuno-fluorescence assay Triage kit, Biosite		50 80 125 150	97 [96,98] 93 [91,95] 87 [85,90] 85 [82,88]	62 [59,66] 74 [70,77] 79 [76,82] 83 [80,85]
(Mueller et al 2005c)	Level III-1 diagnostic	251	Age: 58–82 yrs M/F: 234/17	HF: 139/265	Clinical diagnosis – all data	BNP Enzyme		100 118	96 [92,99] 95 [90,98]	61 [52,70] 64 [55,73]

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
	evidence Q2	[ITT: 276]	Patients with HF hx included	(52%)	(cardiologist)	immunoassay AxSYM assay, Abbott Laboratories		160 295	90 [84,95] 80 [73,87]	73 [64,81] 86 [78,92]

HF = heart failure; hx = history; n/a = not available; NYHA = New York Heart Association classification; ICD = International Classification of Diseases; ITT = intention-to-treat; [a] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'; [b] 430 tests on 334 patients; [c] It is unclear whether this value is a standard error, rather than a standard deviation; [d] Values for cut-off points of 150 and 250 pg/mL also presented in the paper but not included.

NT-proBNP assays (direct evidence of effectiveness)

There was no direct evidence available concerning the effectiveness or impact of NT-proBNP assays on patient health outcomes.

NT-proBNP assays (linked evidence)

Summary – Linked evidence of diagnostic effectiveness of NT-proBNP

The linked diagnostic studies were relatively consistent in their findings that NT-proBNP tests are sensitive, although there was wide variation in the rates. In general, the assays had high negative predictive value. The impact of the NT-proBNP test on patient management and patient health outcomes in the hospital setting was not tested directly, although evidence indicates that treatment for HF has a beneficial impact on health outcomes. It is also likely that earlier and more accurate diagnosis and treatment of alternative conditions (particularly those with acute presentation) would be beneficial for the patient.

Seven studies met the inclusion criteria for providing linked evidence of the effectiveness of NT-proBNP testing in an acute care setting (see Table 12). Six studies provided evidence of the diagnostic accuracy of two NT-proBNP assays—the Elecsys chemiluminescent sandwich immunoassay developed by Roche Diagnostics (5 studies) and an enzyme immunoassay developed by Biomedica (1 study). One further study reported on the impact of NT-proBNP testing on patient management.

The patient populations were primarily adults presenting with acute dyspnoea as the major symptom of suspected HF. The prevalence of HF in all the acute care diagnostic accuracy studies was in the range 34%–83%.

Are NT-proBNP assays accurate?

Seven studies assessed the accuracy of the NT-proBNP tests at correctly identifying or ruling out HF in patients suspected of having the condition. Sensitivity and specificity rates varied considerably between the studies despite the fact that in most cases the same assay was being tested (see Table 13).

In the acute care setting, sensitivity was 68%–100%, specificity was 5%–93%, and negative predictive values for an optimal cut-off point were in the range 90%–100%. The variability in these diagnostic accuracy results is probably a consequence, in part, of the very different thresholds for ruling out HF used in the NT-proBNP studies. Unlike the Triage kit results for BNP testing, no single consistent threshold was presented in the NT-proBNP studies.

In the PRIDE study, Januzzi et al (2005) determined optimal diagnostic thresholds using age categorisation cut-offs of <50 years and >50 years. The corresponding cut-off points for **ruling in** HF for these age groupings were 450 pg/mL (sensitivity 98%, specificity 76%) and 900 pg/mL (sensitivity 90%, specificity 85%), respectively.

This study (level III-1 diagnostic evidence) also found that NT-proBNP testing in addition to clinical judgement (AUC = 0.96) was superior in the diagnosis of HF than either diagnostic method alone (0.94 for NT-proBNP and 0.90 for clinical judgement, p=0.006).

Therefore, like BNP testing, it would appear that in general NT-proBNP assays have high sensitivity at detecting HF but lower specificity. The variability in rates is quite wide, although the negative predictive values are uniformly high (>90%).

Meta-analysis

None of the six NT-proBNP diagnostic accuracy studies had raw data that could be extracted into a 2 x 2 table; thus, pooling of diagnostic accuracy outcomes was not feasible.

Do NT-proBNP assays (linked evidence) change patient management?

In a Danish study of 345 patients referred to a hospital-based clinic by their general practitioners for dyspnoea, echocardiographic confirmation of HF occurred for 81 patients. Nielsen et al (2004) determined that for 68 of these 81 patients, either no or inadequate treatment was being administered at the time of examination—thereby suggesting that earlier diagnosis (possibly through NT-proBNP testing) could enable earlier and more effective treatment (level IV intervention evidence). Nielsen et al (2004) also determined that 51 per cent of 287 patients could have safely forgone echocardiography on the basis of the NT-proBNP rule-out test.

Does treatment on the basis of a NT-proBNP assay change health outcomes?

The review of HF treatment provided in the Background section of this document points to very high-level evidence of treatment effectiveness for some therapies. Therefore, a systematic assessment of all the various HF treatments was considered unnecessary and unproductive in the context of determining the impact of NT-proBNP testing on health outcomes. It is probable that patients 'ruled out' from HF earlier, as a consequence of a NT-proBNP test, would also receive earlier and more accurate treatment than otherwise received when conventional diagnostic strategies are used. The impact of earlier and more accurate treatment on the patient would depend on the nature and severity of the large number of alternative pathologies that present with HF-like symptoms, eg asthma, chronic obstructive pulmonary disease, pneumonia.

Table 12 Summary of included NT-proBNP diagnostic accuracy studies in the hospital setting—characteristics and quality appraisal

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests	Outcomes assessed	Study quality[a]	Applicability[a]
			N	Selection criteria				
(Alibay et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Boulogne Billancourt, France	160	Referred to ED for dyspnoea	Medical hx Clinical exam	Sensitivity Specificity Negative predictive value	Level II diagnostic evidence Q1	P1
(Bayes-Genis et al 2004)	Prospective cohort – cross-classified	Emergency hospital department Barcelona, Spain	89 [ITT: 100]	Patients with symptoms of acute dyspnoea attending ED	Medical hx Clinical exam Other blood tests CXR ECG	Sensitivity Specificity Negative predictive value	Level II diagnostic evidence Q1	P1
(Januzzi et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Boston, Massachusetts, USA	599	Patients presenting to ED with dyspnoea	Medical hx Clinical exam Medication use Blood tests CXR ECG	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Jose et al 2003)	Prospective cohort – cross-classified	Emergency and outpatient hospital departments Vellore, India	119	Patients presenting to emergency or outpatient department with dyspnoea and associated symptoms	Medical hx Clinical exam Medication use CXR ECG	Sensitivity Specificity	Level III-1 diagnostic evidence Q2	P1
(Lainchbury et al 2003)	Prospective cohort – cross-classified	Hospital – ED Christchurch, New Zealand	205	Patients presenting to ED with dyspnoea	Medical hx Clinical exam Blood tests CXR Other diagnostic tests	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1

Study **Author(s)** (Year)	**Study design**	**Setting** Region, site	**Study population**		**Prior tests**	**Outcomes assessed**	**Study quality**[a]	**Applicability**[a]
			N	Selection criteria				
(Mueller et al 2005c)	Prospective cohort – cross-classified	Hospital – ED St John of God Hospital, Linz, Austria	251 [ITT: 276]	Patients presenting to ED with predominant symptom of dyspnoea	Medical hx Clinical exam ECG Blood tests CXR Liver sonography	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P2

HF = heart failure; hx = history; ED = emergency department; CXR = chest X-ray; ECG = electrocardiogram; ITT = intention-to-treat; [a] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

Table 13 Summary of included NT-proBNP diagnostic accuracy studies in the hospital setting—results and precision estimates

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
(Alibay et al 2005)	Level II diagnostic evidence Q1	160	Age: 80±14 yrs M/F: 76/84 Unclear if HF hx	HF: 60/160 (38%)	Consensus clinical diagnosis – all data (cardiologists)	NT-proBNP Chemiluminescent sandwich immuno-assay Elecsys 2010, Roche Diagnostics		280 600 1000 1250	100 [96,100] 100 [96,100] 97 [91,99] 87 [79,93]	5 [2,11] 51 [41,61] 63 [53,72] 66 [56,75]
(Bayes-Genis et al 2004)	Level II diagnostic evidence Q1	89 [ITT: 100]	Age: n/a M/F: 54/35 Unclear if HF hx	HF: 74/89 (83%) – decompensated HF = 58%, masked HF = 25%	Consensus clinical diagnosis – all data (cardiologists)	NT-proBNP Chemiluminescent sandwich immuno-assay Elecsys 1010, Roche Diagnostics		254 423 592 761 973 1099	99 [95,100] 96 [90,99] 94 [87,98] 91 [84,96] 91 [84,96] 90 [82,95]	47 [37,57] 60 [50,70] 73 [63,81] 73 [63,81] 93 [86,97] 93 [86,97]
(Januzzi et al 2005)	Level III-1 diagnostic evidence Q2	599	Age: 22–95 yrs M/F: 51%/49% Patients with HF hx included	HF: 209/599 (35%)	Consensus clinical diagnosis – all data (cardiologists)	NT-proBNP Chemiluminescent immunoassay Elecsys 1010, Roche Diagnostics		300 450 600 900 1000	99 [95,100] 98 [93,100] 96 [90,99] 90 [82,95] 87 [79,93]	68 [58,77] 76 [66,84] 81 [72,88] 85 [76,91] 86 [78,92]
(Jose et al 2003)	Level III-1 diagnostic evidence Q2	119	Age: 54±12 yrs M/F: 78/41 Unclear if HF hx	HF: 73/119 (61%)	Consensus clinical diagnosis Framingham criteria and echo-cardiography	NT-proBNP Enzyme immuno-assay Biomedica		1691	97 [91,99]	89 [81,94]
(Lainchbury et al 2003)	Level III-1 diagnostic evidence Q2	205	Age: 70±14 yrs M/F: 100/105 Patients with HF hx included	HF: 70/205 (34%)	Consensus clinical diagnosis – all data (independent cardiologists)	NT-proBNP Chemiluminescent immunoassay Elecsys 2010, Roche Diagnostics		1184 2030 2875 3721 4567	87 [79,93] 83 [74,90] 80 [71,87] 74 [64,82] 68 [58,77]	71 [61,80] 82 [73,89] 87 [79,93] 90 [82,95] 92 [85,96]

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
(Mueller et al 2005c)	Level III-1 diagnostic evidence Q2	251 [ITT: 276]	Age: 58–82 yrs M/F: 234/17 Patients with HF hx included	HF: 139/265 (52%)	Clinical diagnosis – all data (cardiologist)	NT-proBNP Chemiluminescent sandwich immuno-assay Elecsys 2010, Roche Diagnostics		292 125/450[b] 476 825	95 [90,98] 94 [89,97] 90 [84,95] 87 [80,92]	53 [43,62] 46 [37,56] 65 [55,74] 81 [72,88]

HF = heart failure; n/a = not available; hx = history; ITT = intention-to-treat; [a] The assessment of study quality followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'; [b] Dual cut-off point – 125 pg/mL for patients <75 years and 450 pg/mL for those ≥75 years of age.

Are B-type natriuretic peptide assays effective in the monitoring of heart failure?

Summary – Effectiveness of monitoring with cardiac natriuretic peptides

Good quality level II intervention evidence demonstrated that monitoring patients via NT-proBNP testing resulted in considerably fewer cardiovascular deaths and total cardiovascular events than when patients were monitored via clinical criteria. This reduction was statistically significant for both outcomes—and clinically important for the reduction in cardiovascular events.

An abstract reported similar results for BNP-guided monitoring, but more detail is necessary to determine whether this study could be considered supporting evidence.

Accurate and timely monitoring of HF patients is fundamental in achieving the best possible outcomes. Regular monitoring of clinical status enables the physician to assess treatment effectiveness, or lack thereof. Treatment can consequently be modified according to the patient's clinical response to the medication/intervention. Clinical evaluation can include assessment of functional capacity (NYHA class, maximal or submaximal exercise tests), fluid status (body weight, blood pressure, jugular venous extension), cardiac rhythm (ECG) or biochemical markers (urea, electrolytes and creatinine) (NICE 2003). The aim of this section is to assess whether the addition of B-type natriuretic peptides to the current clinical status workup would provide superior patient outcomes.

There was a substantial evidence base for B-type natriuretic peptides' response to pharmaceutical interventions; however, these studies do not provide any evidence of direct patient benefit (ie increased survival) associated with the introduction of BNP and NT-proBNP assays into monitoring protocols.

Three studies (Inomata et al 2003; Murdoch et al 1999; Troughton et al 2000) met the inclusion criteria for assessing the effectiveness of monitoring HF patients with the addition of B-type natriuretic peptides, compared with standard clinical assessment. Two of the three studies assessed the effectiveness of BNP (Inomata et al 2003; Murdoch et al 1999) while one used NT-proBNP for monitoring HF patients (Troughton et al 2000). Data were only able to be extracted from the good quality randomised controlled trial conducted in New Zealand (Troughton et al 2000). The remaining studies included an abstract (Inomata et al 2003) and an article that did not report on patient relevant outcomes (Murdoch et al 1999).

BNP assays

One abstract provided relevant information on the monitoring of HF patients through BNP testing compared to standard clinical evaluation. Inomata et al (2003) followed 73 consecutive stabilised HF patients, who were randomly allocated to receive successive outpatient management guided by either plasma BNP levels or conventional clinical assessment (level II intervention evidence). In the BNP group, treatment was intensified

for patients with BNP levels above 200 pg/mL. After 2 years of follow-up, significantly (p=0.04) fewer cardiovascular events (death or readmission) occurred in the BNP guided group compared to the group that received conventional treatment. Given that this study was published as an abstract only, there was insufficient detail on study design, baseline patient characteristics and follow-up to establish the reliability of these results.

A single-blinded randomised controlled trial by Murdoch et al (1999) assessed titration of vasodilator therapy according to plasma BNP levels compared with empiric clinical treatment. Patient relevant outcomes were not assessed and the study, with a total of 20 participants (10 in each study arm) and a 2-month follow-up, was underpowered to detect significant differences in the haemodynamic variables measured. They concluded that the BNP tailored vasodilator treatment approach is well tolerated, safe and associated with a more profound inhibition of the renin-angiotensin-aldosterone system. Assessment and specific details of this study are not included in this report due to its lack of patient relevant outcomes.

NT-proBNP assays

Troughton et al (2000) enrolled 69 HF patients in a double-blind randomised controlled trial that attempted to assess the potential benefit of NT-proBNP guided pharmaceutical treatment compared to that guided by clinical criteria (level II intervention evidence). The sample was recruited from a population of 35–85-year-old patients who were admitted to hospital with decompensated HF or were patients of a cardiology outpatient clinic. All patients had a LVEF of <40 per cent and were in New York Heart Association (NYHA) classes II to IV. Randomisation of patients occurred after they were stabilised via medication. Thirty-three patients were allocated to have their therapy guided by NT-proBNP and 36 patients received treatment guided by clinical assessment, which was based on the Framingham criteria for diagnosing decompensated HF. A Framingham score of ≥ 2 units indicates decompensated HF by summing 10 major (1 unit) and minor (0.5 unit) clinical categories. Major symptom categories were paroxysmal nocturnal dyspnoea, positive hepatojugular reflex, basal crackles and third heart sound. Minor categories included orthopnea, reduction in exercise tolerance, resting heart rate >100 beats per minute, jugular venous pressure >4 cm, hepatomegaly and peripheral oedema.

All patients were subject to a baseline clinical assessment in addition to a NT-proBNP assay, echocardiography, functional capacity (6-minute walk test) and quality of life (Minnesota Living with Heart Failure Survey) measurements. In the NT-proBNP group, if patients had NT-proBNP serum concentrations ≥1,691 pg/mL (converted from 200 pmol/L) they were assessed at fortnightly intervals, during which pharmaceutical treatment was intensified in a stepwise protocol (ACE inhibitors → loop diuretics → digoxin → additional diuretic → additional vasodilator) until the NT-proBNP serum concentration dropped below 200 pmol/L. An identical treatment protocol was followed in the clinical group, except that patients with a HF score ≥2 were seen at fortnightly intervals (with identical stepwise increase in treatment) until their HF score was <2. Follow-up every 3 months assessed primary and secondary outcomes, but pharmaceutical treatment was not modified according to HF score or the NT-proBNP concentration during these visits. Troughton et al (2000) reported a complete 6-month follow-up on all patients and an approximate 9.6 month median follow-up in 32 and 29 patients in the NT-proBNP and clinical groups, respectively.

Due to effective randomisation, treatment groups were similar in their baseline characteristics for age, gender distribution, comorbidities, medication and disease severity. Troughton et al (2000) reported a significantly ($p=0.02$) higher rate of cardiovascular death, hospital admission and outpatient HF in the clinical group (54 events) compared to the NT-proBNP group (19 events). The level of significance was higher ($p<0.001$) when using Poisson regression to adjust for small baseline differences in LVEF, NT-proBNP concentration, age, NYHA class, resting heart rate, medication and systolic blood pressure between the treatment arms. The most severe cardiovascular event (death) was also significantly different ($p=0.03$) between the groups, with only one event in the NT-proBNP group compared to seven in the clinical group. When analysed independently, rates of readmission to hospital were not significantly different between the groups. The reduction in total cardiovascular events was presumably achieved by increased dosage of ACE-inhibitor ($p=0.03$) and loop diuretic ($p=0.34$), and a greater prevalence of spironolactone use ($p=0.049$; additional diuretic) at 6 months in the NT-proBNP group. Calculations indicate that there was a 52 per cent reduction in the risk of cardiovascular events in patients monitored with NT-proBNP compared to patients in the clinical group. Therefore, two patients would need to be monitored with NT-proBNP, compared to a clinical evaluation protocol, to prevent one cardiovascular event (Table 14). This is a clinically important result.

Troughton et al (2000) also reported on group differences in the secondary outcomes of LVEF, functional capacity, quality of life and supine blood pressure. The NT-proBNP group's LVEF (determined by echocardiography) improved by 8.3±2.2 per cent compared to 5.3±1.8 per cent in the clinical group, although this difference was not statistically significant ($p=0.23$). Other relevant secondary variables were only commented upon in the text, with Troughton et al (2000) reporting that quality of life scores remained stable and functional capacity improved similarly in both treatment groups.

Table 14 Effectiveness of NT-proBNP guided treatment vs conventional assessment

Study	Quality/ level	Outcomes	BNP group (n=33)	Clinical group (n=36)	Number needed to treat [95%CI]	Relative risk [95%CI]
(Troughton et al 2000)	Random-ised controlled trial Level – II QS - 23/26	Cardiovascular death Clin I – 2/4 Rel –1/5 Stat prec – good	1/33 (3%)	7/36 (19%)	6 [3, 45]	0.16 [0.02, 1.2] p=0.03
		Cardiovascular hospital admission Clin I – 4/4 Rel – 1/5 Stat Prec – poor	7/33 (21%) 7 patients were responsible for 7 admissions	11/36 (31%) 11 patients were responsible for 21 admissions	n/a	0.69 [0.31, 1.6] p=0.38
		Total cardiovascular events (cardiovascular death, cardiovascular hospital admission, any new outpatient episode of decompensated heart failure) Clin I - 1/4 Rel – 1/5 Stat Prec – excellent	15/33 (45%) 15 patients were responsible for 19 total cardiovascular events	34/36 (94%) 34 patients were responsible for 54 total cardiovascular events	2 [1, 3]	0.48 [0.33, 0.71] p<0.001

Clin I = rank scores for the clinical importance of the benefit (1/4 ranked highest and 4/4 ranked lowest); Rel = rank scores for the relevance of the evidence (with 1/5 ranked as a highly relevant outcome and 5/5 as an unproven surrogate outcome); Stat prec = statistical precision; n/a = not applicable

Discussion

In diagnostic studies of either BNP or NT-ProBNP assays, high rates of heart failure (HF) were observed in the acute care setting in the range 33%–83% (median 47%). This is as expected, and probably applicable to that portion of the Australian population presenting to emergency hospital departments with symptoms suggestive of HF.

Safety of NT-proBNP and BNP assays

Studies included in this Assessment Report did not mention physical or psychological harms occurring as a consequence of B-type natriuretic peptide testing. The likelihood of physical harms is low and similar to that of any blood test. However, it is possible that false positive results could engender inappropriate treatment leading to adverse events for the patient, and false negative results could lead to delayed treatment. Psychological harms (eg anxiety) can be associated with the delivery of a diagnosis (whether correct or incorrect); however, this is the case with all diagnostic tests.

Effectiveness of BNP assays in the diagnosis of heart failure in the hospital setting

Diagnostic accuracy

The pooled odds of a positive BNP test in those patients with HF was nearly 47 times that of a positive BNP test in those without HF (sDOR = 46.81, 95%CI 21.5, 102.0; test of OR = 1 : z = 9.68 $p<0.0001$). This is a strong indication of the BNP test's ability to discriminate between the presence and absence of disease. The heterogeneity in results between the studies was possibly explained by differences in test threshold between the studies.

The value of BNP tests is as a highly sensitive test such that a negative result 'rules out' the diagnosis (Sackett et al 1991). This is based on the notion that sensitivity and negative predictive value share the 2 x 2 diagnostic table cell for false negatives (FN) in their denominator. As sensitivity increases toward 100 per cent, FN decreases toward zero. As FN decreases toward zero, negative predictive value increases toward 100 per cent. The negative predictive value at optimal cut-off points for the BNP test was predominantly greater than 90 per cent in the studies included in this Assessment Report. BNP testing therefore appears to act as a 'first line' diagnostic tool to identify patients that should or should not be referred to echocardiography or other diagnostic tests to confirm a clinical diagnosis of HF. Its role, therefore, is not to act as a 'reference standard' test as its specificity is generally not high.

The accuracy of the BNP and NT-proBNP assays has been suggested to be affected by factors such as age, body mass index, history of HF and current treatment with ACE inhibitors or beta-blockers. Several studies have investigated different cut-off points (diagnostic thresholds) for the B-type natriuretic peptides in order to optimise their ability to discriminate—or 'rule out'—HF in certain populations, with mixed results (McCullough et al 2003a; Gustafsson et al 2003; Mueller et al 2005c; Januzzi et al 2005). This is an area requiring further research.

Impact on patient management

Three studies provided direct evidence of the effect of BNP testing on patient management in an emergency setting (Kosowsky et al 2003; Mueller et al 2004b; Teboul et al 2004). Results from the two lower quality pre-test/post-test case series—one of which was an abstract—indicate that the main impact of the test was to change physician diagnoses that were equivocal. Primary diagnoses of HF were rarely altered, with the test result usually only providing confirmation of the clinical diagnosis. It was those situations where the clinical decision was not clear-cut that the BNP test added diagnostic value. A change in the management of the patient subsequently occurred in 28 per cent of cases in one study (Kosowsky et al 2003).

The highest level and quality of evidence available—the randomised controlled trial conducted by Mueller and colleagues (Mueller et al 2004b)—assessed the impact of supplementing clinical diagnosis with BNP testing on those clinical management outcomes such as time to discharge, time to treatment, and hospital and intensive care admission rates. This trial found that BNP-assisted diagnostic assessment significantly shortened the hospital stay of all patients presenting to an ED by a median of 3 days ($p=0.001$). This result was not observed in those patients with kidney disease. There was a statistically significant reduction in time to treatment, by just under a median of half an hour, in the group receiving BNP supplemented diagnostic assessment. However, the clinical importance of such a difference in the commencement of treatment is unclear. Patients receiving BNP-assisted diagnostic assessment were admitted to hospital [RR = 0.88, 95%CI 0.81, 0.97] and intensive care [RR = 0.62, 95%CI 0.42, 0.91] less often than when receiving conventional diagnostic assessment. This is an important difference as only approximately 10 patients would need to be diagnosed with BNP-supplemented assessment, compared to conventional diagnostic testing, to reduce one hospital or ICU admission. Patients with kidney disease (and various comorbidities) were, however, unaffected by the different diagnostic strategies, with admission rates being no different between the two groups. Presumably this is a consequence of the threshold being applied for BNP testing in this trial. Two general cut-off points were used— <100 pg/mL for ruling out HF and >500 pg/mL for ruling in HF. For patients with kidney disease, and for known comorbidities, the threshold may have to be modified—as has been suggested in other studies (McCullough et al 2003a).

The 'ruling in' aspect of the Mueller et al (2004b) trial and delivery of HF treatment on the basis of a high BNP level are unlikely to be replicated in Australia. It is likely that ED clinicians with a high index of suspicion for HF in a particular patient will simply order a confirmatory echocardiogram, given the high number of false positives associated with a BNP test and the impact of comorbidities on test results.

Impact on health outcomes

The impact of BNP testing on health outcomes was measured directly in one good quality randomised controlled trial (Mueller et al 2004b). These health outcomes were pre-specified as secondary outcomes, however, and so the BASEL trial was not necessarily powered to distinguish between the different diagnostic strategies (BNP supplemented clinical diagnosis vs conventional clinical diagnosis) for these outcomes. With respect to in-hospital mortality and 30-day mortality, patients receiving BNP-assisted diagnostic assessment had a reduced rate of death compared to those receiving conventional diagnostic testing, although this was not statistically significant (in-hospital mortality: RR = 0.62, 95%CI 0.32, 1.22, p = 0.21; 30-day mortality: RR = 0.79, 95%CI

0.47, 1.34, p = 0.45). Wide confidence intervals indicate that the analysis was underpowered. However, for elderly patients assessed in a pre-specified subgroup analysis of this trial, there appeared to be a particular benefit with BNP-supplemented diagnostic assessment, with a trend towards a reduction in in-hospital mortality (RR = 0.46, 95%CI 0.21, 1.03, p = 0.051) and a statistically significant and clinically important reduction in 30-day mortality (RR = 0.51, 95%CI 0.26, 0.98, p = 0.039). The latter indicates that 12 elderly patients would require diagnosis with a BNP-supplemented diagnostic strategy, compared to conventional diagnosis, to prevent one death within 30 days.

To definitively determine the direct effect of BNP testing on patient health outcomes, a trial would need to be conducted with sufficient sample size to detect a statistically significant difference in health outcomes (that was also clinically important) for patients receiving BNP-assisted diagnostic assessment compared to conventional assessment. In the absence of such data, the good quality trial by Mueller et al (2004b) indicates that mortality rates were consistently reduced in patients receiving BNP-assisted assessment, and that for the elderly this reduction was both statistically significant and clinically important.

An overall evaluation of the body of evidence supporting BNP testing is provided in Table 15.

Table 15 Assessment of body of diagnostic evidence for BNP assay in hospital setting

Component	A	B	C	D
	Excellent	Good	Satisfactory	Poor
Volume of evidence	Several level I or II studies with low risk of bias			
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact		Substantial[a]	Moderate[b]	
Generalisability		Population(s) studied in the body of evidence are similar to the target population		
Applicability		Applicable to Australian healthcare context with few caveats		

[a] 'Substantial' for the effect on patient management (direct evidence); [b] Moderate for the effect on patient health outcomes, primarily due to limited evidence being available (direct evidence)

Effectiveness of NT-proBNP assays in the diagnosis of heart failure in the hospital setting

Diagnostic accuracy

The evidence base for assessing the diagnostic accuracy of NT-proBNP testing was not as extensive as that available for BNP testing. However, like BNP testing, it would appear that in general NT-proBNP assays have high sensitivity at detecting HF but lower specificity. The variability in accuracy rates is quite wide, although the negative predictive values are uniformly high (>90%). Heterogeneity between the studies may be partly caused by the different test thresholds for ruling in and ruling out HF reported across the studies for the most commonly used assay. Data could not be plotted in sROC space to determine the effect of test thresholds as raw data were not available with which to conduct a meta-analysis. Heterogeneity in results may also be due to the nature of the populations being tested. Test results on those populations with higher levels of drug treatment for suspected HF may have reduced sensitivity in some studies, as ACE inhibitors decrease plasma levels of natriuretic peptides (Gustafsson et al 2003).

NT-proBNP assays, as is the case for BNP assays, appear to be of value as a highly sensitive test such that a negative test result 'rules out' the diagnosis of HF.

Impact on patient management

Given that NT-proBNP testing appears to have a high negative predictive value for ruling out HF, it is likely that many patients could safely forgo further diagnostic testing for HF on the basis of this one blood test. Nielsen et al (2004) determined that 51 per cent of 287 patients could safely forgo echocardiography on the basis of a NT-proBNP test (level IV evidence).

Impact on health outcomes

The direct impact of NT-proBNP testing on health outcomes was not assessed in any of the included studies. High-level evidence of the effect of early treatment for HF (from linked evidence) on patient health outcomes indicates that treatment is beneficial, although this was not investigated systematically in this report. It is probable that similar early detection and treatment of other pathologies would also benefit the patient, although this could depend on the acuteness or nature of the condition at presentation.

An overall assessment of the body of evidence supporting NT-proBNP testing is provided in Table 16

Table 16 Assessment of body of diagnostic evidence for NT-proBNP assay in hospital setting

Component	A	B	C	D
	Excellent	Good	Satisfactory	Poor
Volume of evidence		One or two level II studies with low risk of bias or a SR/multiple level III studies with low risk of bias		
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact				Slight or restricted[a]
Generalisability		Population(s) studied in the body of evidence are similar to the target population		
Applicability		Applicable to Australian healthcare context with few caveats		

[a] Relates to the impact of NT-proBNP testing, as measured *directly*, on patient management and patient health outcomes—there is currently no *direct* evidence available in this setting. *Linked* evidence would suggest that in all probability the effect may be moderate but this would require further trial data in this setting.

Effectiveness of B-type natriuretic peptide assays for monitoring of heart failure

BNP

One abstract (Inomata et al 2003) of a randomised controlled trial reported that outpatients managed by BNP-guided therapy had significantly fewer 'deaths or hospital readmissions' over 2 years of monitoring compared to those who were randomly assigned to a conventional clinical assessment group. It is unclear what 'conventional clinical assessment' involved and the sample characteristics, completeness of follow-up and specific study details are unknown. The evidence for effectiveness of BNP monitoring versus clinical assessment is therefore limited.

NT-proBNP assays

Albeit limited to one good quality randomised controlled study on 69 patients (Troughton et al 2000), there is some evidence to suggest that monitoring via NT-proBNP provides superior outcomes when compared to clinical monitoring. In this trial a statistically significant reduction in cardiovascular deaths and total cardiovascular events was observed. Seven cardiovascular deaths occurred in the clinically monitored group compared to only one in the NT-proBNP group. A *clinically* important reduction in cardiovascular events was also observed in the BNP-guided therapy group. The evidence from this New Zealand study is likely to be relevant to an Australian population.

Overall

The results of Troughton et al (2000), taken together with the fact that B-type natriuretic peptide levels seem to provide prognostic information over and above clinical criteria (see Appendix I of this report), it is reasonable to hypothesise that adjusting pharmaceutical therapy to achieve lower hormone levels would improve outcomes of HF patients compared to clinical monitoring. The potential benefit of adding B-type natriuretic peptide assessment to clinical assessment, compared to clinical assessment alone, for monitoring of HF patients requires further research. The paucity of data on the relative effectiveness of B-type natriuretic peptides for monitoring HF patients highlights the potential for larger controlled studies to verify the results of the small New Zealand trial (Troughton et al 2000) on which our conclusions are predominantly based. Currently an extension of the aforementioned pilot study is being conducted (known as BATTLE SCARRED), with a predicted enrolment of >300 patients, to assess the effectiveness of monitoring HF patients with NT-proBNP. The trial completion date is mid 2007 and thus, publication would be expected by late 2008.

The overall body of evidence for monitoring of HF patients via B-type natriuretic peptides is limited in volume, with currently only two relevant studies contributing to answering this question. However, one of these studies is a well-designed randomised controlled trial and, as such, adds significant weight to a low volume evidence base (Table 17).

Table 17 Assessment of body of evidence on monitoring effectiveness

Component	A	B	C	D
	Excellent	**Good**	**Satisfactory**	**Poor**
Volume of evidence		One or two level II studies with low risk of bias or a SR/multiple level III studies with low risk of bias		
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact		Substantial[a]		
Generalisability		Population(s) studied in the body of evidence are similar to the target population		
Applicability	Directly applicable to Australian healthcare context			

[a] Although only one small NT-proBNP trial reported on patient health outcomes and only one BNP abstract reported on health outcomes

What are the economic considerations?

The purpose of an economic evaluation is to assist decision-makers in ensuring that society's ultimately scarce resources are allocated to those activities from which we will get the most value. That is, it seeks to enhance economic efficiency.

Economic evaluation under the MSAC process focuses on the scarce resources available within the Australian health system. It asks whether these scarce resources would be better spent on producing the amount of health gain obtainable through the intervention in question or through the identified comparator intervention(s).

An expert advisory panel agreed that there was sufficient **diagnostic effectiveness** evidence for B-type natriuretic peptides in the hospital setting to warrant a cost-effectiveness analysis. In contrast, the evidence base for heart failure (HF) monitoring effectiveness was mainly limited to one small controlled trial and, as such, an in-depth economic analysis for this indication was not required. A much larger (300 plus patients) randomised controlled trial that assesses the effectiveness of monitoring HF patients with NT-proBNP (BATTLE SCARRED) is currently underway in New Zealand and should provide more information on this topic. The trial completion date is mid 2007, and thus publication would be expected by late 2008.

Objective

The aim of the present economic evaluation is to review the cost-effectiveness of adding B-type natriuretic peptide testing[11] to current Australian protocols in the diagnosis of (ie ruling out) HF in an ED setting, and to provide an indication of the extent of uncertainty entailed. The perspective of the analysis is that of society. Direct costs of informal care and indirect costs (ie productivity costs) are not considered.

This economic evaluation is thus a trial-based economic analysis of the use of B-type natriuretic peptide tests in an ED setting.

As often occurs in economic evaluation, a paucity of data has necessitated several crucial assumptions. It is important that the results be interpreted in the light of the likely validity of these assumptions.

[11] Due to insufficient evidence for costs and outcomes for each peptide assay, it is assumed that the cost and effectiveness of BNP and NT-proBNP testing are equivalent.

Introduction of B-type natriuretic peptide testing in an emergency department setting: trial-based economic analysis

Summary

Relying on the results of the key trial by Mueller et al (2004b), the incremental costs and outcomes of the management of suspected HF patients presenting to an ED were examined.

Under the conservative assumption that there is no difference in survival beyond 30 days (which was justified by 180-day follow-up results recently published from the key trial), the point estimate of incremental costs and incremental effectiveness suggests that the addition of B-type natriuretic peptide testing to conventional diagnostic strategies **dominates** the use of conventional diagnostic strategies alone. Thus, the point estimate suggests that performing a B-type natriuretic peptide test in the ED setting leads to a superior patient health outcome at a lower cost. Although the 95 per cent confidence interval indicates that the point estimate is subject to some uncertainty, the larger part of the joint probability distribution (78.8%) of incremental costs and effectiveness indicates dominance over conventional diagnostic strategies in this setting.

Regarding financial outlays, the Commonwealth Government would incur an additional net expenditure of $352 thousand under Medicare due to the introduction of B-type natriuretic peptide testing for private patients in private hospital EDs. Although the majority of B-type natriuretic peptide tests will be performed in public hospital EDs, this is unlikely to lead to Australian Government expenditure savings because of capacity constraints, but may make additional public resources available for other patients in need.

Highest level of evidence available

Based on the NHMRC categories, the highest level of evidence available for this economic analysis is one good quality, single-blind (outcome assessment) randomised controlled trial assessing the direct impact of BNP testing on both patient management and patient health outcomes (Mueller et al 2004b). This trial, known as the 'B-type natriuretic peptide for acute shortness of breath evaluation (BASEL) study', was conducted in Basel, Switzerland, on patients presenting to the ED of a university hospital with acute dyspnoea (or breathlessness) as the primary symptom. Dyspnoea is associated with HF and with certain pulmonary conditions.

Note that it is unusual to have direct evidence of effectiveness of a triage diagnostic test, but in the ED setting in this case such a randomised controlled trial (Mueller et al 2004b) was available.

Nevertheless, it must be emphasised that the trial-based economic evaluation under Australian conditions in this section of the report only includes the health gains that were reported over the duration of the trial, that is the difference in mortality at 30 days. This economic evaluation does not include the possible gain in quality of life from the correct

diagnosis and treatment at any time period. A postscript considers the all-cause 180-day mortality rate.

How the test contributes to clinical decision-making

This economic evaluation is not about treating or not treating HF, but rather about shedding light on the choice between using and not using an additional test in the diagnostic workup. Heart failure, when managed, is characterised by a compensated/treated state, interspersed with decompensated (symptomatic) episodes. The use of B-type natriuretic peptide tests is to determine whether symptoms (primarily acute dyspnoea) are a result of a decompensated episode (HF) or not. If not, alternative diagnoses are considered, eg pulmonary conditions. A 30-day period from testing is likely to be sufficient to determine the cost-effectiveness of the test at ruling out a patient with acute dyspnoea symptoms suggestive of HF.

To date, there is no evidence to confirm that the introduction of a B-type natriuretic peptide test in the triage of ED patients has a statistically significant impact on survival beyond 30 days. The test is a triage test—it has a high negative predictive value and a variable positive predictive value. Its main function is to 'rule out' patients from a diagnosis of HF, rather than 'rule in' patients. When the test also indicates suspected HF, this usually only provides additional confirmation of what the rest of the diagnostic workup (symptoms, chest X-ray, laboratory tests) indicates. As reflected in the Mueller et al (2004b) study, B-type natriuretic peptide testing means that fewer patients go on to confirmatory diagnostic testing for HF because more have been 'ruled out', that is triaged. Treatment for HF and its consequent (and recognised) benefits is therefore not a prime consideration, as only a few more patients with symptoms suggestive of HF will receive treatment for HF as a consequence of adding B-type natriuretic peptide testing to the diagnostic workup, than without.

Population

The patients in the trial by Mueller et al (2004b) had a mean age of 70.3 years, with a slightly higher male representation. Prevalence of diagnosed HF in the trial participants, who presented to an ED with acute dyspnoea, was 48 per cent. A large proportion of the trial participants were receiving medications for existing conditions, with approximately one-quarter receiving treatment with beta-blockers, one-half on diuretics, and 40 per cent receiving ACE inhibitors. Common comorbidities included coronary artery disease, chronic obstructive pulmonary disease (COPD) and diabetes in 50, 31 and 23 per cent of patients, respectively. Baseline comorbidities and medication usage suggest that the sample consisted of both new and existing HF patients. The age, gender, comorbidities, presenting symptoms and HF status of the patients in the BASEL study appear to be similar to the target population in Australia.

Intervention and comparator

Patients consenting to participate in the trial were randomised to receive (1) conventional diagnostic assessment or (2) conventional assessment supplemented and guided by BNP testing.

1) The authors reported that, in general, the conventional diagnostic workup included a clinical history, physical examination, ECG, pulse oximetry, blood tests and chest X-ray. This protocol is identical to that recommended by the Cardiac Society of Australia and New Zealand, with one exception. Australian guidelines advocate echocardiograms to either differentially diagnose or confirm an equivocal diagnosis of HF. In the BASEL study, echocardiograms were recommended for all admitted and non-admitted patients (albeit only for those with suspected HF). Patients with differential diagnoses (ruled out from HF by a BNP test and/or clinical judgement) were diagnosed and treated without an echocardiogram. It is unknown what proportions of the 203 echocardiograms were performed on an in- and outpatient basis.

2) In the BASEL study the BNP test results were considered in context with the other clinical information and the physicians' clinical impressions. Hence, BNP was integrated to form a clinical diagnosis rather than overriding existing clinical information. A decision algorithm which was followed by physicians in the BNP diagnostic strategy trial arm indicated that: patients with BNP levels <100 pg/mL were unlikely to have HF and so other diagnoses were investigated; BNP levels >500 pg/mL most likely indicated HF and rapid HF therapy was initiated; intermediate (100–500 pg/mL) levels suggested that clinical judgement be used in conjunction with further diagnostic testing. **Note both the rule-out and rule-in aspects to this decision algorithm**. In the proposed Australian setting, B-type natriuretic peptide testing would be used in a similar fashion with the exception of initiation of therapy based solely on high B-type natriuretic peptide levels. That is, it is likely that an echocardiogram would always be performed as a confirmatory test, prior to treatment, even with BNP levels >500 pg/mL. It should also be noted that the BASEL study used a rapid point-of-care assay, whereas in the Australian setting it is proposed that only certified laboratory assays will be used for the measurement of B-type natriuretic peptides.

Mueller et al (2004b) reported that adding BNP testing to a standard clinical diagnostic workup resulted in a median cost saving of US$1,854 per patient, or a 26 per cent reduction in total cost compared to clinical workup alone. Although it was not stated, this cost saving presumably applied to the primary admission and did not include the costs of readmissions over the 30-day follow-up. The reported costs are equivalent to A$7,195 and A$9,661 for the intervention and control, respectively, when using purchasing power parities (PPP) for 2001 (Organisation for Economic Co-operation and Development 2006).

Although Mueller et al (2004b) reported hospital charges for both arms of their randomised controlled trial, different healthcare and insurance structures may prevent Swiss charges being directly transferable to the Australian setting. Hence, we have used the best available cost estimates for HF and alternative diagnoses, as reported in Australian Refined Diagnosis Related Group (AR-DRG) cost estimates. These cost estimates have been combined with the published and unpublished data (obtained via personal communication with Prof. Mueller) from the key trial (Table 18).

Table 18 Outcomes and process measures used in the economic evaluation from the key randomised controlled trial (Mueller et al 2004b)

Outcome/process measure	Clinical diagnosis[a] + BNP[b]	Clinical diagnosis[a]	Statistic
Primary			
Time to discharge (days)			
median (IQR) all patients	8.0 (1–16)	11.0 (5–18)	p=0.001
mean (95%CI) all patients	10.6[α] (8.9,12.2)	13.7[α] (12.0,15.4)	*NR*
mean (95%CI) admitted patients	13.9[α] (12.0,15.8)	16.0[α] (14.2,17.8)	*NR*
Secondary			
In-hospital mortality – per cent (n/N)	6% (13/225)	9% (21/227)	p=0.21
30-day mortality – per cent (n/N)			
admitted patients	*12% (20/169)*	*14% (27/193)*	*p=0.54*
non-admitted patients	4% (2/56)	3% (1/34)	p=1.0
all patients	10% (22/225)	12% (28/227)	p=0.45
95%CI	*9–10%[c]*	*12–13%[c]*	
difference in 30-day mortality (95%CI)	*2.6% (-3.2,8.3)*		
Other			
Time to treatment (mins) – median (IQR)	63 (16–153)	90 (20–205)	p=0.03
Hospital admission – per cent (n/N)	75% (169/225)	85% (193/227)	p=0.008
95%CI	*70–81%*	*80–90%*	
ICU admission – per cent (n/N)	15% (33/225)	24% (54/227)	p=0.01
30-day readmission rate – per cent (n/N)			
admitted patients	14% (23/169)	10% (20/193)	*p=0.34*
non-admitted patients	5% (3/56)	9%(3/34)	p=0.67
all patients	12% (26/225)	10% (23/227)	p=0.63
Proportions of **alternative** diagnoses for all patients			
COPD	*29.9% (76/254[d])*		*n/a*
pneumonia	*24.4% (62/254)*		*n/a*
pulmonary embolism	*8.3% (21/254)*		*n/a*
pleural effusion or interstitial lung disease or anaemia or sepsis	*23.2% (59/254)*		*n/a*
anxiety disorder	*6.3% (16/254)*		*n/a*
unknown	*7.9% (20/254)*		*n/a*
Personal communication with lead author on key trial			
Heart failure diagnosis – per cent (n/N)			
admitted patient	49.7% (84/169)	55.4% (107/193)	*p=0.28*
non-admitted patients	30.4% (17/56)	26.5% (9/34)	*p=0.69*
Other diagnosis – per cent (n/N)			
admitted patient	50.3% (85/169)	44.6% (86/193)	*p=0.28*
non-admitted patients	69.6% (39/56)	73.5% (25/34)	*p=0.69*

Values in *italics* were calculated during the economic evaluation; ICU = intensive care unit; IQR = inter-quartile range; n/a = not applicable; COPD = chronic obstructive pulmonary disease; [a] Clinical diagnosis, in general, included clinical history, physical exam, ECG, pulse oximetry, blood tests and CXR (personal communication with Prof. Mueller; lead author on the Basel study); [b] BNP measured using a Triage rapid fluorescence assay. The result was considered in context with the other clinical information and the physicians' clinical impressions; [c] Asymmetric 95% confidence intervals are due to rounding; [d] Denominator is 254 because a patient could have more than one diagnosis.

The primary differences between the BNP and control groups were in the admission rates, with 75 and 85 per cent of patients admitted to hospital, respectively; and with 15 and 24 per cent of patients admitted to the intensive care unit, respectively. The statistically lower admission rate ($p=0.008$) was the driving force behind a shortened hospital stay ($p=0.001$) and lower hospital charges ($p=0.006$) for intention-to-treat analyses (ie admitted plus non-admitted patients). Fourteen versus 10 per cent of admitted patients in the BNP and control group, respectively, were readmitted to hospital within 30 days ($p=0.34$). The reduction in primary admission rate attributed to the introduction of BNP testing may also translate to cost savings in the Australian ED setting.

The only patient-relevant health outcome reported in Mueller et al (2004b) was the 30-day mortality rate. No health-related quality of life data were reported.

Estimating the potential costs or cost savings of introducing B-type natriuretic peptide pathology testing

The unit cost of a B-type natriuretic peptide test was obtained through laboratory benchmarking data. Resources taken into account included staff labour costs, reagents, quality control and calibration, overheads and infrastructure, and reasonable profit. The unit cost of performing and reporting the test was $50.59 per test for a batch run of 10 (see Appendix G). The AR-DRG cost estimates from the Round 7 Cost Weight Report (Department of Health and Ageing 2006) were used as a proxy for the real resource use. These AR-DRG cost estimates were applied to episodes of care for HF and alternative diagnoses reported in the key trial (Table 18). The total cost for an admitted episode was weighted in accordance with the relative prevalence of complications or comorbidities associated with each disease category (AR-DRG code suffixes A, B and C). These estimates for HF and each of the alternative diagnoses were again weighted according to their respective numbers of private and public sector episodes. The standard errors for each AR-DRG code were calculated for the weighted estimates to maintain an estimate of error around the mean (see 'Estimating the 95% confidence interval for cost' in Appendix H). An example of this process is provided for HF (Table 19). The costs associated with alternative diagnoses were calculated in a similar way, and their final weighted estimates are reported in Table 21.

Table 19 Patient costs of an admission classified as heart failure

AR-DRG	No. separations (2002–03)	Average length of stay (days)	Cost per episode	Lower limit (95% CI)	Upper limit (95% CI)
Public sector - *estimated*					
F62A	6,710	11.23	$7,629	$7,513	$7,744
F62B	21,934	5.36	$3,483	$3,425	$3,541
Weighted[a]			**$4,454**	*$4,379*	*$4,530*
Private sector - *estimated*					
F62A plus physician costs[a]	1,498	14.96	$8,353	$8,071	$8.635
F62B plus physician costs[a]	6,852	7.95	$4,140	$3,999	$4,280
Weighted[b]			**$4,895**	*$4,721*	*$5,070*
Combined[c]					
			$4,554	*$4,447*	*$4,660*

CI = confidence interval; F62 = heart failure and shock with (A) or without (B) catastrophic complications and comorbidities; [a] Private sector estimates based on the relevant AR-DRG plus physician costs at $128.05 for first day and $64.10 for each subsequent day of average length of stay; [b] Based on the proportion of patients who had catastrophic complications and comorbidities; [c] Using weighted public and private sector estimates and again weighting the average to a 22 per cent private and 78 per cent public sector split; 95%CI in *italics* were calculated during the economic evaluation.

The AR-DRG categories were mapped to the final discharge diagnoses reported by Mueller et al (2004b) as represented in Table 20. Interstitial lung disease, pleural effusion, sepsis and anaemia were grouped together in Mueller's study; therefore, in order to estimate an AR-DRG cost for this disease category, a weighted mean of the four AR-DRG costs was calculated. Similarly, an 'unknown' disease category was reported by Mueller and colleagues. This category was assigned an arbitrary cost estimate of $2,500 plus private physician costs (MBS items 110 and 116). All monetary values reported in Table 21 were indexed to calendar 2005 costs by the Consumer Price Index (Australian Bureau of Statistics 2006).

Table 20 AR-DRG codes used to estimate costs for heart failure and alternative diagnoses reported in Mueller et al (2004b)

Code	Description	Disease category in Mueller et al (2004b)
F62A	Heart Failure & Shock + CCC	Heart failure
F62B	Heart Failure & Shock – CC	
E65A	COPD + CSCC	COPD
E65B	COPD – CSCC	
E61A	Pulmonary Embolism + CSCC	Pulmonary embolism
E61B	Pulmonary Embolism – CSCC	
E62A	Respiratory Infections/Inflammation + CSCC	Pneumonia
E62B	Respiratory Infections/Inflammation + SMCC	
E62C	Respiratory Infections/Inflammation – CC	
E74A	Interstitial Lung Disease + CCC	Interstitial lung disease or pleural effusion or sepsis or anaemia
E74B	Interstitial Lung Disease + SCC	
E74C	Interstitial Lung Disease – CC	
E73A	Pleural Effusion + CCC	
E73B	Pleural Effusion + SCC	
E73C	Pleural Effusion – CC	
T60A	Septicaemia + CCC	
T60B	Septicaemia – CC	
Q61C	Red Blood Cell Disorders – CC	
U65Z	Anxiety Disorders	Anxiety disorder
NA	Arbitrary cost estimate	Unknown

CCC = catastrophic complications and comorbidities; CSCC = catastrophic or severe complications and comorbidities; SMCC = severe or moderate complications and comorbidities; CC = complications and comorbidities; SCC = severe complications and comorbidities; COPD = chronic obstructive pulmonary disease.

The HF AR-DRG cost estimate includes the cost of inpatient echocardiography associated with the diagnosis of HF in admitted patients, but does not account for outpatient echocardiography or lung function tests in non-admitted patients. Using the data of Mueller et al (2004b), 30.4 per cent and 26.5 per cent of non-admitted patients received a diagnosis of HF in the BNP and control arms, respectively. The conservative (high) assumption that 50 per cent of these non-admitted patients with HF have an outpatient echocardiogram (MBS item no. 55113) is integrated into the costing in Table 21. There is also an assumption that 10 per cent of non-admitted patients with 'other' diagnoses undergo a respiratory function test (MBS item no. 11509). A much lower proportion of non-admitted 'other' diagnosis patients are assumed to undergo further testing because lung function tests are not performed for some alternative diagnoses, and are therefore less likely to be used than echocardiography.

An accurate estimate for the rate of echocardiography referral in Australian hospitals is not available. Australian and New Zealand guidelines (NHF & CSANZ 2002) recommend that all admitted HF patients undergo echocardiography, but in practice this does not occur. Nevertheless, the HF AR-DRG cost estimate incorporates the real rate of inpatient referrals to echocardiography in a large sample of Australian hospitals and, as such, is the most appropriate cost estimate for the current clinical setting. The imaging cost bucket of the AR-DRG code may more closely represent the real costs attributed to echocardiography referral, but this subcategory is still not exclusive to a single imaging modality.

The introduction of B-type natriuretic peptide testing is unlikely to influence the rate of inpatient echocardiography referral because it has been shown to decrease the rate of admission to hospital, not the rate of echocardiography use once a patient has been admitted. The costs of HF and 'alternative diagnosis' admissions in Table 21 are therefore the same for the clinical diagnosis and the clinical diagnosis plus BNP columns. Once a patient's symptoms are severe enough to warrant hospital admission, it is improbable that there will be different levels of inpatient echocardiography referral based on triage B-type natriuretic peptide concentrations. In other words, the introduction of B-type natriuretic peptide testing would only serve to filter more suspected HF patients into a discharge status or admission for an 'other' disease scenario. This argument is strongly supported by the data of Mueller et al (2004b), for which the introduction of BNP testing resulted in a decreased rate of hospital admission and, of those admitted patients, a slightly lesser percentage admitted for HF (Table 18).

Table 21 Potential cost savings for 100 patients arriving at an emergency department with acute dyspnoea symptoms suggestive of heart failure, based on service use reported by Mueller et al (2004b) combined with episode cost from AR-DRG estimates.

	Unit cost	Clinical diagnosis + BNP Number of patients	Clinical diagnosis + BNP Total cost	Clinical diagnosis Number of patients	Clinical diagnosis Total cost
N of patients presenting to ED with symptoms suggestive of HF		100		100	
N of primary admissions		75.1		85.0	
N of readmissions		11.6		10.1	
Total N of hospitalisation episodes within 30 days		86.7		95.2	
Admitted patient diagnosis[d]					
HF	$4,842[b]	43.1[a]	$208,581	52.8[a]	$255,438
COPD	$4,503[b]	13.0[a]	$58,730	12.7[a]	$57,128
Respiratory infection/inflammation	$3,147[b]	10.6[a]	$33,486	10.3[a]	$32,572
Pulmonary embolism	$5,089[b]	3.6[a]	$18,339	3.5[a]	$17,838
Pleural effusion Interstitial lung disease Anaemia	$2,439[b,e]	10.1[a]		9.8[a]	
Sepsis			$24,699		$24,025
Anxiety disorder	$3,003[b]	2.7[a]	$8,245	2.7[a]	$8,020
Unknown	$2,658[b]	3.4[a]	$9,124	3.3[a]	$8,875
Cost of non-admitted triage (Category 3)	$338[f]	24.9[a]	$8,412	15.0[a]	$5,063
Cost of outpatient echocardiograms for 100 patients	$231[c]	3.8[g]	$873	2.0	$458
Cost of outpatient lung function tests for 100 patients	$30.85[c]	1.7[g]	$53	1.1	$34
Total B-type natriuretic peptide pathology test cost for all patients	$50.59[h]	100	$5,059	0	$0
Total cost for all patients			$375,601		$409,449
Cost *savings* for 100 patients (95% CI[i])			**$33,849 ($304, $67,393)**		

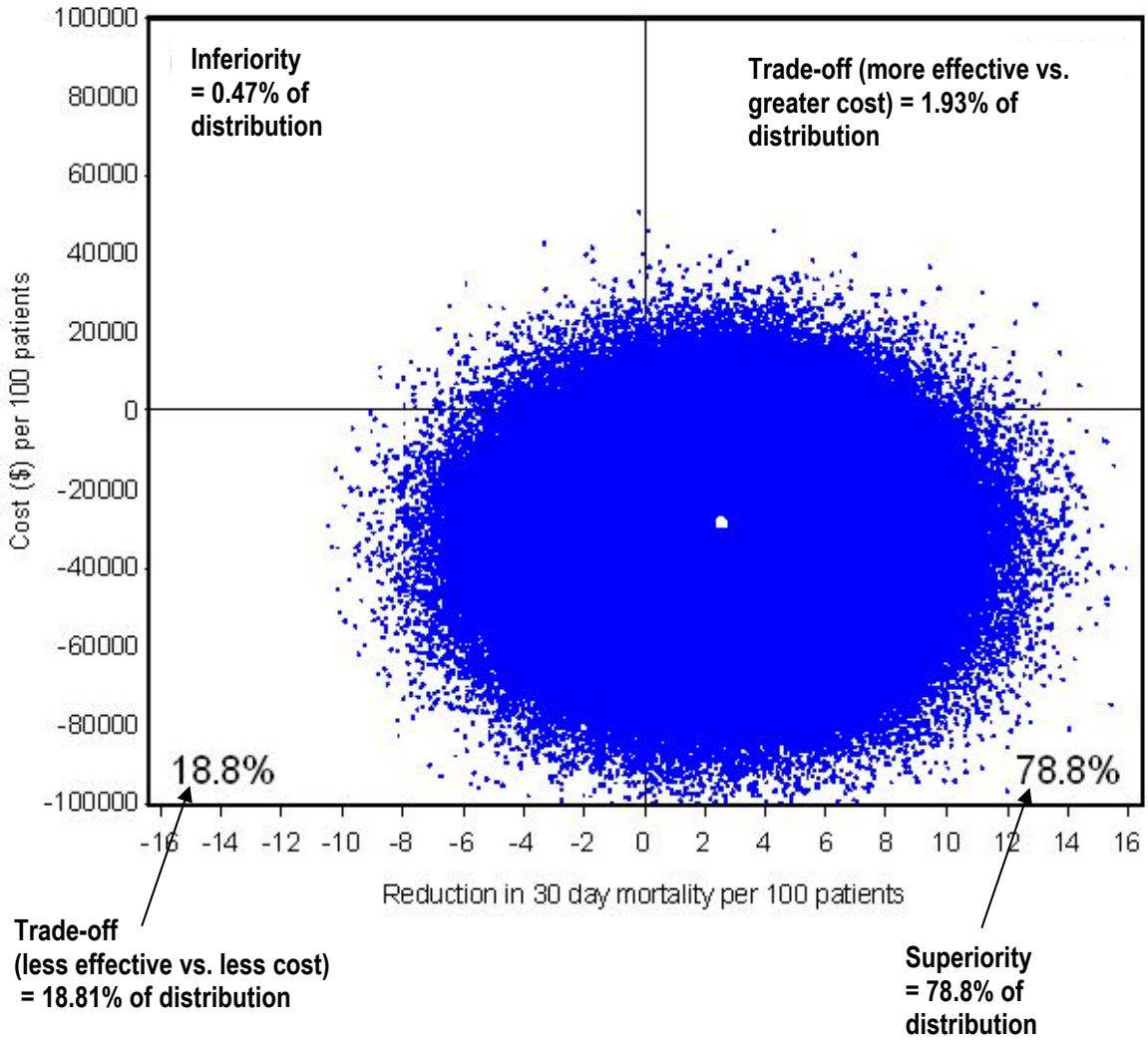
HF = heart failure; [a] Data derived from BASEL study using proportions of admitted patients with HF or other diagnoses in the BNP and control arms (see Table 18), multiplied by the number of admitted patients in each arm – eg Clinical diagnosis + BNP x HF = 0.497 x 86.7 = 43.1; COPD = 0.503 x 0.299 x 86.7 = 13.0; Clinical diagnosis x HF = 0.554 x 95.2 = 52.8; COPD = 0.446 x 0.299 x 95.2 = 12.7; [b] Cost estimated from weighted (public versus private and complications versus no complications) Round 7 AR-DRG (2002–03) costs and adjusted to 2005 costs using CPI indexation (x1.063); [c] Cost estimated from MBS item nos. 55113 and 11509 for echocardiography and lung function tests, respectively; [d] Final discharge diagnoses were only available for the primary admissions. The same proportions were applied to readmissions. Also, the disease proportions for the admitted patient alternative diagnoses are the same as those for all patients (admitted + non-admitted) as these were the only data available; [e] Unknown proportions so therefore an arithmetic mean was calculated, which assumes that the prevalence of pleural effusion, interstitial lung disease, anaemia and sepsis are all equal within this combined disease category; [f] Cost estimated from Round 8 AR-DRG (2003–04) for category 3 (seen within 30 minutes) non-admitted triage cost for a public hospital and adjusted to 2005 costs using CPI indexation (x1.039); [g] Data derived from BASEL study using proportions of non-admitted patients with HF or 'other' diagnoses in the BNP and control arms (see Table 18), multiplied by the number of non-admitted patients in each arm, eg Clinical diagnosis + BNP x Echocardiogram use = (0.304 x 25) x 0.5 (assuming half of non-admitted patients undergo outpatient echocardiography testing) = 3.8; Lung function test = 0.696 x 25 x 0.1 (assuming 10% of non-admitted patients undergo outpatient lung function testing) = 1.7; [h] Assuming batch test; [i] See Appendix H for details on how the 95%CI around the point cost estimate was calculated.

On the basis of resource use data from Mueller et al (2004b) and AR-DRG cost estimates, it appears that the introduction of B-type natriuretic peptide testing in the EDs of Australian hospitals could lead to cost savings of $338 per patient presenting to an ED with acute dyspnoea symptoms (point estimate). This saving represents an approximately 8 per cent reduction in costs currently outlaid for suspected HF patients in this setting. The cost savings are primarily mediated by a lower overall (30-day) admission rate in the clinical diagnosis plus BNP test group compared to the clinical diagnosis group alone.

These estimated savings, taken together with a slightly lower 30-day mortality rate in the BNP test group (Table 18), result in a point estimate for the BNP test plus clinical workup diagnostic strategy that lies within the south-east quadrant of the incremental cost-effectiveness plane, indicating *dominance* (Figure 8). **From the point estimate, the integration of B-type natriuretic peptide testing into current diagnostic strategies is expected to represent both a cost saving and provision of superior patient outcomes when compared to conventional diagnostic workup of patients presenting to an ED with dyspnoea symptoms suggestive of HF.** The 95% confidence area of the joint probability distribution around the point cost-effectiveness estimate lies mainly in the south-west and south-east quadrants, but does cross all quadrants. It is apparent from Figure 8 that it is not possible on the data available to state categorically that the addition of B-type natriuretic peptide testing to conventional diagnostic strategies in an ED setting is a better use of resources. However, the larger part of the joint probability distribution (78.80%) indicates dominance over conventional diagnostic strategies in this setting, while most of the rest (18.81%) indicates uncertainty over whether the test is actually effective in terms of a reduction in the 30-day mortality rate.

Note that effectiveness in the incremental cost-effectiveness plane is usually measured in life-years gained or QALYs gained. Although there is no information available to estimate how an increment in 30-day mortality might translate into an increment in life-years (or QALYs) gained, the variability is contained within the estimate of 30-day mortality and it is likely that calculating the increment in life-years gained would involve multiplying such an increment by a constant amount.

Figure 8 The joint probability distribution of the incremental cost-effectiveness ratio plotted on the incremental cost-effectiveness plane.



Note: Effectiveness and costs are expressed for 100 patients with acute dyspnoea (suggestive of heart failure)

Threshold analysis of the unit cost of a B-type natriuretic peptide pathology test

A threshold analysis suggests that a B-type natriuretic peptide pathology test would have to be charged at $389 for there to be no incremental cost savings. This test unit cost would be the break-even point.

Sensitivity analysis

- The primary factor associated with cost savings attributable to the introduction of B-type natriuretic peptide testing is a reduction in the point estimate of the patient admission rate. Assuming that there is no real difference in 30-day admission rates (primary admissions + readmissions) between the clinical diagnosis and clinical diagnosis plus BNP arms, and that there are identical proportions of HF and other diagnoses in admitted and non-admitted patients, the additional cost per suspected HF episode would be the cost of the pathology test. Under these assumptions of no

effectiveness, the introduction of B-type natriuretic peptide testing would therefore result in a 5.6 per cent increased cost for ruling out patients suspected of HF in an ED setting.

- The rate of echocardiography referral is unlikely to alter in an ED setting due to the introduction of B-type natriuretic peptide testing. However, to test the robustness of this assumption, the episode cost admission in the clinical diagnosis plus BNP arm was increased by $231 (the cost of an echocardiogram). This extreme scenario, which assumes that all admitted patients receive an inpatient echocardiogram in the clinical diagnosis plus BNP arm and that no patients receive an echocardiogram in the clinical diagnosis alone arm, would still result in a point estimate of incremental cost savings of $139 per patient due to the introduction of B-type natriuretic peptide testing.

It should be noted that the key trial used a point-of-care BNP assay, whereas in the Australian setting it would be a certified laboratory test. The time frame in which the B-type natriuretic peptide assay result can be measured may have an important, but currently unknown, effect on patient outcomes.

In addition to their main manuscript, Mueller and colleagues published pre-specified subgroup analyses on women (Mueller et al 2004a), patients with kidney disease (Mueller et al 2005b) and the elderly (Mueller et al 2005a) for the BASEL trial. The latter subgroup analysis in patients aged ≥70 years revealed a statistically significant reduction in 30-day mortality for the diagnostic strategy which used BNP compared to the control. Without additional information from the author, we are unable to calculate the incremental cost-effectiveness ratio point estimate and distribution for this subgroup. Nevertheless, it is clear that the increase in incremental effectiveness in this subgroup would be offset by an increase in costs in the clinical diagnosis plus BNP test arm due to a slight rise in the total number of hospitalisation episodes (BNP: 99 events; control: 97 events), as well as the cost of the pathology test in the clinical diagnosis plus BNP test group. The point cost-effectiveness estimate would therefore lie in the north-east quadrant (more effective but greater cost), representing a trade-off situation.

Strengths and limitations of a trial-based economic analysis

A trial-based economic analysis has the advantage of using information from the study or studies with the highest level of evidence available.

Reasons for proceeding to a health economic model to better understand a proposed healthcare intervention would include: (1) to extrapolate health outcomes beyond the end of the trial in terms of life-years gained and/or quality-adjusted life-years (QALY); (2) to assess the joint probability distribution of costs and effects over the remaining lifetime, and to calculate a cost-effectiveness acceptability curve; and (3) to assess a different management algorithm from that in the key trial(s).

Due to the fact that: (1) the variability in outcome is contained within the estimate of the increment in 30-day mortality and it is likely that calculating the increment in life-years gained would involve multiplying such an increment by a constant amount; (2) the joint probability distribution has been estimated from the 95% confidence intervals of costs and effectiveness; and (3) the trial's management algorithm is a close approximation to that likely to prevail under Australian conditions, it is doubtful whether an economic model would add sufficient further information to influence the decision.

Construction of a full economic model would require an estimate of the incremental life-years saved and their distribution. For rational decision-making it is desirable that different economic analyses be comparable with each other in terms of the outcome measure (ie either in terms of life-years gained or QALYs). To calculate the incremental cost per additional life saved at 30 days would be to generate a non-conventional cost-effectiveness ratio that would not be readily comparable with other studies. An exploration was therefore initiated into whether it was possible to estimate the additional life-years gained from the use of BNP for diagnosis in this population group. The limitation of this approach is that the value of assessing the health outcome over a longer time horizon may be compromised by the nature of the HF disease process, which is characterised by intermittent acute decompensated episodes interspersed with periods of clinical stability. Further, given the variety of alternative diagnoses that are possible for a patient with symptoms of acute dyspnoea who is 'ruled out' from HF in the ED setting, and the lack of available data on survival of these patients beyond the 30-day period, this analysis has proceeded under the conservative assumption that there is no further mortality gain beyond 30 days (ie, survival beyond 30 days is the same for both groups).

No clinical diagnosis plus BNP test survival curve specifically for patients presenting to the ED with acute dyspnoea has been located for Australia or elsewhere. Possibly, a surrogate might be obtained from the survival of patients presenting with their first diagnosed episode of HF coupled with an estimate of the hazard ratio associated with BNP test-informed clinical diagnosis compared to clinical diagnosis alone. Unfortunately, it is doubtful that it would be justifiable to impute a particular hazard ratio from a non-significant difference in 30-day mortality reported from the key trial.

In the absence of strong support for a survival benefit, the conservative approach is to assume that none has been demonstrated. Because 30-day mortality was a secondary outcome in the trial, it was not powered to find a survival benefit should one exist. The 95% confidence interval of the difference in 30-day mortality crossed zero (Table 18). The key trial was also not designed to test equivalence, and if a cost-minimisation approach were to be taken it should be borne in mind that inferiority has not been excluded (Briggs & O'Brien 2001). Furthermore, without a credible estimate of the incremental life-years gained it is not feasible to construct a cost–utility analysis based on QALYs.

Postscript to the economic analysis in the emergency department setting

After the above economic analysis in the Australian setting had been completed, a trial-based economic analysis of the BASEL trial appeared in the literature. Although this was beyond the August 2005 cut-off for the systematic literature review (evidence-based) section of this report, it is included here because it describes the follow-up to the main paper relied upon for the economic analysis in the ED setting. In this more recent paper, Mueller et al (2006) used non-parametric bootstrapping to estimate the distribution of incremental costs and effects on the cost-effectiveness plane during 180 days of follow-up. Apart from the greater length of follow-up, their method of economic analysis appears to be essentially similar to that presented above.

Over the 180 days of follow-up, the difference between the two arms of the trial in all-cause mortality was not statistically significant (see Table 22). Over the same period, both the total days in hospital and the total treatment costs were significantly less in the BNP group (Mueller et al 2006).

Table 22 Outcomes and costs at 180 days from the key randomised controlled trial (Mueller et al 2004b) used in the trial-based economic evaluation

Outcome/process measure	Clinical diagnosis + BNP	Clinical diagnosis	Statistic
Secondary outcome			
180-day all-cause mortality – per cent (n/N)	20 (44/225)	23 (52/227)	0.42
Costs			
Total days in hospital – median (IQR)	10 (2–24)	14 (6–27)	0.005
Total treatment cost, US$ – mean (SD)	$7,930 ($8,805)	$10,503 ($10,176)	0.004

IQR = interquartile range; SD = standard deviation

Source: Mueller et al (2006)

The point estimate of the incremental cost-effectiveness ratio at 180 days was in the south-east quadrant of the incremental cost-effectiveness plane, indicating dominance of the diagnostic algorithm that incorporated BNP testing under Swiss conditions (ie cost saving and superior patient outcomes). The 95% confidence region crossed the boundary into the south-west quadrant but did not cross into either the north-east or north-west quadrants of the incremental cost-effectiveness plane. This suggests that the BNP testing algorithm is less costly but not significantly more effective than its comparator at 180 days follow-up. No health-related quality of life information was reported.

Financial incidence analysis

The Australian National Hospital Morbidity database registered 41,052 separations for HF in the financial year 2002–03, under the International Classification of Disease (ICD) code I50. The estimate is based on almost all Australian public hospitals and a majority of private hospitals. Seventy-eight per cent (n=31,879) of these HF separations were reported in public hospitals, with the remainder (22%; n=9,173) in private hospitals.

In their first paper Mueller et al (2004b) reported that, of the 80 per cent of admitted patients in their sample with acute dyspnoea, 52.6 per cent were diagnosed with HF. This means that 42 per cent (0.8*0.525 = 0.42) of patients with acute dyspnoea arriving at an ED were admitted to hospital with a primary diagnosis of HF. Applying this percentage to the ICD I50 (heart failure) separation data for 2002–03 (n=41,025 separations) above suggests an estimated rate of 97,742 (41,025/0.42) ED presentations due to acute dyspnoea (suggestive of HF) per annum in Australia[12].

The public to private *hospital* proportions for HF admission are 78 and 22 per cent, respectively. The corresponding public to private *patient* split is 64 and 36 per cent, respectively. Assuming that presentation to a public or private ED follows similar

[12] It is worthwhile noting that the Commonwealth Department of Health and Ageing, as part of the casemix data collection, is planning to collate the reasons for ED presentation in a sample of Australian hospitals. These data should more closely reflect the true burden on Australian EDs due to patients presenting with symptoms suggestive of HF.

respective proportions as that of admissions, it is clear that the minority (36 per cent) of events are directly related to MBS reimbursement, whereas the majority (64 per cent) are related to Australian Health Care Agreements between the states/territories and the Commonwealth.

Expenditure by the Australian Government on B-type natriuretic peptide testing for private patients would be $50.59 (assuming that no copayment is involved) for 36 per cent of 97,742 claims per annum. The overall annual additional Australian Government expenditure on B-type natriuretic peptide testing of private patients in Australian EDs would therefore be $1.78 million.

Table 23 describes all MBS item numbers that could be claimed during an ED episode and hospital stay for a private HF patient. Table 24 presents the potential expenditure by the Australian Government for private HF patients.

Table 23 MBS items associated with the diagnosis of heart failure in a private patient

MBS item no.	Schedule fee	Short definition MBS item no. description
110	$128.05	**Professional attendance** Specialist, referred consultation - surgery or hospital (Professional attendance at consulting rooms or hospital by a specialist in the practice of his or her specialty where the patient is referred to him or her) - INITIAL attendance in a single course of treatment, not being a service to which item 106 applies
116	$64.10	**Subsequent professional attendances** Each attendance subsequent to the first in a single course of treatment
507	$58.55	**Level 3 patient initiation fee** Medical practitioner (emergency physician) attendances - emergency department LEVEL 3; Professional attendance on a patient at a recognised emergency department of a private hospital by a medical practitioner who is an emergency physician in the practice of emergency medicine
65070	$17.20	**Full blood count** Erythrocyte count, haematocrit, haemoglobin, calculation or measurement of red cell index or indices, platelet count, leucocyte count and manual or instrument generated differential count - not being a service where haemoglobin only is requested - one or more instrument generated set of results from a single sample; and (if performed) (a) a morphological assessment of a blood film; (b) any service in item 65060 or 65072
66509	$15.75	**Electrolytes** Four tests of any of the following: Quantitation in serum, plasma, urine or other body fluid (except amniotic fluid), by any method except reagent tablet or reagent strip (with or without reflectance meter) of: acetoacetate, acid phosphatase, alanine aminotransferase, albumin, alkaline phosphatase, ammonia, amylase, aspartate aminotransferase, beta-hydroxybutyrate, bicarbonate, bilirubin (total), bilirubin (any fractions), C-reactive protein, calcium (total or corrected for albumin), chloride, creatine kinase, creatinine, gamma glutamyl transferase, globulin, glucose, lactate, lactate dehydrogenase, lipase, magnesium, phosphate, potassium, pyruvate, sodium, total protein, total cholesterol, triglycerides, urea
66515	$19.80	**Renal function** Six tests of any of the above mentioned (66509)
66719	$35.45	**Thyroid function** Thyroid function tests (comprising the service described in item 66716 and 1 or more of the following tests - estimation of free thyroxine index, free thyroxine, free T3, total T3, thyroxine binding globulin) for a patient, if at least 1 of the following conditions is satisfied: (a) the patient has an abnormal level of TSH; (b) the tests are performed: (i) for the purpose of monitoring thyroid disease in the patient; or (ii) to investigate the sick euthyroid syndrome if the patient is an admitted patient; or (iii) to investigate dementia or psychiatric illness of the patient; or (iv) to investigate amenorrhoea or infertility of the patient; (c) the medical practitioner who requested the tests suspects the patient has a pituitary dysfunction; (d) the patient is on drugs that interfere with thyroid hormone metabolism or function
73910	$10.30	**Patient episode initiation** Initiation of a patient episode by collection of a specimen for a service (other than a service described in item 73901, 73903 or 73905 or in Group P9) if the specimen is collected by an approved pathology practitioner or an employee of an approved pathology authority from a person in a residential aged care home or institution
11700	$26.50	**ECG** TWELVE-LEAD ELECTROCARDIOGRAPHY, tracing and report
58500	$35.35	**Chest X-ray** Chest (lung fields) by direct radiography
55113	$230.65	**Echocardiography** M-MODE and 2 DIMENSIONAL REAL TIME ECHOCARDIOGRAPHIC EXAMINATION of the heart from at least 2 acoustic windows, with measurement of blood flow velocities across the cardiac valves using pulsed wave and continuous wave Doppler techniques, and real time colour flow mapping from at least 2 acoustic windows, with recordings on video tape or digital medium, not being a service associated with a service to which an item in Subgroups 1 (with the exception of item 55054) or 3, or another item in this Subgroup (with the exception of items 55118 and 55130), applies, for the investigation of symptoms or signs of cardiac failure, or suspected or known ventricular hypertrophy or dysfunction, or chest pain (R)

Table 24 Potential Australian Government expenditure on 100 private patients arriving at an emergency department with acute dyspnoea (suggestive of heart failure), using service use reported by Mueller et al (2004b) combined with MBS costs

	Schedule fee	Proportion reimbursed	Clinical diagnosis + BNP		Clinical diagnosis	
			N	Expenditure	N	Expenditure
N of private patients presenting to an ED with acute dyspnoea suggestive of HF			100		100	
Level 3 patient initiation fee; clinical history and physical exam	$58.55	0.85	100[a]	$4,977	100[a]	$4,977
Full blood count	$17.20	0.85	100[a]	$1,462	100[a]	$1,462
Electrolytes	$15.75	0.85	100[a]	$1,339	100[a]	$1,339
Renal function	$19.80	0.85	100[a]	$1,683	100[a]	$1,683
Thyroid function	$35.45	0.85	100[a]	$3,013	100[a]	$3,013
ECG; cardiomegaly, electrolyte disturbances, bundle branch blocks, myocardial infarction	$26.50	0.85	100[a]	$2,253	100[a]	$2,253
CXR; cardiomegaly and fluid in lungs	$35.35	0.85	100[a]	$3,005	100[a]	$3,005
Patient episode initiation	$10.30	0.85	100[a]	$1,030	100[a]	$1,030
BNP indicative cost estimate	$50.59		100[a]	$5,059	100[a]	$0
Admitted patients			75[b]		85[b]	
HF diagnosis			37.28[c]		47.1[c]	
Echocardiogram[d]	$230.65	0.75	37.28	$6,449	47.1[c]	$8,148
Professional attendance	$128.05	0.75	37.28	$3,580	47.1[c]	$4,523
Subsequent professional attendance (x3)	$192.30	0.75	37.28	$5,377	47.1[c]	$6,793
Other diagnosis			37.72	n/a[e]	37.9[e]	n/a[e]
Total Aust Govt expenditure for all patients[f]				$39,071		$38,071
Additional expenditure for 100 patients				+$1,000		

HF = heart failure; n/a = not available; ED = emergency department; ECG = electrocardiogram; CXR = chest X-ray; [a] Assuming all patients presenting to the private ED undergo all tests; [b] Difference in admission rate to hospital reported by Mueller et al (2004b); [c] Calculated by using the proportion of admitted patients with HF diagnosis (personal communication with lead author on key trial); [d] Assumes that all echocardiograms are performed after admission to hospital; [e] MBS claimed costs are not presented for 'other' because the reclaim for a large range of alternative diagnoses is difficult to estimate, and the numbers of patients in each group are similar so they are not likely to affect the relative cost savings/expenditure from the introduction of BNP; [f] Savings related to any difference in pharmaceutical use have been ignored. Also, the cost breakdown does not take into account safety net issues (although changes in safety net would only change absolute cost values in each arm, not the incremental cost difference between them).

For private patients in private hospitals, the Australian Government's $1.78 million outlay on B-type natriuretic peptide testing would be *partially* offset by a reduced number of hospital admissions (Table 24; from Mueller et al 2004b) and thus reduced MBS item claims for private inpatient services. The financial analysis for private HF patients arriving at an ED (Table 24) shows that the introduction of B-type natriuretic peptide testing would be associated with an additional net cost to the Commonwealth of $10 per patient. This additional cost would only be incurred in 36 per cent of the estimated number of relevant ED arrivals per annum (97,742). The total annual incremental expenditure by the Australian Government due to the introduction of B-type natriuretic peptide testing for private patients in private hospital EDs would therefore be $352,000 thousand, which incorporates the additional outlay of $1.78 million required for B-type natriuretic peptide testing and cost savings due to fewer echocardiograms and a lower private inpatient MBS reclaim.

In the public sector the introduction of B-type natriuretic peptide testing in EDs is likely to reduce the number of related hospital admissions. This may lead to resource savings per patient with acute dyspnoea suspected of HF, but it is unlikely that these savings will be seen by the Australian Government if public hospitals are already operating at close to capacity. A reduction in admissions due to B-type natriuretic peptide testing would, however, make public hospital beds available for other patients in need.

Conclusions

Safety

The likelihood of adverse events as a consequence of the B-type natriuretic peptide testing procedure is low and similar to that of any blood test. Psychological or physical harms are a possibility due to the inevitability of false positive or false negative results being associated with a test, and sequelae such as inappropriate or delayed treatment. None of the included studies in this report, however, reported on patient physical or psychological harms as a consequence of B-type natriuretic peptide testing.

Diagnostic effectiveness

BNP testing

The effectiveness of supplementing conventional diagnostic assessment with BNP testing was evaluated by a large volume of evidence, with the highest quality evidence being one good quality level II direct intervention study as well as two good quality level II diagnostic accuracy studies. Overall, the body of evidence was relatively consistent in its findings that BNP tests are sensitive with a high negative predictive value which effectively 'rules out' heart failure (HF) in a patient with a negative test. A strong pooled diagnostic odds ratio indicates that BNP tests effectively discriminate between the presence or absence of HF in patients. Variation in diagnostic accuracy between studies was possibly due to the different test thresholds for ruling out HF employed in the studies. The clinical impact of the test was most noticeable in terms of patient management by the clinician. The test appeared to have its main impact in situations where the clinical diagnosis was initially uncertain. Time to discharge, time to treatment and hospital and intensive care unit admissions were reduced in patients receiving BNP-assisted diagnostic assessment compared to conventional assessment. The impact on patient health outcomes was in the right direction (a reduction in mortality rate). However, the main trial was limited by a lack of statistical power and so the result was only statistically significant in the pre-specified subgroup of elderly patients. The populations studied in the evidence base are applicable to the target population in Australia, that is patients presenting to an ED with acute dyspnoea (suggestive of HF). The results of the studies are largely generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in triaging and managing suspected HF patients.

In conclusion, on the basis of the evidence presented, BNP testing appears to be a valuable 'first line' diagnostic test that, when added to conventional diagnostic assessment, assists the acute care physician to correctly 'rule out' HF in patients presenting with acute dyspnoea. It also appears to benefit the patient by reducing or preventing hospital stay, and decreasing the time to treatment. It may also have the potential to reduce mortality rates within 30 days in some patients.

NT-proBNP testing

The effectiveness of NT-proBNP testing added to conventional diagnostic assessment was evaluated by a reasonable volume of evidence, with the highest quality evidence being three good quality level II and several average quality level III diagnostic accuracy studies. Overall, the body of evidence was relatively consistent in its findings that NT-proBNP tests are sensitive, although there was wide variation in rates that is probably a result of differences in test threshold (which could not be evaluated independently). In general, the assays had high negative predictive values which effectively 'rule out' HF in a patient with a negative test. The clinical impact of the test, in terms of patient management by the physician, is unknown in this setting. The impact of the NT-proBNP test on the health outcomes of patients with acute dyspnoea in this setting is also unknown. The populations studied in the evidence base are applicable to the target population in Australia, that is patients presenting to an ED with acute dyspnoea. The results of the studies are largely generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in diagnosing HF.

In conclusion, on the basis of the evidence presented, NT-proBNP testing appears to be a valuable diagnostic test that, when added to conventional diagnostic assessment, may assist the acute care physician to correctly 'rule out' HF in patients presenting with acute dyspnoea.

Effectiveness for monitoring

One good quality, but small, randomised controlled trial demonstrated that monitoring patients via NT-proBNP resulted in fewer cardiovascular deaths and total cardiovascular events than in patients monitored via clinical criteria. The beneficial effect of the hormone-guided monitoring was presumably mediated, through more predominant ACE-inhibitor and spironolactone use, to achieve the target NT-proBNP concentrations. An abstract of a randomised controlled trial reported similar results for BNP assay-guided monitoring, but more detail is necessary to determine whether this study could be considered supporting evidence. A larger randomised controlled trial currently being conducted in New Zealand by Richard Troughton and colleagues should shed further light on this poorly researched area of HF patient monitoring via B-type natriuretic peptides.

Economic considerations

Emergency department setting

Relying on the results of the key trial by Mueller et al (2004b), the incremental costs and outcomes of the management of suspected HF patients presenting to an ED were examined.

Given no statistically significant difference in survival at 30 days and 180 days, the point estimate of incremental costs and incremental effectiveness suggests that the addition of B-type natriuretic peptide testing dominates its comparator. Thus, the point estimate suggests that performing a B-type natriuretic peptide test in the ED setting leads to a superior health outcome at a lower cost. However, the 95% confidence interval of the

joint probability distribution of incremental costs and effectiveness crosses all four quadrants of the incremental cost-effectiveness plane, indicating that the point estimate is subject to some uncertainty but that 78.8 per cent of the joint probability distribution is in the dominant quadrant.

Regarding financial outlays, the Commonwealth Government will incur an additional expenditure of $352,000 under Medicare due to the introduction of B-type natriuretic peptide testing for private patients in private hospital EDs. Although the majority of B-type natriuretic peptide tests will be performed in public hospital EDs, this is unlikely to lead to Australian Government expenditure savings because of capacity constraints, but may make additional public resources available for other patients in need.

Recommendations

'MSAC has considered the safety, effectiveness and cost-effectiveness of the use of assays of brain natriuretic peptides (BNP) in the diagnosis of heart failure in patients presenting with dyspnoea in the hospital emergency setting and the use of the assays in monitoring the progress of patients with heart failure.

MSAC finds that there is sufficient evidence of the safety, effectiveness and cost-effectiveness of the use of these assays in the diagnosis of heart failure but insufficient evidence of effectiveness and cost-effectiveness for their use in monitoring the progress of patients with heart failure.

MSAC recommends that public funding be provided for the use of assays of BNP in the diagnosis of heart failure in the hospital emergency setting.'

The Minister for Health and Ageing accepted this recommendation on 5 February 2007.

Appendix A MSAC terms of reference and membership

MSAC's terms of reference are to:

- advise the Minister for Health and Ageing on the strength of evidence pertaining to new and emerging medical technologies and procedures in relation to their safety, effectiveness and cost-effectiveness and under what circumstances public funding should be supported;
- advise the Minister for Health and Ageing on which new medical technologies and procedures should be funded on an interim basis to allow data to be assembled to determine their safety, effectiveness and cost-effectiveness;
- advise the Minister for Health and Ageing on references related either to new and/or existing medical technologies and procedures; and
- undertake health technology assessment work referred by the Australian Health Ministers' Advisory Council (AHMAC) and report its findings to AHMAC.

The membership of MSAC comprises a mix of clinical expertise covering pathology, nuclear medicine, surgery, specialist medicine and general practice, plus clinical epidemiology and clinical trials, health economics, consumers, and health administration and planning:

Member	**Expertise or affiliation**
Dr Stephen Blamey (Chair)	general surgery
Associate Professor John Atherton	cardiology
Professor Syd Bell	pathology
Associate Professor Michael Cleary	emergency medicine
Dr Paul Craft	clinical epidemiology and oncology
Dr Jane Cook	Medical Officer, Department of Health and Ageing (November 2006)
Dr Kwun Fong	thoracic medicine
Dr David Gillespie	gastroenterology
Dr Debra Graves	medical administrator
Professor Jane Hall	health economics
Professor John Horvath	Chief Medical Officer, Department of Health and Ageing
Associate Professor Terri Jackson	health economics
Professor Brendon Kearney	health administration and planning
Dr Ray Kirk	health research
Associate Professor Frederick Khafagi	nuclear medicine
Associate Professor Donald Perry-Keene	endocrinology
Dr Ewa Piejko	general practice

Dr Brian Richards	Principle Medical Adviser Department of Health and Ageing (May 2007)
Ms Sheila Rimmer	consumer health issues
Ms Catherine Farrell	Department of Health and Ageing representative
Professor Ken Thomson	radiology
Dr Douglas Travis	urology
Dr Mary Turner	Australian Health Ministers' Advisory Council representative
Dr David Wood	orthopaedics

Appendix B Advisory panel and evaluators

Advisory Panel for MSAC Application 1087 - B-type natriuretic peptide assays

Prof Syd Bell Director, Pathology Department SEALS Prince of Wales Hospital, New South Wales	Chair (MSAC member)
Dr John Beilby Bioclinical Biochemistry, PathCentre, Western Australia	Nominee, Australasian Association of Clinical Biochemists
Ms Paula Calcino Consumer representative, Australian Capital Territory	Nominee, Consumers' Health Forum of Australia (HF)
Dr Jenny Doust PhD, FRACGP, BMBS, BEcons, BA, Grad Dip Clin Epi Senior Research Fellow in Clinical Epidemiology Discipline of General Practice/Division of Health Systems, Policy and Practice, University of Queensland	Nominee, Royal Australian College of General Practitioners
Dr Debra Graves MMB MHA FRACMA Chief Executive Officer, Royal College of Pathologists of Australasia, New South Wales	MSAC member
A/Prof Terri Jackson PhD Associate Professor, Australian Centre for Economics Research in Health, University of Queensland	MSAC member/Health economist
Prof Julia M Potter B Med Sc, MB BS, PhD, FRCPA Professor of Pathology, ANU Medical School Executive Director, ACT Pathology, Canberra Hospital, Australian Capital Territory	Nominee, Royal College of Pathologists of Australasia
A/Prof David Sullivan MBBS FRACP FRCPA Senior Staff Specialist, Department of Clinical Biochemistry, Royal Prince Alfred Hospital, New South Wales	Nominee, Cardiac Society of Australia and New Zealand

Advisory Panel for MSAC Application 1087 - B-type natriuretic peptide assays (cont)

Evaluators

Ms Tracy Merlin
BA(Hons), MPH
Lead researcher and Manager, AHTA

A/Prof John Moss
M Soc Sci, B Ec, MB BS, FCHSE
Head and Senior Lecturer,
Discipline of Public Health
The University of Adelaide

Dr Anthony Brooks
BSc(Hons), PhD
Research Officer

Ms Skye Newton
B Psych(Hons)
Research Officer

Ms Hedyeh Hedayati
B Biotech(Hons), GDPH
Research Officer

Mr Tom Sullivan
BSc(Hons), B Soc Sci
Research Officer

Professor Janet Hiller
MPH, PhD
Director, AHTA

Adelaide Health Technology Assessment (AHTA),
Discipline of Public Health,
School of Population Health and Clinical Practice,
University of Adelaide

Appendix C Search strategies

Bibliographic databases used to identify literature

Electronic database	Time period
AustHealth	1997 – 08/2005
Australian Medical Index	1996 – 08/2005
Australian Public Affairs Information Service (APAIS) – Health (Informit)	1990 – 08/2005
Cinahl	1988 – 08/2005
Cochrane Library – including, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database, the NHS Economic Evaluation Database	1988 – 08/2005
Current Contents	1993 - 08/2005
Embase	1988 – 08/2005
Pre-Medline and Medline	1988 – 08/2005
ProceedingsFirst	1998 – 08/2005
PsycInfo	1988 – 08/2005
Web of Science – Science Citation Index Expanded	1995 – 08/2005
EconLit	1988 – 08/2005

Other sources of evidence (1988 – 08/2005)

Electronic database	Source
Aetna	http://www.aetna.com/index
Australian Department of Health and Ageing	http://www.health.gov.au/
Current Controlled Trials metaRegister	http://controlled-trials.com/
Google Scholar	http://www.scholar.google.com/
Health Technology Assessment international	http://www.htai.org
International Network for Agencies for Health Technology Assessment	http://www.inahta.org/
NHMRC – National Health and Medical Research Council (Australia)	http://www.health.gov.au/nhmrc/
National Library of Medicine Health Services / Technology Assessment Text	http://text.nlm.nih.gov/
National Library of Medicine Locator Plus database	http://locatorplus.gov
New York Academy of Medicine Grey Literature Report	http://www.nyam.org/library/greylit/index.shtml
Specialty websites	See Appendix D
Trip database	http://www.tripdatabase.com
UK National Research Register	http://www.update-software.com/national/
US Department of Health and Human Services (reports and publications)	http://www.os.dhhs.gov/
US Medicare	http://www.medicare.gov/
Websites of Health Technology Agencies	See Appendix D
Hand searching (journals from 2004–05)	
American Heart Journal	Library or electronic access

American Journal of Cardiology	Library or electronic access
Annals of Clinical Biochemistry	Library or electronic access
Circulation	Library or electronic access
Circulation Research	Library or electronic access
Clinical Chemistry	Library or electronic access
European Journal of Heart Failure	Library or electronic access
Journal of the American College of Cardiology	Library or electronic access
Journal of Cardiac Failure	Library or electronic access
Heart (British Cardiac Society)	Library or electronic access
Heart, Lung and Circulation	Library or electronic access
Hypertension	Library or electronic access
Stroke	Library or electronic access
Thorax	Library or electronic access
Expert clinicians	
Any information provided by expert clinicians associated with this review was assessed as to whether it met the inclusion criteria	MSAC Advisory Panel
Pearling	
All included articles had their reference lists searched for additional relevant source material	

Search terms used

Area of inquiry	Search terms
Safety, effectiveness and cost-effectiveness of B-type natriuretic peptide assays	**MeSH** Heart failure, congestive; ventricular dysfunction; dyspnoea; edema, cardiac; peptide fragments; natriuretic peptide, brain **Text words** Congestive heart failure, HF, heart failure, ventricular dysfunction, heart decompensation, cardiac *edema, dyspn*ea pro-BNP, pro-brain natriuretic peptide (1-76), NT-BNP, proBNP (1-76), Amino-terminal pro-brain natriuretic peptide, NT-proBNP, N-terminal pro-BNP, BNP, brain natriuretic peptide, pro?bnp, bnp, B-type natriuretic peptide **Limits** Human, publication year [1988–2005]

Appendix D Internet sites searched

Websites of health technology assessment groups

AUSTRALIA

- Australian Safety and Efficacy Register of New Interventional Procedures – Surgical (ASERNIP-S) http://www.surgeons.org/open/asernip-s.htm
- Centre for Clinical Effectiveness, Monash University http://www.med.monash.edu.au/healthservices/cce/evidence/
- Health Economics Unit, Monash University http://chpe.buseco.monash.edu.au

AUSTRIA

- Institute of Technology Assessment / HTA unit http://www.oeaw.ac.at/ita/e1-3.htm

CANADA

- Agence d'Evaluation des Technologies et des Modes d'Intervention en Santé (AETMIS) http://www.aetmis.gouv.qc.ca/fr/
- Alberta Heritage Foundation for Medical Research (AHFMR) http://www.ahfmr.ab.ca/publications.html
- Canadian Coordinating Office for Health Technology Assessment (CCHOTA) http://www.ccohta.ca/entry_e.html
- Canadian Health Economics Research Association (CHERA/ACRES) – Cabot database http://www.mycabot.ca
- Centre for Health Economics and Policy Analysis (CHEPA), McMaster University http://www.chepa.org
- Centre for Health Services and Policy Research (CHSPR), University of British Columbia http://www.chspr.ubc.ca
- Health Utilities Index (HUI) http://www.fhs.mcmaster.ca/hug/index.htm
- Institute for Clinical and Evaluative Studies (ICES) http://www.ices.on.ca

DENMARK

- Danish Institute for Health Technology Assessment (DIHTA) http://www.dihta.dk/publikationer/index_uk.asp
- Danish Institute for Health Services Research (DSI) http://www.dsi.dk/engelsk.html

FINLAND

- FINOHTA http://www.stakes.fi/finohta/e/

FRANCE

- L'Agence Nationale d'Accréditation et d'Evaluation en Santé (ANAES) http://www.anaes.fr/

GERMANY

- German Institute for Medical Documentation and Information (DIMDI) / HTA http://www.dimdi.de/en/hta/index.html

THE NETHERLANDS

- Health Council of the Netherlands Gezondheidsraad http://www.gr.nl/adviezen.php

NEW ZEALAND

- New Zealand Health Technology Assessment (NZHTA) http://nzhta.chmeds.ac.nz/

NORWAY

- Norwegian Centre for Health Technology Assessment (SMM) http://www.oslo.sintef.no/smm/Publications/Engsmdrag/FramesetPublications.htm

SPAIN

- Agencia de Evaluación de Tecnologias Sanitarias, Instituto de Salud "Carlos III"I/Health Technology Assessment Agency (AETS) http://www.isciii.es/aets/

- Catalan Agency for Health Technology Assessment (CAHTA) http://www.aatrm.net/html/en/dir393/doc7921.html

SWEDEN

- Swedish Council on Technology Assessment in Health Care (SBU) http://www.sbu.se/www/index.asp

- Center for Medical Health Technology Assessment http://www.cmt.liu.se/English/Engstartsida.html

SWITZERLAND

- Swiss Network on Health Technology Assessment (SNHTA) http://www.snhta.ch/

UNITED KINGDOM

- Health Technology Board for Scotland http://www.htbs.org.uk/

- National Health Service Health Technology Assessment (UK) / National Coordinating Centre for Health Technology Assessment (NCCHTA) http://www.hta.nhsweb.nhs.uk/

- University of York NHS Centre for Reviews and Dissemination (NHS CRD) http://www.york.ac.uk/inst/crd/

- National Institute for Clinical Excellence (NICE) http://www.nice.org.uk

UNITED STATES

- Agency for Healthcare Research and Quality (AHRQ) http://www.ahrq.gov/clinic/techix.htm

- Harvard School of Public Health – Cost-Utility Analysis Registry http://www.hsph.harvard.edu/cearegistry/

- U.S. Blue Cross/ Blue Shield Association Technology Evaluation Center (TEC) http://www.bcbs.com/tec/index.html

Specialty websites

American Heart Association www.americanheart.org

American College of Cardiology www.acc.org

British Heart Foundation www.bhf.org.uk

Cardiac Society of Australia and New Zealand www.csanz.edu.au

European Society of Cardiology www.escardio.org

Heart Rhythms Society www.hrpatients.org

National Academy of Clinical Biochemistry www.nacb.org

National Heart Foundation of Australia www.heartfoundation.com.au

National Heart, Lung and Blood Institute (US) www.nhlbi.gov

New Zealand Heart Foundation www.nhf.org.nz

Appendix E Studies included in this review

Diagnosis in hospital setting

Studies on BNP assays

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Alibay et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Ambroise Pare Hospital, Boulogne Billancourt, France	160	Referred to ED for dyspnoea	Age: 80±14 yrs M/F: 76/84 HF: 60/160 (38%) CAD: 45/160 (28%) Pulmonary disease: 55/160 (34%) Unclear if HF hx	Consensus clinical diagnosis Two senior cardiologists using all data[b] including ECG, echo-cardiogram, CXR and the effect of therapy	BNP Single-use fluorescence immunoassay Triage kit, Biosite Diagnostics		CX P1 Level II diagnostic evidence Q1 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Apple et al 2003)	Retrospective cohort – cross-classified	Two hospitals – hospital-wide usage Minneapolis, MN and Hartford, CT. USA	334	Chart reviews of patients ruled in/out for HF or being monitored for HF therapy decisions	Age: HF – mean of 67 yrs; non-HF – mean of 61 yrs M/F: 52%/48% HF: 172/334 (52%) Unclear if HF hx		BNP Single-use fluorescence immunoassay Triage kit, Biosite Diagnostics	Clinical diagnosis Physician discharge dictations based on NYHA clinical criteria; and ICD-9 codings	CX P2 Level III-2 diagnostic evidence Q3 [Quadas = 5/14]	Sensitivity Specificity Negative predictive value False positive rate False negative rate

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Dao et al 2001) *Possible partial overlap with Maisel 2002 and Knudsen 2004. Raw data available*	Prospective cohort – cross-classified	Hospital – ED San Diego Veteran's Healthcare System, USA	250	Convenience sample presenting to ED with symptoms of dyspnoea. Patients with dyspnoea or acute coronary syndromes clearly not secondary to HF were excluded	Age: 63±0.9[c] yrs M/F: 94/6 HF: 97/250 (39%) Patients with HF hx included	Consensus clinical diagnosis Two senior cardiologists using all data[b] including ECG, echo-cardiogram, CXR, the effect of therapy and further cardiac testing	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		CX P2 Level III-1 diagnostic evidence Q2 [Quadas = 10/14]	Sensitivity Specificity Negative predictive value False positive rate False negative rate
(Dokainish et al 2004)	Prospective cohort – cross-classified	Hospital Baylor College of Medicine, Houston, Texas, USA	122 [ITT: 145]	Hospital inpatients referred to cardiology consult service for suspected HF	Age: 56±13 yrs M/F: 62/60 HF: 70/122 (57%) Patients with HF hx included	Clinical diagnosis Cardiologist using all data[b] including patient history, clinical exam, laboratory and radiographic tests. Applied Framingham criteria	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		CX P2 Level III-1 diagnostic evidence Q2 [Quadas = 11/14]	Sensitivity Specificity
(Fleischer et al 1997)	Prospective cohort – cross-classified	Hospital Christchurch Hospital, Christchurch, New Zealand	123	Patients requiring urgent admission to hospital for acute dyspnoea	Age: 68 (23–90) yrs M/F: 69/54 HF: 43/123 (35%) Patients with		BNP In-house assay	Clinical diagnosis Based on intention-to-treat HF with diuretic therapy within 24 hours	CX P1 Level III-2 diagnostic evidence Q3 [Quadas = 7/14]	Sensitivity Specificity Negative predictive value False positive rate False negative rate

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
					HF hx included			of admission		
(Knudsen et al 2004) *Same study as Maisel 2002 and McCullough 2002 – although more complete raw data set. Additional independent outcome information reported. Possible overlap with Krishnaswamy 2001 and Dao 2001.*	Prospective cohort – cross-classified	Multicentre study of five US and two European teaching hospitals – EDs 'Breathing Not Properly' multinational study	880 [ITT: 1586]	Patients presenting to ED with sudden onset of dyspnoea or worsening of chronic dyspnoea Excluded: dyspnoea obviously not related to HF, myocardial infarct or advanced renal failure; patients without complete case history	Age: 64±16 yrs M/F: 482/398 HF: 447/880 (51%) Patients with HF hx included	Consensus clinical diagnosis Two independent cardiologists using all data[b] <30 days from ED visit, including patient history, clinical exam, clinical tests and CXR. Applied Framingham criteria and NHANES for HF	BNP Fluorescence immunoassay Triage kit, Biosite Diagnostics		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 11/14]	Sensitivity Specificity Negative predictive value False positives False negatives
(Kosowsky et al 2003)	Pre-test/post-test case series	Hospital – ED Brigham and Women's Hospital, Boston, Massachusetts, USA	88	Consecutive patients presenting to ED with dyspnoea Excluded: <55 yrs	Not stated	Clinical diagnosis (± echo-cardiogram)	Clinical diagnosis (± echo-cardiogram) + BNP BNP – Point-of-care commercial kit		C1 P1 Level IV intervention evidence [Abstract]	Change in diagnosis Change in management

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Koulouri et al 2004)	Prospective cohort – cross-classified	Hospital – ED, paediatric ICU and wards, cardiothoracic ICU Childrens Hospital Los Angeles, Los Angeles, California, USA	49 [ITT: 51]	Infants and children presenting with objective findings of respiratory distress Excluded: chronic lung disease, renal dysfunction, premature babies, single ventricle physiology, comorbid renal or cardiac disease, incomplete echo-cardiographic data	Age: n/a for overall group M/F: n/a HF: 23/49 (47%) Unclear if HF hx	Clinical diagnosis According to New York University Pediatric Heart Failure criteria and echocardio-graphy to establish cardiac diagnosis	BNP Fluorescence immunoassay Triage kit, Biosite		CX P2 Level III-2 diagnostic evidence Q3 [Quadas = 10/14]	Sensitivity Specificity Negative predictive value
(Krishnaswamy et al 2001) *Possible partial overlap with Maisel 2002 and*	Prospective cohort – cross-classified	Hospital San Diego Veteran's Healthcare System, USA	400	Inpatients and outpatients referred for echo-cardiography to evaluate left ventricular function.	Age: n/a for whole group M/F: 385/15 HF: 132/400 (33%) Patients with HF hx included	Clinical diagnosis Cardiologists using Framingham criteria and confirmed by echo-	BNP Fluorescence immunoassay Triage kit, Biosite Diagnostics		CX P2 Level III-1 diagnostic evidence Q2 [Quadas = 10/14]	Sensitivity Specificity Negative predictive value

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
Knudsen 2004				Excluded: patients with referral to assess valve disease, vegetation or cardiac cause of stroke		cardiography				
(Lainchbury et al 2003)	Prospective cohort – cross-classified	Hospital – ED Christchurch Hospital, Christchurch, New Zealand	205	Patients presenting to ED with dyspnoea Excluded: blood sample unable to be obtained within 8 hrs of visit	Age: 70±14 yrs M/F: 100/105 HF: 70/205 (34%) Patients with HF hx included	Consensus clinical diagnosis Two independent cardiologists using all data[b], including ED and inpatient records and all results of investigations. Applied Framingham criteria and European Society of Cardiology Guidelines for HF	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics [+ 2 in-house BNP assays – clinical + research assay]		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Logeart et al 2002)	Prospective cohort – cross-classified	Hospital Hôpital Beaujon, Clichy, France	163 [ITT= 235]	Patients presenting to ED with acute severe dyspnoea Excluded: patients with AMI, chest	Age: n/a for whole group M/F: 109/54 HF: 115/163 (71%) Patients with HF hx included	Consensus clinical diagnosis Two cardiologists and a pneumologist using all data[b], including hospital course and	BNP Immuno-fluorescence assay Triage kit, Biosite Diagnostics		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 11/14]	Sensitivity Specificity Negative predictive value

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
				injury, recent surgery, receiving Rx 2+ hrs before arrival, or where echo-cardiography not feasible		cardiac tests. Applied Framingham criteria for HF				
(Maisel et al 2002) (McCullough et al 2002) *Knudsen 2004 also provided subgroup analysis on complement-ary data; possible partial overlap with Dao 2001 and Krishnaswa my 2001*	Prospective cohort – cross-classified	Multicentre study of five US and two European teaching hospitals – EDs 'Breathing Not Properly' multinational study	1586	Patients presenting to ED with predominant symptom of dyspnoea Excluded: dyspnoea obviously not related to HF, myocardial infarct or advanced renal failure; unstable angina, <18 yrs	Age: 64±17 yrs M/F: 883/703 HF: 744/1586 (47%) Patients with HF hx included	Consensus clinical diagnosis Two cardiologists using all data[b] <30 days from ED visit, including patient history, clinical exam, clinical tests (ie echo-cardiogram), hospital course and CXR. Applied Framingham criteria and NHANES for HF	BNP Immuno-fluorescence assay Triage kit, Biosite		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 13/14]	Sensitivity Specificity Negative predictive value
(Mueller et al 2004b) (Mueller et al 2004a)	Randomised, single-blind controlled trial	University hospital – ED Basel,	452 BNP arm:	Consecutive patients presenting to ED with acute dyspnoea as	BNP arm: Age: 70 [95%CI 68,72] yrs M/F: 132/93	Clinical diagnosis (± echo-cardiogram)	Clinical diagnosis + BNP (± echo-cardiogram)		C1 P1 Level II intervention evidence	Change in management • Admission - hospital

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Mueller et al 2005b) (Mueller et al 2005a)	+ 3 pre-specified subgroup analyses based on: • Gender • Age • Renal disease status	Switzerland 'B-Type Natriuretic Peptide for Acute Shortness of Breath Evaluation (BASEL) Study'	n=225 Control arm: n=227	primary symptom Excluded: patients with dyspnoea caused by trauma; severe renal disease; cardiogenic shock; requesting early transfer to another hospital	HF: (45%) Patients with HF hx included Control arm: Age: 71 [95%CI 69,73] yrs M/F: 130/97 HF: (51%) Patients with HF hx included		[BNP – Fluorescence immunoassay Biosite Diagnostics]		Q1 [Downs & Black = 23/27 + adequately powered for primary outcome (time to discharge)]	- ICU • Time to treatment • Time to discharge Change in health outcome • In-hospital mortality • 30-day mortality • Readmission Costs • Treatment costs • ICU costs
(Mueller et al 2005c)	Prospective cohort – cross-classified	Hospital – ED St John of God Hospital, Linz, Austria	251 [ITT: 276]	Patients presenting to ED with predominant symptom of dyspnoea Excluded: patients with myocardial infarction, acute coronary syndrome, troponin positive, trauma,	Age: 58–82 yrs M/F: 234/17 HF: 139/265 (52%) Patients with HF hx included	Clinical diagnosis A cardiologist using all data[b] from patient records based on clinical evaluation (history, physical exam, ECG, CXR, liver sonography, echo-cardiography) and applying Framingham	BNP Enzyme immunoassay AxSYM assay, Abbott Laboratories		CX P2 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
				clinical evaluation not complete in 3 days, index test not complete within 4 hrs of blood withdrawal		criteria				
(Teboul et al 2004)	Pre-test/post-test case series	Mobile intensive care unit (MICU), Emergency Medical Service Paris, France	52	All emergency calls for patients with acute dyspnoea, which led to the dispatch of a MICU Excluded: patients with dyspnoea caused circumstantial-ly, patients <50 years	Age: 79 (53–96) yrs M/F: 29/23 HF: not stated Unclear if HF hx		Clinical diagnosis + BNP BNP – Immuno-fluorescence assay Triage kit, Biomedical Diagnostics	Clinical diagnosis (mobile emergency physician) – based on medical history, physical exam, signs and symptoms, ECG anomalies	C1 P1 Level IV intervention evidence Q3 [Young et al = 3/3]	Change in diagnosis Change in management

M/F = male/female; hx = history; NYHA = New York Heart Association classifications for heart failure; BP = blood pressure; HF = heart failure; ITT = intention-to-treat; CAD = coronary artery disease; LVD = left ventricular dysfunction; n/a = not available; NHANES = National Health and Nutrition Examination Survey; ED = emergency department; AMI = acute myocardial infarction; CXR = chest X-ray; ECG = electrocardiogram; [a] For an explanation of the appraisal system used in this assessment, refer to section on 'Strength of the evidence in individual studies' in chapter on 'Approach to assessment'; [b] Excluding BNP data; [c] Unclear if this is a standard error, instead of a standard deviation

Studies on NT-proBNP assays

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Alibay et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Ambroise Pare Hospital, Boulogne Billancourt, France	160	Referred to ED for dyspnoea	Age: 80±14 yrs M/F: 76/84 HF: 60/160 (38%) CAD: 45/160 (28%) Pulmonary disease: 55/160 (34%) Unclear if HF hx	Consensus clinical diagnosis Two senior cardiologists using all data[b] including ECG, echo-cardiogram, CXR and the effect of therapy	NT-proBNP Chemi-luminescent sandwich immuno-assay Elecsys 2010, Roche Diagnostics		CX P1 Level II diagnostic evidence Q1 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Bayes-Genis et al 2004)	Prospective cohort – cross-classified	Emergency hospital department Hospital de la Santa Creu i Sant Pau, Barcelona, Spain	89 [ITT: 100]	Patients with symptoms of acute dyspnoea attending ED Excluded: NYHA class I & II, <40 yrs, dyspnoea secondary to chest trauma or cardiac tamponade, acute coronary syndromes (except with HF	Age: overall not available M/F: 54/35 HF: 74/89 (83%) – decompensated HF = 58%, masked HF = 25% Unclear if HF hx	Consensus clinical diagnosis Two senior cardiologists using all data[b] <7 days from ED visit; including ECG, echo-cardiogram, spirometry, CXR, pulmonary volumes and arterial blood gases	NT-proBNP Chemi-luminescent sandwich immuno-assay Elecsys 1010, Roche Diagnostics		CX P1 Level II diagnostic evidence Q1 [Quadas = 14/14]	Sensitivity Specificity Negative predictive value

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
				as cause), severe renal insufficiency and liver cirrhosis						
(Januzzi et al 2005)	Prospective cohort – cross-classified	Emergency hospital department Massachusetts General Hospital, Boston, Massachusetts, USA	599	Patients presenting to ED with dyspnoea Excluded: <21 yrs; severe renal insufficiency; dyspnoea after chest trauma or secondary to severe coronary ischaemia; >2 hrs delay after urgent IV loop diuretic Rx; unblinded BNP measurement	Age: 22–95 yrs M/F: 51%/49% HF: 209/599 (35%) Patients with HF hx included	Consensus clinical diagnosis – Cardiologists using all data[b] up to 60-day follow-up, including laboratory and cardiac tests, echo-cardiogram, CXR. When necessary adjudicated dx re Framingham Heart Study criteria	NT-proBNP Chemi-luminescent immuno-assay Elecsys 1010, Roche Diagnostics		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Jose et al 2003)	Prospective cohort – cross-classified	Emergency and outpatient hospital departments Christian Medical College and Hospital, Vellore, India	119	Patients presenting to emergency or outpatient dept with dyspnoea and associated symptoms such as oedema,	Age: 54±12 yrs M/F: 78/41 HF: 73/119 (61%) Unclear if HF hx	Consensus clinical diagnosis Framingham criteria and echo-cardiography	NT-proBNP Enzyme immuno-assay Biomedica		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 11/14]	Sensitivity Specificity

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
				weight gain, cough or wheezing Excluded: patients with acute coronary syndromes						
(Lainchbury et al 2003)	Prospective cohort – cross-classified	Hospital – ED Christchurch Hospital, Christchurch, New Zealand	205	Patients presenting to ED with dyspnoea Excluded: blood sample unable to be obtained within 8 hrs of visit	Age: 70±14 yrs M/F: 100/105 HF: 70/205 (34%) Patients with HF hx included	Consensus clinical diagnosis Two independent cardiologists using all data[b], including ED and inpatient records and all results of investigations. Applied Framingham criteria and European Society of Cardiology Guidelines for HF	NT-proBNP Chemi-luminescent immunoassay Elecsys 2010, Roche Diagnostics [+ an in-house radioimmuno-assay]		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Nielsen et al 2004)	Post-test case series	Hospital Hospital-based clinic, Denmark	345	Consecutive patients referred by 74 GPs with dyspnoea symptoms	Age: median 65 (18–89) yrs M/F: 51%/49% HF: 81/345 (24%) Patients with HF hx included		NT-proBNP Chemi-luminescent sandwich immuno-assay Roche Diagnostics	Echo-cardiogram	C1 P1 Level IV intervention evidence Q3	*(potential)* Change in management

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
				Excluded: missing blood sample or other diagnostic data						
(Mueller et al 2005c)	Prospective cohort – cross-classified	Hospital – ED St John of God Hospital, Linz, Austria	251 [ITT: 276]	Patients presenting to ED with predominant symptom of dyspnoea Excluded: patients with myocardial infarction, acute coronary syndrome, troponin positive, trauma, clinical evaluation not complete in 3 days, index test not complete within 4 hrs of blood withdrawal	Age: 58–82 yrs M/F: 234/17 HF: 139/265 (52%) Patients with HF hx included	Clinical diagnosis A cardiologist using all data[b] from patient records based on clinical evaluation (history, physical exam, ECG, CXR, liver sonography, echo-cardiography) and applying Framingham criteria	NT-proBNP Chemi-luminescent sandwich immuno-assay Elecsys 2010, Roche Diagnostics		CX P2 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value

M/F = male/female; HF = heart failure; CAD = coronary artery disease; ECG = electrocardiography; CXR = chest X-ray; ITT = intention-to-treat; n/a = not available; Rx = treatment; IV = intravenous; dx = diagnosis; GP = general practitioner; [a] For an explanation of the appraisal system used in this assessment, refer to section on 'Strength of the evidence in individual studies' in 'Approach to assessment' chapter; [b] Excluding NT-proBNP data; [c] Unclear if this is a standard error, instead of a standard deviation.

Monitoring

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics BNP group	Patient and study characteristics Clinical group	Outcome (number of outcomes)	Assay details
(Inomata et al 2003) Sagamihara, Japan	Randomised controlled trial Level - II Quality score - NA Cons. Rec. - Y Blinded - NR Objective - Y Follow-up – 2 years	Population Patients with LVEF <40% discharged from hospital after decompensated HF (n=73)	Not reported	Not reported	Primary Rate of mortality Secondary Readmission	Not reported
(Troughton et al 2000) Christchurch, NZ	Randomised controlled trial Level - II Quality score - 23/26 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up – 9.7 months for BNP 9.5 months for clinical group	Population Patients with impaired systolic function (LVEF <40%) and symptomatic HF Exclude Recent acute coronary syndrome (within 3 months), pending cardiac transplant or revascularisation, severe stenotic valvular heart disease or by severe pulmonary hepatic or renal disease	Sample description N 33 Mean age 68 % male 78 % IHD 73 Disease severity (admission) NT-BNP(pmol/L) 217 % NYHA class II 72 LVEF 28 Comorbidities Diabetes 12 Hypertension 64	Sample description N 36 Mean age 72 % male 75 % IHD 75 Disease severity(admission) NT-BNP (pmol/L) 251 % NYHA class II 67 LVEF 26 Comorbidities Diabetes 14 Hypertension 67	Primary Total cardiovascular events Functional capacity Secondary Hospital readmission Left-ventricular function	Not reported

(Murdoch et al 1999) Glasgow, Scotland	Randomised controlled trial Level – II Quality score - ? Cons. Rec. - Y Blinded - Y Objective - Y Follow-up – 8 weeks	Population Mild to moderate CHF patients receiving stable conventional therapy attending a specialist chronic HF clinic	Sample description N 10 Mean age 62 % male 80 Disease severity (admission) BNP 112 pg/mL NYHA class 2.5 (0.17) LVEF (%) 25 (3.1)[a]	Sample description N 10 Mean age 64 % male 100 Disease severity(admission) BNP 140 pg/mL NYHA class 2.4 (0.16) LVEF (%) 25 (1.8)[a]	Hemodynamic changes	Direct, specific, monoclonal antibody radioimmunoassay kit supplied by Shionogi & Co (Settsu-shi, Osaka, Japan)

Level = level of evidence; HF = heart failure; Y = yes; N = no; Cons. Rec. = Was there consecutive patient recruitment?; Blinded = Was the outcome determined by assessors blinded to peptide concentration and other potentially prognostic variables?; Objective = Was assessment of the outcome objective?; Follow-up has been converted to days using the assumption that there are 30 days in each month; % gender/disease/medication; % male = percentage of the sample which were male, had the specified disease or were taking the specified medication; IHD = ischaemic heart disease; NA = not applicable; NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction; round brackets ( ) enclose a range; square brackets [ ] enclose an inter-quartile range.

Appendix F Excluded studies

Diagnosis in hospital setting

Unable to extract relevant data

Alehagen, U., Lindstedt, G. et al (2003). 'Utility of the amino-terminal fragment of pro-brain natriuretic peptide in plasma for the evaluation of cardiac dysfunction in elderly patients in primary health care', *Clinical Chemistry*, 49 (8), 1337–1346.

Bettencourt, P., Ferreira, A. et al (2000). 'Evaluation of brain natriuretic peptide in the diagnosis of heart failure', *Cardiology*, 93 (1–2), 19–25.

Reasons for referral for echocardiography included suspected HF and other reasons (therefore unclear if LVD rates relate solely to HF)

Atisha, D., Bhalla, M.A. et al (2004). 'A prospective study in search of an optimal B-natriuretic peptide level to screen patients for cardiac dysfunction', *American Heart Journa*l, 148 (3), 518–523.

Bhalla, V., Isakson, S. et al (2005). 'Diagnostic ability of B-type natriuretic peptide and impedance cardiography: Testing to identify left ventricular dysfunction in hypertensive patients', *American Journal of Hypertension*, 18 (2), 73S–81S.

Epshteyn, V., Morrison, K. et al (2003). 'Utility of B-type natriuretic peptide (BNP) as a screen for left ventricular dysfunction in patients with diabetes', *Diabetes Care*, 26 (7), 2081–2087.

Hutcheon, S.D., Gillespie, N.D. et al (2002). 'B-type natriuretic peptide in the diagnosis of cardiac disease in elderly day hospital patients', *Age Ageing*, 31 (4), 295–301.

Lubien, E., DeMaria, A. et al (2002). 'Utility of B-natriuretic peptide in detecting diastolic dysfunction - Comparison with Doppler velocity recordings', *Circulation*, 105 (5), 595–601.

Maisel, A.S., Koon, J. et al (2001). 'Utility of B-natriuretic peptide as a rapid, point-of-care test for screening patients undergoing echocardiography to determine left ventricular dysfunction', *American Heart Journal*, 141 (3), 367–374.

McLean, A.S., Tang, B. et al (2003). 'Increased B-type natriuretic peptide (BNP) level is a strong predictor for cardiac dysfunction in intensive care unit patients', *Anaesthesia and Intensive Care*, 31 (1), 21–27.

Mueller, T., Gegenhuber, A. et al (2004). 'Head-to-head comparison of the diagnostic utility of BNP and NT-proBNP in symptomatic and asymptomatic structural heart disease', *Clinica Chimica Acta*, 341 (1–2), 41–48.

Talwar, S., Squire, I.B. et al (1999). 'Plasma N-terminal pro-brain natriuretic peptide and the ECG in the assessment of left-ventricular systolic dysfunction in a high risk population', *European Heart Journal*, 20 (23), 1736–1744.

Vourvouri, E.C., Schinkel, A.F. et al (2003). 'Screening for left ventricular dysfunction using a hand-carried cardiac ultrasound device', *European Journal of Heart Failure*, 5 (6), 767–774.

Yamamoto, K., Burnett, J.C., Jr. et al (1996). 'Superiority of brain natriuretic peptide as a hormonal marker of ventricular systolic and diastolic dysfunction and ventricular hypertrophy', *Hypertension*, 28 (6), 988–994.

Yamamoto, K., Burnett J.C, Jr. et al (2000). 'Clinical criteria and biochemical markers for the detection of systolic dysfunction', *Journal of Cardiac Failure*, 6 (3), 194–200.

Data included in another paper

Knudsen, C.W., Riis, J.S. et al (2004). 'Diagnostic value of a rapid test for B-type natriuretic peptide in patients presenting with acute dyspnoea: effect of age and gender', *European Journal of Heart Failure*, 6 (1), 55–62.

Maisel, A.S., Clopton, P. et al (2004). 'Impact of age, race, and sex on the ability of B-type natriuretic peptide to aid in the emergency diagnosis of heart failure: results from the Breathing Not Properly (BNP) multinational study', *American Heart Journal*, 147 (6), 1078–1084.

Maisel, A.S., McCord, J. et al (2003). 'Bedside B-type natriuretic peptide in the emergency diagnosis of heart failure with reduced or preserved ejection fraction - Results from the breathing not properly multinational study', *Journal of the American College of Cardiology*, 41 (11), 2010–2017.

McCullough, P.A., Hollander, J.E. et al (2003). 'Uncovering heart failure in patients with a history of pulmonary disease: rationale for the early use of B-type natriuretic peptide in the emergency department', *Academic Emergency Medicine*, 10 (3), 198–204.

Morrison, L.K., Harrison, A. et al (2002). 'Utility of a rapid B-natriuretic peptide assay in differentiating congestive heart failure from lung disease in patients presenting with dyspnea', *Journal of the American College of Cardiology*, 39 (2), 202–209.

Steg, P.G., Joubin, L. et al (2005). 'B-type natriuretic peptide and echocardiographic determination of ejection fraction in the diagnosis of congestive heart failure in patients with acute dyspnea', *Chest*, 128 (1), 21–29.

Wu, A.H.B., Omland, T. et al (2004). 'The effect of diabetes on B-type natriuretic peptide concentrations in patients with acute dyspnea - An analysis from the breathing not properly multinational study', *Diabetes Care*, 27 (10), 2398–2404.

The non-hospital setting

Cowie, M.R., Struthers, A.D. et al (1997). 'Value of natriuretic peptides in assessment of patients with possible new heart failure in primary care', *Lancet*, 350 (9088), 1349–1353.

Gustafsson, F., Badskjaer, J. et al (2003). 'Value of N-terminal proBNP in the diagnosis of left ventricular systolic dysfunction in primary care patients referred for echocardiography', *HeartDrug*, 3 (3), 141–146.

Hobbs, F.D., Davis, R C. et al (2002). 'Reliability of N-terminal pro-brain natriuretic peptide assay in diagnosis of heart failure: cohort study in representative and high risk community populations', *British Medical Journal*, 324 (7352), 1498.

Hobbs, F.D., Davis, R.C. et al (2004). 'Reliability of N-terminal proBNP assay in diagnosis of left ventricular systolic dysfunction within representative and high risk populations', *Heart*, 90 (8), 866–870.

Landray, M.J., Lehman, R. & Arnold, I. (2000). 'Measuring brain natriuretic peptide in suspected left ventricular systolic dysfunction in general practice: Cross-sectional study', *British Medical Journal*, 320 (7240), 985–986.

Sim, V., Hampton, D. et al (2003). 'The use of brain natriuretic peptide as a screening test for left ventricular systolic dysfunction - Cost-effectiveness in relation to open access echocardiography', *Family Practice*, 20 (5), 570–574.

Wright, S.P., Doughty, R.N. et al (2003a). 'Plasma amino-terminal pro-brain natriuretic peptide and accuracy of heart-failure diagnosis in primary care - A randomized, controlled trial', *Journal of the American College of Cardiology*, 42 (10), 1793–1800.

Zaphiriou, A., Robb, S., Murray-Thomas, T., Mendez, G., Fox, K., Mcdonagh, T., Hardman, S.M.C., Dargie, H.J. & Cowie, M.R. (2005). The diagnostic accuracy of plasma BNP and NTproBNP in patients referred from primary care with suspected heart failure: Results of the UK natriuretic peptide study. *European Journal of Heart Failure,* 7**,** 537–541.

Abstracts – where a higher level of evidence was available in full text

Hobbs, F.D.R., Davis, R.C. et al. (2000). 'Plasma N-terminal pro-brain natriuretic peptide has similar predictive value to brain natriuretic peptide in diagnosis of heart failure in the community', *European Heart Journal*, 21, 133.

Hobbs, F.D.R., Davis, R.C. et al. (2001). 'Performance characteristics of N terminal pro-brain natriuretic peptide (NT-ProBNP) and BNP assays in the diagnosis of heart failure in community settings', J*ournal of the American College of Cardiology*, 37 (2), 147A.

Inomata, T., Nakano, H. et al. (2004). 'Exceptions of the diagnostic value for heart failure of plasma brain-type natriuretic peptide levels/disproportion between brain natriuretic peptide and norepinephrine levels implies constrictive pericarditis', *European Heart Journal*, 25, 627.

Jourdain, R., Logeart, D. et al. (2003). 'Brain natriuretic peptide usefulness for diagnosing heart failure in elderly patients (BUD study): about 300 patients over 75', *European Heart Journal*, 24, 116.

Luchner, A., Hengstenberg, C. et al. (2002). 'Automated measurement of N-terminal proBNP for biochemical detection of left ventricular dysfunction', *European Heart Journal*, 23, 572.

Luong, M.V., Auziere, L. et al. (2003). 'Echocardiographic determinants of brain natriuretic peptide levels in patients suspected of heart failure', *European Heart Journal*, 24, 412.

Maisel, A., Harrison, A. et al. (2001). 'B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnea', *European Heart Journal*, 22, 377.

Passino, C., Bramanti, F. et al. (2003). 'Eighteen-minute N terminal pro-brain natriuretic peptide versus 24-hour brain natriuretic peptide assay as diagnostic markers in congestive heart failure', *European Heart Journal*, 24, 360.

Triepels, R.H., Busscher, S. et al. (2003). 'N-terminal pro-brain natriuretic peptide (NT-proBNP) as screening test for early stage heart failure', *Clinical Chemistry*, 49 (6), A37.

Trochu, J.N., Pattier, S. et al. (2003). 'Accuracy of N terminal pro-brain natriuretic peptide for the diagnosis of congestive heart failure in patients admitted with dyspnea in the emergency department', *European Heart Journal*, 24, 260.

Vinereanu, D., Lim, P. et al. (2003). 'Echocardiographic screening for heart failure must analyse left-ventricular longitudinal function. Comparison versus brain natriuretic peptide', *European Heart Journal*, 24, 351.

Wieczorek, S.J., Ferrier, A. et al. (2000). 'B-type natriuretic peptide for the evaluation of congestive heart failure as determined from the New York Heart Association Classification system and the six minute walk test', *American Journal of Clinical Pathology*, 114 (2), 312.

Zaninotto, M., Mion, M. et al. (2003). 'N-Terminal pro-brain natriuretic peptide in the differential diagnosis of dyspnea in an emergency department', *Clinical Chemistry*, 49 (6), A41.

Zaphiriou, A., Robb, S. et al. (2003). 'Using brain natriuretic peptide and N terminal pro-brain natriuretic peptide to rule out heart failure: does it work in clinical practice? Results of the UK natriuretic peptide study', *European Heart Journal*, 24, 260.

Monitoring

Not enough information

Gackowski, A., Isnard, R. et al (2002). 'Brain natriuretic peptide plasma level falls early in severe acute heart failure responding to treatment and is a strong prognostic marker', *Circulation,* 106 (19), 564–565.

Groenning, B.A., Hildebrandt, P. et al (2003). 'Brain natriuretic peptide and N terminal pro-brain natriuretic peptide for treatment monitoring in patients with left-ventricular systolic heart failure: a substudy of the CARMEN trial', *European Heart Journal,* 24, 361.

Havranek, E.P., Masoudi, F.A. et al (2003). 'Changes in BNP level are not associated with outcomes in outpatients with heart failure', *Circulation,* 108 (17), 692–693.

Ishii, J., Nakamura, Y. et al (2003). 'Prognostic utility of N-terminal Pro-BNP versus BNP in patients hospitalized for worsening chronic heart failure', *Circulation,* 108 (17), 343.

Ishii, J., Wang, J.H. et al (1999). 'Early risk stratification using cardiac troponin T and brain natriuretic peptide in patients with congestive heart failure', *Circulation,* 100 (18), 679.

O'Neill, J.O., Bott-Silverman, C. et al (2003). 'B-type natriuretic peptide levels are not a surrogate marker for invasive haemodynamics during management of patients with severe heart failure', *Circulation,* 108 (17), 557.

Seino, Y., Fukushima, M. et al (2002). 'Plasma concentrations of N-terminal pro-BNP versus BNP in patients with chronic heart failure; More discerning marker for the progression', *Circulation,* 106 (19), 683.

Appendix G Unit cost of test

Indicative calculation of the unit cost of the actual NT-proBNP test under Australian conditions

Unit cost of consumables per patient

The major single consumable cost in cardiac peptide assays is the reagent. A 100 test kit costs $2,187.50, which amounts to $21.87 per aliquot.

The quality control (QC) samples contain two different known levels of NT-proBNP and are assayed in the same way as the patient sample. The cost of the QC sample is $12 per run (based on 13 assays per $150 box). In each run one aliquot of reagent is used for each patient sample and two aliquots are used for the QC samples.

Therefore, the cost of the QC assay is $12 + (2 * $21.87)

Calibrator samples cost $21 per set (2 per set) (based on 10 calibrations per $210 box). Calibrators are samples that contain a known amount of NT-proBNP and relate to an international standard. A set of calibrators is included in one run every 7 days and is assayed as if it is a patient sample except that each of the two levels in the calibrator set is assayed in duplicate. Therefore, every time a calibration is done, four aliquots of the reagent are used.

Therefore, the cost of a calibration assay is $21 + (4 * $21.87)

Unit cost of testing per patient

Total costs include reagent, calibrators and QC costs plus other laboratory costs (salaries, assets and overheads). An example showing costs for a batch specimen run (10 specimens per run) is provided.

1. Reagent costs

The number of patient samples that can be assayed in one run ranges from 1 to 80 (approximate maximum and varies with the platform).

The cost of an assay of a patient specimen will depend on the number of patient specimens in each run (A) and the number of runs each week (B).

Hence the formula:

$$\text{Reag/calib/QC cost (per specimen)} = \$21.87 + \frac{\$12 + (2 * \$21.87)}{A} + \frac{\$21 + (4 * \$21.87)}{A * B}$$

Assuming B = seven runs per week, then:

For 10 specimens per run, reagent cost = $28.99

2. Cost per test of salaries, assets, overheads etc

= $14.25 per test

3. Therefore, total costs

for 10 specimens per run = $43.24

Indicative calculation for performing and reporting the test

Allowing for a 17 per cent profit margin, and assuming that no copayment is involved, the indicative unit cost of performing and reporting the test is:

$43.24 x 1.17 = $50.59

Appendix H Statistical methods for economic considerations

Estimation of the 95% confidence interval for cost

The 95% confidence interval for overall cost savings per 100 patients was calculated using AR-DRG cost/activity estimates and data from Mueller et al (2004b). Each AR-DRG code for heart failure and the alternative diagnoses provided point cost estimates and standard errors for each complication category (ie suffix A, B or C) in both the public and private health sectors. A weighted cost estimate across complication categories for each disease was calculated, using the number of separations within each complication category, for both the public and private sectors. An overall combined cost for each disease was then calculated by weighting estimates across public and private sectors, again using the relevant number of separations. As the overall combined cost estimate for each disease was a linear combination of original AR-DRG cost estimates (using appropriate weights), the associated standard error was readily calculated from original AR-DRG standard errors using standard statistical techniques. Standard errors were multiplied by 1.96 and then subtracted and added to the point estimate to calculate a 95% confidence interval wherever needed.

The other source of error in cost estimation was the number of primary admissions per 100 patients with or without BNP testing in the diagnostic workup (Mueller et al 2004b). For primary admission rates, 95% confidence intervals were calculated using standard statistical techniques. It was assumed that the number of readmissions remained a constant fraction of primary admissions (which had variance) and that the point cost estimate for non-admitted patients was a constant (ie had no variance).

The 95% confidence interval for cost savings per 100 patients varied from $304 to $67,393 (ie, it did not cross zero), indicating a strong probability that use of the test in these circumstances is cost saving.

Estimation of the 95% confidence interval for effectiveness

The incremental effectiveness of the use of B-type natriuretic peptide tests as part of the diagnostic strategy was calculated from Mueller's data, that is the difference in 30-day mortality rates between the two diagnostic arms of the trial. The associated standard error was derived by approximating 30-day mortality rates for the two diagnostic arms as being from a normal distribution, and then calculating the error of their difference. As before, the standard error was multiplied by 1.96 and then subtracted and added to the point estimate to calculate a 95% confidence interval. This varied from –3.2 to 8.3 per 100 patients and thus crossed zero, indicating that the sign of the point estimate is subject to uncertainty.

Modelling cost-effectiveness

Assessment of the joint probability distribution around the point estimate was achieved by simulating a large number of cost and effectiveness values, both from normal distributions, whose means and variances were defined by the final estimates and their associated confidence intervals. It was assumed that there was no dependence between cost and effectiveness. These points were plotted on the incremental cost-effectiveness

plane so that a confidence area which captured 95 per cent of the data could be estimated and the percentage of data points in each quadrant could be calculated.

Appendix I B-type natriuretic peptide assays as a prognostic tool for patients with heart failure

Summary

Prognostic use

Heart failure (HF) prognosis assists decisions regarding the type and aggressiveness of therapy that should be instituted, and allows patients to adapt accordingly and plan for the future. B-type natriuretic peptides are indicated to be prognostic predictors of death and/or rehospitalisation. In the proposed Australian setting, B-type natriuretic peptides would not replace current prognostic indicators of HF. To be of use, B-type natriuretic peptides must provide additional prognostic information over and above that already provided by the battery of currently measured variables.

It is important to highlight that the prognostic ability of B-type natriuretic peptides has been tested in stabilised HF populations. Measuring B-type natriuretic peptides on discharge from hospital is therefore likely to yield a more accurate estimate of prognosis because it represents the chronic disease state rather than an acute episode.

Effectiveness

The body of evidence available for the prognostic potential of B-type natriuretic peptides was good to excellent. Sixty prognostic studies met the criteria for inclusion in the review. Studies included for assessment recruited patients with varying disease severity and adjusted for different confounders in addition to different length of follow-up, data input (dichotomous, continuous) and outcomes (death, death or cardiovascular events, cardiovascular events). Despite this, results are remarkably consistent, with the vast majority of studies reporting that B-type natriuretic peptides provide additional prognostic information over that already provided by a variety of existing clinical strategies. The evidence base was more extensive for BNP (n=36) compared to NT-proBNP (n=16) but results were consistent across studies. Eight studies provided insufficient information for formal inclusion in the meta-analyses but were assessed narratively and used in sensitivity analysis to test the robustness of the effect. Seven studies included both BNP and NT-proBNP in their multivariate analyses.

NT-proBNP assays

Only one of the 16 studies that formed the evidence base for NT-proBNP reported a non-significant effect. Statistically significant or close to significant pooled estimates were present for all outcomes (death, death or cardiovascular event, cardiovascular event) which highlighted the independent prognostic value of NT-proBNP. In most cases a random effects model was used, which demonstrates that the magnitude of effect varied between studies. Non-significant ($p>0.05$) pooled estimates were calculated on a limited

number of studies (ie <3). The lack of significance may be due to the small study numbers rather than a lack of real effect.

In conclusion, there is strong evidence of the independent prognostic potential of NT-proBNP on a moderate volume evidence base.

BNP assays

Thirty-six studies reported on at least one outcome (death, death or cardiovascular event, cardiovascular event) in assessing the independent prognostic potential of BNP. Eight of these studies reported that BNP was not independently predictive of outcome after controlling for chosen variables.

Statistically significant pooled estimates or non-significant trends were present for all outcomes (death, death or cardiovascular event, cardiovascular event) and random effect models dominated the analyses. This demonstrates that BNP plays an independent role in HF patient prognosis and that the magnitude of effect varies between studies. Reasons for the heterogeneity have not been investigated due to the relatively small study numbers for each subcategory (continuous and dichotomous) for each of three outcomes. In three out of four cases, non-significant pooled estimate trends were reported on less than three studies. The limitations of meta-analysing small study numbers, rather than lack of real effect, may explain the lack of statistical significance of a pooled result.

In conclusion, there is strong evidence of the independent prognostic potential of BNP on a large volume evidence base.

Approach to assessment

Background

Patient prognosis is often established simultaneously with diagnosis to aid clinical decisions regarding the type and aggressiveness of therapy (Cardarelli & Lumicao 2003). For example, invasive treatments might be deemed inappropriate in patients with a very poor prognosis; and it may be of value to optimise drug treatments for those patients identified as most at risk of death (Bouvy et al 2003). Prognostic information may also allow patients to adapt and plan for the future (Glasziou et al 2001).

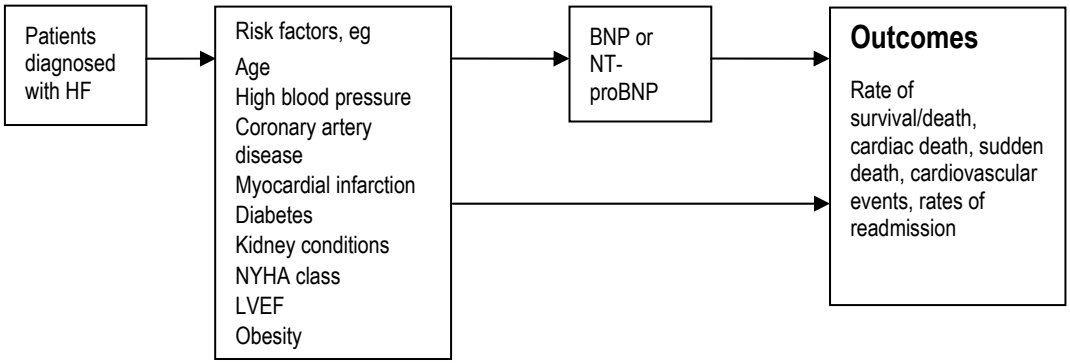
Similar to diagnosis, there is no consensus as to the method of risk stratification for patients with HF. Current methods of determining prognosis include the use of demographics (age, sex, marital status), medical history (previous admissions for HF), clinical presentation (blood pressure, renal dysfunction), comorbidities (depression, diabetes), electrocardiography, measurement of cardiac performance and biochemical markers of myocardial damage (Bouvy et al 2003; Jernberg et al 2004).

Traditionally, assessing functional impairment is the primary method used to determine clinical status in HF (Bettencourt 2004). Some of the common methods include the New York Heart Association (NYHA) classification, quality of life scales, peak oxygen uptake (VO_2) and the HF survival score (Bettencourt 2004; Doust et al 2005). The limitation of the aforementioned measurements is that the relationship between symptoms and the severity of cardiac dysfunction or prognosis is often poor (Remme & Swedberg 2001). Other methods have included measurement of left ventricular ejection fraction (through an echocardiogram), serum sodium concentrations, age and history of diabetes mellitus (Doust et al 2005). The role of left ventricular ejection fraction (systolic function) for prognosis is currently being debated. A recent European study stratified patients by age group (<75 and ≥ 75 years) and found no significant difference in survival between patients with preserved or deteriorated systolic function (Varela-Roman et al 2005).

Obesity has been alluded to as a major risk factor for developing HF; however, once HF has developed, a lower body mass index is related to worse prognosis (Mehra et al 2004).

To assess the effectiveness of the BNP and NT-proBNP assays as supplemental prognostic tests, the effect of the *addition* of B-type natriuretic peptides to current risk factors on patient relevant outcomes would need to be compared to the current risk factors alone (eg age, blood pressure, diabetes; see Figure 9).

Figure 9 Clinical pathway for use of B-type natriuretic peptide assays for heart failure prognosis



HF = heart failure; NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction

Objective

The objective of this assessment is to determine whether there is sufficient evidence that the NT-proBNP and/or the BNP assay can effectively predict clinical outcomes for patients with HF (ie its function as a prognostic tool). This is linked to the assays' role in monitoring and guiding the treatment of patients with HF.

Research questions

Effectiveness

- Does BNP accurately predict health outcomes in patients with heart failure over and above other known risk factors (eg left ventricular ejection fraction, New York Heart Association class, renal insufficiency, hypertension, coronary artery disease, history of myocardial infarction, obesity, serum sodium concentrations, age, history of diabetes mellitus, peak oxygen uptake (VO_2) or the heart survival score)?

- Does NT-proBNP accurately predict health outcomes in patients with heart failure over and above other known risk factors (eg left ventricular ejection fraction, New York Heart Association class, renal insufficiency, hypertension, coronary artery disease, history of myocardial infarction, obesity, serum sodium concentrations, age, history of diabetes mellitus, peak oxygen uptake (VO_2) or the heart survival score)?

Expert advice

An advisory panel with expertise in pathology, clinical biochemistry, cardiology and consumer issues was established to evaluate the evidence from this systematic review of the literature and to provide advice to the MSAC from a clinical perspective. In selecting members for advisory panels, the MSAC's practice is to approach the appropriate medical colleges, specialist societies and associations and consumer bodies for nominees. Membership of the advisory panel associated with this MSAC assessment is provided at Appendix A.

Review of the literature

Literature sources and search strategies

The medical literature was searched to identify relevant studies concerning B-type natriuretic peptides for the period between 1988 and August 2005, as B-type natriuretic peptide assays were first reported in 1988. Appendix C describes the electronic databases that were used for this search and other sources of evidence—particularly grey literature—that were investigated. Grey literature was included in the search strategy. Unpublished literature, however, was not canvassed as it is difficult to search for this literature exhaustively and systematically, and trials that are difficult to locate are often smaller and of lower methodological quality (Egger et al 2003). It is, however, possible that these unpublished data could impact on the results of this review.

The search terms used to identify literature in electronic databases on the safety and effectiveness of using B-type natriuretic peptide assays to predict clinical outcomes in patients with HF are also presented in Appendix C.

Inclusion/Exclusion criteria

In general, studies were excluded if they:

- did not address the research question;
- did not provide information on the pre-specified target population;
- did not include one of the pre-specified interventions;
- did not compare results to the pre-specified comparator;
- did not address one of the pre-specified outcomes and/or provided inadequate data on these outcomes (in some instances, a study was included to assess one or more outcomes but had to be excluded for other outcomes due to data inadequacies); or
- did not have the appropriate study design.

Where two (or more) papers reported on different aspects of the same study, such as the methodology in one and the findings in the other, they were treated as one study. Similarly, if the same data were duplicated in multiple articles, results from the most comprehensive, or most recent article only were included.

The criteria for including studies relevant to determining the *effectiveness* of B-type natriuretic peptide assays as prognostic tools are provided in Box 6.

Box 6 Study selection criteria for prognostic effectiveness

Selection criteria	Inclusion criteria
Population	Patients with heart failure
Intervention	NT-proBNP or BNP assays + other known risk factors
Comparator(s)	Risk factors – Left ventricular ejection fraction, New York Heart Association class, serum sodium concentrations, age, history of diabetes mellitus, renal insufficiency, hypertension, obesity, coronary artery disease, history of myocardial infarction, peak oxygen uptake or heart survival score
Outcomes	Rate of survival/death, cardiac death, sudden death, cardiovascular events, readmission rate
Study design	Prospective or retrospective cohort studies or systematic reviews of cohort studies
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base.

Search results

The process of study selection for this report went through six phases:

1. All reference citations from all literature sources were collated into an Endnote 8.0 database;

2. Duplicate references were removed;

3. Studies were excluded, on the basis of the complete citation information, if it was obvious that they did not meet the inclusion criteria. All other studies were retrieved for full-text assessment;

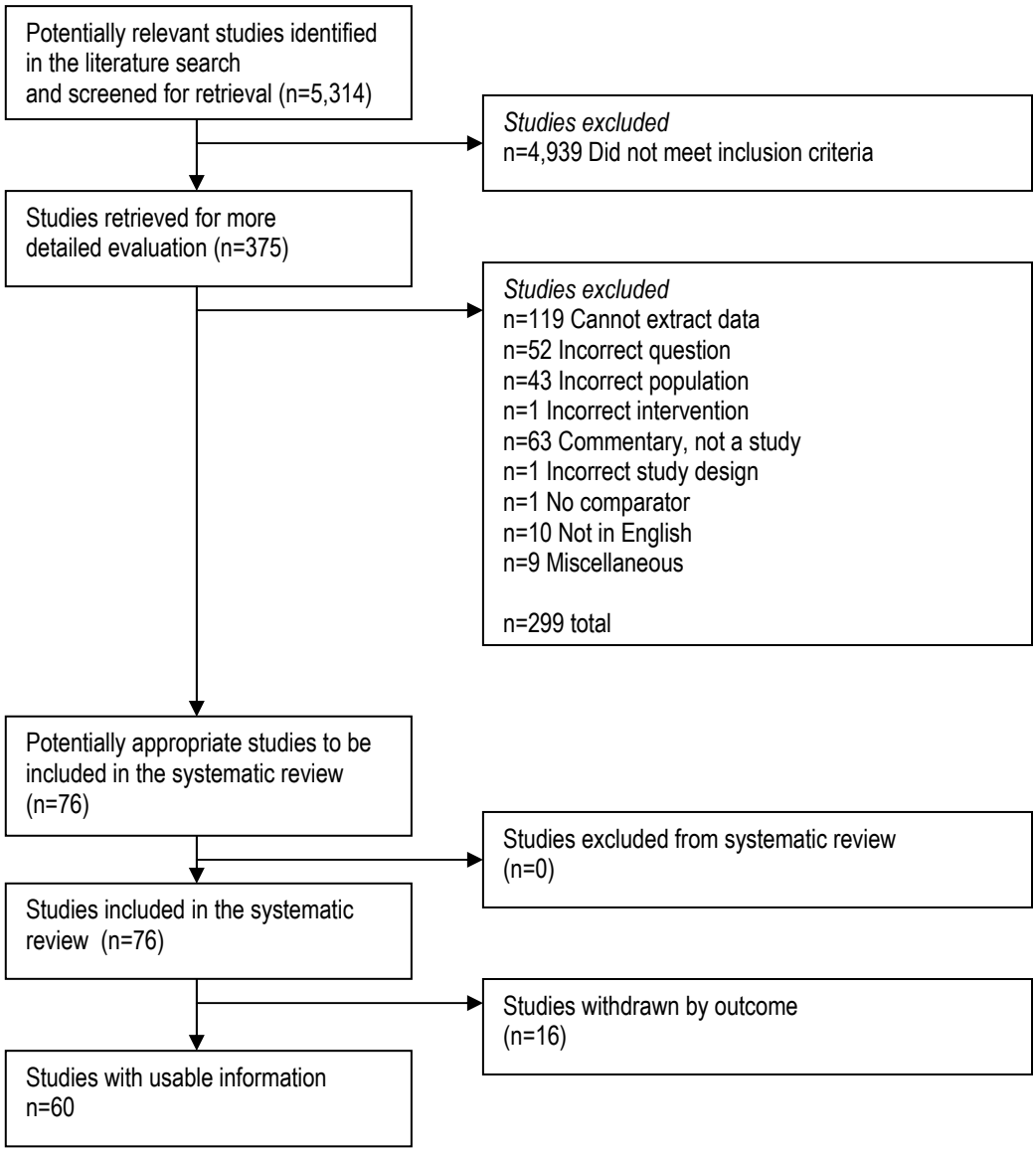
4. Inclusion criteria were independently applied to the full-text articles by two or more researchers. Those articles meeting the criteria formed part of the evidence base. The remainder provided background information;

5. The reference lists of the included articles were pearled for additional relevant studies. These were retrieved and assessed according to phase 4; and

6. The evidence base consisted of articles from phases 4 and 5 that met the inclusion criteria.

Any doubt concerning inclusions at phase 4 was resolved by group consensus. The results of this study selection process are provided in Figure 10.

Figure 10 Summary of the process used to identify and select prognostic studies for the assessment



Adapted from Moher et al (1999)

Data extraction and analysis

A study profile was developed for each included prognostic study (Appendix J). Studies that were unable to be retrieved or that met the inclusion criteria but contained insufficient or inadequate data are provided in Appendix K. Definitions of all technical terms and abbreviations are provided in the Glossary.

The ability of B-type natriuretic peptide assays to predict clinical outcomes (ie death) in HF patients was predominantly investigated via univariate analysis, followed by multivariate Cox proportional hazards analysis. Usually, only statistically significant predictors were selected to enter the multivariate model, which was typically built in a stepwise manner. Only a few studies conducted a multivariate Cox proportional hazards

regression, entering all relevant potential prognostic variables. In the context of the question asked (ie whether or not B-type natriuretic peptides added prognostic information to current clinical practice), the latter non-stepwise approach is superior. This is because all relevant variables are in a multivariate model and therefore the hazard ratio for B-type natriuretic peptide is truly adjusted for other clinical variables. In contrast, the stepwise technique removes non-contributing variables from the multivariate model, so that if all variables except B-type natriuretic peptide are non-significant the multivariate hazard ratio for B-type natriuretic peptide is not adjusted for other clinical variables. This analysis shows that B-type natriuretic peptide may be a stronger prognostic indicator than other variables, but it does not highlight the adjusted risk of having a high B-type natriuretic peptide level after adjustment of potential clinically accepted confounders (eg age, gender, New York Heart Association class, left ventricular ejection fraction). The effect estimates reported for B-type natriuretic peptide in most studies may therefore be inflated due to the lack of adjustment of all potential prognostic variables.

Effect estimates were reported in the literature as hazard or odds ratios with 95% confidence intervals. Inferential statistics were usually reported using Wald's χ^2 test. Statistical significance was assumed at $p<0.05$. For each report outcome category of: (1) death; (2) death or cardiovascular event or (3) cardiovascular event, hazard and odds ratios were tabulated and meta-analysed separately for continuous variables (ie converted to effect estimate per 1,000 pg/mL) and dichotomous variables. Studies that suited the inclusion criteria but did not report an effect estimate (because they usually only report a chi-squared statistic) were tabulated under 'insufficient information'. These studies were unable to be meta-analysed independently and therefore they were assessed narratively.

The evidence base was generally sufficient to allow a quantitative synthesis of the results. Pooled event ratios were calculated using the meta-analytic method reported by Knapp & Hartung (2003) using a restricted maximum likelihood (REML) estimate of between study variability. This method uses a t-test statistic that adjusts for the underestimation in the usual DerSimonian and Laird (1986) method.

Statistical heterogeneity in event ratios across the included studies was investigated using a Cochran Q test. As this test for heterogeneity is underpowered, the test was considered statistically significant at $p<0.1$. In the absence of heterogeneity the results from a fixed effects model were reported. When heterogeneity was present, reasons for this heterogeneity were commented on. All statistical calculations and meta-analysis were undertaken using the statistical computer package *Stata version 8.2* (Stata Corporation 2004). Egger's test (Egger et al 2003) was used to examine the possibility of publication bias. A more conservative p-value of $p<0.1$ was used again to somewhat account for the small number of studies in some outcomes.

Appraisal of the evidence

The evidence presented in the selected studies was assessed and classified using the dimensions of evidence defined by the National Health and Medical Research Council (NHMRC 2000).

These dimensions (Table 25) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from the

literature identified as informing a particular intervention. The last two require expert clinical input as part of their determination.

Table 25 Evidence dimensions

Type of evidence	Definition
Strength of the evidence	
Level	The study design used, as an indicator of the degree to which bias has been eliminated by design.[a]
Quality	The methods used by investigators to minimise bias within a study design.
Statistical precision	The *p*-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect.
Size of effect	The distance of the study estimate from the 'null' value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

[a]See Table 26

Strength of the evidence

Level

The three subdomains (level, quality and statistical precision) are collectively a measure of the strength of the evidence.

The new version of the NHMRC evidence hierarchy provides a ranking of various study designs ('levels of evidence') by the type of research question being addressed (NHMRC 2005). Table 26 provides an abbreviated version of this hierarchy detailing the ranking of studies for a prognosis research question.

Table 26 Designation of prognostic levels of evidence

Level	Prognosis
I *	A systematic review of level II studies
II	A prospective cohort study ***
III-1	All or none §§§
III-2	Analysis of prognostic factors amongst untreated control patients in a randomised controlled trial
III-3	A retrospective cohort study
IV	Case series, or cohort study of patients at different stages of disease

* A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence; *** At study inception the cohort is either non-diseased or all at the same stage of the disease; §§§ All or none of the people with the risk factor(s) experience the outcome. For example, no smallpox develops in the absence of the specific virus; and clear proof of the causal link has come from the disappearance of small pox after large-scale vaccination.

Note 1: Assessment of comparative harms/safety should occur according to the hierarchy presented for each of the research questions, with the proviso that this assessment occurs within the context of the topic being assessed. Some harms are rare and cannot feasibly be captured within randomised controlled trials; physical harms and psychological harms may need to be addressed by different study designs; harms from diagnostic testing include the likelihood of false positive and false negative results; harms from screening include the likelihood of false alarm and false reassurance results.

Note 2: When a level of evidence is attributed in the text of a document, it should also be framed according to its corresponding research question, eg level II intervention evidence, level IV diagnostic evidence, level III-2 prognostic evidence etc.

Source: NHMRC (2005)

Quality

The quality appraisal of prognostic studies (ie those concerned with the predictive ability of B-type natriuretic peptide assays at determining clinical outcomes in HF patients) was conducted using a checklist developed by the NHMRC (2000).

Study quality was presented in this assessment report both in terms of the components of quality (eg selection bias, measurement or misclassification bias) and the overall quality score.

Statistical precision

Statistical precision was determined using standard statistical principles. Small confidence intervals and p-values give an indication as to the probability that the reported effect is real (NHMRC 2000).

Size of effect

It is important to establish whether statistically significant differences are also clinically important. The size of the effect needs to be determined, as well as whether the 95% confidence interval includes only clinically important effects. Rank scoring methods were used to assess the clinically important benefit of the individual and pooled effect sizes in the available prognostic studies (NHMRC 2000).

Relevance of evidence

Similarly, the outcome being measured in the studies should be appropriate and clinically relevant. Inadequately validated (predictive) surrogate measures of a clinically relevant outcome should be avoided (NHMRC 2000). Rank scoring methods were used to determine the clinical relevance of the outcome being predicted in the prognostic studies (NHMRC 2000).

Results of assessment

Prognostic effectiveness

B-type natriuretic peptides are indicated to be prognostic predictors of death and/or rehospitalisation. In the proposed Australian setting, B-type natriuretic peptides will not replace current prognostic indicators of HF. To be of use, B-type natriuretic peptides must therefore provide additional prognostic information over and above that already provided by the battery of currently measured variables.

All included studies (n=60) followed a similar protocol, which involved: (1) recruiting a sample of clinically diagnosed HF patients; (2) measuring a variety of prognostic factors (including B-type natriuretic peptides) at baseline; (3) following patients for a period of time while monitoring for outcomes (mortality and/or cardiovascular event); (4) analysing data in such a way that enabled the prognostic potential of B-type natriuretic peptides to be isolated from other prognostic variables (ie multivariable analysis). In some cases data were analysed from existing registries, larger cohorts or randomised controlled trials on pharmaceutical interventions for HF. Studies that reported only univariate relationships between B-type natriuretic peptide and prognostic outcome were excluded as they did not answer the proposed question. Similarly, for studies that reported subset analyses in addition to overall analyses (eg Berger 2003), only results of the latter were extracted because the prognostic ability of B-type natriuretic peptide was assessed within a general HF population and not specific subsets of patients. These results are more likely to be generalisable to the HF population in Australia. The population of interest was clinically diagnosed HF, be it of systolic, diastolic, ischaemic or non-ischaemic origin. Studies that reported on the prognostic potential of B-type natriuretic peptide for myocardial infarction or acute coronary syndromes were excluded.

The report is segregated firstly into evidence of prognostic potential for NT-proBNP and BNP. The assessment of each biomarker is further subdivided into studies that reported on continuous data or dichotomous data, and those that reported insufficient information (continuous or dichotomous data). The results of the latter group cannot, by definition, be included in our primary meta-analysis to obtain an overall effect estimate. However, 'insufficient information' studies that reported non-significance and significance (but not associated hazard or odds ratios) were assigned dummy variables calculated from reported or assumed p-values and estimated standard errors. These dummy variables were introduced into a secondary meta-analysis to examine the effect, if any, of their inclusion.

The strength of the effect for studies that dichotomised variables is somewhat dependent on the cut-off point assigned. The majority of studies dichotomised about the sample median. Obviously, these cut-off points vary significantly between studies, depending on the disease severity of the cohort, whether patients had their B-type natriuretic peptide measured in a stabilised or decompensated state and the types of assays used for sample analysis. The combined effect estimate calculated for dichotomised data must therefore be seen in the context of the varying sample dependent cut-off points used. Ideal cut-off points (via ROC analysis) or 75th percentile cut-off points result in inflated effect estimates. In comparison, studies that report effect estimates on continuous data (ie per unit B-type natriuretic peptide) are less likely to be biased by arbitrary selection of cut-off values.

There is no consensus opinion as to the method of risk stratification for HF patients. Hence, a wide variety of prognostic indicators were examined and adjusted for in multivariable analyses. Stepwise multivariate analysis was the most commonly performed procedure. Stepwise analysis selects the most dominant factors for prognosis while removing weaker predictors from the model. Hence, studies often report the effect estimate for B-type natriuretic peptide adjusted for only a few other significant variables, even though many accepted prognostic variables were made available to the stepwise protocol. Only a few studies reported a hazard ratio for B-type natriuretic peptide adjusted for other accepted prognostic indicators regardless of their statistical significance.

NT-proBNP

Mortality

Summary – NT-proBNP predicting mortality

Taken together, the eight included studies reported that NT-proBNP was a significant independent predictor of all-cause mortality after the adjustment of a variety of confounding variables. The magnitude of the effect was borderline for limited continuous data but significant for the dichotomous data.

Eight studies reported on the ability of NT-proBNP to predict all-cause mortality in HF patients. Two studies reported log transformed continuous hazard and odds ratios, two studies reported dichotomised hazard and odds ratios, and four studies reported significant chi-squared statistics.

Two studies reported log hazard and odds ratios on continuous data, which were converted to our standard units by estimating the linear effect of a 1,000 pg/mL increase from their sample mean (Rossig et al 2004) and median (Kirk et al 2004). An estimated hazard ratio of one study (George et al 2005) that reported a significant chi-squared statistic using continuous data (insufficient information) was included in a secondary meta-analysis.

There is limited value in conducting a meta-analysis on the results of two studies; however, to be consistent in our approach to evidence assessment, we have combined the estimates (Table 27). A random effects model yielded a non-significant (p=0.36) point estimate of 1.39, 95%CI [0.096, 20.08] per 1,000 pg/mL increase in NT-proBNP.

On inclusion of the calculated data (Table 29; *italics*) from George et al (2005), the random effects model yielded a slightly more significant (p=0.23) effect estimate of 1.28, 95%CI [0.69, 2.38]. Egger's test for publication bias was not significant.

Table 27 Mortality—NT-proBNP; hazard or odds ratio; *continuous data*

Author Year Location Sample size	Outcome Ratio type	Significant univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
(Rossig et al 2004) Germany N=48	All-cause HR	NT-proBNP <0.001 Serum pro-apoptotic activity 0.002 MAP 0.002 Creatinine 0.002 LVEF 0.060 NYHA 0.004 Beta-blocker medication 0.022 Age 0.050 Inclusion @ p<0.005	MAP 0.041 Serum pro-apoptotic activity 0.008	per 1,000 pg/mL Stepwise Cox proportional hazards regression	1.72 0.001	1.05	1.24
(Kirk et al 2004) Denmark N=161	All-cause OR	NR Inclusion @ NR	LVEF NR NYHA NR Gender NR Age NR	per 1,000 pg/mL Multivariate logistic regression model	1.13	1.01	1.05
Combined effect estimate Random effects model (REML) Cochran Q = 98; 1 degree of freedom; p<0.001 Egger's test – n/a				per 1,000 pg/mL	**1.39** 0.36	**0.096**	**20.08**

HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); MAP = mean arterial pressure; NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction

Three studies reported dichotomous hazard and odds ratios for NT-proBNP as a prognostic indicator of all-cause mortality (Table 28). The combined fixed effect estimate was significant (2.45; 95%CI [1.08, 5.55]), with no publication bias. When data from two additional insufficient information studies using dichotomous variables (Table 29; Gardner et al 2005, Richards et al 2001) were included, the combined effect estimate increased to 2.70, 95%CI [1.76, 4.13] and was highly statistically significant (p=0.003). As before, there was no significant publication bias.

Table 28 Mortality—NT-proBNP; hazard or odds ratio; *dichotomised data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Hartmann et al 2004) Europe N=1011	All-cause HR	NT-proBNP <0.001 Age 0.006 SBP <0.001 HF aetiology 0.05 Creatinine clearance 0.0001 Recent hospitalisation 0.01 High risk combination 0.005 Treatment group (placebo/carvedilol) 0.04 All relevant variables included	Treatment group (placebo/carvedilol) 0.006 Age 0.021 HF aetiology 0.201 SBP <0.001 Recent hospitalisation 0.019 High risk combination 0.043	>1,767 pg/mL Median Cox proportional hazards regression	2.17 0.002	1.33	3.54
(Fisher et al 2003) UK N=87	All-cause OR	NT-proBNP NR Age NR Gender NR NYHA NR LVEF NR Comorbidity NR History of HF NR Creatinine NR Inclusion @ p<0.05	NR	>2,994 pg/mL Median Stepwise multivariate analysis	2.22 0.03	1.08	4.56
(Kellett 2005) Ireland N=342	In-hospital mortality OR Dichot	SBP <0.0001 Urea <0.0001 NT-proBNP <0.0001 Age <0.004 O_2 saturation <0.004 MEW score <0.002 Leukocytosis <0.02 Respiratory rate <0.02 Hb <0.03 Abnormal ECG <0.03 Abnormal chest X-ray NS Disease history NS Gender NS Inclusion @ NR	SBP <0.0002 Urea level <0.011 Leukocytosis <0.008	>11,500 pg/mL Optimal cut-off point Logistic regression	4.62 <0.002	1.77	12.07
Combined effect estimate Fixed effects model Cochran Q = 2; 1 degree of freedom; p=0.37 Egger's test = 2.54; p=0.36				NA	**2.45** 0.04	**1.08**	**5.55**

HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); SBP = systolic blood pressure; HF = heart failure; Hb = haemoglobin; ECG = electrocardiogram; NA = not applicable; NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction

Table 29 Mortality—NT-proBNP; insufficient information; *dichotomised and continuous data*

Author Year Location Sample size	Outcome Ratio type Data type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units/cut-off point Position Analysis type *Assumptions*	Effect estimate Reported *Estimated*	95%CI Low	95%CI High
(Gardner et al 2005) UK N=182	All-cause HR Dichot	NT-proBNP NR Peak VO_2 NR Sodium NR Creatinine NR HFSS NR Haemoglobin NR Inclusion @ p<0.10	None	>1,505 pg/mL Median Stepwise Cox proportional hazards regression analysis *Assume SE is the median of known SEs (SE = 0.37)*	χ^2 = 14.2 <0.001 *4.03*	NR *1.95*	NR *8.33*
(Richards et al 2001) Multicentre Australia and New Zealand N=297	All-cause HR Dichot	NT-proBNP 0.00005 ADM 0.0002 LVEF 0.014 Treatment group (placebo/carvedilol) 0.02 Inclusion @ NR	ADM <0.05 Treatment group (placebo/carvedilol) NR NYHA NR LVEF NR Prior MI NR Age NR Prior hospital admission NR	>837 pg/mL Median Cox proportional hazards regression *Assume SE is the median of known SEs (SE = 0.37) and p-value of 0.01*	NR <0.05 *2.59*	NR *1.26*	NR *5.36*
(George et al 2005) Israel N=188	All-cause HR Contin	Age <0.05 Gender <0.05 NYHA <0.05 LVEF <0.05 Creatinine <0.05 Haemoglobin <0.05 Erythropoietin <0.05 C-reactive protein <0.05 NT-proBNP <0.05 Inclusion @ p<0.1	LVEF 0.04 Haemoglobin 0.003 Erythropoietin 0.02	per 1-pg/mL NA Stepwise Cox proportional hazards regression analysis *Assume SE is the median of known SEs (SE=0.000026)*	χ^2 = 13.6 0.001 *1.10*	NR *1.05*	NR *1.16*

HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); ADM = Adrenomedullin; Peak VO_2 = peak oxygen consumption; LVEF = left ventricular ejection fraction; MI = myocardial infarction; HFSS = heart failure survival score (weighted combination of ejection fraction, heart rate, sodium, mean arterial pressure, peak VO_2, conduction delay, and coronary artery disease); NA = not applicable; NYHA = New York Heart Association classification; SE = standard error; *Italics* = the assumptions made to approximate hazard ratios and confidence intervals for inclusion into secondary meta-analysis (all values are converted to 1,000 pg/mL).

Mortality or cardiovascular event

Summary – NT-proBNP predicting death or cardiovascular event

Based on a small evidence base (n=5 studies) and a borderline level of statistical significance, it appears that NT-proBNP has potential as a prognostic indicator for death or cardiovascular event.

Five studies formed the evidence base for the prognostic ability of NT-proBNP to predict mortality or cardiovascular events (predominantly rehospitalisation). All studies reported that NT-proBNP significantly and independently predicted this outcome. One and two studies reported dichotomous hazard and odds ratios, respectively. Two studies reported only chi-squared statistics and therefore could not be included in the primary meta-analysis. No studies reported sufficient information on the prognostic ability of NT-proBNP tests (continuous data) for this outcome.

Three studies, two of which reported dichotomous odds ratios and the other a hazard ratio, all reported NT-proBNP as a significant independent predictor of mortality or cardiovascular event (Table 30). However, due to the limited sample size for the meta-analysis the combined effect estimate (3.12) was associated with only borderline significance (p=0.11).

Table 30 Mortality or cardiovascular event—NT-proBNP; hazard or odds ratio; *dichotomised data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Hartmann et al 2004) Europe N=1011	All-cause death or hospital-isation for HF HR	NT-proBNP <0.001 Age 0.02 LVEF 0.003 SBP 0.03 HF aetiology 0.008 Creatinine clearance <0.001 Recent hospitalisation <0.001 High-risk combination <0.001 Treatment group (placebo/carvedilol) 0.098 All relevant variables included	Treatment group (placebo/carvedilol) 0.016 Age 0.065 HF aetiology 0.087 SBP 0.008 Recent hospitalisation <0.001 High risk combination 0.02	>1,767 pg/mL Median Cox proportional hazards regression analysis	2.11 0.0001	1.54	2.9
(Fisher et al 2003) UK N=87	All-cause death or readmissi on OR	NT-proBNP NR Age NR Gender NR NYHA NR LVEF NR Comorbidity NR History of HF NR Creatinine NR Inclusion @ p<0.05	NR	>2,994 pg/mL Median Stepwise multivariate analysis	4.15 0.003	1.62	10.62
(O'Brien et al 2003) UK N=34	Death or readmissi on or outpatient worsening of HF OR	NT-proBNP$_{(DIS)}$ NR Age NR Killip class NR History of HF NR Creatinine NR Inclusion @ p<0.10	NT-proBNP$_{(ADM)}$ 0.082 Age NS Killip class NS History of HF NS Creatinine NS	>1,664 pg/mL Median Binary logistic regression analysis	15.30 0.026	1.4	168.9
Combined effect estimate Fixed effects model Cochran Q = 4; 2 degrees of freedom; p=0.12 Egger's test = 1.96; p=0.04				NA	**3.12** 0.11	**0.54**	**18.15**

Position = where the cut-off point was located within the sample distribution; HF = heart failure; HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); SBP = systolic blood pressure; LVEF = left ventricular ejection fraction; (DIS) = discharge value, (ADM) = admission value; NA = not applicable; NYHA = New York Heart Association classification

The inclusion of the study by Gardner et al (2005), which reported NT-proBNP as a significant independent predictor of mortality or cardiovascular event but did not report an effect estimate, resulted in an increase in the pooled estimate to 4.5, 95%CI [1.14, 17.99] and was statistically significant (p=0.04). The inclusion of Gardner's results (an estimated effect estimate of 9.4; Table 31) also caused the random effects model to be chosen for the calculation of the combined effect estimate (Cochran Q = 11.9; p=0.008). Zugck et al (2002) were the only group to report on continuous data; however,

insufficient information was reported and therefore data could not be integrated into the qualitative assessment of effect.

Table 31 Mortality or cardiovascular event—NT-proBNP; insufficient information; *dichotomised and continuous data*

Author **Year** **Location** **Sample size**	**Outcome** **Ratio type** **Data type**	**Univariate predictors of outcome (p≤0.05)** **Inclusion p-value**	**Confounding factors controlled for in multivariate analysis**	**Units/cut-off point** **Position** **Analysis type** ***Assumptions***	**Effect estimate** **Reported** ***Estimated***	**95%CI**	
						Low	**High**
(Gardner et al 2005) UK N=182	All-cause death or urgent CTx HR Dichot	NT-proBNP NR Peak VO_2 NR Sodium NR Creatinine NR HFSS NR Hb NS Inclusion @ p<0.10	None	>1,505 pg/mL Median Stepwise Cox proportional hazards regression analysis *Assume SE is median of known SEs (SE=0.48)*	χ^2 = 21.8 <0.001 *9.40*	NR *3.67*	NR *24.09*
(Zugck et al 2002) Germany N=408	Cardiac death or hospital-isation for heart failure HR Cont	NT-proBNP 0.0001 Peak VO_2 0.0001 LVEF 0.0001 Treatment group (placebo/beta-blocker) 0.0002 Norepinephrine 0.0017 Inclusion @ NR	Peak VO_2 0.0005 LVEF 0.0021	per 1 pg/mL NA Stepwise Cox proportional hazards regression analysis	χ^2 = 8.1 0.005	NR	NR

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); CTx = cardiac transplantation; NA = not applicable; Peak VO_2 = peak oxygen consumption; LVEF = left ventricular ejection fraction; HFSS = heart failure survival score (weighted combination of ejection fraction, heart rate, sodium, mean arterial pressure, peak VO_2, conduction delay, and coronary artery disease); Hb = haemoglobin; SE = standard error; *Italics* = the assumptions made to approximate hazard ratios and confidence intervals for inclusion into secondary meta-analysis.

Cardiovascular event

Summary – NT-proBNP predicting cardiovascular events

Two out of three studies reported that NT-proBNP was a significant independent predictor of either HF mortality or hospitalisation. The remaining study, which did not report significance, adjusted for a variable that is not part of standard clinical care. Hence, qualitatively there is evidence to suggest that NT-proBNP would be useful for predicting future cardiovascular events in HF patients.

The prognostic ability of NT-proBNP for cardiovascular events (hospitalisation due to HF) was poorly represented, with only three studies reporting this outcome (Table 32). Two studies found that NT-proBNP was a significant independent predictor of cardiovascular hospitalisation after adjustment for various other prognostic indicators. The other study reported that, despite NT-proBNP being a strong univariate predictor of HF hospitalisation, it was not significant after adjustment for plasma surfactant protein-B.

No studies reported adjusted effect estimates; therefore, a quantitative assessment of this outcome was not possible.

Table 32 Cardiovascular event—NT-proBNP; insufficient information; *dichotomised and continuous data*

Author Year Location Sample size	Outcome Ratio type	Significant univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units/cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Richards et al 2001) Multicentre Australia and New Zealand N=297	HF mortality HR Dichot	NT-proBNP 0.00005 ADM 0.0002 LVEF 0.014 Treatment group (placebo/carvedilol) 0.02 Inclusion @ NR	ADM <0.05 Treatment group (placebo/carvedilol) NR NYHA NR LVEF NR Prior MI NR Age NR Prior hospital admission NR	>837 pg/mL Median Cox proportional hazards regression	NR <0.05	NR	NR
(George et al 2005) Israel N=188	HF hospital-isation HR Contin	Age <0.05 Gender <0.05 NYHA <0.05 LVEF <0.05 Creatinine <0.05 Hb <0.05 Erythropoietin <0.05 C-reactive protein <0.05 NT-proBNP <0.05 Include @ p<0.1	NYHA 0.01 Haemoglobin 0.001 Erythropoietin 0.003	NT-proBNP NA Stepwise Cox proportional hazards regression analysis	χ^2 = 11.2 <0.001	NR	NR
(De Pasquale et al 2004) Australia N=53	Hospital-isation for HF OR Cont	Dyspnoea score NR Left ventricular failure score NR 6-MWT NR SP-B NR NT-proBNP NR Inclusion @ NR	SP-B 0.005	ln NT-proBNP Conditional logistic regression	NR 0.24	NR	NR

Position = where the cut-off point was located within the sample distribution; HF = heart failure; HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); NA = not applicable; ADM = adrenomedullin; LVEF = left ventricular ejection fraction; MI = myocardial infarction; 6-MWT = 6-minute walk test; SP-B = plasma surfactant protein B; Hb = haemoglobin.

BNP

Mortality

Summary – BNP predicting mortality

An evidence base of nine studies appears to indicate that there is a trend for BNP providing independent prognostic information over and above other clinical, echographic, demographic and medication use variables. The significance of combined estimates is robust to the inclusion of studies that reported insufficient information and no effect.

Of nine studies that reported on the prognostic potential of BNP for all-cause mortality, six showed that BNP was an independent prognostic factor in multivariate analyses (Table 33). Two of the three studies that reported non-significant results were categorised to 'insufficient information' due to the removal of non-significant predictors in the stepwise analysis model, and were therefore not reported. The remaining study, which reported a non-significant hazard ratio, did not use a stepwise selection procedure.

Table 33 Mortality—BNP; hazard or odds ratio; *continuous data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
(Bettencourt 2000) Portugal N=139	All-cause HR	BNP <0.0001 MAP 0.03 NYHA 0.03 HF aetiology 0.01 LVEF 0.001 ECG abnormality[a] 0.0009 Atrial fibrillation 0.004 6-MWT <0.0001 Na^+ 0.02 Uric acid 0.004 Inclusion @ p<0.05	6-MWT 0.0001 ECG abnormality[a] 0.01 Atrial fibrillation 0.01 HF aetiology 0.02	per 1,000 pg/mL Stepwise Cox proportional hazards regression analysis	1.105 0.002	1.037	1.178
(Latini et al 2004) Multicentre 302 centres; 16 countries N=4305 Val-HeFT	All-cause HR	NR Inclusion @ NR	Aldosterone 0.37 Norepinephrine 0.02 Renin activity 0.01 Demographic and clinical/echographic variables NR	per 1,000 pg/mL Cox proportional hazards regression	3.297 <0.0001	2.705	4.016
(Shiba et al 2004) Japan Multicentre – 26 hospitals in Tohoku region N=684	All-cause HR	Age <0.001 HF aetiology <0.001 History of admission 0.01 Diabetes 0.001 Tachycardia 0.01 NYHA <0.001 LVEF 0.001 LVDD 0.001 BNP <0.001 ACE-I/ARB use 0.035 Inclusion @ NR Did not include variables which correlated >0.7	Age 0.004 Diabetes 0.003 Tachycardia 0.021 NYHA 0.001 Rural resident 0.008	per 1,000 pg/mL Stepwise Cox proportional hazards regression	2.717 <0.001	1.492	3.002
(Wijeysundera et al 2003) US and Canada Multicentre – 26 centres PRAISE-2 N=181	All-cause HR	Age 0.029 Gender NS Creatinine 0.028 NYHA 0.011 LVEF 0.073 NE 0.035 Dopamine 0.016 NT-proANP <0.001 BNP <0.001 Inclusion @ p<0.05	Age 0.432 Creatinine 0.659 NYHA 0.151 NE 0.350 Dopamine 0.562 NT-proANP 0.001	per 1,000 pg/mL Cox proportional hazards regression	0.990 0.17	0.969	1.004

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
(Maisel et al 2004) US Multicentre – 10 centres N=464	All-cause OR	NR Include @ NR	NYHA 0.648 Initial disposition [b] 0.889 Actual disposition 0.735	per 1,000 pg/mL Logistic regression	3.619[c] 0.001	1.817	7.206
Combined effect estimate Random effects model (REML) Cochran Q = 195; 4 degrees of freedom; p<0.001 Egger's test – Intercept = 6.96; p=0.05				per 1,000 pg/mL	**1.96**	**0.90**	**4.28**

HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or NS = not significant or NR = not reported); ADM = adrenomedullin; MAP = mean arterial pressure; LVEF = left ventricular ejection fraction; ECG = electrocardiogram; 6-MWT = 6-minute walk test; NE = norepinephrine; NYHA = New York Heart Association classification; LVDD = left ventricular diastolic diameter; ACE-I/ARB = Angiotensin-converting enzyme inhibitors/Angiotensin II receptor blockers; [a] E wave deceleration time representative of restrictive filling of the left ventricle; [b] Physician opinion, based on estimated NYHA class and whether or not they believed the patient would ultimately be hospitalised; [c] Converted from log BNP odds ratio, only approximate

Four studies reported adjusted hazard ratios on continuous data for the outcome of mortality (Table 33). One study reported an odds ratio per log BNP concentration, which exponentiated to an approximate odds ratio of 3.6 for every 1,000 pg/mL increase in BNP from the study's median concentration of 764 pg/mL. The meta-analysis was run with and without the odds ratio of Maisel et al (2004) due to the uncertainty associated with the odds ratio conversion.

Inclusive of the converted effect estimate of Maisel et al (2004), the meta-analysis was associated with a highly significant heterogeneity test. Hence, using a random intercept model, the overall effect estimate was 1.96 per 1,000 pg/mL increase in BNP (p=0.07) for all-cause mortality. Egger's test demonstrated borderline significance (p=0.05), suggesting the possibility for publication bias; however, bias was in the opposite direction to what would be expected. The results were similar when we excluded the data of Maisel et al (2004), with the random intercept model (to account for heterogeneity) resulting in an overall effect estimate of 1.74, 95%CI [0.66, 4.66]. Egger's test for publication bias was not significant (p=0.09).

To capture the most conservative overall effect estimate for BNP on all-cause mortality, the meta-analysis was re-run with an assumed null effect (HR = 1.00; 95%CI [0.99, 1.01]; assuming standard error was the same as unadjusted) for one study (van der Meer et al 2004) that reported no effect (Table 35, continuous data). This study did not report sufficient information to be included in the original meta-analysis. The inclusion of null effect reduced the overall effect estimate to 1.74, 95%CI [0.91, 3.34] (p=0.08).

Table 34 Mortality—BNP; hazard or odds ratio; *dichotomised data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Anand et al 2003) Multicentre 302 centres; 16 countries N=4305 Val-HeFT	All-cause HR	NR Inclusion – all	NYHA NR LVEF NR ACE-I (baseline) NR beta-blocker (baseline) NR HF aetiology NR Age NR	>97 pg/mL Median Cox proportional hazards regression	2.1 NR	1.79	2.42
(Vrtovec et al 2003) US Texas N=241	All-cause HR	Prolonged QTc interval <0.0001 BNP 0.0001 QRS duration 0.01 Digoxin 0.01 Age NS Gender NS HF aetiology NS NYHA NS LVEF NS Inotropes, diuretics, ACE inhibitors and beta-blockers NS Inclusion @ p<0.05	Prolonged QTc interval 0.0001	>1,000 pg/mL 75th percentile Stepwise Cox proportional hazards regression	1.99 0.0005	1.18	3.36
Combined effect estimate Fixed effects model (REML) Cochran Q = 0.04; 1 degree of freedom; p=0.846 Egger's test – NA				NA	**2.09**	**0.82**	**5.35**

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); LVEF = left ventricular ejection fraction; NYHA = New York Heart Association classification; ACE-I = Angiotensin-converting enzyme inhibitors; NA = not applicable

Two studies reported adjusted dichotomous hazard ratios for BNP on all-cause death. A meta-analysis of the two studies provides limited combined estimate information (Table 34). Nevertheless, a fixed effects REML meta-analysis resulted in a combined effect estimate of 2.09, 95%CI [0.82, 5.35] (p=0.06). Inclusion of the study by Watanabe et al (2005), which reported insufficient information on dichotomised data (Table 35), and a re-analysis resulted in a non-significant combined effect estimate of 1.74; 95%CI [0.68, 4.42] (p=0.13).

Table 35 Mortality—BNP; insufficient information; *dichotomised and continuous data*

Author Year Location Sample size	Outcome Ratio type Data type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units/cut-off point Position Analysis type *Assumptions*	Effect estimate Reported *Estimated*	95%CI	
						Low	High
(van der Meer et al 2004) Netherlands N=74	All-cause HR Cont	BNP 0.001 Erythropoietin 0.002 Hb 0.003 LVEF 0.059 GFR 0.096 Age 0.11 Inclusion @ p<0.05	Erythropoietin 0.03 Hb 0.005 Age 0.06	per 1 pg/mL NA Stepwise Cox proportional hazards regression *Assume SE of unadjusted analysis and null effect*	NR NS *1.00*	NR *0.99*	NR *1.01*
(Watanabe et al 2005) Multicentre, Tohoku district N=417 CHART study	All-cause death HR Dichot	Age NR NYHA NR HF history NR Diabetes NR LVDD NR ACE-I/ARB NR beta-blocker NR Gender NR HF aetiology NR LVEF NR Inclusion @ NR	Age 0.005 NYHA 0.003 Non-sustained ventricular tachycardia 0.0010 LVDD 0.0002 Diabetes <0.0001 ACE-I/ARB 0.0112	>132 pg/mL Median Stepwise Cox proportional hazards regression *Assume SE of the death or CV event outcome and null effect*	NR NS	NR	NR

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); NA = not applicable; LVEF = left ventricular ejection fraction; NYHA = New York Heart Association classification; ACE-I = Angiotensin-converting enzyme inhibitors; ARB = Angiotensin II receptor blockers; LVDD = left ventricular end diastolic volume; GFR = glomerular filtration rate; Hb = haemoglobin; SE = standard error; CV = cardiovascular SE = standard error; *Italics* = the assumptions made to approximate hazard ratios and confidence intervals for inclusion into secondary meta-analysis.

Mortality or cardiovascular event

Summary – BNP predicting death or cardiovascular event

Twenty-five studies (13 reporting continuous data, 6 with continuous data but insufficient information, 5 with dichotomous data, and 1 with dichotomous data but insufficient information) reported on the ability of BNP to predict death or cardiovascular event. Only three reported a non-significant adjusted relationship, two of which could possibly be explained. The level of consistency between studies, in addition to a statistically significant range of effect estimates (2.9 to 8.1) using the relatively conservative REML meta-analytic methodology, infers that BNP predicts mortality or cardiovascular events in HF patients independent of several demographic, risk factor, clinical and echographic variables.

Twenty-five studies formed the evidence base for the prognostic ability of BNP for mortality or cardiovascular event. Only three studies reported a non-significant association between BNP and outcome, two of which reported non-significant hazard ratios and one which had insufficient information.

Thirteen studies reported adjusted hazard or odds ratios that were appropriate for a meta-analysis (Table 36). Three studies had logarithmic units converted to a linear scale by assuming an increase of 1,000 pg/mL from their reported sample medians (Tamura et al 2001; Maisel et al 2004) and means (Ishii et al 2003). The meta-analysis was run with and without these converted results.

With all studies included the overall effect estimate was a significant 4.6 per 1,000 pg/mL increase in BNP ($p=0.015$), within a 95% confidence interval of 1.4 to 14.1. The point effect estimate increased to 8.1, 95%CI [1.4, 48.2] ($p=0.03$) when the log converted results were excluded from the analysis. Due to a highly significant heterogeneity test, all aforementioned values were calculated using random effects meta-analytic models. In both analyses Egger's test for publication bias was significant.

Table 36 Mortality or cardiovascular event—BNP; hazard or odds ratio; *continuous data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
(Maeda et al 2000) Japan N=102	Cardiac death or hospitalisation HR	Age NS Gender NS HF aetiology NS NYHA ADM NS NYHA STAB NS LVEF ADM NS LVEF STAB 0.0005 ANP ADM NS ANP STAB <0.0001 BNP ADM NS BNP STAB <0.0001 NE ADM NS NE STAB 0.0072 ET-1 ADM NS ET-1 STAB 0.0303 IL-6 ADM NS IL-6 STAB 0.0043 TNF-α ADM NS TNF-α STAB 0.0163 Inclusion – all	IL-6 STAB 0.0059	per 1,000 pg/mL Stepwise Cox proportional hazards regression	2.72 <0.0001	2.72	7.37
(Imamura et al 2001) Japan Multicentre; 6 institutions N=171 Ehime MIBG HF study	Cardiac death or rehospitalisation HR	Age NS Gender NS HF aetiology NS NYHA <0.0001 Cardiothoracic ratio <0.001 LVEF NS NE <0.0001 ANP <0.0001 BNP <0.0001 MIBG data <0.029 Inclusion @ p<0.05	MIBG washout <0.0001	per 1,000 pg/mL Stepwise Cox proportional hazards regression	147 <0.0001	16.03	1340
(Hamada et al 2005) Japan N=52	Cardiac death or rehospitalisation HR	NR Inclusion – all	Doppler echo-cardiographic variables (ADM and DIS) NS Haemodynamic variables (ADM and DIS) NS BNP (ADM) 0.314	per 1,000 pg/mL Cox proportional hazards regression	147 0.086 DIS	3.55	6059
(Latini et al 2004) Multicentre; 302 centres; 16 countries	All-cause death or morbid event HR	Aldosterone NR Norepinephrine NR Renin activity NR Demographic and	Aldosterone 0.13 Norepinephrine 0.01 Renin activity 0.003 Demographic and	per 1,000 pg/mL Cox proportional hazards regression	3.30 <0.0001	2.71	3.64

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
N=4305 Val-HeFT		clinical/echographic variables NR Inclusion – all	clinical/echographic variables NR				
(Logeart et al 2004) France N=105	All-cause death or hospitalisation HR	Age NS Gender NS HF aetiology NS Diabetes NS AF NS LVEF NS Inotropic use 0.007 Echocardiographic abnormality <0.005 BNP (ADM) 0.0001 BNP (DIS) 0.0001 %ΔBNP 0.0001 Include variables associated with outcome or known predictors of mortality	Doppler echocardiography variables NR Age NR Diabetes NR LVEF NR Inotropic drug use NR %ΔBNP NR BNP (ADM)	per 1,000 pg/mL Cox proportional hazards regression	3.70 0.027 DIS	1.22	11.81
(Sakatani et al 2004) Japan N=70	Cardiac death or hospitalisation HR	NR Inclusion @ NR	Hypertension 0.786 Smoking 0.138 NYHA 0.162 NE 0.084 Dopamine 0.101 ANP 0.188 % plasma lymphocytes 0.001	per 1,000 pg/mL Cox proportional hazards regression	2.6^{E12} 0.213	9.9^{E-8}	2.6^{E31}
(Setsuta et al 2002) Japan N=56	Cardiac death or hospitalisation HR	Age 0.85 Gender 0.09 NYHA 0.0001 cTnT 0.004 Fatty acid binding protein 0.0001 ANP 0.0001 BNP 0.0001 NE 0.006 Cardiothoracic ratio 0.10 LVEF 0.002 Inclusion – all	Age 0.68 Gender 0.02 NYHA 0.06 cTnT 0.04 Fatty acid binding protein 0.04 ANP 0.99 NE 0.12 Cardiothoracic ratio 0.13 LVEF 0.02	per 1,000 pg/mL Cox proportional hazards regression	0.81 0.93	0.09	12.13
(Ishii et al 2003) Japan N=100	Cardiac death or hospitalisation HR	Age 0.032 Gender NS HF aetiology 0.054 NYHA 0.0002 LVEF 0.008 Creatinine 0.0013	LVEF 0.070 cTnT 0.0001	per 1,000 pg/mL Stepwise Cox proportional hazards regression	1.20 0.0005	1.08	1.34

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
		cTnT <0.0001 logBNP <0.0001 Inclusion @ NR					
(Tamura et al 2001) Japan N=48	Cardiac death or hospitalisation HR	Left ventricular mass index P<0.05 NYHA P<0.05 LVEF P<0.05 logBNP P<0.05 Inclusion @ p<0.05	Left ventricular mass 0.816 LVEF 0.224 NYHA 0.560	per 1,000 pg/mL Cox proportional hazards regression	8.16 0.015	1.50	44.28
(Koseki 2003) Japan N=194	All-cause death or hospitalisation OR	**DCM subgroup** NYHA NR Age NR LVDD NR LVEF NR BNP NR Diabetes NR History of HF NR Medication NR Inclusion – all	NYHA NS Age NS LVDD p<0.05 LVEF NS Diabetes NS	per 1,000 pg/mL Multivariate logistic regression	3.8E5 <0.05	2.19	6.4E10
(Koseki 2003) N=163	"	**MI subgroup** as above	NYHA NS Age NS LVDD NS LVEF NS Diabetes p<0.05	per 1,000 pg/mL Multivariate logistic regression	1.3E6 <0.05	1.22	1.3E12
(Koseki 2003) N=245	"	**VHD subgroup** as above	NYHA NS Age NS LVDD NS LVEF NS Diabetes p<0.05	per 1,000 pg/mL Multivariate logistic regression	4.1E4 <0.05	1.49	1.2E9
(Koseki 2003) N=100	"	**LVH subgroup** as above	NYHA NS Age NS LVDD NS LVEF NS Diabetes NS	per 1,000 pg/mL Multivariate logistic regression	3.7E8 <0.05	1.81	7.1E16
(Weinberg et al 2003) Multicentre; 26 centres PRAISE-2 N=161	Cardiac death or transplantation OR	ΔST-2 0.048 BNP <0.0001 ΔBNP 0.34 NT-proANP <0.0001 NE 0.056 Dopamine 0.043 Epinephrine 0.66 Angiotensin II 0.79 Creatinine 0.053	ΔST-2 0.039	per 1,000 pg/mL Stepwise multiple variable logistic regression	1.03 0.0001	1.02	1.05

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
		LVEF 0.09 Gender 0.61 Age 0.01 Race 0.08 Inclusion – potential outcome predictors (p-value NR)					
(Maisel et al 2004) US Multicentre; 10 centres N=464	All-cause death, hospital admission or ED visit OR	NR Inclusion @ NR	NYHA 0.507 Initial disposition 0.130 Actual disposition 0.229	per 1,000 pg/mL Logistic regression	1.81	1.19	2.74
(Gackowski et al 2004) France N=95	Death, rehospitalisation or urgent CTx HR	BNP ADM and DIS <0.08 Clinical score ADM and DIS <0.002 Previous HF treatment 0.0002 Dobutamine infusion <0.0001 Right ventricle systolic pressure >50 or restrictive ADM and DIS <0.06 LVEF 0.029 Inclusion – all	Previous CHF treatment 0.0012 Dobutamine infusion 0.0001 Right ventricle systolic pressure >50 or restrictive 0.0026	per 1,000 pg/mL Stepwise Cox regression analysis	9.88 0.0026	2.20	50.20
Combined effect estimate Random effects model (REML) Cochran Q = 320.2; 15 degrees of freedom; p<0.001 Egger's test – p=0.005				per 1,000 pg/mL	**4.46 0.015**	**1.41**	**14.13**

HR = hazard ratio; OR = odds ratio; E = exponent; Superscript text (author p-values or NS = not significant or NR = not reported); LVEF = left ventricular ejection fraction; NYHA = New York Heart Association classification; ACE-I = Angiotensin-converting enzyme inhibitors; ARB = Angiotensin II receptor blockers; GFR = glomerular filtration rate; Hb = haemoglobin; LVDD = left ventricular diastolic diameter; DCM = dilated cardiomyopathy; MI = myocardial infarction; VHD = valvular heart disease; LVH = left ventricular hypertrophy; ET-1 = Endothelin-1; IL-6 = Interleukin-6; TNF-α = tumour necrosis factor alpha; NE = norepinephrine; %Δ = per cent change; ADM = admission; DIS = discharge; AF = atrial fibrillation; COPD = chronic obstructive pulmonary disease; CTx = cardiac transplantation; ST-2 = a stress protein, specifically an interleukin-1 receptor family member, that is induced by mechanical strain in cardiac myocytes; cTnT = cardiac troponin T; MIBG = ^{123}I-metaiodobenzylguanidine

Of the six studies that reported insufficient information (continuous data) to be included in a secondary meta-analysis, five reported that BNP was a significant independent predictor of mortality or cardiovascular event. In order to include all studies several assumptions were made. P-values (assumed or reported) along with standard errors (from relevant sources) were used to calculate a test statistic, adjusted point effect estimate and associated 95% confidence interval. These values, together with the assumptions made to calculate them, are contained in Table 37 in *italics*. The meta-analysis with all possible studies included resulted in a significant pooled point effect estimate of 6.4, 95%CI [2.63, 15.59]. Egger's test for publication bias remained significant.

Table 37 Mortality or cardiovascular event—BNP; insufficient information; *dichotomised and continuous data*

Author **Year** **Location** **Sample size**	**Outcome** **Ratio type** **Data type**	**Univariate predictors of outcome (p≤0.05)** **Inclusion p-value**	**Confounding factors controlled for in multivariate analysis**	**Units/cut-off point** **Position** **Analysis type** ***Assumptions***	**Effect estimate**	**95%CI**	
						Low	**High**
(Koglin 2001) Germany N=78	Cardiac death or deterioration of physical activity Cont	HFSS $^{<0.0001}$ BNP $^{<0.05}$ Inclusion – all	HFSS $^{<0.05}$	per 1 pg/mL NA Proportional hazards regression model *Assume SE of unadjusted analysis*	NR $^{0.748}$ *1.39*	NR	NR
(Matsui et al 2002) Japan N=74	Cardiac death or readmission Cont	NYHA (6 months) $^{<0.05}$ LVEF (6 months) $^{<0.05}$ LVEF (% change) $^{<0.05}$ BNP $^{ADM <0.05}$ BNP (6 months) $^{<0.05}$ NE ADM $^{<0.05}$ NE (6 months) $^{<0.05}$ MIBG variables $^{<0.05}$ All above variables were measured for admission, 6 months and % change Age ^{NS} Gender ^{NS} Inclusion – all	Δ delayed HM $^{<0.05}$	per 1 pg/mL NA Stepwise Cox proportional hazards regression A*ssume median of known SE (SE = 0.001) and p=0.01*	BNP (6 months) $^{<0.05}$ *13.14*	NR *1.85*	NR *93.3*
(Isnard et al 2003) France N=250	Death or urgent CTx Cont	NYHA $^{<0.0001}$ Resting heart rate $^{<0.0001}$ SBP $^{0.0004}$ LVEF $^{0.0003}$ LVDD $^{0.0002}$ Na^+ $^{0.002}$ VO_2 peak $^{<0.0001}$ % of predicted VO_2 peak $^{<0.0001}$ ANP $^{<0.0001}$ NT-proANP $^{<0.0001}$ BNP $^{<0.0001}$ NE $^{0.003}$ ET-1 $^{<0.0001}$ Inclusion – all	Resting heart rate $^{0.006}$ Na^+ $^{0.005}$ BNP $^{0.0006}$ ET-1 $^{0.007}$ Peak VO_2 $^{0.012}$	per 1 pg/mL NA Stepwise Cox proportional hazards regression A*ssume median of known SE (SE = 0.001)*	$\chi^2 = 11.9$ $^{0.0006}$ *30.93*	NR *4.4*	NR *220*

Author **Year** **Location** **Sample size**	**Outcome** **Ratio type** **Data type**	**Univariate predictors of outcome (p≤0.05)** **Inclusion p-value**	**Confounding factors controlled for in multivariate analysis**	**Units/cut-off point** **Position** **Analysis type** ***Assumptions***	**Effect estimate**	**95%CI**	
						Low	**High**
(Tsutamoto et al 1999) Japan N=290	Cardiac death or readmission HR Cont	Age NS Gender NS NYHA 0.008 HF aetiology NS LVEF <0.0001 Right heart catheterisation variables <0.05 NE 0.004 ET-1 NS Angiotensin II 0.0002 ANP <0.0001 BNP <0.0001 Treatment variables NS Inclusion – all	LVEDP 0.03 LVEDVI 0.004 NE 0.04 ACE-I 0.013	per 1 pg/mL NA Stepwise Cox proportional hazards regression *Assume SE of the CV death outcome*	χ^2 = 23.83 <0.0001 *41.2*	NR *9.3*	NR *184*
(Tsutsui et al 2002) Japan N=84	Cardiac death or rehospitalisation Cont	NR Inclusion @ NR	oxLDL <0.05	per 1 pg/mL NA Stepwise Cox proportional hazards regression A*ssume median of known SE (SE = 0.001) and p=0.01*	NR <0.05 *13.14*	NR *1.85*	NR *93.3*
(Ishii et al 2003) Japan N=100	Cardiac death or rehospitalisation HR Cont	Age NS Gender NS NYHA 0.024 LVEF 0.027 cTnT <0.0001 cTnI 0.0024 logBNP 0.003 logANP 0.032 Inclusion @ NR	cTnT 0.0063	ln BNP pg/mL Stepwise Cox proportional hazards regression *Convert from log; assume median of known SE (SE = 0.001)*	NR 0.0095 *13.4*	NR *1.9*	NR *95.0*

Author Year Location Sample size	Outcome Ratio type Data type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units/cut-off point Position Analysis type *Assumptions*	Effect estimate	95%CI	
						Low	High
(Dokainish et al 2005) US N=110	Cardiac death or hospitalisation HR Dichot	CHF history 0.02 LVEF 0.004 Mitral E/Ea; early diastolic velocity/tissue early diastolic annular velocity 0.0001 Other Doppler echographic variables <0.01 Age NS HF aetiology NS Diabetes NS Inclusion @ NR	Mitral E/Ea; early diastolic velocity/tissue early diastolic annular velocity 0.0001	>250 pg/mL Optimal cut-off point Stepwise Cox proportional hazards regression *Assume SE = 0.55[a] and p=0.01*	NR <0.05 *4.1*	NR *1.4*	NR *12.1*

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); NA = not applicable; MIBG = ^{123}I-metaiodobenzylguanidine; SBP = systolic blood pressure; LVEF = left ventricular ejection fraction; LVDD = left ventricular end diastolic diameter; CTx = cardiac transplantation; NE = norepinephrine; Peak VO_2 = peak oxygen consumption; HFSS = heart failure survival score, which is calculated via components of heart failure aetiology, resting heart rate, LVEF, mean blood pressure, intraventricular conduction delay, VO_2 peak and serum sodium; MIBG = ^{123}I-metaiodobenzylguanidine; ADM = admission; LVEDP = left ventricle end diastolic pressure (mmHg); LVEDVI = left ventricular end diastolic volume index; oxLDL = oxidised low-density lipoprotein; cTnT = cardiac troponin T; cTnI = cardiac troponin I; SE = standard error; *Italics* = the assumptions made to approximate hazard ratios and confidence intervals for inclusion into secondary meta-analysis (all values are converted to 1000 pg/mL); [a] Predicted via a regression on sample size and mean BNP.

All five studies that analysed the predictive ability of BNP for mortality or cardiovascular event (dichotomised data) reported that it was a significant independent predictor of outcome (Table 38). One study that did not report sufficient information for dichotomised data (Dokainish et al 2005, Table 37) also found that BNP was an independent prognostic indicator of death or a cardiovascular event after adjustment for a cardiac catheterisation variable.

The combined effect estimate was 2.91, 95%CI [1.39, 6.05] without the estimated effect of Dokainish et al (2005) and 2.97, 95%CI [1.66, 5.29] with the converted insufficient information study included. Both effect estimates were significant ($p<0.02$) and both were developed using random effects models due to the borderline significance of the Q statistic when the fixed effects model was fitted.

Table 38 Mortality or cardiovascular event—BNP; hazard or odds ratio; *dichotomised data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Anand et al 2003) Multicentre; 302 centres; 16 countries N=4305 Val-HeFT	All-cause mortality or morbid event HR	NR Inclusion – all	NYHA NR LVEF NR ACE-I (baseline) NR Beta-blocker (baseline) NR HF aetiology NR Age NR	>97 pg/mL Median Cox proportional hazards regression	2.2 NR	1.98	2.52
(Bertinchant et al 2005) France N=63	Cardiac death or readmission HR	BNP 0.007 ANP 0.071 LVEF 0.11 Inclusion @ p<0.15	None	>254 pg/mL Optimised Stepwise Cox proportional hazards regression	3.23 0.01	1.32	7.94
(de Groote et al 2004a) The Netherlands N=407	Cardiac death or urgent transplantation HR	Age NR NYHA NR Atrial fibrillation NR HF aetiology NR LVEF NR RVEF NR Peak VO_2 NR % of predicted maximal VO_2 NR Left atrial diameter NR LVDD NR Mitral valve echocardiographic characteristics NR Sodium NR Creatinine NR NE NR BNP NR Aldosterone NR ET-1 NR Inclusion @ p<0.1	% of predicted maximal VO_2 <0.00001 Left atrial diameter 0.004 Age 0.005 Aldosterone 0.015	>110 pg/mL Median Stepwise Cox proportional hazards regression	3.17 0.0004	1.68	5.96
(Watanabe et al 2005) Multicentre, Tohoku district N=417 CHART study	Heart failure death (excludes sudden death) or hospitalisation HR	Age NR NYHA NR CHF history NR Diabetes NR LVDD NR ACE-I/ARB NR Beta-blocker NR Gender NR HF aetiology NR LVEF NR	Age 0.005 NYHA 0.007 CHF history 0.007 LVDD 0.027	>132 pg/mL Median Stepwise Cox proportional hazards regression	2.1 0.0168	1.14	3.85

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
		Inclusion @ NR					
(Verdiani et al 2005) Italy N=100	Cardiac death or readmission HR	Age NR Gender NR HF history NR Length of stay NR HF aetiology NR COPD NR Diabetes NR Hypertension NR Anxiety disorder NR Left atrial size NR LVEF NR Sodium NR Creatinine NR Hb NR Beta-blocker treatment NR Inclusion – all	NYHA <0.05	>696 pg/mL 75th percentile Stepwise Cox proportional hazards regression	15 <0.0001	4.2	53.8
Combined effect estimate Random effects model (REML) Cochran Q = 10.4; 4 degrees of freedom; p=0.034 Egger's test – p=0.144				NA	**2.91** **0.016**	**1.39**	**6.05**

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction; RVEF = right ventricular ejection fraction; ACE-I = Angiotensin-converting enzyme inhibitor; ARB = Angiotensin II receptor blockers; Hb = haemoglobin; LVDD = left ventricular diastolic diameter; Peak VO_2 = peak oxygen consumption; ET-1 = Endothelin-1; NE = norepinephrine; ADM = admission; DIS = discharge; COPD = chronic obstructive pulmonary disease; LVDD = left ventricular end diastolic volume; NA = not applicable

Cardiovascular event

Summary – BNP predicting cardiovascular event

It is clear from the evidence that BNP is significantly correlated with the outcome of cardiovascular death, independent of various demographic, neurohormonal and clinical variables. With and without studies that report insufficient information, the pooled effect estimate for continuous data is significant. The limited data on dichotomous BNP suggests that it may predict cardiovascular death.

The prognostic ability of BNP for the outcome of a cardiovascular event was assessed in 14 studies, of which all but two reported a significant independent relationship. Five, two and seven studies were categorised under continuous, dichotomous and insufficient information subheadings, respectively.

All studies that reported sufficient information on continuous data found that BNP was a significant predictor of cardiac death, independent of a range of clinical, echographic, biochemical marker and cardiac catheterisation variables (Table 39). The combined estimate, calculated via a random effects model with a REML estimate of between-study variation, was significant ($p=0.02$) at 7.81 per 1,000 pg/mL increase in BNP. The addition of estimated hazard ratios for seven additional studies (two of which reported a non-significant effect estimate) that reported insufficient information resulted in a reduced, but highly significant ($p<0.001$), combined estimate of 4.9 per 1,000 pg/mL of BNP, 95%CI [2.64, 9.01] (Table 40).

Table 39 Cardiovascular event—BNP; hazard or odds ratio; *continuous data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
(Kyuma et al 2004) Japan N=158	Cardiac death (pump failure + sudden) HR	BNP 0.0002 Renal dysfunction 0.003 Age 0.007 Cardiac I-MIBG activity 0.03 Use of nitrates 0.03 Diabetes 0.04 Inclusion @ p<0.05	Renal dysfunction NS Age NS Cardiac I-MIBG activity NS Use of nitrates NS Diabetes NS	per 1,000 pg/mL Cox proportional hazards regression	2.72 0.024	1.11	6.67
(Maeda et al 2000) Japan N=102	Cardiac death HR Cont	Age NS Gender NS HF aetiology 0.043 NYHA ADM NS NYHA STAB 0.0006 LVEF ADM NS LVEF STAB 0.011 ANP ADM NS ANP STAB <0.0001 BNP ADM 0.0161 BNP STAB <0.0001 NE ADM NS NE STAB 0.0005 ET-1 ADM NS ET-1 STAB 0.0354 IL-6 ADM 0.044 IL-6 STAB <0.0001 TNF-α ADM 0.043 TNF-α STAB 0.014 Inclusion – all	IL-6 STAB 0.0002	per 1,000 pg/mL Stepwise Cox proportional hazards regression	2.72 <0.0001	2.72	7.37
(Tsutamoto et al 1997) Japan N=158	Cardiac death HR	BNP <0.0001 NYHA <0.0001 ANP <0.0001 NE <0.0001 PCWP <0.0001 Mean pulmonary arterial pressure <0.0001 LVEF <0.0001 Right atrial pressure 0.002 Age 0.032 cGMP 0.039 Gender NS Cardiac index NS	PCWP 0.003	per 1,000 pg/mL Stepwise Cox proportional hazards regression	20.0 <0.0001	2.72	54.2

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units Analysis type	Effect estimate	95%CI	
						Low	High
		MAP NS Inclusion – all					
(Tsutamoto et al 1999) Japan N=290	Cardiac death HR	Age NS Gender NS NYHA NS HF aetiology NS LVEF 0.005 Right heart catheterisation variables <0.05 NE NS ET-1 NS Angiotensin II 0.002 ANP <0.0001 BNP <0.0001 Treatment variables NS Inclusion – all	None	per 1,000 pg/mL Stepwise Cox proportional hazards regression	54.2 <0.0001	20.0	396
(Tsutamoto et al 2001) Japan N=96	Cardiac death HR	Age 0.06 Gender 0.06 NYHA <0.0001 HF aetiology 0.741 LVEF <0.0001 NE <0.001 ET-1 <0.001 sFas 0.01 TNF-α 0.01 ANP <0.0001 BNP <0.0001 Inclusion – all	LVEF 0.02 sFas 0.009 HF aetiology 0.038	per 1,000 pg/mL Stepwise Cox proportional hazards regression	7.39 <0.0001	3.37	16.2
Combined effect estimate Random effects model (REML) Cochran Q = 21.1; 4 degrees of freedom; p<0.001 Egger's test – 4.46; p=0.065				NA	**7.81 0.022**	**1.63**	**37.5**

HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); NA = not applicable;
MIBG = ^{123}I-metaiodobenzylguanidine; MAP = mean arterial pressure; ADM = admission value; STAB = stabilisation after 3 months of treatment; PCWP = pulmonary capillary wedge pressure; cGMP = cyclic guanosine 3',5'-cyclic monophosphate; sFas = soluble Fas, an inhibitor of apoptosis; ET-1 = endothelin-1; NE = norepinephrine; TNF-α = tumour necrosis factor alpha; LVEF = left ventricular ejection fraction; IL-6 = interleukin 6; NYHA = New York Heart Association classification

Table 40 Cardiovascular event—BNP; insufficient information; *dichotomised and continuous data*

Author **Year** **Location** **Sample size**	**Outcome** **Ratio type**	**Univariate predictors of outcome (p≤0.05)** **Inclusion p-value**	**Confounding factors controlled for in multivariate analysis**	**Units / cut-off point** **Position** **Analysis type** ***Assumptions***	**Effect estimate** **Reported** ***Estimated***	**95%CI**	
						Low	**High**
(Selvais et al 2000) Belgium N=109	Cardiac death HR Cont	ET-1 <0.0001 NYHA <0.0001 NT-proANP <0.0001 BNP <0.0001 LVEF 0.0001 Age 0.01 Gender 0.17 Inclusion – all	ET-1 0.001 NT-proANP 0.0003	per 1 pg/mL Stepwise Cox proportional hazards regression *Assume SE equals that of unadjusted analysis*	χ^2 = 0.3 0.56 *1.31*	NR *0.53*	NR *3.22*
(Imamura et al 2001) Japan Multicentre; 6 institutions N=171 Ehime MIBG HF study	Cardiac death HR Cont	Age NS Gender NS HF aetiology NS NYHA 0.0003 Cardiothoracic ratio 0.001 LVEF 0.01 LVDD 0.002 LVESD 0.0001 NE 0.0031 ANP 0.0002 BNP 0.0002 MIBG data <0.005 Inclusion @ p<0.05	MIBG washout <0.0001	per 1 pg/mL NA Stepwise Cox proportional hazards regression *Assume SE equals that of unadjusted analysis*	NR NS	NR *7.37*	NR *2888*
(Isnard et al 2003) France N=250	Sudden death HR Cont	NYHA 0.013 Resting heart rate 0.039 LVEF 0.13 VO_2 peak 0.24 % of predicted VO_2 peak 0.049 ANP 0.001 NT-proANP 0.06 BNP <0.0001 NE 0.014 ET-1 0.02 Inclusion – all	None	per 1 pg/mL NA Stepwise Cox proportional hazards regression *Assume median of known SE (SE = 0.00046)*	χ^2 = 19.9 <0.0001 *7.78*	NR *3.16*	NR *19.2*
(Matsui et al 2002) Japan N=74	Cardiac death Cont	NYHA (6 months) <0.0002 LVEF (6 months) 0.03 LVEF (% change) 0.04 BNP ADM <0.0001 BNP (6 months) <0.0001	Delayed HM <0.05	per 1 pg/mL NA Stepwise Cox proportional hazards regression *Assume*	χ^2 = 7.92 0.005 *3.65*	NR *1.48*	NR *8.99*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Units / cut-off point Position Analysis type *Assumptions*	Effect estimate Reported *Estimated*	95%CI	
						Low	High
		NE ADM 0.0009 NE (6 months) <0.0001 MIBG variables <0.1 All above variables were measured for admission, 6 months and % change Age NS Gender NS Inclusion – all		*median of known SE (SE = 0.00046)*			
(Tsutsui et al 2002) Japan N=84	Cardiac death Cont	Age 0.25 Gender 0.32 HF aetiology 0.06 Diabetes 0.72 LVEF 0.041 NYHA <0.0001 BNP <0.0001 NE <0.0001 oxLDL <0.0001 Inclusion – all	oxLDL 0.0006	per 1 pg/mL Stepwise Cox proportional hazards regression A*ssume median of known SE (SE = 0.00046)*	χ^2 = 13.65 0.0002 *5.47*	NR 2.22	NR *13.48*
(Yu & Sanderson 1999) China N=91	Cardiac death HR Cont	Age 0.0007 Gender 0.43 NYHA 0.02 LVDD 0.54 LVEF 0.001 PASP 0.006 ANP 0.0005 BNP <0.0001 Inclusion – all	Age 0.003 Gender 0.32 NYHA 0.55 LVDD 0.50 LVEF 0.38 PASP 0.88 ANP 0.48	per 1 pg/mL Cox proportional hazards regression A*ssume median of known SE (SE = 0.00046)*	χ^2 = 18.3 <0.0001 *7.16*	NR *2.9*	NR *17.63*
(Ishii et al 2003) Japan N=100	Cardiac death HR Cont	Age NS Gender NS NYHA 0.04 LVEF 0.04 cTnT <0.0001 cTnI 0.003 logBNP <0.0001 logANP NS Inclusion @ NR	cTnT 0.016	ln BNP pg/mL Stepwise Cox proportional hazards regression *Convert from log; assume median of known SE (SE = 0.00046)*	NR 0.034 *2.64*	NR *1.07*	NR *6.50*

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; Superscript text (author p-values or NS = not significant or NR = not reported); LVDD = left ventricular end diastolic diameter; LVESD = left ventricular end systolic diameter; ADM = admission value; STAB = stabilisation after 3 months of treatment; oxLDL = oxidised low-density lipoprotein; PASP = pulmonary arterial systolic pressure; cTnT = cardiac troponin T; cTnI = cardiac troponin I; MIBG = ^{123}I-metaiodobenzylguanidine; LVEF = left ventricular ejection fraction; NA = not applicable; NYHA = New York Heart Association classification; NE = norepinephrine; ET-1 = endothelin-1; SE = standard error; *Italics* = the assumptions made to approximate hazard ratios and confidence intervals for inclusion into secondary meta-analysis.

Only two studies contributed to the dichotomised data evidence base (Table 41). Both reported that BNP was a significant independent predictor of cardiac mortality adjusted for an electrocardiogram variable (Vrtovec et al 2003) and echocardiographic/exercise test variables (De Groote et al 2004b). The combined estimate of 1.99 was associated with borderline significance (p=0.21) given the fact that only two studies were included in a meta-analysis. No other studies reported on the prognostic potential of BNP for cardiovascular event.

Table 41 Cardiovascular event—BNP; hazard or odds ratio; *dichotomised data*

Author Year Location Sample size	Outcome Ratio type	Univariate predictors of outcome (p≤0.05) Inclusion p-value	Confounding factors controlled for in multivariate analysis	Cut-off point Position Analysis type	Effect estimate	95%CI	
						Low	High
(Vrtovec et al 2003) US Texas N=241	Cardiac death (pump failure + sudden) HR	Prolonged QTc interval $^{<0.0001}$ BNP $^{0.0003}$ QRS duration $^{0.01}$ Digoxin $^{0.01}$ Age ^{NS} Gender ^{NS} HF aetiology ^{NS} NYHA ^{NS} LVEF ^{NS} Inotropes, diuretics, ACE inhibitors and beta-blockers ^{NS} Inclusion @ p<0.05	Prolonged QTc interval $^{0.0001}$	>1,000 pg/mL 75th percentile Stepwise Cox proportional hazards regression	1.76 $^{0.0007}$	1.01	3.07
(De Groote et al 2004b) The Netherlands N=150	Cardiac death HR	Age $^{0.002}$ NYHA $^{0.002}$ Atrial fibrillation $^{0.02}$ Exercise test variables (duration, blood pressure, VO_2, RER) $^{\leq 0.001}$ ANP, BNP and NE at rest and at peak exercise $^{\leq 0.006}$ LVEF $^{0.022}$ Left atrial diameter $^{<0.0001}$ Mitral DT $^{0.001}$ Inclusion @ p<0.10	Left atrial diameter $^{0.006}$ % of predicted VO_{2max} $^{0.002}$	>260 pg/mL Optimal cut-off point Stepwise Cox proportional hazards regression	2.5 $^{0.01}$	1.2	5.6
Combined effect estimate Fixed effects model Cochran Q = 0.52; 1 degree of freedom; p=0.47 Egger's test – NA				NA	**1.99 $^{0.21}$**	**0.11**	**36.9**

Position = where the cut-off point was located within the sample distribution; HR = hazard ratio; OR = odds ratio; Superscript text (author p-values or ^{NS} = not significant or ^{NR} = not reported); LVEF = left ventricular ejection fraction; RER = respiratory exchange ratio; Mitral DT = mitral delay time; NE = norepinephrine; NYHA = New York Heart Association classification; NA = not applicable.

Discussion

Effectiveness of B-type natriuretic peptide assays for prognosis in heart failure patients

As outlined in Table 42, the body of evidence available for the prognostic potential of B-type natriuretic peptides was good to excellent. Sixty prognostic studies met the criteria for inclusion in the review. Studies included for assessment recruited patients with varying disease severity and adjusted for different confounders in addition to different length of follow-up, data input (dichotomous, continuous) and outcomes (death, death or cardiovascular events, cardiovascular events). Despite this, results are remarkably consistent, with the vast majority of studies reporting that B-type natriuretic peptides were independent prognostic indicators of outcome. The evidence base was more extensive for BNP (n=36) compared to NT-proBNP (n=16) but results were consistent across studies.

Table 42 Assessment of body of prognostic evidence

Component	A Excellent	B Good	C Satisfactory	D Poor
Volume of evidence	Several level I or II studies with low risk of bias			
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact		Substantial		
Generalisability	Population(s) studied in body of evidence are the same as the target population			
Applicability		Applicable to Australian healthcare context with few caveats		

It is important to highlight that the prognostic ability of B-type natriuretic peptide has been tested in stabilised heart failure (HF) populations. A measurement taken on admission for diagnosis with the patient in a decompensated state may not provide prognostic information to the extent suggested by the prognostic meta-analyses. Admission levels of B-type natriuretic peptide reflect the severity of the acute decompensation superimposed on the chronic HF condition of the patient. Measuring B-type natriuretic peptide on discharge from hospital is therefore likely to yield a more accurate estimate of prognosis because it represents the chronic disease state rather than an acute episode. In a decompensated state, B-type natriuretic peptide levels are likely to reflect the extent to which the heart is stressed, but superior prognostic ability could be achieved by measuring B-type natriuretic peptide in a stabilised patient on discharge.

NT-proBNP

Mortality

Three studies reported significant prognostic ability of NT-proBNP for all-cause mortality using continuous data. One of these studies reported insufficient information for the first meta-analysis, but was included in the secondary meta-analysis with an estimated hazard ratio. Five studies reported on the same outcome for NT-proBNP using dichotomised variables, two of which reported insufficient information. Again, their respective hazard ratios were estimated using assumed p-values and standard errors from appropriate sources.

Continuous data

Due to the paucity of qualitative data for this outcome, a narrative analysis of the three studies that reported the prognostic ability of NT-proBNP for all-cause mortality (continuous) is presented.

Rossig et al (2004) reported that NT-proBNP, along with mean arterial blood pressure and serum pro-apoptotic activity, were independent predictors of all-cause mortality in a sample (n=48) of stabilised but advanced HF patients. A stringent p-value of <0.005 was used to select univariate predictors to enter a multivariate model. Even though age, LVEF and beta-blocker use were significantly ($p<0.06$) related to outcome they were not included in a multivariate analysis; hence, the hazard ratio only adjusts for one haemodynamic variable (mean arterial blood pressure) and one biochemical marker (serum pro-apoptotic activity).

Kirk and colleagues (2004) published a moderate, but significant, odds ratio of 1.66 for log converted NT-proBNP in relation to all-cause mortality. This converted to an approximate hazard ratio increase of 1.13 per 1,000 pg/mL. This study monitored 161 diagnosed HF patients (from 2,230 patients who were admitted to hospital) for 1-year using national registers. NT-proBNP was measured the morning after hospital admission and therefore the concentration may be indicative of a decompensated state. Nevertheless, it was concluded that NT-proBNP was a strong prognostic predictor of 1-year mortality in HF patients regardless of NYHA class, age, gender and systolic dysfunction in objectively diagnosed HF patients.

George et al (2005) reported that NT-proBNP was a significant independent predictor of all-cause mortality after adjusting for LVEF, haemoglobin and erythropoietin levels in a sample of 182 stabilised patients with advanced HF. However, they did not report an adjusted effect estimate. Assuming a median standard error of known values (ie 0.000026), the estimated hazard ratio to be included in a secondary meta-analysis was calculated as 1.10, 95%CI [1.05, 1.16].

Although there are consistent findings of significance for the three studies, the magnitude of effect varies. Reasons for heterogeneity include different patient groups (Rossig et al (2004)—stable compensated end stage HF; Kirk et al (2004)—a wide range of decompensated HF patients) and adjustment for different additional prognostic variables. Given the small number of studies (n=3) and the fact that there were consistent findings between them, a combined effect estimate of 1.28 per 1,000 pg/mL increase in NT-proBNP ($p=0.23$) could be suggested to be borderline significant, despite not reaching the standard level of statistical significance.

Dichotomous data

Three studies (Hartmann et al 2004; Fisher et al 2003; Kellet 2005) reported on the prognostic potential for NT-proBNP over and above other clinical variables, with similar significant results. Hartmann et al (2004) prospectively followed a large sample of 1,011 severe HF patients (LVEF = 20±4%) for a median of 5 months who were randomly assigned to beta-blocker or placebo treatment. Accordingly, the patient group was heavily medicated with digitalis (58%), diuretic (99%), ACE inhibitor or AT_{II} receptor antagonist (99%), spironolactone (26%) and amiodarone (18%). The univariate relative risk of an NT-proBNP serum concentration above the median for all-cause death was 2.7, 95%CI [1.7, 4.3]. The hazard ratio was slightly reduced (2.2; 95%CI [1.1, 3.5]) but still significant (p=0.0001) after adjustment for the treatment group, age, aetiology of HF, systolic blood pressure, recent hospitalisation or high risk combination (ascites, rales, oedema at randomisation). Other potential prognostic clinical variables such as LVEF, creatinine and gender were not significant in univariate analyses, and were therefore not entered into the stepwise multiple regression procedure.

Fisher et al (2003) studied 87 HF patients admitted to hospital on an emergency basis with decompensated HF due to left ventricular systolic dysfunction. Discharge NT-proBNP (ie stabilised patients) was assessed for its predictive ability on all-cause mortality. Fisher et al (2003) reported that a multivariate analysis revealed NT-proBNP as a significant independent predictor of death. Even though the authors stated that significant (p<0.05) univariate variables of age, gender, NYHA class, LVEF, heart rhythm, comorbidity, prior HF hospitalisation and creatinine were adjusted for in multivariate analyses, they did not report the adjusted variables. It is therefore unclear what variables were adjusted for with regards to the odds ratio reported.

Kellett (2005) reported on 342 consecutive patients admitted to the ED with HF. NT-proBNP, measured within 2 hours of admission, was found to be one of four (out of 21 clinical, X-ray and ECG variables) independent predictors of in-hospital mortality. The reported effect estimate may be inflated due to the fact that the cut-off NT-proBNP level to dichotomise patient groups was selected via ROC curve analysis, rather than the arbitrary selection of the middle value.

The significant findings of these three studies were supported by two additional studies that reported NT-proBNP significance (Gardner et al 2005; Richards et al 2001) but did not give details of an effect estimate. Gardner et al (2005) concluded that NT-proBNP was a superior predictor of all-cause mortality because it was the only selected variable out of a variety of potential predictors (Table 29) in a forward stepwise multivariate regression. However, the stepwise statistical procedure, in addition to the fact that there were no other significant independent predictors of outcome, resulted in the effect estimate not being adjusted for any other variables. Richards et al (2001) recruited a cohort of stabilised HF patients from multiple centres around Australia and New Zealand to investigate the effect of beta-blocker (carvedilol) treatment on patient relevant outcomes. The secondary hypothesis examined the prognostic potential of adrenomedullin and NT-proBNP for treatment effectiveness and patient relevant outcomes. Richards et al (2001) reported a univariate hazard ratio of 4.7 for all-cause mortality and concluded that NT-proBNP was a significant prognostic indicator after the adjustment for adrenomedullin, treatment group, NYHA class, LVEF, prior MI, age and prior hospital admission. However, an adjusted hazard ratio was not reported. Consequent analysis revealed that the prognostic potential of NT-proBNP was weaker in the subgroup receiving beta-blocker treatment compared to those who were not.

The combined effect estimates from the three studies which reported sufficient information and the five studies (including insufficient information) combined were significant and ranged from 2.45 to 2.70.

Overall

Taken together, the eight included studies reported that NT-proBNP was a significant independent predictor of all-cause mortality after the adjustment of a variety of confounding variables. The magnitude of the effect was borderline for limited continuous data but significant for the dichotomous data.

Death or cardiovascular event

In contrast to the evidence base (26 studies) for BNP, only five NT-proBNP studies answered the inclusion criteria for prognosis with regard to death or cardiovascular event.

Continuous data

Only one study, which was categorised to insufficient information, provided evidence of NT-proBNP for prognosis with regard to death or cardiovascular event. Zugck et al (2002) recruited 408 consecutive stabilised HF patients to assess whether beta-blocker treatment affected the prognostic utility of commonly used clinical markers and hormonal biomarkers. In addition to their current pharmaceutical regime, 243 patients were medicated with beta-blockers; the remainder maintained their usual treatment. Overall, NT-proBNP provided prognostic information additional to that provided by LVEF and peak VO_2 for cardiac death or rehospitalisation. The investigators highlighted that the independent prognostic significance of NT-proBNP was lost in subset analyses for patients taking beta-blockers. Nevertheless, this review focuses on the prognostic potential of NT-proBNP for HF patients overall, not just those medicated with beta-blockers. This study therefore provides additional qualitative support for the independent prognostic significance of NT-proBNP over other existing clinical variables.

Dichotomous data

Three studies reported the independent significance of NT-proBNP for the prediction of death or cardiovascular event in HF patients. One study reported significance but did not report an associated effect estimate.

The pooled effect estimate via a fixed effects model of 3.12 was associated with borderline significance (p=0.11). The use of a fixed effects model seems contradictory to the data, which included effect estimates of 2.1, 4.2 and 15.3. However, the latter estimate was based on only 34 patients and was consequently associated with an extremely large confidence interval (95%CI [1.4, 169]). This effect estimate, reported by O'Brien et al (2003), therefore did not carry much weight in the meta-analysis, and its inclusion did not result in a significant heterogeneity test.

Overall

Based on a small evidence base (n=5) and a borderline level of statistical significance, it appears that NT-proBNP has prognostic potential for death or cardiovascular event.

Cardiovascular event

Three studies, two of which are previously described for the outcome of death (Richards et al 2001; George et al 2005), assessed the ability of NT-proBNP to predict a cardiovascular event. Richards et al (2001) reported that NT-proBNP was a significant prognostic indicator of HF mortality after adjustment for adrenomedullin, treatment group (some patients were treated with beta-blockers in addition to their normal pharmaceutical regime), NYHA class, LVEF, prior MI, age and prior hospital admission. George et al (2005) found that NT-proBNP levels were predictive of further hospitalisation due to HF, even after adjustment for NYHA class, haemoglobin and erythropoietin prognostic indicators. Neither group stated an adjusted effect estimate.

De Pasquale et al (2004) reported, in a sample of 53 stabilised HF patients, that NT-proBNP did not provide additional prognostic information over that provided by surfactant protein B (SP-B). A conditional logistic regression model was used, and reported that for each 25 per cent increase in the median SP-B there was a 4.7 times risk of cardiovascular hospitalisation. Even though NT-proBNP was significantly higher in patients who experienced further HF hospitalisation compared to those who did not ($p=0.02$), it did not provide independent prognostic information when included in a multivariate analysis with SP-B. It is unclear whether other variables measured in this study were significant univariate predictors of outcome and/or whether they were adjusted for in the final multivariate analysis. Furthermore, plasma surfactant protein B is not measured in routine clinical practice.

Overall

Two out of three studies reported that NT-proBNP was a significant independent predictor of either HF mortality or hospitalisation. The remaining study, which did not report significance, adjusted for a variable that is not part of standard clinical care. Hence, qualitatively there is evidence to suggest that NT-proBNP would be useful for predicting future cardiovascular events in HF patients.

BNP

Mortality

Nine studies reported on the prognostic potential of BNP for all-cause mortality, of which five, two and two studies were categorised under continuous, dichotomous and insufficient information subheadings, respectively. Seven of nine studies indicated that BNP was an independent prognostic indicator of all-cause mortality.

Continuous data

Sources of heterogeneity (Cochran $Q = 195$; $p<0.001$) between the studies using continuous data may be due to different populations at baseline, different medications at baseline or throughout follow-up, and the adjustment of different confounding variables when reporting adjusted effect estimates.

Of the studies that reported sufficient information, effect estimates ranged from no effect (Wijeysundera et al 2003); HR = 0.99) to a highly significant effect (Maisel et al 2004; OR = 3.6). It is worthwhile highlighting that the former study adjusted for NT-proANP in their multivariate analyses, which may have introduced issues with co-linearity and masked the significance of BNP. Furthermore, because NT-proANP

measurement is not part of standard clinical practice, the non-significant adjusted effect estimate of this particular study should be interpreted with caution.

The other study that reported BNP non-significance was conducted by van der Meer and colleagues (2004) and aimed to investigate the prognostic importance of erythropoietin in HF. This study reported that although BNP serum concentration was the most significant univariate predictor of death, its prognostic significance was removed after adjusting for erythropoietin and haemoglobin levels. To form a conservative combined effect estimate the meta-analysis was re-run with a null value assigned to this study.

On the basis of the evidence at hand, the effect estimate ranges from 1.74 (p=0.08; six studies—inclusive of insufficient information) to 1.96 (p=0.07; five studies) per 1,000 pg/mL increase in BNP. The upper estimate suggests that an increase in BNP concentration of 1,000 pg/mL confers almost twice the risk of dying. This conclusion should be tempered by the fact that 95% confidence intervals around these point estimates suggest that they only represent effect trends (ie borderline significance).

Dichotomous data

Two studies reported similar adjusted hazard ratios for dichotomised BNP for the prediction of all-cause mortality. Anand et al (2003) used Val-HEFT trial data to examine the prognostic potential of BNP in 4,305 stabilised HF patients (NYHA class II-IV) followed for 2–3 years. They reported a significant hazard ratio of 2.1 adjusted for clinical, demographic, medication use and echographic variables. In comparison, Vrtovec et al (2003) recruited a more homogenous group of severely diseased (NYHA class III-IV) and more heavily medicated (beta-blocker = 73%, ACE-I = 87%) stabilised HF patients. In contrast to the consensus, Vrtovec et al (2003) used the 75th percentile of their sample BNP distribution to dichotomise their patients into low and high groups. They reported a highly significant hazard ratio (HR = 2.1; 95%CI [1.2, 3.4]) adjusted for an electrocardiogram variable (prolonged QTc interval). LVEF, NYHA class, age and medication usage were adjusted for because they did not reach univariate significance.

These studies combined, via REML fixed effect meta-analysis, resulted in a borderline significant (p=0.06) overall effect point estimate of 2.09, 95%CI [0.82, 5.34].

In an additional study conducted by Watanabe et al (2005), BNP was reported not to be a significant predictor of all-cause mortality after adjustment for clinical, demographic, echographic and medication use variables. However, the same study reported that adjusted BNP dichotomised about the sample median predicted cardiovascular death or hospitalisation. In this study only 70 per cent of all-cause mortality was of cardiac origin. The lack of correlation for BNP may be due to the fact that 30 per cent of mortality was not related to cardiac causes and consequently had little relation to BNP at baseline. When this study was included in a secondary meta-analysis with a null effect (assuming a standard error the same as that for death or cardiovascular event), the combined estimate via the random effects model was not significant (1.74; 95%CI [0.68, 4.42]; p=0.13). This lower effect estimate should be viewed with caution due to the aforementioned limitations of the non-significant effect estimate reported by Watanabe et al (2005).

Overall

Using an evidence base of nine studies it appears that there is a trend for BNP providing independent prognostic information over and above other clinical, echographic, demographic and medication use variables. The significance of combined estimates is robust to the inclusion of studies that reported insufficient information and no effect.

Mortality or cardiovascular event

Twenty-five studies reported on the prognostic potential of BNP for mortality or cardiovascular event, of which 13, five and seven studies were categorised under continuous, dichotomous and insufficient information subheadings, respectively. Twenty-three of 26 studies indicated that BNP was an independent prognostic indicator of mortality or cardiovascular event.

Continuous data

Even though there was consistency in the reporting of BNP's prognostic significance, considerable heterogeneity in the magnitude of reported effect sizes was apparent. Hence, a random effects REML model was chosen for all analyses. Regardless of what studies were included or excluded in analyses, the combined effect estimate was significant and ranged from 4.46, 95%CI [1.41, 14.13] to 8.07, 95%CI [1.4, 48.2].

It should be noted, however, that the meta-analyses included some outliers, namely data from Koseki et al (2003) and Sakatani et al (2004). The study by Koseki et al contributed four hazard ratios to the meta-analyses, and when converted to 1,000 pg/mL increments the reported effect estimates were extremely large. When data were checked, the only explanation for this aberration was that this study included a less severely diseased patient group which, as a consequence, did not have a large range of (or very high) BNP concentrations. Thus, adjusting Koseki's effect estimate to a 1,000 pg/mL increment may have extrapolated the data beyond the limits on which the original effect estimate was established (Koseki et al 2003). The other large point estimate from Sakatani et al (2004) may have been due to a misprint in the manuscript, but this assumption is difficult to confirm. Nevertheless, these studies were included in meta-analyses but, because of their extremely large standard errors (see 95% confidence intervals associated with point estimates in Table 36), they did not appreciably influence the pooled estimate.

Three continuous data studies (Koglin et al 2001; Setsuta et al 2002; Sakatani et al 2004) reported that BNP was not a significant independent predictor of mortality or cardiovascular events. However, two of these studies reported moderate to strong intercorrelations between BNP and other variables included in the multivariate analysis, which may have masked the peptide's significant prognostic potential. Koglin et al (2001) reported a correlation of –0.71 between HF survival score and BNP, and Setsuta et al (2002) reported a correlation of 0.66 between BNP and fatty acid binding protein. Both fatty acid binding protein and HF survival score were significant in multivariate analyses; thus, without these variables the significance of BNP on prognosis may have been realised. However, the inclusion of highly correlated variables in multivariate regression does not always lead to BNP being ruled out, as is evidenced by Hamada et al (2005) and Logeart et al (2004), who included related variables but still reported BNP significance.

Dichotomous data

All available studies (n=6) that used dichotomous BNP reported BNP as a significant independent predictor of mortality or a cardiovascular event independent of a range of clinical, cardiac catheterisation, exercise test, medication use and echographic variables. With the exception of the effect estimate from Verdiani et al (2005), the magnitude of effect for BNP on the outcome was reasonably consistent, even though studies adjusted for different variables. The reason behind an aberrantly large effect estimate of 15 reported by Verdiani et al (2005), and hence a significant heterogeneity test for the pooled estimate, was almost certainly the fact that the 75th percentile was chosen as a cut-off point to dichotomise BNP values, rather than the median used in other studies.

Overall

Taken together, 25 (13 continuous; 6 continuous but insufficient information; 5 dichotomous; 1 dichotomous but insufficient information) studies reported on the prognostic ability of BNP. Only three reported a non-significant adjusted relationship, two of which could possibly be explained. The level of consistency between studies, in addition to a statistically significant range of effect estimates (2.9 to 8.1) using the relatively conservative REML meta-analytic methodology, infers that BNP predicts mortality or cardiovascular events in HF patients independent of several demographic, risk factor, clinical and echographic variables.

Cardiovascular event

Fourteen studies were included for assessing the prognostic potential of BNP for cardiovascular event. Half of the studies failed to report effect estimates, but estimated hazard ratios using assumed standard errors and reported (or assumed) p-values were included in the secondary meta-analysis. In the majority of studies the 'cardiovascular event' was that of cardiac death. Only two studies reported that BNP was a non-significant indicator of cardiac death.

Continuous data

A relatively large (effect estimate 7.8 per 1,000 pg/mL) and significant (p=0.022) combined effect estimate was reported for the five studies which reported adjusted effect estimates on continuous data for the outcome of cardiovascular event. Conversion of effect estimates reported by Tsutamoto et al (1997, 1999) to 1,000 pg/mL increments yielded very high effect estimates of 20 and 54.2. It is unknown why these two studies reported large effect estimates—it is possible that, because of the stepwise procedure, the effect estimates were adjusted only for pulmonary capillary wedge pressure or not adjusted at all. It is worthwhile noting that the large point estimate of 54 was contained within a large confidence interval ranging from 20 to 396 and, as such, would not have significantly impacted upon the pooled estimate.

The study by Selvais et al (2000) was one of the few studies to report BNP as a non-significant prognostic marker of cardiovascular death. However, NT-proANP was adjusted for in the stepwise analysis, which may have masked the independent prognostic ability of BNP. This effect is clear in the studies which included both NT-proBNP and BNP in multivariate regression (see 'NT-proBNP versus BNP' section), where inevitably only one marker proves to be significant because of intercorrelations between these peptides of the same molecular family. Potentially, if BNP was compared to demographic and echographic measures alone in a multivariate equation, it may reveal the significance of BNP. Nevertheless, the inclusion of this and the null effect reported by Imamura et al (2001) did not affect the significance of the overall pooled estimate.

Dichotomous data

Two studies reported significant prognostic ability of BNP to predict cardiac death (De Groote et al 2004b; Vrtovec et al 2003). The study by de Groote et al (2004) found that in a sample of 150 patients, BNP was correlated with cardiac death, independent of left atrial diameter and the percentage of predicted VO_{2max} achieved during a peak exercise test. The separating value for dichotomisation of BNP was chosen via ROC curve analysis so the effect estimate reported could be inflated. De Groote also contributed to the 'death or cardiovascular event' outcome in a separate study on 407 HF patients.

Vrtovec et al (2003) recruited a homogenous group of severely diseased (NYHA class III-IV) and heavily medicated (beta-blocker = 73%, ACE-I = 87%) stabilised HF patients. In contrast to the consensus, this study used the 75th percentile of the sample BNP distribution to dichotomise their patients into low and high BNP groups. A highly significant hazard ratio was reported (HR = 1.76; 95%CI [1.01, 3.07]; p=0.0007) adjusted for a prolonged QTc interval in an electrocardiogram. Other clinical and echocardiographic variables such as LVEF, NYHA class, age and medication usage were not adjusted for, univariate significance was not reached, and they were therefore not included in the stepwise Cox proportional hazards regression.

Even though both studies reported that BNP was a significant predictor of cardiac death, a meta-analysis resulted in only borderline significance (p=0.21). However, this is more a reflection of the limitations of using only two studies in a meta-analysis rather than a true non-significant effect. Hence, the combined effect of the two studies (effect estimate = 1.99) should be treated as supporting evidence for the effectiveness of BNP as a prognostic marker.

Overall

It is clear from the evidence that BNP is significantly correlated with the outcome of cardiovascular death, independent of various demographic, neurohormonal and clinical variables. With and without studies that report insufficient information, the pooled effect estimate for continuous data is significant. The limited data on dichotomised BNP suggests that it may be related to cardiovascular death.

Conclusions

Prognostic effectiveness

Fifty-two studies were included for the assessment of prognostic effectiveness for NT-proBNP (n=16) and BNP (n=36). Eight studies provided insufficient information for formal inclusion in the meta-analyses but were assessed narratively and used in sensitivity analyses to test the robustness of the effects. Studies were consistent in their design, but patient groups, follow-ups, data types (dichotomous, continuous), adjusted variables in multivariate analyses and outcomes (death and/or cardiovascular event) varied considerably between studies. The consistency for which BNP and NT-proBNP were selected as significant independent predictors of outcome suggests that these peptides provide additional prognostic information over that already provided by a variety of existing clinical strategies.

NT-proBNP assays

Only one out of 16 studies that formed the evidence base for NT-proBNP reported a non-significant effect.

Significant or close to significant pooled estimates were present for all outcomes (death, death or cardiovascular event, cardiovascular event), which highlighted the independent prognostic value of NT-proBNP. In most cases a random effects model was used, which

demonstrates that the magnitude of effect varied between studies. Non-significant ($p<0.05$) pooled estimates were calculated on a limited number of studies (ie <3). The lack of significance may be due to the small study numbers rather than a lack of real effect.

In conclusion, there is strong evidence of the independent prognostic potential of NT-proBNP on a moderate volume evidence base.

BNP assays

Thirty-six studies reported on at least one outcome (death, death or cardiovascular event, cardiovascular event) in assessing the independent prognostic potential of BNP. Eight of these studies reported that BNP was not independently predictive of outcome after controlling for chosen variables.

Significant pooled estimates or non-significant trends were present for all outcomes (death, death or cardiovascular event, cardiovascular event) and random effect models dominated the analyses. This demonstrates that BNP plays an independent role in HF patient prognosis and that the magnitude of effect varies between studies. Reasons for the heterogeneity have not been investigated due to the relatively small study numbers for each subcategory (continuous and dichotomous) for each of three outcomes. In three out of four cases, non-significant pooled estimate trends were reported on less than three studies. The limitations of meta-analysing small study numbers, rather than lack of real effect, therefore may explain the lack of pooled statistical significance.

In conclusion, there is strong evidence of the independent prognostic potential of BNP on a large volume evidence base.

Appendix J Prognostic studies included in this review

Studies included for BNP assays

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
Anand et al (2003) Multicentre; 302 centres; 16 countries Val-HeFT (Same cohort as Latini et al (2004)	Prospective cohort study Level - II Quality score 2/4 Cons. Rec. - N Blinded Y Objective - Y Follow-up 2–3 years	Population Participants in the Val-HeFT trial with stable, symptomatic HF Inclusion criteria History of HF for at least 3 months, LVEF <40% and LV internal diastolic diameter/body surface area ≥2.9 cm/m^2 Exclusion criteria <18 years	Sample description N 4305 Mean age 63±11 % male 80 % IHD 57 [CHD] Disease severity BNP[STAB] 181±230 % NYHA I/II/III/IV 0/62/36/2 LVEF 27±7 PeakVO_2 SBP 124 Creatinine	Comorbidities % Atrial fibrillation 12 Diabetes Hypertension COPD Medication usage % Beta-blockers 36 ACE inhibitors 93 Diuretics 86 Digitalis Spironolactone Amiodarone	All-cause death (n=832) Death or morbid event, the latter defined as non-fatal cardiac arrest, HF hospitalisation or intravenous medication for more than 4 hours (n=NR; 30.5%)	Type IRMA (Shionogi, Schering CIS, Milan, Italy) Intra-assay COV Appropriate within- and between-laboratory controls were performed Inter-assay COV Not reported Time sample was drawn Prior to randomisation to an angiotensin receptor blocker
Bertinchant et al (2005) Nimes, France	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - NR Blinded Y Objective - Y Follow-up days 660[M] (30–1350)	Population Acute (unstable; decompensated) and chronic stable congestive HF Inclusion criteria LVEF <45% and symptomatic within previous 12 months Exclusion criteria Patients with acute or recent (<2 months) myocardial infarction, acute transient symptomatic ischaemia, cardiac surgery within previous 2 months, acute or chronic end-stage renal or liver disease, severe pulmonary or systemic illness, chronic	Sample description N 63 Median age 54 (36–70) % male 55 % IHD 19 Disease severity BNP[ADM and STAB] 331±453 % NYHA I or II/III/IV 37/49/14 LVEF 24±10 PeakVO_2 SBP Creatinine	Comorbidities % Atrial fibrillation 18 Diabetes 10 Hypertension COPD Medication usage % Beta-blockers 16 ACE-I/ARBs 75 Diuretics Digitalis Spironolactone 10 Amiodarone 8	Cardiac death or readmission (n=21)	Type Shionoria (Shionogi, Osaka, Japan) Intra-assay COV 5.7% @ 20 pg/mL Inter-assay COV 4.2% @ 20 pg/mL Time sample was drawn During scheduled outpatients visits (n=44) or within 12 hours of admission to hospital (n=19)

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
		inflammatory disease, malignancy or skeletal muscle diseases were excluded						
Bettencourt et al (2000) Porto, Portugal	Prospective cohort study Level - II Quality score - 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 541±347	Population Patients with mild to moderate HF Inclusion criteria Diagnosis based on clinical evaluation and echocardiography Exclusion criteria NR	Sample description N 139 Mean age 70±9 % male 59 % IHD 53 Disease severity BNP^{STAB} 398±429 % NYHA I/II/III/IV 12/83/6/0 LVEF 34±13 $PeakVO_2$ SBP Creatinine	Comorbidities % Atrial fibrillation 30 Diabetes Hypertension COPD Medication usage % Beta-blockers 25 ACE inhibitors 89 Diuretics 96 Digitalis Spironolactone Amiodarone		All-cause death (n=39)	Type IRMA (Shianogi, Osaka, Japan) Intra-assay COV <8% Inter-assay COV <8% Time sample was drawn Within 1 week of first appointment
De Groote et al (2004a) Lille, France; Rotterdam, The Netherlands	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 787^{M} 100% follow-up	Population NR Inclusion criteria After optimisation of therapy, Ambulatory, stable for at least 2 months, LVEF ≤45% Exclusion criteria Myocardial infarction, an episode of unstable angina, or undergone coronary revascularisation in the previous 3 months	Sample description N 407 Mean age 57±11 % male % IHD 45 Disease severity BNP^{STAB} 110 % NYHA I/II/III/IV NR/NR /26/NR LVEF 33±13 $PeakVO_2$ 15±5 SBP 117±20 Creatinine 11±3	Comorbidities % Atrial fibrillation 13 Diabetes 28 Hypertension 41 COPD Medication usage % Beta-blockers 93 ACE inhibitors 95 Diuretics 81 Digitalis Spironolactone 26 Amiodarone		Cardiac death or urgent transplantation (n=78)	Type Shionoria (Shionogi, Osaka, Japan) Intra-assay COV 2.7% @ 21pg/mL 2% @ 520 pg/mL Inter-assay COV 4.2% @ 21 pg/mL 2.1% @ 520 pg/mL Time sample was drawn After patients were stabilised on medication
De Groote et al (2004b) Lille, France; Rotterdam, the Netherlands	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y	Population Stable patients with moderate congestive HF Inclusion criteria Ambulatory patients stable for at least 6 months, LVEF ≤45%	Sample description N 150 Mean age 55±13 % male % IHD 50 Disease severity $BNP^{STAB\ M}$ 107	Comorbidities % Atrial fibrillation 12 Diabetes Hypertension COPD Medication usage % Beta-blockers 91		Cardiac death (n=35)	Type Shionoria (Shionogi, Osaka, Japan) Intra-assay COV 2.7% @ 21pg/mL 2% @ 520 pg/mL

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics	Outcome (number of outcomes)	Assay details
	Blinded N Objective - Y Follow-up days 1171 M 100% follow-up	Exclusion criteria Chronic renal failure, cardiac transplant planned, participation in previous study, or technical reasons	% NYHA I/II/III/IV LVEF 33±10 PeakVO2 16±6 SBP 119±23 Creatinine ACE inhibitors 93 Diuretics 80 Digitalis Spironolactone Amiodarone		Inter-assay COV 4.2% @ 21 pg/mL 2.1% @ 520 pg/mL Time sample was drawn After patients were stabilised on medication
Dokainish et al (2005) US, Texas	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 527±47	Population Patients admitted to hospital for CHF Inclusion criteria Diagnosed with CHF using the Framingham criteria Exclusion criteria Non-sinus rhythm, mitral valve disease, unstable angina, acute MI, other terminal diseases	Sample description N 110 Mean age 57 Weighted % male 53 % IHD 24 Disease severity BNP STAB 400 Weighted % NYHA I/II/III/IV LVEF 40 Weighted PeakVO2 SBP Creatinine Comorbidities % Atrial fibrillation Diabetes 45 Hypertension 77 COPD Medication usage % Beta-blockers 51 ACE inhibitors 78 Diuretics 92 Digitalis Spironolactone Amiodarone	Cardiac death or readmission (n=54)	Type Biosite Triage BNP test Intra-assay COV NR Inter-assay COV NR Time sample was drawn 24 hours prior to discharge after stabilisation.
Gackowski et al (2004) Paris, France	Prospective cohort study Level - II Quality score – 4/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 60	Population Patients admitted for acute decompensated HF Inclusion criteria Progressive resting dyspnoea with clinical signs of pulmonary and/or peripheral congestion based on Framingham criteria, requiring urgent hospitalisation and treatment with an intravenous diuretic and/or dobutamine Exclusion criteria Inability to give informed consent, severe pulmonary, hepatic or renal disease; acute coronary	Sample description N 95 Mean age 57±12 % male 60 % IHD 48 Disease severity BNP ADM 346±177 BNP STAB 300 % NYHA I/II/III/IV LVEF 34±16 PeakVO2 SBP 120±24 Creatinine 129±61 Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 26 ACE inhibitors 63 Diuretics 47 Digitalis Spironolactone Amiodarone	Death or rehospitalisation, or urgent CTx (n=37)	Type IRMA Intra-assay COV 8% Inter-assay COV 11% Time sample was drawn On admission to the ED, 24 hours and on the 7th day or discharge

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics	Outcome (number of outcomes)	Assay details
		syndromes unless HF the predominant manifestation or impossibility to obtain Doppler echo measure of good quality			
Hamada et al (2005) Ube, Japan	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR Blinded NR Objective - Y Follow-up days 177±111	Population Acutely decompensated CHF Inclusion criteria LVEF ≤40% and NYHA III-IV Exclusion criteria Acute and recent myocardial infarction within a month, unstable angina pectoris, severe mitral regurgitation, atrial fibrillation, renal failure on hemodynamic dialysis, post-pacemaker implantation	Sample description N 52 Mean age 64 Weighted % male % IHD 60 Disease severity BNP ADM 727 Weighted BDN DIS 259 Weighted % NYHA I/II/III/IV LVEF 31 Weighted PeakVO_2 SBP Creatinine Comorbidities % Atrial fibrillation 0 Diabetes Hypertension COPD Medication usage DIS % Beta-blockers 55 ACE inhibitors 90 Diuretics 100 Digitalis 26 Spironolactone Amiodarone	Cardiac death or readmission (n=20)	Type IRMA (Shionogi, Japan) Intra-assay COV NR Inter-assay COV NR Time sample was drawn On admission to hospital and pre-discharge
Imamura et al (2001) Japan Multicentre; 6 institutions Ehime MIBG HF study	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR Blinded NR Objective - Y Follow-up days 810±240	Population Participants in Ehime MIBG Heart Failure Study; 31 in patients and 140 outpatients Inclusion criteria LVEF <40% Exclusion criteria Patients with severe diabetes mellitus whose haemoglobin A1c was more than 8% or those who had chronic renal failure and were on hemodialysis	Sample description N 171 Mean age 63±11 % male 73 % IHD 44 Disease severity BNP 158±188 % NYHA I/II/III/IV LVEF 27±10 PeakVO_2 SBP Creatinine Comorbidities % Atrial fibrillation Diabetes 0 Hypertension COPD Medication usage % Beta-blockers ACE-I/ARB 100 Diuretics 78 Digitalis 73 Spironolactone Amiodarone	Cardiac death (n=11) Cardiac death or hospitalisation (n=27)	Type IRMA (Shionoria) Intra-assay COV NR Inter-assay COV NR Time sample was drawn NR
Ishii et al (2003) Tokyo, Japan	Prospective cohort study Level - II Quality score – 4/4	Population Patients admitted for worsening chronic HF Inclusion criteria NYHA III-IV Exclusion criteria	Sample description N 100 Mean age 68±11 % male 56 % IHD 37 Disease severity Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage %	Death or worsening of HF (n=44)	Type RIA (Shiono RIA BNP assay) Intra-assay COV 5.2% Inter-assay COV

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
	Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 391 (16–884) 100% follow-up	Patients who had clinical or ECG evidence of ACS after 2 months of treatment, patients with percutaneous coronary intervention or coronary artery bypass graft surgery during the 2 months, patients with history of recent myocardial infarction or coronary revascularisation within 3 months, myocarditis, renal failure or pulmonary diseases	BNP^{ADM} BNP^{STAB} % NYHA I/II/III/IV LVEF $PeakVO_2$ SBP Creatinine	753±598 249±276 0/57/40/3 43±13 12±5	Beta-blockers ACE-I /ARB Diuretics Digitalis Spironolactone Amiodarone	54 100 91 94 		6.1% Time sample was drawn On admission to hospital and 2 months after the patient had been stabilised
Isnard et al (2003) Paris, France	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 584^{M} (12–1368)	Population Patients attending an outpatient clinic Inclusion criteria LVEF <45% and NYHA II-III Exclusion criteria Pulmonary, renal or liver disease or recent MI	Sample description N Mean age % male % IHD Disease severity BNP^{STAB} % NYHA I/II/III/IV LVEF $PeakVO_2$ SBP Creatinine	 250 54±12 NR 25 198±195 2/76/20/2 29±10 17±6 125±23 	Comorbidities Atrial fibrillation Diabetes Hypertension COPD Medication usage Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	% 12 % 21 89 75 14 32	Cardiac death or urgent CTx transplantation (n=47) Sudden death (n=19)	Type IRMA (Peninsula Labs, California) Intra-assay COV 8% Inter-assay COV 11% Time sample was drawn Sample was drawn in patients who had stabilised HF for at least 1 month
Koglin et al (2001) Munich, Germany	Prospective cohort study Level – II Quality score – 2/4 Cons. Rec. - NR Blinded Y Objective - N Follow-up days 398^{M} (248–493) 100% follow-up	Population Ambulatory patients with congestive HF treated as outpatients Inclusion criteria After optimisation of medical therapy Exclusion criteria NR	Sample description N Mean age % male % IHD Disease severity $BNP^{STAB\ M}$ % NYHA I/II/III/IV LVEF $PeakVO_2$ SBP Creatinine	 78 51±9 (24–65) 89 31 105 (5–686) 3/42/33/12 36±15 15±4 	Comorbidities Atrial fibrillation Diabetes Hypertension COPD Medication usage Beta-blockers ACE-I/ ARB Diuretics Digitalis Spironolactone Amiodarone	% % 80 90 96 82 	Cardiac death or deterioration of physical activity (n=18)	Type IRMA (Shionoria BNP, CIS Biointernational, France) Intra-assay COV NR Inter-assay COV 7% @ 20 pg/mL 5% @ 291 pg/mL Time sample was drawn On entry into the study, after stabilisation

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
Koseki et al (2003) Tokyo, Japan Multicentre, CHART registry	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR Blinded N Objective - Y Follow-up days 365	Population CHF- dilated cardiomyopathy, Both primary and secondary myopathies, left ventricular pump failure Inclusion criteria LVEF ≥50%, left ventricular diastolic dimension (LVDD) ≥55 mm or past history of congestive HF Exclusion criteria Ischaemic cardiomyopathy, ischaemic heart disease, severe valvular heart disease, systemic hypertension and cor pulmonale	Sample description N 194 Mean age 82±14 % male 72.7 % IHD 0 Disease severity BNP STAB 226±419 % NYHA I/II/III/IV LVEF 38±12 PeakVO_2 SBP Creatinine	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE-I/ARB 100 Diuretics 80 Digitalis 55 Spironolactone Amiodarone	Death or readmission (n=31)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn After stabilisation for at least 3 weeks
		Population CHF- myocardial infarction Inclusion criteria Documented history of MI and/or typical findings of MI on ECG and echocardiography, ischaemic cardiomyopathy Exclusion criteria	Sample description N 163 Mean age 70±10 % male 75.5 % IHD 100 Disease severity BNP STAB 305±389 % NYHA I/II/III/IV LVEF 39±12 PeakVO_2 SBP Creatinine	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors 80 Diuretics 78 Digitalis 26 Spironolactone Amiodarone	Death or readmission (n=NR)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn After stabilisation for at least 3 weeks

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
		Population CHF- valvular heart disease Inclusion criteria Past history, (rheumatic fever and/or the long lasting heart murmur), physical examination and echo findings of mitral and/or aortic valve disorders, including those who had undergone surgical repair Exclusion criteria	Sample description N 245 Mean age 72±11 % male 60.4 % IHD 0 Disease severity BNP STAB 216±302 % NYHA I/II/III/IV LVEF 51±14 PeakVO2 SBP Creatinine	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors 61 Diuretics 86 Digitalis 64 Spironolactone Amiodarone	Death or readmission (n=NR)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn After stabilisation for at least 3 weeks
		Population CHF- left ventricular hypertrophy Inclusion criteria Hypertensive heart disease and hypertrophic cardiomyopathy Exclusion criteria	Sample description N 100 Mean age 71±14 % male 63 % IHD 0 Disease severity BNP STAB 280±304 % NYHA I/II/III/IV LVEF 55±15 PeakVO2 SBP Creatinine	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors 84 Diuretics 75 Digitalis 46 Spironolactone Amiodarone	Death or readmission (n=26)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn After stabilisation for at least 3 weeks
Kyuma et al (2004) Japan	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded NR Objective - Y Follow-up 16±9 months	Population Patients with CHF classified ass NYHA II-IV Inclusion criteria Recent history of HF and ability to undergo all testing procedures Exclusion criteria Valvular heart disease or congenital heart disease that required surgical repair	Sample description N 158 Mean age 64±13 % male 70 % IHD 28 Disease severity BNPSTAB 353±449 % NYHA I/II/III/IV 12/49/25/14 LVEF 41±17 PeakVO2 SBP	Comorbidities % Atrial fibrillation 40 Diabetes 28 Hypertension 9 COPD Medication usage % Beta-blockers 55 ACE-I/ARB 64 Diuretics 67 Digitalis 28 Spironolactone	Cardiac death via pump failure (n=15)	Type IRMA (Shionogi) Intra-assay COV NR Inter-assay COV NR Time sample was drawn On the same or within a few days after CHF had been stabilised

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
			Creatinine		Amiodarone	11		
Latini et al (2004) Multicentre, 302 centres; 16 countries Val-HeFT (Same cohort as Anand et al (2003)	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - N Blinded Y Objective - Y Follow-up 23 months	Population Participants in the Val-HeFT trial with stable, symptomatic HF Inclusion criteria History of HF for at least 3 months, LVEF <40% and LV internal diastolic diameter/body surface area ≥2.9 cm/m^2 Exclusion criteria <18 years	Sample description N 4305 Mean age 63±11 % male 80 % IHD 57 CHD Disease severity BNP STAB NR % NYHA I/II/III/IV 0/62/36/2 LVEF 27±7 PeakVO_2 SBP 124 Creatinine		Comorbidities % Atrial fibrillation 12 Diabetes Hypertension COPD Medication usage % Beta-blockers 36 ACE inhibitors 93 Diuretics 86 Digitalis Spironolactone Amiodarone	All-cause death (n=832) Death or morbid event, the latter defined as non-fatal cardiac arrest, HF hospitalisation or intravenous medication for more than 4 hours (n=NR; 30.5%)	Type IRMA (Shionogi, Schering CIS, Milan, Italy) Intra-assay COV Appropriate within and between laboratory controls were performed Time sample was drawn Prior to randomisation to an angiotensin receptor blocker
Logeart et al (2004) France	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 182	Population Admitted to cardiology department for severely decompensated CHF (NYHA IV) Inclusion criteria Clinical diagnosis (Framingham criteria) Exclusion criteria Acute myocardial infarction, severe valve disease, surgical patients, poor adherence to therapy	Sample description N 105 Mean age 69±14 % male 69 % IHD 39 Disease severity BNP ADM 1015±604 BNP STAB 457±451 % NYHA I/II/III/IV 0/0/0/100 LVEF 38±15 PeakVO_2 SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	Death or CV hospitalisation (n=51)	Type Triage BNP test (Biosite Diagnostics) Intra-assay COV NR Inter-assay COV NR Time sample was drawn On admission and serially until discharge
Maeda et al (2000) Otsu, Japan	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days	Population Patients admitted with decompensated CHF that were optimised for treatment in the following 3 months Inclusion criteria LVEF <45% and NYHA III-IV Exclusion criteria NR	Sample description N 102 Mean age 64±2 % male 63 % IHD 35 Disease severity BNP ADM BNP STAB % NYHA I/II/III/IV 0/0/57/45 LVEF 23±1		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 29 ACE inhibitors 71 Diuretics 79 Digitalis 55	Cardiac death (n=26) Cardiac death or readmission (n=47)	Type IRMA (Shionoria, Japan) Intra-assay COV NR Inter-assay COV NR Time sample was drawn On admission to hospital and

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
	807±42.3		PeakVO$_2$ SBP Creatinine		Spironolactone Amiodarone			3 months after treatment had been optimised
Maisel et al (2004) US Multicentre; 10 sites	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - NR Blinded Y Objective - Y Follow up days 30 and 90	Population Congestive HF Inclusion criteria BNP >100 pg/mL, aged over 18 years, presenting to ED with CHF and who received treatment Exclusion criteria Current myocardial infarction or acute coronary syndrome with ST-segment deviation of ≥1 mm, renal failure requiring dialysis, or patients with a baseline BNP concentration of ≤100 pg/mL	Sample description N 464 Mean age 64 (51–76) % male 54 % IHD 33 Disease severity BNP ADM % NYHA I/II/III/IV 3/29/45/23 LVEF PeakVO$_2$ SBP 141 [121–166] Creatinine		Comorbidities % Atrial fibrillation 25 Diabetes 41 Hypertension 77 COPD 22 Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	All-cause death (n=36) All-cause death or cardiac related hospital admission or ED visit (n=NR)	Type Triage BNP test (Biosite, San Diego, California) Intra-assay COV Previously described (5,12,17,18) Inter-assay COV NR Time sample was drawn Drawn during the ED visit
Matsui et al (2002) Otsu, Japan	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 731±65	Population Congestive HF resulting from dilated cardiomyopathy on optimised treatment Inclusion criteria LVEF <45% and NYHA II-IV, Dilated cardiomyopathy diagnosed from cardiac catheterisation, coronary angiography, endomyocardial biopsy Exclusion criteria Patients with ischaemic heart disease, diabetes mellitus, or autonomic failure related to neurologic disease	Sample description N 74 Mean age 55±1 % male 74 % IHD 0 Disease severity BNP STAB % NYHA I/II/III/IV 0/10/77/14 LVEF 31±1 PeakVO$_2$ SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes 0 Hypertension COPD Medication usage % Beta-blockers 20 ACE inhibitors 27 Diuretics 85 Digitalis 66 Spironolactone Amiodarone	Cardiac death (n=12) Cardiac death or readmission (n=23)	Type IRMA (Shionogi, Japan) Intra-assay COV Reported previously [5,31] Inter-assay COV Reported previously [5,31] Time sample was drawn Within 1 week of admission to hospital and after 6 months of optimal drug treatment
Sakatani et al (2004)	Prospective cohort study	Population Patients admitted too hospital for CHF, who were subsequently stabilised	Sample description N 70 Mean age 72±12 (33–91) % male 49		Comorbidities % Atrial fibrillation 24 Diabetes 27 Hypertension 41		Cardiac death or readmission (n=18)	Type IMRA (3–4) Intra-assay COV

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
Kyoto, Japan	Level - II Quality score – 2/4 Cons. Rec. - Y Blinded NR Objective - Y Follow p days 510±270	Inclusion criteria Clinical diagnosis, fatigue, exertional dyspnoea, orthopnea, activity limitation, a third heart sound, elevated jugular venous pressure, rales and leg oedema Exclusion criteria Infection, trauma, surgery, MI or glucocorticoid use within 8 weeks, malignancy, radiation therapy, chemotherapy	% IHD 26 Weighted Disease severity BNP STAB 310 Weighted % NYHA I/II/III/IV 6/43/43/9 LVEF PeakVO_2 SBP 125 Weighted Creatinine		COPD Medication usage % Beta-blockers 20 ACE inhibitors 54 Diuretics 83 Digitalis Spironolactone Amiodarone			NR Inter-assay COV NR Time sample was drawn After stabilisation in hospital
Selvais et al (2000) Brussels, Belgium	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - NR Blinded Y Objective - Y Follow-up days Up to 36 months	Population Congestive HF Inclusion criteria LVEF <35% and NYHA II-IV Exclusion criteria NR	Sample description N 109 Mean age 63 Weighted % male 84 % IHD 88 Disease severity BNP NR NR % NYHA I/II/III or IV 0/60/40 LVEF 26 Weighted PeakVO_2 SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 16 ACE inhibitors 96 Diuretics 40 Digitalis 13 Spironolactone Amiodarone	Cardiac death (n=32)	Type Radio-immunoassays Intra-assay COV 7% @ 350 pg/mL Inter-assay COV 6% @ 350 pg/mL (not a misprint) Time sample was drawn Not described; assumedly in stabilised state
Setsuta et al (2002) Tokyo, Japan	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 480±360 (150–1050)	Population Patient with CHF Inclusion criteria NYHA II-IV Exclusion criteria Patients who had a history of recent myocardial infarction (within 3 months), angina pectoris, myocarditis, renal failure (serum creatinine level >2.0 mg/mL), active hepatic or pulmonary disease, or elevated levels of creatine kinase	Sample description N 56 Mean age 66 Weighted % male NR % IHD Disease severity BNP STAB 366 Weighted % NYHA I/II/III/IV 0/71/25/4 LVEF 38 Weighted PeakVO_2 SBP Creatinine (mg/dl) 1.1 Weighted		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	Cardiac death or readmission (n=18)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn Blood samples were drawn in patients with stable CHF

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
		(>250 IU/l) were excluded. Active myocardial ischaemia were also excluded				
Shiba et al (2004) Multicentre – 26 hospitals in Tohoku region; registry Sendai, Japan	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR Blinded NR Objective - Y Follow-up days 386±336 97.4% (1 year) 93.0% (2 years) 87.4% (3 years)	Population Chronic HF Inclusion criteria LVEF <50% or LV end-diastole diameter ≥55 mm, or at least one episode of HF Exclusion criteria Less than 18 years of age or clinically unstable	Sample description N 1154 (684[Z]) Mean age 68±13 % male 66.5 % IHD Disease severity BNP [STAB] 265±349 % NYHA I/II/III/IV 21/63/15/1 LVEF 49±16 PeakVO_2 SBP Creatinine	Comorbidities % Atrial fibrillation Diabetes 19 Hypertension 39 COPD Medication usage % Beta-blockers 25 ACE-I/ARB 69 Diuretics Digitalis Spironolactone 21 Amiodarone 3	All-cause death (n=107)	Type NR Intra-assay COV NR Inter-assay COV NR Time sample was drawn At entry into registry – all patients were stabilised
Tamura et al (2001) Japan	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 324±33 (30–750)	Population CHF >65 years of age Inclusion criteria Clinical diagnosis Exclusion criteria Chronic inflammatory disease, carcinoma, or renal failure	Sample description N 48 Mean age 78±1 (67–92) % male 48 % IHD 38 Disease severity BNP [STAB] 225 % NYHA I/II/III/IV 25/38/29/8 LVEF 46 [Weighted] PeakVO_2 SBP 120 [Weighted] Creatinine	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 17 ACE inhibitors 38 Diuretics 75 Digitalis 31 Spironolactone 27 Amiodarone	Cardiac death or hospitalisation (n=12)	Type RIA, Shionoria BNP assay kit; Shionogi Intra-assay COV NR Inter-assay COV NR Time sample was drawn Just prior to hospital discharge
Tsutamoto et al (1997) Otsu, Japan	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR	Population Hospitalised for chronic CHF Inclusion criteria LVEF <45% Exclusion criteria Infection, chronic inflammatory disease, malignancy, or renal	Sample description N 85 Mean age 60±1 (22–84) % male 72 % IHD 41 Disease severity BNP [STAB]	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 25	Cardiac death (n=25)	Type IRMA (Shionoria) Intra-assay COV 5.2% Inter-assay COV 6.1%

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
	Blinded N Objective - Y Follow-up days 720	failure	% NYHA I/II/III/IV 0/54/21/25 LVEF PeakVO$_2$ SBP Creatinine		ACE inhibitors 68 Diuretics 85 Digitalis 75 Spironolactone Amiodarone			Time sample was drawn All patients were clinically stable
Tsutamoto et al (1999) Otsu, Japan	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 812^{M}	Population Patients with asymptomatic or newly symptomatic left ventricular dysfunction who underwent cardiac catheterisation for clinical indications Inclusion criteria LVEF <45% and NYHA I-II Exclusion criteria Patients who had infection, chronic inflammatory disease, malignancy or renal failure	Sample description N 290 Mean age 59 (18–82) % male 77 % IHD 64 Disease severity BNP $^{STAB\ M}$ 56 % NYHA I/II/III/IV 32/68/0/0 LVEF 37±1 PeakVO$_2$ SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 23 ACE inhibitors 53 Diuretics 48 Digitalis 34 Spironolactone Amiodarone	Cardiac death or readmission (n=49) Cardiac death (n=24)	Type IRMA (Shionogi, Japan) Intra-assay COV NR Inter-assay COV NR Time sample was drawn All patients were clinically stable
Tsutamoto et al (2001) Otsu, Japan	Prospective cohort study Level - II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 1147 (682–1798)	Population Patients with symptomatic left ventricular dysfunction Inclusion criteria LVEF <45% and NYHA II-IV Exclusion criteria Patients who had signs of renal failure, infection, malignancy or collagen disease	Sample description N 96 Mean age 59 (23–79) % male 79 % IHD 42 Disease severity BNP ^{STAB} % NYHA I/II/III/IV 0/65/23/13 LVEF 28.8 ^{Weighted} PeakVO$_2$ SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes Hypertension COPD Medication usage % Beta-blockers 28 ACE inhibitors 79 Diuretics 78 Digitalis 72 Spironolactone Amiodarone	Cardiac death (n=29)	Type IRMA (Shionogi, Japan) Intra-assay COV 5.2% Inter-assay COV 6.1% Time sample was drawn All patients were clinically stable
Tsutsui et al (2002) Otsu, Japan	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y	Population Chronic congestive HF Inclusion criteria LVEF <45% and NYHA II-IV Exclusion criteria Patients with infection,	Sample description N 84 Mean age 63±2 % male 75 % IHD 58 Disease severity		Comorbidities % Atrial fibrillation Diabetes 18 Hypertension 29 COPD Medication usage %	Cardiac death (n=14) Cardiac death or rehospitalisation (n=12)	Type IRMA (Shionogi, Japan) Intra-assay COV NR Inter-assay COV

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
	Blinded Y Objective - Y Follow-up days 780±16 (589–984)	inflammatory diseases, malignancy, renal failure, congenital malformations of the heart or vessels, angina pectoris or a history of acute myocardial infarction within the past 3 months	BNP NR 334±42 % NYHA I/II/III/IV 0/63/31/5 LVEF 31±1 PeakVO2 SBP Creatinine		Beta-blockers 37 ACE-I/ARB 82 Diuretics 69 Digitalis 40 Spironolactone Amiodarone			NR Time sample was drawn Not stated – on entry into the study in presumably stabilised patients
Van der Meer et al (2004)	Retrospective cohort study Level – II Quality score – 2/4 Cons. Rec. - Y Blinded N Objective - Y Follow-up days 1100 (844–1934)	Population Stable (for 3 months) mild to advanced CHF patients admitted to tertiary referral centre Inclusion criteria ESC guidelines Exclusion criteria Isolated diastolic dysfunction, valvular disease, MI (within 12 weeks), cerebrovascular accident (within 12 weeks), or severe renal failure	Sample description N 74 Mean age 61±2 % male 73 % IHD 47 Disease severity BNP STAB 380±47 % NYHA I/II/III/IV 0/37/32/31 LVEF 31±1 PeakVO2 19± 2 SBP Creatinine		Comorbidities % Atrial fibrillation Diabetes 12 Hypertension 41 COPD Medication usage % Beta-blockers 68 ACE inhibitors 81 Diuretics 64 Digitalis Spironolactone Amiodarone	All-cause death (n=22)	Type IRMA (Shionoria, Japan) Intra-assay COV NR Inter-assay COV NR Time sample was drawn NR
Verdiani et al (2005) Florence, Italy	Retrospective cohort study Level – II Quality score – 4/4 Cons. Rec. - Y Blinded Y Objective - Y Follow-up days 30 days	Population Patients admitted to hospital with decompensated HF Inclusion criteria Exacerbation of previous HF or new onset as defined by Framingham criteria, NYHA III or IV on admission Exclusion criteria Presence on non-cardiac illness which could affect short term prognosis	Sample description N 100 Mean age 78±10 % male 71 % IHD 62 Disease severity BNP STAB M 739 [355–1333] BNP DIS M 414 [220–696] % NYHA I/II/III or IV 0/0//100 LVEF 38±11 PeakVO2 SBP Creatinine 1.4±0.7		Comorbidities % Atrial fibrillation NR Diabetes 25 Hypertension 50 COPD 25 Medication usage % Beta-blockers 22 ACE-I/ARB 94 Diuretics 95 Digitalis Spironolactone Amiodarone	Readmission or death (n=17)	Type IRMA (Shionoria, Japan) Intra-assay COV NR Inter-assay COV NR Time sample was drawn Within 2 hours of admission and just before discharge
Vrtovec et al (2003) US,	Cohort study (not recorded if retrospective or prospective)	Population A severe HF group selected from patients in an outpatient HF clinic Inclusion criteria >400 pg/mL BNP and NYHA III-IV	Sample description N 241 Mean age 67±14 % male 59 % IHD 66		Comorbidities % Atrial fibrillation 0 Diabetes Hypertension COPD		All-cause death (n=46) Cardiac death (sudden death or pump failure	Type Triage BNP test (Biosite Diagnostics) Intra-assay COV

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
Texas	Level - II Quality score – 1/4 Cons. Rec. - N Blinded N Objective - Y Follow-up days – 182	for at least 2 months <u>Exclusion criteria</u> Patients with pacemakers or implantable cardioverter-defibrillators, atrial fibrillation, pacemaker rhythm, and patients taking type III antiarrhythmic drugs	<u>Disease severity</u> BNP STAB M 850 % NYHA I/II/III/IV 0/0/74/26 LVEF 27±9 PeakVO_2 SBP 115±22 Creatinine 1.44±0.7 mg/dL		<u>Medication usage %</u> Beta-blockers 73 ACE inhibitors 87 Diuretics 97 Digitalis Spironolactone Amiodarone		death) (n=42) Sudden cardiac death (n=18) Pump failure death (multi-organ failure caused by HF progression) (n=24)	NR <u>Inter-assay COV</u> NR <u>Time sample was drawn</u> At the start of the study ('at the time of evaluation')
Watanabe et al (2005) Japan Multicentre N=417 CHART – Chronic Heart failure Analysis and Registration in the Tohoku district	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - NR Blinded NR Objective - Y Follow-up days 780	<u>Population</u> Hospital based sample of CHF patients with diastolic and systolic origin <u>Inclusion criteria</u> CHF diagnosed via Framingham criteria or LVEF <50% or LVDD >55 mm – patients had to be stable for at least 3 weeks <u>Exclusion criteria</u> NR	<u>Sample description</u> N 417 Mean age 64±14 % male 69 % IHD 48 <u>Disease severity</u> BNP STAB 274±380 % NYHA I or II/III or IV 81/19 LVEF 38±12 PeakVO_2 SBP Creatinine		<u>Comorbidities %</u> Atrial fibrillation Diabetes 22 Hypertension COPD <u>Medication usage %</u> Beta-blockers 43 ACE-I/ARB 77 Diuretics Digitalis Spironolactone Amiodarone	All-cause mortality (n=66) HF death or readmission (n=74)	<u>Type</u> RIA (Shionoria; CIS, France) <u>Intra-assay COV</u> NR <u>Inter-assay COV</u> NR <u>Time sample was drawn</u> At registration, in a stabilised CHF sample
Weinberg et al (2002) US and Canada Multicentre; 26 centres PRAISE-2	Prospective cohort study Level - II Quality score -1/4 Cons. Rec. - N Blinded N Objective - Y	<u>Population</u> Severe chronic HF of non-ischaemic pathogenesis <u>Inclusion criteria</u> LVEF <30% and NYHA III-IV <u>Exclusion criteria</u> Over 18 years of age and had HF of non-ischaemic pathogenesis, symptoms at rest or on minimal	<u>Sample description</u> N 161 Median age 60 % male 73 % IHD 0 <u>Disease severity</u> BNP STAB 56 % NYHA I/II/III/IV 0/0/88/12 LVEF 22 (11–30)		<u>Comorbidities %</u> Atrial fibrillation Diabetes Hypertension 15 COPD <u>Medication usage %</u> Beta-blockers ACE inhibitors Diuretics	Death or transplantation (n=NR)	<u>Type</u> Previously described [11,12] <u>Intra-assay COV</u> NR <u>Inter-assay COV</u> NR <u>Time sample was drawn</u> After stabilisation on ACE

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
substudy	Follow-up days NR	exertion and LVEF <30%. All patients undergoing treatment with angiotensin-converting enzyme (ACE) inhibitors and digoxin for at least 3 months	$PeakVO_2$ SBP Creatinine	 1.1 (0.8–1.9)	Digitalis Spironolactone Amiodarone			inhibitors and digoxin for 3 months but before randomisation to amlodipine
Wijeysundera et al (2003) US and Canada Multicentre; 26 centres PRAISE-2 substudy	Prospective cohort study Level - II Quality score – 1/4 Cons. Rec. - N Blinded N Objective - Y Follow-up days 711	Population Non-ischaemic cardiomyopathy Inclusion criteria HF of non-ischaemic aetiology, NYHA III-IV and LVEF <30%, over 18 years, on treatment with ACE inhibitors and digoxin for at least 3 months Exclusion criteria Recent or remote history of angina. History of MI, coronary bypass surgery, or coronary angioplasty, active myocarditis, known congenital heart disease, sudden death, or untreated sustained or symptomatic ventricular tachycardia. Severe haematological, primary renal, hepatic, endocrine (other than diabetes mellitus), collagen vascular (other than rheumatoid arthritis), or neurological disease. Treatment with angiotensin-II receptor antagonists, calcium-channel blockers, oestrogen, pentoxyfilline or beta-blockers within 4 weeks	Sample description N Mean age % male % IHD Disease severity BNP STAB % NYHA I/II/III/IV LVEF $PeakVO_2$ SBP Creatinine	 181 58±13 74 0 0/0/88/12 22±6	Comorbidities % Atrial fibrillation Diabetes Hypertension 14 COPD Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone		All-cause death (n=53)	Type Previously described Intra-assay COV NR Inter-assay COV NR Time sample was drawn After stabilisation on ACE inhibitors and digoxin for 3 months but before randomisation to amlodipine
Yu & Sanderson (1999) Hong Kong,	Prospective cohort study Level - II	Population Acute HF Inclusion criteria LVEF <50% and clinical features	Sample description N Mean age % male % IHD	 91 62±1 68 52	Comorbidities % Atrial fibrillation Diabetes Hypertension COPD		Cardiac death (n=22)	Type IRMA (in-house measurement) Intra-assay COV NR

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
China	Quality score – 2/4 Cons. Rec. - Y Blinded - N Objective - Y Follow-up days 365[M] 94.5% follow-up	Exclusion criteria Patients with significant systemic disease, major organ failure or malignancy were excluded	Disease severity BNP [STAB M] % NYHA I/II/III/IV LVEF $PeakVO_2$ SBP Creatinine	 165 0/47/47/6 35±1	Medication usage Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	% 21 84 98		Inter-assay COV NR Time sample was drawn After pharmaceutical stabilisation, usually 7–10 days after admission to hospital

Level = level of evidence; Y = yes; N = no; Cons. Rec.= Was their consecutive patient recruitment?; Blinded = Was the outcome determined by assessors blinded to peptide concentration and other potentially prognostic variables?; Objective = Was the assessment of outcome objective?; Follow-up has been converted to days using the assumption that there are 30 days in each month; % gender/disease/medication = percentage of the sample which were male, had the specified disease or were taking the specified medication; HF = heart failure; IHD = ischaemic heart disease; NT-proBNP = Plasma concentration of N-terminal B-type natriuretic peptide (mg/dL); NYHA = New York heart association classification; LVEF = left ventricular ejection fraction; $PeakVO_2$ = peak oxygen consumption ($mLO_2/kg/min$); SBP = systolic blood pressure (mmHg); Creatinine = as a marker of renal function (μmol/l) AF = atrial fibrillation; CTx = cardiac transplantation; COPD = chronic obstructive pulmonary disease; [DIS] = discharge values; [M] = median; [Z] = a subset, with a complete baseline data workup, of the larger cohort was used for multivariate regression analyses; round brackets ( ) enclose a range; square brackets [ ] enclose an inter-quartile range; COV = coefficient of variation; CHD = coronary heart disease.

Studies included for NT-proBNP assays

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
Bettencourt et al (2004) Porto, Portugal	Prospective cohort study Level - II Quality score - 3/4 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up days 180	Population Patients admitted to Internal Medicine department of Servico de Medicina Hospital with decompensated HF between Oct 2002 and Mar 2003 Inclusion criteria Not specified Exclusion criteria Patients with acute coronary syndromes	Sample description N 156 Mean age 73±11 % male 47 % IHD 47 Disease severity NT-proBNP ADM M 6779 NT-proBNP STAB M 4137 % NYHA I/II/III/IV 0/0/33/67 % LVEF <25% 23 % 25%<LVEF <45% 40 % LVEF>45% 20 PeakVO2 SBP 129 Creatinine 124	Comorbidities % AF 46 Diabetes 52 Hypertension 44 COPD Medication usage % Beta-blockers 39 ACE inhibitors 87 Diuretics 98 Digitalis Spironolactone 37 Amiodarone	Death (n=28) Death or readmission (n=67)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV 0.9% @ 474 pg/mL 1.1% @ 8005 pg/mL 0.9% @ 13682 pg/mL Inter-assay COV Not stated Time sample was drawn Within 24 hours of hospital admission and before discharge
De Pasquale et al (2004) South Australia, Australia	Prospective cohort study Level - II Quality score - 1/4 Cons. Rec. - N Blinded - NR Objective - Y Follow-up days 540	Population Patients from the Flinders Medical Centre HF clinic Inclusion criteria Not specified Exclusion criteria Primary lung disease, with the exclusion of COPD, or an inability to provide consent	Sample description N 53 Mean age 67M % male % IHD 70 Disease severity NT-proBNP STAB % NYHA I/II/III/IV 0/32/42/26 LVEF PeakVO2 SBP 129 Creatinine 124	Comorbidities % AF Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	Death or readmission (n=25)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn Sample drawn in a stabilised population of HF patients

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
Fisher et al (2003) Glasgow, United Kingdom	Prospective cohort study Level - II Quality score - 2/4 Cons. Rec. - N Blinded - Y Objective - Y Follow-up days 365	Population Patients taking part in RCT of nurse intervention Inclusion criteria Not stated Exclusion criteria Not stated	Sample description N 87 Mean age 75 % male 59 % IHD Disease severity NT-proBNP STAB M 2994 (134–35000) % NYHA I/II/III/IV 0/24/32/44 LVEF PeakVO2 SBP Creatinine 132±7	Comorbidities % AF 29 Diabetes 15 Hypertension COPD 28 Medication usage % Beta-blockers 10 ACE inhibitors 43 Diuretics 69 Digitalis Spironolactone Amiodarone	Death (n=28) Death or readmission with HF (n=42)	Type Immunoassay (Roche) Intra-assay COV Not reported Inter-assay COV Not reported Time sample was drawn Just prior to discharge
Hartmann et al (2004) Multicentre, Europe	Prospective cohort study Level - II Quality score - 1/4 Cons. Rec. - N Blinded - NR Objective - Y Follow-up days 157M(1–488)	Population Patients with chronic HF (ischaemic or non-ischaemic cardiomyopathy), being randomised to treatment via Carvedilol or placebo Inclusion criteria All patients were pretreated and stabilised with diuretics and ACE inhibitors for at least 2 months Exclusion criteria Patients with substantial fluid retention need for intensive care, or treatment with IV inotropic agents or vasodilators within 4 days of screen for NT-proBNP	Sample description N 1048 Mean age 63±11 % male 81 % IHD 66 Disease severity NT-proBNP STAB M 2727 [1277–5920] % NYHA I/II/III/IV LVEF 20±4 PeakVO2 SBP 127±19 Creatinine	Comorbidities % AF Diabetes Hypertension COPD Medication usage % Beta-blockers 50 ACE inhibitors 93 Diuretics 99 Digitalis 58 Spironolactone 26 Amiodarone 18	All-cause death (n=83) All-cause death or readmission with HF (n=187)	Type Roche (ELISA prototype) Intra-assay COV 5.7% @ 423 pg/mL 6.1% @ 2114 pg/mL Inter-assay COV 15.8% @ 127 pg/mL 8.2% @ 2114 pg/mL Time sample was drawn On the day of randomisation to Carvedilol or placebo treatment (ie after 2 months of treatment)
Hartmann et al (2004) Multicentre, Europe	Prospective cohort study Level - II Quality score - 1/4	Population Patients with chronic HF (ischaemic or non-ischaemic cardiomyopathy), being randomised to treatment via Carvedilol or placebo	Sample description N 1011 Mean age 63±11 % male 81 % IHD 66 Disease severity	Comorbidities % AF Diabetes Hypertension COPD Medication usage %	All-cause death (n=78) All-cause death or readmission with HF	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
	Cons. Rec. - N Blinded - NR Objective - Y Follow-up days 159[M] (1–488)	Inclusion criteria All patients were pre-treated with diuretics and ACE inhibitors for at least 2 months Exclusion criteria Patients with substantial fluid retention need for intensive care, or treatment with IV inotropic agents or vasodilators within 4 days of screen for NT-proBNP	NT-proBNP [STAB M] 1767 [748–3927] % NYHA I/II/III/IV LVEF 20±4 PeakVO2 SBP 127±19 Creatinine Beta-blockers 50 ACE inhibitors 92 Diuretics 99 Digitalis 59 Spironolactone 26 Amiodarone 18	(n=180)	Intra-assay COV 2.4% @ 355 pg/mL 1.8% @ 4962 pg/mL Inter-assay COV 2.9% @ 355 pg/mL 2.3% @4962 pg/mL Time sample was drawn On the day of randomisation to Carvedilol or placebo treatment (ie after 2 months of treatment)
Gardner et al (2005) Glasgow, United Kingdom	Prospective cohort study Level - II Quality score - 3/4 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up days 554[M] (1–1115)	Population Patients referred to the Scottish Cardiopulmonary Transplant Unit for cardiac transplant assessment between April 2001 and April 2004 Inclusion criteria Patients with CHF secondary to a LV systolic dysfunction (LVEF ≤35%) in NYHA II-IV Exclusion criteria Age <16 years, pregnancy or concurrent malignancy	Sample description N 182 Mean age 51±11 % male 79 % IHD 45 Disease severity NT-proBNP [ADM M] 1505 [517–4015] % NYHA I/II/III/IV 0/18/63/20 LVEF 15±7 PeakVO2 11±4 SBP 110±19 Creatinine 107±54 Comorbidities % AF 18 Diabetes Hypertension COPD Medication usage % Beta-blockers 72 ACE-I/ARB 100 Diuretics Digitalis Spironolactone 61 Amiodarone	All-cause death (n=30) All-cause death or urgent heart transplantation (n=34)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV NR Inter-assay COV NR Time sample was drawn On the first screening visit to the cardiopulmonary transplant unit
Gardner et al (2005) Glasgow, United Kingdom	Prospective cohort study Level - II Quality score - 3/4 Cons. Rec. - Y Blinded - Y Objective - Y	Population Patients referred to the Scottish Cardiopulmonary Transplant Unit for cardiac transplant assessment between April 2001 and 2003 Inclusion criteria Patients with CHF secondary to a LV systolic dysfunction (LVEF	Sample description N 150 Mean age 50±10 % male 83 % IHD 31 Disease severity NT-proBNP [ADM M] 1494 [530-3930] % NYHA I/II/III/IV 0/14/66/19 LVEF 15±7 Comorbidities % AF 19 Diabetes Hypertension COPD Medication usage % Beta-blockers 70 ACE inhibitors 79 Diuretics	All-cause death (n=25) All-cause death or urgent heart transplantation (n=29)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV 6% maximum Inter-assay COV

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
	Follow-up days 666[M] (1–1047)	≤35%) in NYHA II-IV Exclusion criteria Age <16 years, pregnancy or concurrent malignancy	PeakVO$_2$ 12±4 SBP Creatinine		Digitalis Spironolactone 57 Amiodarone			6% maximum Time sample was drawn On the first screening visit to the cardiopulmonary transplant unit
Gardner et al (2005) Glasgow, United Kingdom	Prospective cohort study Level - II Quality score - 3/4 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up days 370[M] (1–660)	Population Patients referred to the Scottish Cardiopulmonary Transplant Unit for cardiac transplant assessment between April 2001 and 2003 Inclusion criteria Patients with CHF secondary to a LV systolic dysfunction (LVEF ≤35%) in NYHA II-IV Exclusion criteria Age <16 years, pregnancy or concurrent malignancy	Sample description N 97 Mean age 51±11 % male 87 % IHD 51 Disease severity NT-proBNP [STAB M] 1548 [604–4127] % NYHA I/II/III/IV 0/11/62/27 LVEF 14±7 PeakVO$_2$ 11±3 SBP Creatinine		Comorbidities % AF 14 Diabetes Hypertension COPD Medication usage % Beta-blockers 70 ACE inhibitors 77 Diuretics Digitalis Spironolactone 61 Amiodarone	All-cause death (n=17) All-cause death or urgent heart transplantation (n=21)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn On the first screening visit to the cardiopulmonary transplant unit
Gardner et al (2003) Glasgow, United Kingdom	Prospective cohort study Level - II Quality score – 3/4 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up days 374[M] (1–660)	Population Patients referred to the Scottish Cardiopulmonary Transplant Unit for cardiac transplant assessment between April 2001 and Dec 2002 Inclusion criteria Patients with CHF secondary to a LV systolic dysfunction (LVEF ≤35%) in NYHA II-IV Exclusion criteria Age <16 years, pregnancy or concurrent malignancy	Sample description N 142 Mean age 50±11 % male 82 % IHD 46 Disease severity NT-proBNP [STAB M] 1490 [511–3887] % NYHA I/II/III/IV 0/15/66/19 LVEF 15±7 PeakVO$_2$ 12±4 SBP 107 [98–120] Creatinine 120 [100–140]		Comorbidities % AF 18 Diabetes Hypertension COPD Medication usage % Beta-blockers 69 ACE inhibitors 79 Diuretics Digitalis Spironolactone 60 Amiodarone	All-cause death (n=20) All-cause death or urgent heart transplantation (n=24)	Type Electrochemiluminescence immunoassay (Roche GmbH), analysed by the Elecsys 2010 analyser Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn On the first screening visit to the cardiopulmonary transplant unit
George et al	Prospective cohort	Population	Sample description		Comorbidities %		All-cause death	Type

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics				Outcome (number of outcomes)	Assay details
(2005) Tel Aviv, Israel	study Level - II Quality score - 3/4 Cons. Rec. - Y Blinded - Y Objective - Y Follow-up days 374[M] (1–660)	Patients with advanced CHF attending an outpatient clinic Inclusion criteria Patients with CHF in NYHA II to IV Exclusion criteria Age <18 yrs, pregnancy, therapeutic use of EPO, known concurrent malignancy	N Mean age % male % IHD Disease severity NT-proBNP [STAB] % NYHA I/II/III/IV LVEF PeakVO_2 SBP Creatinine	188 71±12 77 73 1881 [Weighted] 2.8±0.6 38±14 45% with renal failure	AF Diabetes Hypertension COPD Medication usage Beta-blockers ACE-I/ARB Diuretics Digitalis Spironolactone Amiodarone	31 34 59 % 69 78 72 42 	(n=38) All-cause death or hospitalisation due to CHF (n=67)	Automated immunoassay (Elecsys proBNP test; Roche) Intra-assay COV NR Inter-assay COV NR Time sample was drawn Baseline serum samples were drawn at the initial visit
Kirk et al (2004) Copenhagen, Denmark	Prospective cohort study Level - II Quality score - 2/4 Cons. Rec. - N Blinded - Y Objective - Y Follow-up days 365	Population Patients admitted to Amager Hospital (general city hospital) between April 1998 and March 1999 Inclusion criteria Patients over the age of 40 years Exclusion criteria Discharged within 24 hours after admission (n=155); death before inclusion (n=56); mental or physical status not allowing examination (n=68); refused consent (n=129); echographic findings suggesting immediate intervention (n=13)	Sample description N Mean age % male % IHD Disease severity NT-proBNP [ADM M] % NYHA I/II/III/IV LVEF PeakVO_2 SBP Creatinine	 161 78 53 48 3797 (9–77762) 45±1 146±3 	Comorbidities AF Diabetes Hypertension COPD Medication usage Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone	% 16 13 % 	All-cause death (n=51)	Type ELISA Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn The day after admission to hospital
Kellett (2005) Tipperary, Ireland	Prospective cohort study Level - II Quality score - 3/4	Population Patients admitted as acute medical emergencies with suspected heart disease based on clinical criteria, chest X-ray and ECG	Sample description N Mean age % male % IHD	 342 74±11 54 7	Comorbidities AF Diabetes Hypertension COPD	% 17 12 	Within-hospital stay death (n=31)	Type Elecsys NT-proBNP assay (Roche) Intra-assay COV NR

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics		Outcome (number of outcomes)	Assay details
	Cons. Rec. - Y Blinded - NR Objective - Y Follow-up days Within-hospital stay; 9±7 days	Inclusion criteria NR Exclusion criteria NR	Disease severity NT-proBNP ADM 4510 Weighted % NYHA I/II/III/IV LVEF 45±1 PeakVO2 SBP 146±3 Creatinine	Medication usage % Beta-blockers ACE inhibitors Diuretics Digitalis Spironolactone Amiodarone		Inter-assay COV NR Time sample was drawn Within 2 hours of admission
O'Brien et al (2003) Leicester, United Kingdom	Prospective cohort study Level - II Quality score - 2/4 Cons. Rec. - Y Blinded - NR Objective - Y Follow-up days 350M (2–762)	Population Clinical diagnosis of acute LVF, the severity meriting coronary care management Inclusion criteria 96 consecutive patients (no particular exclusion criteria, so to reflect normal clinical practice) Exclusion criteria Patients with acute MI, diagnosed via ECG or creatine kinase levels two times normal levels	Sample description N 96 Mean age 74 % male 53 % IHD 48 Disease severity NT-proBNP ADM M 2532 (12–24837) NT-proBNP STAB (N=34) 1644 (12–6897) % NYHA I/II/III/IV LVEF PeakVO2 SBP Creatinine 114	Comorbidities % AF Diabetes 23 Hypertension 44 COPD Medication usage % Beta-blockers 43 ACE inhibitors 70 Diuretics 96 Digitalis Spironolactone Amiodarone	Death or readmission with HF (n=37)	Type ELISA, non-competitive technique Intra-assay COV 2.3% Inter-assay COV 4.8% Time sample was drawn Immediately on arrival at the coronary care unit and again when the patient was stabilised and ready for discharge home
Richards et al (2001) Multicentre, Australia and New Zealand	Prospective cohort study Level – II Quality score - 2/4 Cons. Rec. - Y Blinded - NR Objective - Y Follow-up days 540	Population Patients with chronic stable HF caused by IHD, LVEF <45%, current NYHA functional class II or III or previous functional class II-IV who were randomised to Carvedilol or placebo treatment Inclusion criteria Not stated Exclusion criteria NYHA functional class IV, HR <50 bpm, second or third degree heart block, blood pressure issues,	Sample description N 297 Mean age % male % IHD Disease severity NT-proBNP STAB 837 [465–1514] % NYHA I 30 %NYHA II or III 70 LVEF 29 [22–35] PeakVO2 SBP Creatinine	Comorbidities % AF Diabetes Hypertension COPD Medication usage % Beta-blockers ACE inhibitors 85 Diuretics 75 Digitalis Spironolactone Amiodarone	All-cause mortality (n=35) Admission with HF (n=41)	Type Published methods (11) Intra-assay COV Inter-assay COV Time sample was drawn Sample drawn in a stabilised population of HF patients

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics	Outcome (number of outcomes)	Assay details
		treadmill exercise duration <2 or >18 minutes, coronary event within 4 weeks, primary myocardial or valve disease, diabetes, chronic airway disease, hepatic disease, renal impairment (creatinine >250 µmol/L) and life-threatening non-cardiac disease or current treatment with a beta-blocker, beta-agonist or Verapamil			
Rossig et al (2004) Germany	Prospective cohort study Level - II Quality score - 1/4 Cons. Rec. - N Blinded - NR Objective - Y Follow-up days 1254	Population Patients either admitted to, or current patients of, an outpatient HF clinic between June 1998 and August 2000 Inclusion criteria Not specified Exclusion criteria Patients with a history of MI within 3 months of blood sampling, renal insufficiency (>191 µmol/L) or concomitant infectious or primary pulmonary disease	Sample description N 48 Mean age 57±1 % male 77 % IHD 46 Disease severity NT-proBNP STAB 3666±595 % NYHA I/II/III/IV 0/42/44/15 LVEF 25±1 PeakVO2 SBP Creatinine Comorbidities % AF 25 Diabetes Hypertension COPD Medication usage % Beta-blockers 42 ACE inhibitors 100 Diuretics 90 Digitalis 75 Spironolactone 23 Amiodarone 38	All-cause mortality (n=16)	Type ELISA (Roche) Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn Sample drawn at start of study in a stabilised HF population
Zugck et al (2002) Heidelberg, Germany	Prospective cohort study Level - II Quality score - 2/4 Cons. Rec. - Y Blinded - NR Objective - Y Follow-up days 365	Population Patients referred to the Department of Cardiology for assessment of their HF status and/or evaluation of their potential candidacy for heart transplantation between Nov 1994 to Jan 2000 165 patients have stable beta-blocker treatment; 243 patients have no beta-blocker treatment Inclusion criteria Treatment with ACE inhibitor or	Sample description N 408 Mean age 55±11 % male 84 % IHD 70 Disease severity NT-proBNP STAB 3341±4026 NYHA 2.3±0.7 LVEF 22±10 PeakVO2 14±5 SBP Creatinine Comorbidities % % AF % Diabetes % Hypertension % COPD Medication usage % Beta-blockers 40 ACE inhibitors 97 Diuretics 86 Digitalis 70 Spironolactone Amiodarone 7	Cardiac death or readmission due to worsening HF (n=119)	Type ELISA Intra-assay COV Not stated Inter-assay COV Not stated Time sample was drawn Sample drawn in a stabilised population of HF patients

Study Location	Study design	Population Inclusion criteria Exclusion criteria	Patient and study characteristics	Outcome (number of outcomes)	Assay details
		AT1 receptor agonist. All patients had to be in a stable condition for at least 4 weeks, with medication individually optimised <u>Exclusion criteria</u> Not specified			

Level = level of evidence; Y = yes; N = no; Cons. Rec. = Was their consecutive patient recruitment?; Blinded = Was the outcome determined by assessors blinded to peptide concentration and other potentially prognostic variables?; Objective = Was the assessment of outcome objective?; Follow-up has been converted to days using the assumption that there are 30 days in each month; % gender/disease/medication; % male = percentage of the sample which were male, had the specified disease or were taking the specified medication; IHD = ischaemic heart disease; NT-proBNP = Plasma concentration of N-terminal B-type natriuretic peptide (mg/dL); NYHA = New York Heart Association classification; LVEF = left ventricular ejection fraction; PeakVO_2 = peak oxygen consumption (mLO_2/kg/min); SBP = systolic blood pressure (mmHg); Creatinine = as a marker of renal function (μmol/L)AF = atrial fibrillation; COPD = chronic obstructive pulmonary disease; ^{DIS} = discharge values; ^{M} = median; round brackets ( ) enclose a range; square brackets [ ] enclose an inter-quartile range; COV = coefficient of variation; LA = last available NT-proBNP measurement before death; ^{Z} = a subset, with a complete baseline data workup, of the larger cohort was used for multivariate regression analyses; LVD_d = left ventricular end-diastolic diameter; Weighted = calculated as a weighted average between two reported groups

Appendix K Excluded prognostic studies

Could not extract data

Akioka, K., Takeuchi, K. et al (2000). 'Prognostic value of Doppler transmitral flow patterns and cardiac natriuretic peptides in patients with chronic congestive heart failure admitted for episodes of acute decompensation', *Heart and Vessels*, 15 (2), 53–60.

Anand, I.S., Fisher, L. et al (2003). 'Changes in plasma brain natriuretic peptide and norepinephrine over time and subsequent mortality and morbidity in heart failure: Results from Val-HeFT', *Journal of the American College of Cardiology*, 41 (6), 222A.

Barretto, A.C.P., Oliveira, M.T. et al (2004). 'NT-ProBNP serum levels are an excellent prognostic predictor in patients with advanced heart failure hospitalized for compensation', *Journal of Cardiac Failure*, 10 (4), S95.

Berger, R., Gore, O. et al (2004). 'Natriuretic peptide levels (BNP, N-BNP, N-ANP) for risk stratification in patients with systolic left ventricular dysfunction under optimised therapy evaluation of cut-off levels in 910 patients', *European Heart Journal*, 25, 2856.

Berger, R., Huelsmann, M. et al (2002). 'B-type natriuretic peptide predicts sudden death in patients with chronic heart failure: a selection mode for ICD implantation?' *European Heart Journal*, 23, 570.

Bettencourt, P., Azevedo, A. et al (2004). 'N-terminal-pro-brain natriuretic peptide predicts outcome after hospital discharge in heart failure patients', *Circulation*, 110 (15), 2168–2174.

Bettencourt, P., Ferreira, S. et al (2002). 'Preliminary data on the potential usefulness of B-type natriuretic peptide levels in predicting outcome after hospital discharge in patients with heart failure', *American Journal of Medicine*, 113 (3), 215–219.

Bluestein, B.I., Belenky, A. et al (2002). 'Development of an automated test for B-type natriuretic peptide(BNP)as an aid in the diagnosis and evaluation of congestive heart failure on the Bayer ADVIA Centaur chemiluminescent system', *Clinical Chemistry*, 48 (6), A85.

Bluestein, B.I., Belenky, A. et al (2003). 'ADVIA Centaur (R) B-type natriuretic peptide (BNP): A new routine and rapid immunoassay for use in the diagnosis and management of patients with heart failure', *Clinical Chemistry*, 49 (6), A68.

Boomsma, F., Tjeerdsma, G. et al (2002). 'Evaluation of a rapid bedside brain natriuretic peptide assay for clinical applicability in patients with congestive heart failure', *Journal of Hypertension*, 20, S308.

Braunschweig, F., Fahrleitner, A. et al (2003). 'Relationship between serial measurements of N-terminal pro-brain natriuretic peptide and ambulatory cardiac filling pressures in outpatients with heart failure', *Journal of the American College of Cardiology*, 41 (6), 141A.

Brunner-La Rocca, H.P., Kaye, D.M. et al (2001). 'Resting cardiac production of natriuretic peptides determines the effects of infusion of brain natriuretic peptide in heart failure', *Journal of the American College of Cardiology*, 37 (2), 148A.

Caidahl, K., Omland, T. et al (2002). 'N-terminal pro B-type natriuretic peptide independently predicts long-term survival in patients with acute coronary syndromes without clinical signs of heart failure', *Circulation*, 106 (19), 455–456.

Campana, C., Pasotti, M. et al (2005). 'Change in plasma brain natriuretic peptide is a strong predictor of prognosis in patients with advanced heart failure awaiting for cardiac transplantation', Conference Proceeding: The International Society for Heart and Lung Transplantation twenty-fifth anniversary meeting and scientific sessions, Philadelphia, PA, In: Change in plasma brain natriuretic peptide is a strong predictor of prognosis in patients with advanced heart failure awaiting for cardiac transplantation, S131.

Cerisano, G., Bolognese, L. et al (2002). 'Plasma BNP versus clinical, echo and angiographic variables for predicting early left ventricular remodelling and death after reperfused acute myocardial infarction and left ventricular dysfunction', *European Heart Journa*l, 23, 615.

Cheng, V., Kasenagra, R. et al (2000). 'A rapid bedside test for brain natriuretic peptide accurately predicts therapeutic response and clinical outcomes in patients admitted with decompensated heart failure', *Journal of Investigative Medicine*, 48 (1), 66A.

Cheng, V., Kazanagra, R. et al (2001). 'A rapid bedside test for B-type peptide predicts treatment outcomes in patients admitted for decompensated heart failure: A pilot study', *Journal of the American College of Cardiology*, 37 (2), 386–391.

Coletta, A., Thackray, S. et al (2002). 'Clinical trials update: highlights of the scientific sessions of the American College of Cardiology 2002: LIFE, DANAMI 2, MADIT-2, MIRACLE-ICD, OVERTURE, OCTAVE, ENABLE 1 & 2, CHRISTMAS, AFFIRM, RACE, WIZARD, AZACS, REMATCH, BNP trial and HARDBALL', *European Journal of Heart Failure*, 4 (3), 381–388.

Colombo, M.G., Clerico, A. et al (1999). 'Endothelin and cardiac natriuretic peptides in heart failure', *American Journal of Hypertension*, 12 (4), 1A.

Cowie, M., Fox, K. et al (2002). 'Chromogranin and BNP predict prognosis but not hospitalisation in patients presenting with heart failure for the first time', *European Heart Journal*, 23, 326.

Cowie, M., Fox, K.F. et al (2000). 'N-terminal Pro-BNP better predicts prognosis in patients presenting with heart failure for the first time than NT-pro-ANP or BNP', *European Heart Journal*, 21, 536.

Cowie, M.R., Metcalfe, C. et al (2005). 'N-terminal brain natriuretic peptide and subsequent hospital admission for worsening heart failure', *Heart*, 91 (3), 371–372.

De Groote, P., Dagom, J. et al (2003). 'BNP and peak V02: A powerful combination for risk stratification in patients with congestive heart failure', *Circulation*, 108 (17), 557.

Dixen, U., Wallevik, L. et al (2003). 'Prolonged signal-averaged P wave duration as a prognostic marker for morbidity and mortality in patients with congestive heart failure', *Scandinavian Cardiovascular Journal*, 37 (4), 193–198.

Dokainish, H., Ambriz, E. et al (2004). 'Tissue Doppler echocardiography and B-type natriuretic peptide in the prognosis of congestive heart failure', *Journal of the American College of Cardiology*, 43 (5), 346A.

Farra, Y.M. & Binkley, P.F. (2004). 'The relation between plasma brain natriuretic peptide concentration and length of hospital stay in patients with congestive heart failure', *Journal of Cardiac Failur*e, 10 (4), S52.

Feola, M., Deorsola, A. et al (2000). 'Brain natriuretic peptide correlates with VO2max and O2 pulse in patients with heart failure', *Journal of Hypertension*, 18, S207.

Fisher, C., Berry, C. et al (2003a). 'N terminal B-type natriuretic peptide, but not the FP437 putative cardiac hormone relaxin, predicts adverse outcome in patients with chronic heart failure', *European Heart Journal*, 24, 61.

Fisher, C., Berry, C. et al (2003b). 'N-terminal B-type natriuretic peptide, but not the putative cardiac hormone relaxin, predicts adverse outcome in patients with chronic heart failure', *Heart*, 89, A38.

Fujimoto, S., Amano, H. et al (2004). 'Usefulness of I-123-metaiodobenzylguanidine myocardial scintigraphy in the prediction of cardiac events in patients with cardiomyopathy showing stabilization of symptoms or preserved cardiac function', *Annals of Nuclear Medicine*, 18 (7), 591–598.

Gackowski, A., Isnard, R. et al (2002). 'Brain natriuretic peptide plasma level falls early in severe acute heart failure responding to treatment and is a strong prognostic marker', *Circulation*, 106 (19), 564–565.

Gardner, R.S., Chong, V. et al (2004a). 'N-terminal brain natriuretic peptide is a more powerful predictor of mortality than endothelin-1, adrenomedullin and tumour necrosis factor in patients with advanced heart failure', *European Heart Journal*, 25, 291.

Gardner, R.S., Chong, V. et al (2004b). 'N-terminal brain natriuretic peptide (NT-proBNP) is a more powerful predictor of mortality than endothelin-1 (ET-1), adrenomedullin (ADM), and tumour necrosis factor (TNF) alpha in patients with advanced heart failure', *Heart*, 90, A19.

Gardner, R.S., Martin, W. et al (2003). 'N terminal pro-brain natriuretic peptide is the most powerful predictor of mortality in patients with advanced heart failure', *European Heart Journal*, 24, 33.

Ghani, F., Wu, A. et al (2003). 'Circulating levels of B-type natriuretic peptide (BNP) are useful for diagnosis and are associated with severity of symptoms in chronic heart failure: preliminary data from a multicenter study with Bayer ADVIA Centaur (R) BNP assay', *Clinical Chemistry*, 49 (6), A67–A68.

Goedel-Meinen, L., Hofmann, M. et al (2004). 'Prognostic value of N-terminal-pro brain natriuretic peptide - Comparison of patients with diastolic and systolic heart failure', *European Heart Journal*, 25, 286.

Goncalves, E., Lesny, P. et al (2003). 'Brain natriuretic peptide is powerful predictor of elevation of ventricular filling pressures in patients with advanced congestive heart failure', *European Heart Journal*, 24, 532.

Groenning, B.A., Riegger, G.A. et al (2002). 'Association between N-terminal pro brain natriuretic peptide and left ventricular function in mild to moderate heart failure', *European Heart Journal*, 23, 269.

Hamanaka, I., Saito, Y. et al (1998). 'Plasma BNP concentration is the most sensitive marker for management of chronic congestive heart failure', *Journal of the American College of Cardiology*, 31 (2), 155A.

Hartmann, F., Packer, M. et al (2002a). 'The prognostic significance of NT-proBNP in severe chronic heart failure is independent of aetiology: a substudy of the COPERNICUS trial', *European Heart Journal*, 23, 269.

Hartmann, F.S., Amann-Zalan, I. et al (2003). 'Influence of creatinine on plasma concentrations and the prognostic value of N-terminal pro-BNP in severe chronic heart failure: Data from a substudy of the COPERNICUS trial', *Journal of the American College of Cardiology*, 41 (6), 143A.

Hartmann, F.S., Richardt, G. et al (2002b). 'Plasma N-terminal pro-BNP predicts the occurrence of major clinical events in severe chronic heart failure: Results of a substudy of the COPERNICUS trial', *Journal of the American College of Cardiology*, 39 (5), 200A.

Havranek, E.P., Masoudi, F.A. et al (2003). 'Changes in BNP level are not associated with outcomes in outpatients with heart failure', *Circulation*, 108 (17), 692–693.

Hiroyuki, N., Junnich, I. et al (2004). 'Prognostic value of combination of N-terminal Pro-BNP and troponin T after initiation of treatment in patients with chronic heart failure', *Journal of Cardiac Failure*, 10 (5), S176.

Horwich, T.B., Patel, J. et al (2003). 'Cardiac troponin I is associated with impaired hemodynamics, progressive left ventricular dysfunction, and increased mortality rates in advanced heart failure', *Circulation*, 108 (7), 833–838.

Hullin, R., Tanner, H. et al (2003). 'In advanced heart failure elevated BNP and CRP serum levels identify high risk patients', *Circulation*, 108 (17), 601.

Ishii, J., Nakamura, Y. et al (2003). 'Prognostic utility of N-terminal Pro-BNP versus BNP in patients hospitalized for worsening chronic heart failure', *Circulation*, 108 (17), 343.

Ishii, J., Nomura, M. et al (2002). 'Risk stratification using a combination of cardiac troponin T and brain natriuretic peptide after optimized treatment in patients with advanced heart failure', *Journal of the American College of Cardiology*, 39 (5), 130A–131A.

Ishii, J., Wang, J.H. et al (1999). 'Early risk stratification using cardiac troponin T and brain natriuretic peptide in patients with congestive heart failure', *Circulation*, 100 (18), 679.

Isnard, R., Chafirovskaia, O. et al (2001a). 'Mitral flow velocity patterns and plasma brain natriuretic peptide changes in patients with chronic heart failure', *Circulation*, 104 (17), 738–739.

Isnard, R., Gackowski, A. et al (2004). 'Plasma brain natriuretic peptide level changes in acutely decompensated heart failure are associated with outcome in patients with both reduced and preserved left ventricle systolic function', *European Heart Journal*, 25, 285.

Isnard, R., Pousset, F. et al (2001b). 'Value of plasma brain natriuretic peptide for risk stratification in patients with chronic heart failure and intermediate value of peak oxygen consumption', *European Heart Journal*, 22, 377.

Johnston, N., Green, M. et al (2002). 'NT-pro BNP predicts toleration of beta-blocker therapy in patients with chronic heart failure', *European Heart Journal*, 23, 270.

Jourdain, P., Funck, F. et al (2003). 'Brain natriuretic peptide plasma level may predict acute decompensation in chronic heart failure patients during carvedilol titration', *European Heart Journal*, 24, 530.

Jourdain, P., Funck, F. et al (2002a). 'Determination of a prognostic threshold for bedside B type natriuretic peptide in chronic heart failure', *European Heart Journal*, 23, 575.

Jourdain, P., Funck, F. et al (2002b). 'BNP variations under diuretic treatment predicts short term mortality in acute or decompensated heart failure', *European Heart Journal*, 23, 571.

Junnichi, I., Tadashf, N. et al (2004). 'Prognostic value of combination of N-terminal pro-BNP and troponin T in patients hospitalized for worsening chronic heart failure', *Journal of Cardiac Failure*, 10 (5), S176.

Kampa, I.S. & Keffer, P. (2003). 'Evaluation of the B-type natriuretic peptide (BNP) and pro-brain natriuretic peptide (NT-pro-BNP) in congestive heart failure (CHF) patients', *Clinical Chemistry*, 49 (6), A32–A33.

Kitagawa, F., Ishii, J. et al (2004). 'Prognostic value of combination of N-terminal pro-BNP and cardiac troponin T in patients hospitalized for worsening chronic heart failure', *Clinical Chemistry*, 50 (6), A2.

Kordick, M. F., Aleman, C. & Chan, S. B. (2004). 'Correlation between chest radiograph and B-type natriuretic peptide in congestive heart failure', *Annals of Emergency Medicine*, 44 (4), S55.

Kruger, S., Graf, J. et al (2001). 'High brain natriuretic peptide levels are associated with decreased exercise capacity in patients with chronic heart failure', *European Heart Journal*, 22, 33.

Kruger, S., Janssens, U. et al (2002a). 'Brain natriuretic peptide levels predict exercise capacity in patients with chronic heart failure', *Journal of the American College of Cardiology*, 39 (5), 183A.

Kruger, S., Janssens, U. et al (2002b). 'Value of cardiopulmonary exercise testing and brain natriuretic peptide to predict prognosis of patients with chronic heart failure', *European Heart Journal*, 23, 575.

Kuster, G.M., Tanner, H. et al (2002). 'B-type natriuretic peptide for diagnosis and treatment of congestive heart failure', *Swiss Medical Weekly*, 132 (43–44), 623–628.

Lainchbury, J.G., Swanney, M.P. et al (2001). 'Change in plasma brain natriuretic peptide during exercise is an important predictor of survival in systolic heart failure', *European Heart Journal*, 22, 377.

Latin, R., Masson, S. et al (2002). 'Brain natriuretic peptide is an important predictor of outcome in patients with heart failure in Val-HeFT', *Circulation*, 106 (19), 613–614.

Latini, R., Masson, S. et al (2002). 'Comparative prognostic significance of plasma aldosterone, BNP, norepinephrine and renin in 4305 patients with heart failure in Val-HeFT', *European Heart Journal*, 23, 271.

Latini, R., Masson, S. et al (2003a). 'The relationship between B-type natriuretic peptide, heart structure, and function in 4,300 patients with heart failure enrolled to Val-HeFT', *Journal of the American College of Cardiology*, 41 (6), 221A.

Latini, R., Wong, M. et al (2003b). 'Predicting outcomes in chronic heart failure from short-term change in BNP as a surrogate end-point: Val-HeFT data', *Circulation*, 108 (17), 557–558.

Latini, R., Wong, M. et al (2003c). 'Brain natriuretic peptide, but not norepinephrine, predicts improvements in left-ventricular structure and function in patients with chronic heart failure enrolled to Val-HeFT', *European Heart Journal*, 24, 34.

Lensky, M.M., Horwich, R.B. & Fonarow, G.C. (2004). 'The value of B-type natriuretic peptide and cardiac troponin I as predictors of sudden death in patients with advanced heart failure', *Journal of the American College of Cardiology*, 43 (5), 172A.

Lewin, J., O'Loughlin, C. et al (2004). 'Poor sensitivity of weight and brain natriuretic peptide changes in predicting clinical deterioration of heart failure', *European Heart Journal*, 25, 137.

Logeart, D., Thabut, G. et al (2002). 'Predischarge measurement of B-type natriuretic peptide predicts outcome of patients admitted for decompensated heart failure', *Circulation*, 106 (19), 472.

Luther, S.A., McCullough, P.A. et al (2004). 'B-type natriuretic peptide levels are not associated with health status in patients with heart failure', *Circulation*, 109 (20), E233.

Macnab, A., Jenkins, N.P. et al (2002). 'Left ventricular dysfunction following mitral valve repair can be predicted by pre-operative brain natriuretic peptide', *Circulation*, 106 (19), 506.

Maeda, K., Tsutamoto, T. et al (1999). 'High levels of plasma brain natriuretic peptide and interleukin-6 after treatment are independent risk factors for mortality in patients with congestive heart failure', *Circulation*, 100 (18), 679.

Maisel, A., Cheng, V. et al (2000). 'A rapid bedside test for B-type natriuretic peptide predicts treatment outcomes in patients admitted with decompensated heart failure', *European Heart Journal*, 21, 585.

Maisel, A.S., Bhalla, M.A. et al (2004). 'Utility of B-natriuretic peptide levels in predicting outcome of hospitalized patients with congestive heart failure: Results of the Breathing Not Properly (BNP) Multinational Study', *Journal of the American College of Cardiology*, 43 (5), 234A.

Mant, D. 'Measurement of BNP (B-type natriuretic peptide) to guide treatment of elderly patients at risk of heart failure: a randomised trial in primary care'.

Mathur, V.S., Butler, J. et al (2002). 'Efficacy of B-type natriuretic peptide (BNP) in patients with renal insufficiency and acute heart failure: Results from the VMAC trial', *Journal of the American Society of Nephrology*, 13, 442A–443A.

Matsuda, N., Endo, Y. et al (2000). 'Plasma brain natriuretic peptide levels predict tolerance to carvedilol in patients with severe heart failure', *Circulation*, 102 (18), 627–628.

Matsuo, S., Nakamura, Y. et al (2003). 'Prognostic value of iodine-123 metaiodobenzylguanidine imaging in patients with heart failure', *Experimental and Clinical Cardiology*, 8 (2), 95–98.

Missov, E., Wieczorek, S. et al (2000). 'Brain natriuretic peptide levels are associated with improved clinical status in patients with chronic heart failure after treatment with Enbrel (R) (Etanercept, TNF receptor)', *Circulation*, 102 (18), 532.

Mitrovic, V., Anand, I. et al (2004). 'Nonterminal pro B-type natriuretic peptide in patients with chronic heart failure: Correlation with B-type natriuretic peptide and left ventricular structure and function', *Journal of the American College of Cardiology*, 43 (5), 235A.

Nicholls, G. (1999) In New Zealand Journal of Medical Laboratory Science, Vol. 53 *New Zealand AMI*, pp. 145.

Nicholls, M.G., Lainchbury, J.G. et al (2001). 'Brain natriuretic peptide-guided therapy for heart failure', *Annals of medicine*, 33 (6), 422–427.

O'Brien, R.J., Demme, B. et al (2002). 'Pre-discharge but not admission levels of NT-proBNP independently predict death and heart failure after admission with acute heart failure', *European Heart Journal*, 23, 270.

Osca, J.Q., Zorio, E. et al (2001). 'Brain natriuretic peptide as an indicator of long term survival in isolated diastolic heart failure', *Circulation*, 104 (17), 501.

Pacher, R., Schenk, K. et al (2002). 'The efficacy of new interventional programmes depends on the stage of heart failure, evaluated by BNP plasma concentration', *European Heart Journal*, 23, 271.

Pacher, R., Stanek, B. et al (2000). 'Prognostic value of plasma brain natriuretic peptide versus 1-98 proatrial natriuretic peptide in chronic heart failure (Neurohumoral Substudy Elite II)', *Circulation*, 102 (18), 532.

Perna, E.R., Canella, J.P.C. et al (2004a). 'Combined determination of natriuretic peptides and necrosis markers in decompensated heart failure: Should NT-proBNP and troponin T be measured on admission, predischarge or both?' *Journal of Cardiac Failure*, 10 (4), S102–S103.

Perna, E.R., Macin, S.M. et al (2004b). 'Combining measurements of N-terminal-pro brain natriuretic peptide and troponin T for long-term risk stratification in decompensated heart failure: should they be measured on admission, predischarge or both?' *European Heart Journal*, 25, 653.

Pieske, B., Herrmann-Lingen, C. et al (2002). 'NT-pro brain natriuretic peptide as marker for severity and prognosis of left ventricular dysfunction', *Circulation*, 106 (19), 683.

Pousset, F., Isnard, R. et al (2003). 'Brain natriuretic peptide plasma levels is a better predictor of sudden death than heart rate variability in mild to moderate heart failure', *European Heart Journal*, 24, 63.

Price, J.F., Little, K. et al (2004). 'Plasma B-type natriuretic peptide levels in infants and children with chronic compensated and acute decompensated heart failure', *Journal of Cardiac Failure*, 10 (4), S48.

Ryoji, T., Yukihit, S. et al (2004). 'Combined measurements of cardiac troponin T and N-terminal pro-brain natriuretic peptide in patients with chronic heart failure', *Journal of Cardiac Failure*, 10 (5), S177.

Sandmann, H. 'Can Serum NT-BNP improve the management of acute dyspnea in the Medical Assessment Unit?'

Sato, Y., Taniguchi, R. et al (2002). 'Serum cardiac troponin T and plasma brain natriuretic peptide in patients with cardiac decompensation', *Heart*, 88 (6), 647–648.

Scardovi, A.B., Aspromonte, N. et al (2004). 'Correlation between brain natriuretic peptide and cardiopulmonary exercise test parameters in patients with chronic heart failure', *Journal of the American College of Cardiology*, 43 (5), 170A.

Seino, Y., Fukushima, M. et al (2002). 'Plasma concentrations of N-terminal pro-BNP versus BNP in patients with chronic heart failure; More discerning marker for the progression', *Circulation*, 106 (19), 683.

Selvais, P.L., Donckier, J.E. et al (1998). 'Cardiac natriuretic peptides for diagnosis and risk stratification in heart failure: Influences of left ventricular dysfunction and coronary artery disease on cardiac hormonal activation', *European Journal of Clinical Investigation*, 28 (8), 636–642.

Stanek, B., Frey, B. et al (2000). 'Prognostic value of repetitive plasma big endothelin versus plasma brain natriuretic peptide levels in chronic heart failure', *Circulation*, 102 (18), 531.

Stevens, T.L., Mahoney, D.W. et al (1997). 'Brain natriuretic peptide is a marker of New York Heart Association classification in out-patient follow-up of left ventricular dysfunction', *Journal of the American College of Cardiology*, 29 (2), 30177.

Takeda, Y., Fukutomi, T. et al (2003). 'Relationship between iodine-123 metaiodobenzylguanidine myocardial imaging and plasma brain natriuretic peptide concentration in patients with isolated diastolic heart failure', *European Heart Journal*, 24, 53.

Taniguchi, R., Sato, Y. et al (2004). 'Combined measurements of cardiac troponin T and N-terminal pro-brain natriuretic peptide in patients with heart failure', *Circulation Journal*, 68 (12), 1160–1164.

Tanner, H., Mohacsi, R. et al (2004). 'The combination of brain natriuretic peptide and C-reactive protein identifies chronic heart failure patients at high cardiovascular risk', *European Heart Journal*, 25, 291.

Toney, L.J., Gloria, C.T. et al (2004). 'Assessment of technical and clinical performance of N-terminal proBNP and BNP in congestive heart failure', *Clinical Chemistry*, 50 (6), A6.

Troughton, R.W., Prior, D.L. et al (2003). 'Relationship of plasma BNP levels to tissue Doppler and color M-mode indices of diastolic function and left ventricular filling pressure in systolic heart failure', *Journal of the American College of Cardiology*, 41 (6), 448A.

Tsutamoto, T., Maeda, K. et al (1996). 'Plasma brain natriuretic peptide concentration as a prognostic predictor in patients with symptomatic left ventricular dysfunction', *Circulation*, 94 (8), 1490.

Tsutamoto, T., Maeda, K. et al (2000). 'Plasma brain natriuretic peptide concentration as a useful predictor of mortality and morbidity in patients with diastolic heart failure', *Circulation*, 102 (18), 624.

Tsutamoto, T., Maeda, Y. et al (1993). 'Plasma brain natriuretic peptide concentration as a prognostic predictor in patients with chronic congestive-heart-failure', *Circulation*, 88 (4), 26.

Vanderheyden, M., Geelen, P. et al (2003). 'Myocardial production of brain natriuretic peptide in patients with advanced heart failure', *European Heart Journal*, 24, 657.

Vijayaraghavan, K., Singh-Khalsa, M. & Crum, S. (2002). 'Utility of brain natriuretic peptide in chronic decompensated heart failure in an outpatient heart failure clinic', *Circulation*, 106 (16), E82.

Wei, C.M., Rodeheffer, R.J. et al (1996). 'Temporal responses in ventricular atrial natriuretic peptide (ANP) and brain natriuretic peptide (BNP) secretion in human heart failure', *Journal of Investigative Medicine*, 44 (7), A351.

Wieczorek, S.J., Bailly, K.R. et al (2000). 'Clinical evaluation of the Triage B-Type Natriuretic Peptide Assay for point care testing of patients with congestive heart failure', *Clinical Chemistry*, 46 (6), A77.

Wong, M., Latini, R. et al (2004). 'Brain natriuretic peptide and symptoms as markers of severity in patients with chronic heart failure: Val-HeFT data', *European Heart Journal*, 25, 136–137.

Yoo, B. S., Jung, H. S. et al (2004). 'The clinical value of brain natriuretic peptide (BNP) level monitoring after acute exacerbation in congestive heart failure patients', *Journal of Cardiac Failure*, 10 (4), S51.

Yoshimura, M. (2003). 'Plasma BNP as a marker of heart failure status in the treatment with carvedilol', *Journal of Molecular And Cellular Cardiology*, 35 (11), A40.

Zucchelli, G., Prontera, C. et al (2004). 'IRMA BNP and chemiluminescence BNP immunoassay in patients with heart failure', *Clinical Chemistry*, 50 (6), A15.

Retrieved after end-date for included papers

Ishii, J., Nomura, M. et al (2002). 'Risk stratification using a combination of cardiac troponin T and brain natriuretic peptide in patients hospitalized for worsening chronic heart failure', *American Journal of Cardiology,* 89 (6), 691–695.

References

AIHW (2003). *Heart failure. What of the future?* Canberra, Australian Institute of Health and Welfare.

AIHW (2004a). AIHW National Hospital Morbidity Database. Australian Institute of Health and Welfare.

AIHW (2005). *General practice activity in Australia 2004–05*. Australian Institute of Health and Welfare (General Practice Series No. 18), AIHW Cat. No. GEP 18. Canberra.

AIHW & NHF (2004). *Heart, stroke and vascular diseases - Australian facts 2004*. (Cardiovascular Disease Series No. 22). Canberra, Australian Institute of Health and Welfare and National Heart Foundation of Australia.

Alibay, Y., Beauchet, A., El Mahmoud, R., Schmitt, C., Brun-Ney, D., Benoit, M.O., Dubourg, O., Boileau, C., Jondeau, G. & Puy, H. (2005). 'Plasma N-terminal pro-brain natriuretic peptide and brain natriuretic peptide in assessment of acute dyspnea', *Biomedicine & Pharmacotherapy,* 59, 20–24.

American Heart Association (2005). *Congestive heart failure.*

Anand, I.S., Fisher, L.D., Chiang, Y.T., Latini, R., Masson, S., Maggioni, A.P., Glazer, R.D., Tognoni, G. & Cohn, J.N. (2003). 'Changes in brain natriuretic peptide and norepinephrine over time and mortality and morbidity in the Valsartan Heart Failure Trial (Val-HeFT)', *Circulation,* 107, 1278–1283.

Apple, F.S., Trinity, E., Steen, J., Prawer, S. & Wu, A.H. (2003). 'BNP test utilization for CHF in community hospital practice', *Clinica Chimica Acta,* 328, 191–193.

Armitage, P., Berry, G. & Matthews, J.N.S. (2002). *Statistical methods in medical research,* Oxford, Blackwell Science.

Aurbach, A., Russ, W., Battegay, E., Bucher, H.C., Brecht, J.G., Schadlich, P.K. & Sendi, P. (2004). 'Cost-effectiveness of ramipril in patients at high risk for cardiovascular events: a Swiss perspective', *Swiss Medical Weekly,* 134, 399–405.

Australian Bureau of Statistics (2006). *Consumer Price Index, Australia, Dec 2005* [Internet]. Available from: http://www.abs.gov.au/ausstats/abs%40.nsf/mf/6401.0 [Accessed 22 February 2006].

Azzazy, H.M.E. & Christenson, R.H. (2003). 'B-type natriuretic peptide: Physiologic role and assay characteristics', *Heart Failure Reviews,* 8, 315–320.

Banerjee, P., Clark, A.L. & Cleland, J.G. (2004). 'Diastolic heart failure: a difficult problem in the elderly', *American Journal of Geriatric Cardiology,* 13, 16–21.

Bayes-Genis, A., Santalo-Bel, M., Zapico-Muniz, E., Lopez, L., Cotes, C., Bellido, J., Leta, R., Casan, P. & Ordonez-Llanos, J. (2004). 'N-terminal probrain natriuretic peptide (NT-proBNP) in the emergency diagnosis and in-hospital monitoring of patients with dyspnoea and ventricular dysfunction', *European Journal of Heart Failure,* 6, 301–308.

Berger, R., Strecker, K., Huelsman, M., Moser, P., Frey, B., Bojic, A., Stanek, B. & Pacher, R. (2003). 'Prognostic power of neurohumoral parameters in chronic heart failure depends on clinical stage and observation period', *Journal of Heart and Lung Transplantation,* 22 (9), 1037–1045.

Bertinchant, J.P., Combes, N. et al (2005). 'Prognostic value of cardiac troponin T in patients with both acute and chronic stable congestive heart failure: comparison with atrial natriuretic peptide brain natriuretic peptide and plasma norepinephrine', *Clinica Chimica Acta,* 352 (1–2), 143–153.

Bettencourt, P. (2004). 'NT-proBNP and BNP: biomarkers for heart failure management', *European Journal of Heart Failure,* 6, 359–363.

Bouvy, M.L., Heerdink, E.R., Leufkens, H.G. & Hoes, A.W. (2003). 'Predicting mortality in patients with heart failure: a pragmatic approach', *Heart,* 89, 605–609.

Briggs, A. & O'Brien, B. (2001). 'The death of cost-minimization analysis?', *Health Economics*, 10, 179–184.

Cardarelli, R. & Lumicao, T.G., Jr. (2003). 'B-type natriuretic peptide: a review of its diagnostic, prognostic, and therapeutic monitoring value in heart failure for primary care physicians', *Journal of the American Board of Family Practice,* 16, 327–333.

Chenevier-Gobeaux, C., Claessens, Y.E., Voyer, S., Desmoulins, D. & Ekindjian, O.G. (2005). 'Influence of renal function on N-terminal pro-brain natriuretic peptide (NT-proBNP) in patients admitted for dyspnoea in the Emergency Department: Comparison with brain natriuretic peptide (BNP)', *Clinica Chimica Acta,* in press.

Collinson, P.O., Barnes, S.C., Gaze, D.C., Galasko, G., Lahiri, A. & Senior, R. (2004). 'Analytical performance of the N terminal pro B type natriuretic peptide (NT-proBNP) assay on the Elecsys 1010 and 2010 analysers', *European Journal of Heart Failure,* 6, 365–368.

Cowie, M.R., Struthers, A.D., Wood, D.A., Coats, A.J., Thompson, S.G., Poole-Wilson, P.A. & Sutton, G.C. (1997). 'Value of natriuretic peptides in assessment of patients with possible new heart failure in primary care', *Lancet,* 350, 1349–1353.

Craig, J., Bradbury, I., Cummins, E., Downie, S., Foster, L. & Stout, A. (2004). *The use of B-type natriuretic peptides in the investigation of patients with suspected heart failure,* NHS Quality Improvement, Scotland.

Dao, Q., Krishnaswamy, P., Kazanegra, R., Harrison, A., Amirnovin, R., Lenert, L., Clopton, P., Alberto, J., Hlavin, P. & Maisel, A.S. (2001). 'Utility of B-type natriuretic peptide in the diagnosis of congestive heart failure in an urgent-care setting', *Journal of the American College of Cardiology,* 37, 379–385.

De Denus, S., Pharand, C. & Williamson, D.R. (2004). 'Brain natriuretic peptide in the management of heart failure: the versatile neurohormone', *Chest,* 125, 652–668.

De Groote, P., Dagorn, J., Soudan, B., Lamblin, N., Mcfadden, E. & Bauters, C. (2004a). 'B-type natriuretic peptide and peak exercise oxygen consumption provide independent information for risk stratification in patients with stable congestive heart failure', *Journal of the American College of Cardiology,* 43, 1584–1589.

De Groote, P., Soudan, B., Lamblin, N., Rouaix-Emery, N., Mc Fadden, E., Meurice, T., Mouquet, F. & Bauters, C. (2004b). 'Is hormonal activation during exercise useful for risk stratification in patients with moderate congestive heart failure?', *American Heart Journal,* 148, 349–355.

De Pasquale, C.G., Arnolda, L.F., Doyle, I.R., Aylward, P.E., Chew, D.P. & Bersten, A.D. (2004). 'Plasma surfactant protein-B - A novel biomarker in chronic heart failure', *Circulation,* 110, 1091–1096.

Deeks, J.J. (2001). 'Systematic reviews of evaluations of diagnostic and screening tests', *British Medical Journal,* 323, 157–162.

Department of Health and Ageing (2006). *Round 7 (2002-2003) Cost report* [Internet]. Available from: http://www.health.gov.au/internet/wcms/publishing.nsf/Content/health-casemix-costing-fc_r7.htm [Accessed 15 January 2006].

DerSimonian, R. & Laird, N. (1986). 'Meta-analysis in clinical trials', *Controlled Clinical Trials,* 7, 177–188.

Dhir, M., Arora, U. & Nagueh, S.F. (2004). 'The role of echocardiography in the diagnosis and prognosis of patients with heart failure', *Expert Review of Cardiovascular Therapy,* 2, 141–144.

Dokainish, H., Zoghbi, W.A., Lakkis, N.M., Ambriz, E., Patel, R., Quinones, M.A. & Nagueh, S.F. (2005). 'Incremental predictive power of B-type natriuretic peptide and tissue Doppler echocardiography in the prognosis of patients with congestive heart failure', *Journal of the American College of Cardiology,* 45, 1223–1226.

Dokainish, H., Zoghbi, W.A., Lakkis, N.M., Quinones, M.A. & Nagueh, S.F. (2004). 'Comparative accuracy of B-type natriuretic peptide and tissue Doppler echocardiography in the diagnosis of congestive heart failure', *American Journal of Cardiology,* 93, 1130–1135.

Doust, J.A., Glasziou, P.P., Pietrzak, E. & Dobson, A.J. (2004). 'A systematic review of the diagnostic accuracy of natriuretic peptides for heart failure', *Archives of Internal Medicine,* 164, 1978–1984.

Doust, J.A., Pietrzak, E., Dobson, A. & Glasziou, P. (2005). 'How well does B-type natriuretic peptide predict death and cardiac events in patients with heart failure: systematic review', *British Medical Journal,* 330, 625.

Downs, S.H. & Black, N. (1998). 'The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions', *Journal of Epidemiology and Community Health,* 52, 377–384.

Egger, M., Juni, P., Bartlett, C., Holenstein, F. & Sterne, J. (2003). 'How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study', *Health Technology Assessment,* 7, 1–76.

Fisher, C., Berry, C., Blue, L., Morton, J.J. & Mcmurray, J. (2003). 'N-terminal pro B type natriuretic peptide, but not the new putative cardiac hormone relaxin, predicts prognosis in patients with chronic heart failure', *Heart (British Cardiac Society),* 89, 879–881.

Fleischer, D., Espiner, E.A., Yandle, T.G., Livesey, J.H., Billings, J., Town, I. & Richards, A.M. (1997). 'Rapid assay of plasma brain natriuretic peptide in the assessment of acute dyspnoea', *New Zealand Medical Journal,* 110, 71–74.

Gackowski, A., Isnard, R., Golmard, J.L., Pousset, F., Carayon, A., Montalescot, G., Hulot, J.S., Thomas, D., Piwowarska, W. & Komajda, M. (2004). 'Comparison of echocardiography and plasma B-type natriuretic peptide for monitoring the response to treatment in acute heart failure', *European Heart Journal,* 25, 1788–1796.

Gardner, R.S., Chong, K.S., Morton, J.J. & Mcdonagh, T.A. (2005). 'N-terminal brain natriuretic peptide, but not anemia, is a powerful predictor of mortality in advanced heart failure', *Journal of Cardiac Failure,* 11, S47–S53.

Garg, R. & Yusuf, S. (1995). 'Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials', *Journal of the American Medical Association,* 273, 1450–1456.

George, J., Patal, S., Wexler, D., Abashidze, A., Shmilovich, H., Barak, T., Sheps, D. & Keren, G. (2005). 'Circulating erythropoietin levels and prognosis in patients with congestive heart failure comparison with neurohormonal and inflammatory markers', *Archives of Internal Medicine,* 165, 1304–1309.

Glasziou, P., Irwig, L.M., Bain, C. & Colditz, G. (2001). *Systematic reviews in health care: A practical guide,* Cambridge, Cambridge University Press.

Gustafsson, F., Badskjaer, J., Hansen, F.S., Poulsen, A.H. & Hildebrandt, P. (2003). 'Value of N-terminal proBNP in the diagnosis of left ventricular systolic dysfunction in primary care patients referred for echocardiography', *HeartDrug,* 3, 141–146.

Hamada, Y., Tanaka, N., Murata, K., Takaki, A., Wada, Y., Oyama, R., Liu, J., Harada, N., Okuda, S., Hadano, Y. & Matsuzaki, M. (2005). 'Significance of predischarge BNP on one-year outcome in decompensated heart failure - Comparative study with echo-Doppler indexes', *Journal of Cardiac Failure,* 11, 43–49.

Hammerer-Lercher, A., Neubauer, E., Muller, S., Pachinger, O., Puschendorf, B. & Mair, J. (2001). 'Head-to-head comparison of N-terminal pro-brain natriuretic peptide, brain natriuretic peptide and N-terminal pro-atrial natriuretic peptide in diagnosing left ventricular dysfunction', *Clinica Chimica Acta,* 310, 193–197.

Hartmann, F., Packer, M., Coats, A.J.S., Fowler, M.B., Krum, H., Mohacsi, P., Rouleau, J.L., Tendera, M., Castaigne, A., Anker, S.D., Amann-Zalan, I., Hoersch, S. & Katus, H.A. (2004). 'Prognostic impact of plasma N-terminal pro-brain natriuretic peptide in severe chronic congestive heart failure - A substudy of the carvedilol prospective randomized cumulative survival (COPERNICUS) trial', *Circulation,* 110, 1780–1786.

Hobbs, F.D. (2002). 'Unmet need for diagnosis of heart failure: the view from primary care', *Heart,* 88 Suppl. 2, 9–11.

Hobbs, F.D., Davis, R.C., Roalfe, A.K., Hare, R. & Davies, M.K. (2004). 'Reliability of N-terminal proBNP assay in diagnosis of left ventricular systolic dysfunction within representative and high risk populations', *Heart,* 90, 866–870.

Hobbs, F.D., Davis, R.C., Roalfe, A.K., Hare, R., Davies, M.K. & Kenkre, J.E. (2002). 'Reliability of N-terminal pro-brain natriuretic peptide assay in diagnosis of heart failure: cohort study in representative and high risk community populations', *British Medical Journal,* 324, 1498.

Imamura, Y., Fukuyama, T., Mochizuki, T., Miyagawa, M. & Watanabe, K. (2001). 'Prognostic value of Iodine-123-metaiodobenzylguanidine imaging and cardiac natriuretic peptide levels in patients with left ventricular dysfunction resulting from cardiomyopathy', *Japanese Circulation Journal,* 65, 155–160.

Inomata, T., Nishii, M., Takehana, H., Nakahata, J.I., Koitabashi, T., Nakano, H., Takeuchi, I., Kurokawa, S. & Izumi, T. (2003). 'Brain natriuretic peptide-guided treatment reduces cardiovascular events of heart failure in outpatient management', *Circulation,* 108, 446–446.

Ishii, J., Cui, W., Kitagawa, F., Kuno, T., Nakamura, Y., Naruse, H., Mori, Y., Ishikawa, T., Nagamura, Y., Kondo, T., Oshima, H., Nomura, M., Ezaki, K. & Hishida, H. (2003). 'Prognostic value of combination of cardiac troponin T and B-type natriuretic peptide after initiation of treatment in patients with chronic heart failure', *Clinical Chemistry,* 49, 2020–2026.

Isnard, R., Pousset, F., Chafirovskaia, O., Carayon, A., Hulot, J.S., Thomas, D. & Komajda, M. (2003). 'Combination of B-type natriuretic peptide and peak oxygen consumption improves risk stratification in outpatients with chronic heart failure', *American Heart Journal,* 146, 729–735.

Januzzi, J.L., Camargo, C.A., Anwaruddin, S., Baggish, A.L., Chen, A.A., Krauser, D.G., Tung, R., Cameron, R., Nagurney, J.T., Chae, C.U., Lloyd-Jones, D.M., Brown, D.F., Foran-Melanson, S., Sluss, P.M., Lee-Lewandrowski, E.L. & Lewandrowski, K.B. (2005). 'The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study', *American Journal of Cardiology,* 95, 948–954.

Januzzi, J.L. & Maisel, A.S. (2004). 'Routine measurement of natriuretic peptide to guide the diagnosis and management of chronic heart failure', *Circulation,* 109, e325–e326; author reply e325–e326.

Jernberg, T., James, S., Lindahl, B., Stridsberg, M., Venge, P. & Wallentin, L. (2004). 'NT-proBNP in unstable coronary artery disease--experiences from the FAST, GUSTO IV and FRISC II trials', *European Journal of Heart Failure,* 6, 319–25.

Jose, J.V., Gupta, S.N. & Selvakumar, D. (2003). 'Utility of N-terminal pro-brain natriuretic peptide for the diagnosis of heart failure', *Indian Heart Journal,* 55, 35–39.

Kellett, J. (2005). 'The prediction of in-hospital mortality by amino terminal pro-brain natriuretic peptide (NT-proBNP) levels and other independent variables in acutely ill patients with suspected heart disease', *European Journal of Internal Medicine,* 16, 195–199.

Kenchaiah, S., Evans, J.C., Levy, D., Wilson, P.W., Benjamin, E.J., Larson, M.G., Kannel, W.B. & Vasan, R.S. (2002). 'Obesity and the risk of heart failure', *New England Journal of Medicine,* 347, 305–313.

Kirk, V., Bay, M., Parner, J., Krogsgaard, K., Herzog, T.M., Boesgaard, S., Hassager, C., Nielsen, O.W., Aldershvile, J. & Nielsen, H. (2004). 'N-terminal proBNP and mortality in hospitalised patients with heart failure and preserved vs. reduced systolic function: data from the prospective Copenhagen Hospital Heart Failure Study (CHHF)', *European Journal of Heart Failure,* 6, 335–341.

Knapp, G. & Hartung, J. (2003). 'Improved tests for a random effects meta-regression with a single covariate', *Statistics in Medicine,* 22, 2693–2710.

Knudsen, C.W., Omland, T., Clopton, P., Westheim, A., Abraham, W.T., Storrow, A.B., Mccord, J., Nowak, R.M., Aumont, M.C., Duc, P., Hollander, J.E., Wu, A.H., Mccullough, P.A. & Maisel, A.S. (2004). 'Diagnostic value of B-Type natriuretic peptide and chest radiographic findings in patients with acute dyspnea', *American Journal of Medicine,* 116, 363–368.

Koglin, J., Pehlivanli, S., Schwaiblmair, M., Vogeser, M., Cremer, P. & vonScheidt, W. (2001). 'Role of brain natriuretic peptide in risk stratification of patients with congestive heart failure', *Journal of the American College of Cardiology,* 38 (7), 1934–1941.

Koseki, Y., Watanabe, J., Shinozaki, T., Sakuma, M., Komaru, T., Fukuchi, M., Miura, M., Karibe, A., Kon-no, Y. and Numaguchi, H. (2003). 'Characteristics and 1-year prognosis of medically treated patients with chronic heart failure in Japan', *Circulation Journal,* 67 (5), 431–436.

Kosowsky, J.M., Weiner, C. & Morrissey, J.H. (2003). 'Impact of B-type natriuretic peptide testing on medical decision making for older patients with dyspnea', *Annals of Emergency Medicine,* 42, S11.

Koulouri, S., Acherman, R.J., Wong, P.C., Chan, L.S. & Lewis, A.B. (2004). 'Utility of B-type natriuretic peptide in differentiating congestive heart failure from lung disease in pediatric patients with respiratory distress', *Pediatric Cardiology,* 25, 341–346.

Krishnaswamy, P., Lubien, E., Clopton, P., Koon, J., Kazanegra, R., Wanner, E., Gardetto, N., Garcia, A., Demaria, A. & Maisel, A.S. (2001). 'Utility of B-natriuretic peptide levels in identifying patients with left ventricular systolic or diastolic dysfunction', *American Journal of Medicine,* 111, 274–279.

Krum, H. (2001). 'Guidelines for management of patients with chronic heart failure in Australia', *Medical Journal of Australia,* 174, 459–466.

Krum, H., Tonkin, A.M., Currie, R., Djundjek, R. & Johnston, C.I. (2001). 'Chronic heart failure in Australian general practice. The Cardiac Awareness Survey and Evaluation (CASE) Study', *Medical Journal of Australia,* 174, 439–444.

Kyuma, M., Nakata, T., Hashimoto, A., Nagao, K., Sasao, H., Takahashi, T., Tsuchihashi, K. & Shimamoto, K. (2004). 'Incremental prognostic implications of brain natriuretic peptide, cardiac sympathetic nerve innervation, and noncardiac disorders in patients with heart failure', *Journal of Nuclear Medicine,* 45, 155–163.

Lainchbury, J.G., Campbell, E., Frampton, C.M., Yandle, T.G., Nicholls, M.G. & Richards, A.M. (2003). 'Brain natriuretic peptide and N-terminal brain natriuretic peptide

in the diagnosis of heart failure in patients with acute shortness of breath', *Journal of the American College of Cardiology,* 42, 728–735.

Landray, M.J., Lehman, R. & Arnold, I. (2000). 'Measuring brain natriuretic peptide in suspected left ventricular systolic dysfunction in general practice: Cross-sectional study', *British Medical Journal,* 320, 985–986.

Latini, R., Masson, S., Anand, I., Salio, M., Hester, A., Judd, D., Barlera, S., Maggioni, A.P., Tognoni, G. & Cohn, J.N. (2004). 'The comparative prognostic value of plasma neurohormones at baseline in patients with heart failure enrolled in Val-HeFT', *European Heart Journal,* 25, 292–299.

Leibovitch, E.R. (2005). 'Once upon a time, heart failure was easy', *Geriatrics,* 60, 31–33.

Littenberg, B. & Moses, L.E. (1993). 'Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method', *Medical Decision Making,* 13, 313–321.

Logeart, D., Saudubray, C., Beyne, P., Thabut, G., Ennezat, P.V., Chavelas, C., Zanker, C., Bouvier, E. & Solal, A.C. (2002). 'Comparative value of Doppler echocardiography and B-type natriuretic peptide assay in the etiologic diagnosis of acute dyspnea', *Journal of the American College of Cardiology,* 40, 1794–1800.

Logeart, D., Thabut, G., Jourdain, P., Chavelas, C., Beyne, P., Beauvais, F., Bouvier, E. & Solal, A.C. (2004). 'Predischarge B-type natriuretic peptide assay for identifying patients at high risk of re-admission after decompensated heart failure', *Journal of the American College of Cardiology,* 43, 635–641.

Maeda, K., Tsutamoto, T., Wada, A., Mabuchi, N., Hayashi, M., Tsutsui, T., Ohnishi, M., Sawaki, M., Fujii, M., Matsumoto, T. & Kinoshita, M. (2000). 'High levels of plasma brain natriuretic peptide and interleukin-6 after optimized treatment for heart failure are independent risk factors for morbidity and mortality in patients with congestive heart failure', *Journal of the American College of Cardiology,* 36, 1587–1593.

Maisel, A.S. (2003). 'The diagnosis of acute congestive heart failure: Role of BNP measurements', *Heart Failure Reviews,* 8, 327–334.

Maisel, A., Hollander, J.E., Guss, D., McCullough, P., Nowak, R., Green, G., Saltzberg, M., Ellison, S.R., Bhalla, M.A., Bhalla, V., Clopton, P., Jesse, R. & Invest, R. (2004). 'Primary results of the rapid emergency department heart failure outpatient trial (REDHOT)', *Journal of the American College of Cardiology,* 44, 1328–1333.

Maisel, A.S., Krishnaswamy, P., Nowak, R.M., Mccord, J., Hollander, J.E., Duc, P., Omland, T., Storrow, A.B., Abraham, W.T., Wu, A.H., Clopton, P., Steg, P.G., Westheim, A., Knudsen, C.W., Perez, A., Kazanegra, R., Herrmann, H.C. & Mccullough, P.A. (2002). 'Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure', *New England Journal of Medicine,* 347, 161–167.

Maisel, A.S. & McCullough, P.A. (2003). 'Cardiac natriuretic peptides: a proteomic window to cardiac function and clinical management', *Reviews in Cardiovascular Medicine,* 4 Suppl 4, S3–12.

Maisel, A. & Zoorob, R. (2005). *CME Bulletin*. American Academy of Family Physicians.

Matsui, T., Tsutamoto, T., Maeda, K., Kusukawa, J. & Kinoshita, M. (2002). 'Prognostic value of repeated 123I-metaiodobenzylguanidine imaging in patients with dilated cardiomyopathy with congestive heart failure before and after optimized treatments - Comparison with neurohumoral factors', *Circulation Journal,* 66, 537–543.

McCullough, P.A., Duc, P., Omland, T., Mccord, J., Nowak, R.M., Hollander, J.E., Herrmann, H.C., Steg, P.G., Westheim, A., Aumont, M.C., Knudsen, C.W., Storrow, A.B., Abraham, W.T., Lamba, S., Wu, A.H., Perez, A., Clopton, P., Krishnaswamy, P., Kazanegra, R. & Maisel, A.S. (2003a). 'B-type natriuretic peptide and renal function in the diagnosis of heart failure: An analysis from the BNP multinational study', *Journal of the American College of Cardiology,* 41, 222A.

McCullough, PA., Nowak, R. M., Mccord, J., Hollander, J.E., Herrmann, H. C., Steg, P.G., Due, P., Westheim, A., Omland, T., Knudsen, C.W., Storrow, A.B., Abraham, W.T., Lamba, S., Wu, A.H.B., Perez, A., Clopton, P., Krishnaswamy, P., Kazanegra, R. & Maisel, A.S. (2002). 'B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure - Analysis from breathing not properly (BNP) multinational study', *Circulation,* 106, 416–422.

McCullough, P.A., Omland, T. & Maisel, A.S. (2003b). 'B-type natriuretic peptides: A diagnostic breakthrough for clinicians', *Reviews in Cardiovascular Medicine,* 4, 72–80.

McCullough, P.A. & Sandberg, K.R. (2003). 'Sorting out the evidence on natriuretic peptides', *Reviews in Cardiovascular Medicine,* 4 Suppl 4, S13–19.

Mcdonagh, T.A., Holmer, S., Raymond, I., Luchner, A., Hildebrant, P. & Dargie, H.J. (2004). 'NT-proBNP and the diagnosis of heart failure: a pooled analysis of three European epidemiological studies', *European Journal of Heart Failure,* 6, 269–273.

Mehra, M.R., Uber, P.A., Park, M.H., Scott, R.L., Ventura, H.O., Harris, B. C. & Frohlich, E.D. (2004). 'Obesity and suppressed B-type natriuretic peptide levels in heart failure', *Journal of the American College of Cardiology,* 43, 1590–1595.

Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D. & Stroup, D.F. (1999). 'Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses', *Lancet,* 354, 1896–1900.

Mottram, P.M. & Marwick, T.H. (2005). 'Assessment of diastolic function: what the general cardiologist needs to know', *Heart,* 91, 681–695.

MSAC (2005). *Guidelines for the assessment of diagnostic technologies*, Canberra, ACT, Commonwealth of Australia.

Mueller, C., Laule-Kilian, K., Frana, B., Rodriguez, D., Rudez, J., Scholer, A., Buser, P., Pfisterer, M. & Perruchoud, A. P. (2005a). 'The use of B-type natriuretic peptide in the management of elderly patients with acute dyspnoea', *Journal of Internal Medicine,* 258, 77–85.

Mueller, C., Laule-Kilian, K., Schindler, C., Klima, T., Frana, B., Rodriguez, D., Scholer, A., Christ, M., Perruchoud, A. P. (2006). 'Cost-effectiveness of B-type natriuretic peptide testing in patients with acute dyspnea', *Archives of Internal Medicine,* 166 (10),1081–1087.

Mueller, C., Laule-Kilian, K., Scholer, A., Frana, B., Rodriguez, D., Schindler, C., Marsch, S. & Perruchoud, A. P. (2004a). 'Use of B-type natriuretic peptide for the management of women with dyspnea', *American Journal of Cardiology,* 94, 1510–1514.

Mueller, C., Laule-Kilian, K., Scholer, A., Nusbaumer, C., Zeller, T., Staub, D. & Perruchoud, A. P. (2005b). 'B-type natriuretic peptide for acute dyspnea in patients with kidney disease: Insights from a randomized comparison', *Kidney International,* 67, 278–284.

Mueller, C., Scholer, A., Laule-Kilian, K., Martina, B., Schindler, C., Buser, P., Pfisterer, M. & Perruchoud, A. P. (2004b). 'Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea', *New England Journal of Medicine,* 350, 647–654.

Mueller, T., Gegenhuber, A., Poelz, W. & Haltmayer, M. (2005c). 'Diagnostic accuracy of B type natriuretic peptide and amino terminal proBNP in the emergency diagnosis of heart failure', *Heart,* 91, 606–612.

Murdoch, D.R., Mcdonagh, T.A., Byrne, J., Blue, L., Farmer, R., Morton, J.J. & Dargie, H.J. (1999). 'Titration of vasodilator therapy in chronic heart failure according to plasma brain natriuretic peptide concentration: randomized comparison of the hemodynamic and neuroendocrine effects of tailored versus empirical therapy', *American Heart Journal,* 138, 1126–1132.

New York Heart Association. *New York Heart Association classification for congestive heart failure,* New York Heart Association.

NHF & CSANZ (2002). *Guidelines on the Contemporary Management of the Patient with Chronic Heart Failure in Australia*, National Heart Foundation of Australia and Cardiac Society of Australia & New Zealand.

NHMRC (2000). *How to use the evidence: assessment and application of scientific evidence,* Canberra, National Health and Medical Research Council.

NHMRC (2005). *NHMRC additional levels of evidence and grades for recommendations for developers of guidelines, Pilot Program 2005*, National Health and Medical Research Council, Australian Government.

NICE (2003). *Chronic heart failure: Management of chronic heart failure in adults in primary and secondary care,* London, National Institute for Clinical Excellence.

Nielsen, L.S., Svanegaard, J., Klitgaard, N.A. & Egeblad, H. (2004). 'N-terminal pro-brain natriuretic peptide for discriminating between cardiac and non-cardiac dyspnoea', *European Journal of Heart Failure,* 6, 63–70.

O'Brien, R.J., Squire, I.B., Demme, B., Davies, J.E. & Ng, L.L. (2003). 'Pre-discharge, but not admission, levels of NT-proBNP predict adverse prognosis following acute LVF', *European Journal of Heart Failure,* 5, 499–506.

Organisation for Economic Co-operation and Development (2006). *Purchasing power parities (PPPs) for OECD countries 1980-2005* [Internet]. Available from: http://www.oecd.org/statisticsdata/0,2643,en_2649_34357_1_119656_1_1_1,00.html [Accessed 25 February 2006].

Penn State College of Medicine (2004). *Transthoracic echocardiography*.

Remme, W.J. & Swedberg, K. (2001). 'Guidelines for the diagnosis and treatment of chronic heart failure', *European Heart Journal,* 22, 1527–1560.

Richards, A.M., Doughty, R., Nicholls, M.G., Macmahon, S., Sharpe, N., Murphy, J., Espiner, E.A., Frampton, C. & Yandle, T.G. (2001). 'Plasma N-terminal pro-brain natriuretic peptide and adrenomedullin: Prognostic utility and prediction of benefit from carvedilol in chronic ischemic left ventricular dysfunction', *Journal of the American College of Cardiology,* 37, 1781–1787.

Roche Diagnostics (2002). Press release.

Rossig, L., Fichtlscherer, S., Heeschen, C., Berger, J., Dimmeler, S. & Zeiher, A.M. (2004). 'The pro-apoptotic serum activity is an independent mortality predictor of patients with heart failure', *European Heart Journal,* 25, 1620–1625.

Sackett, D.L., Hayes, R.B., Guyatt, G.H. & Tugwell, P. (1991). *Clinical Epidemiology: A Basic Science for Clinical Medicine,* Boston, MA, Little, Brown and Company.

Sakatani, T., Hadase, M., Kawasaki, T., Kamitani, T., Kawasaki, S. & Sugihara, H. (2004). 'Usefulness of the percentage of plasma lymphocytes as a prognostic marker in patients with congestive heart failure - Comparison with brain natriuretic peptide', *Japanese Heart Journal,* 45, 275–284.

Sallinen, J. (2004) Abbott AxSYM (R) BNP assay. Abbott Diagnostics.

Seino, Y., Ogawa, A., Yamashita, T., Fukushima, M., Ogata, K., Fukumoto, H. & Takano, T. (2004). 'Application of NT-proBNP and BNP measurements in cardiac care: a more discerning marker for the detection and evaluation of heart failure', *European Journal of Heart Failure,* 6, 295–300.

Selvais, P.L., Robert, A., Ahn, S., Van Linden, F., Ketelslegers, J.M., Pouleur, H. & Rousseau, M.F. (2000). 'Direct comparison between endothelin-1, N-terminal proatrial natriuretic factor, and brain natriuretic peptide as prognostic markers of survival in congestive heart failure', *Journal of Cardiac Failure,* 6, 201–207.

Setsuta, K., Seino, Y., Ogawa, T., Arao, M., Miyatake, Y. & Takano, T. (2002). 'Use of cytosolic and myofibril markers in the detection of ongoing myocardial damage in patients with chronic heart failure', *American Journal of Medicine,* 113, 717–722.

Shamsham, F. & Mitchell, J. (2000). 'Essentials of the diagnosis of heart failure', *American Family Physician,* 61, 1319–1328.

Shiba, N., Watanabe, J., Shinozaki, T., Koseki, Y., Sakuma, M., Kagaya, Y. & Shirato, K. (2004). 'Analysis of chronic heart failure registry in the Tohoku district - Third year follow-up', *Circulation Journal,* 68, 427–434.

Shibata, M.C., Flather, M.D. & Wang, D. (2001). 'Systematic review of the impact of beta blockers on mortality and hospital admissions in heart failure', *European Journal of Heart Failure,* 3, 351–357.

Sikaris, K. (2004). 'BNP - Considering a heartfelt message', *Heart Lung and Circulation,* 13S, S31-S37.

Sim, V., Hampton, D., Phillips, C., Lo, S.N., Vasishta, S., Davies, J. & Penney, M. (2003). 'The use of brain natriuretic peptide as a screening test for left ventricular systolic dysfunction - cost-effectiveness in relation to open access echocardiography', *Family Practice,* 20, 570–574.

Stata Corporation (2004). *Intercooled Stata 8.2 for Windows,* College City, Texas, Stata Corporation.

Swedberg, K., Cleland, J., Dargie, H., Drexler, H., Follath, F., Komajda, M., Tavazzi, L., Smiseth, O.A., Gavazzi, A., Haverich, A., Hoes, A., Jaarsma, T., Korewicki, J., Levy, S., Linde, C., Lopez-Sendon, J.L., Nieminen, M.S., Pierard, L. & Remme, W.J. (2005). 'Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology', *European Heart Journal,* 26, 1115–1140.

Tamura, K., Takahashi, N., Nakatani, Y., Onishi, S. & Iwasaka, T. (2001). 'Prognostic impact of plasma brain natriuretic peptide for cardiac events in elderly patients with congestive heart failure', *Gerontology,* 47, 46–51.

Teboul, A., Gaffinel, A., Meune, C., Greffet, A., Sauval, P. & Carli, P. (2004). 'Management of acute dyspnoea: use and feasibility of brain natriuretic peptide (BNP) assay in the prehospital setting', *Resuscitation,* 61, 91–96.

Troughton, R.W., Frampton, C.M., Yandle, T.G., Espiner, E.A., Nicholls, M.G. & Richards, A.M. (2000). 'Treatment of heart failure guided by plasma aminoterminal brain natriuretic peptide (N-BNP) concentrations', *Lancet,* 355, 1126–1130.

Tsutamoto, T., Wada, A., Maeda, K., Hisanaga, T., Mabuchi, N., Hayashi, M., Ohnishi, M., Sawaki, M., Fujii, M., Horie, H., Sugimoto, Y. & Kinoshita, M. (1999). 'Plasma brain natriuretic peptide level as a biochemical marker of morbidity and mortality in patients with asymptomatic or minimally symptomatic left ventricular dysfunction - Comparison with plasma angiotensin II and endothelin-1', *European Heart Journal,* 20, 1799–1807.

Tsutamoto, T., Wada, A., Maeda, K., Hisanaga, T., Maeda, Y., Fukai, D., Ohnishi, M., Sugimoto, Y. & Kinoshita, M. (1997). 'Attenuation of compensation of endogenous cardiac natriuretic peptide system in chronic heart failure - Prognostic role of plasma brain natriuretic peptide concentration in patients with chronic symptomatic left ventricular dysfunction', *Circulation,* 96, 509–516.

Tsutamoto, T., Wada, A., Maeda, K., Mabuchi, N., Hayashi, M., Tsutsui, T., Ohnishi, M., Fujii, M., Matsumoto, T., Yamamoto, T., Takayama, T. & Kinoshita, M. (2001). 'Relationship between plasma levels of cardiac natriuretic peptides and soluble Fas: plasma soluble Fas as a prognostic predictor in patients with congestive heart failure', *Journal of Cardiac Failure,* 7, 322–328.

Tsutsui, T., Tsutamoto, T., Wada, A., Maeda, K., Mabuchi, N., Hayashi, M., Ohnishi, M. & Kinoshita, M. (2002). 'Plasma oxidized low-density lipoprotein as a prognostic predictor in patients with chronic congestive heart failure', *Journal of the American College of Cardiology,* 39, 957–962.

Van Der Meer, P., Voors, A.A., Lipsic, E., Smilde, T.D., Van Gilst, W.H. & Van Veldhuisen, D.J. (2004). 'Prognostic value of plasma erythropoietin on mortality in patients with chronic heart failure', *Journal of the American College of Cardiology,* 44, 63–67.

Varela-Roman, A., Grigorian, L., Barge, E., Bassante, P., De La Pena, M.G. & Gonzalez-Juanatey, J.R. (2005). 'Heart failure in patients with preserved and deteriorated left ventricular ejection fraction', *Heart,* 91, 489–494.

Verdiani, V., Nozzoli, C., Bacci, F., Cecchin, A., Rutili, M.S., Paladini, S. & Olivotto, I. (2005). 'Pre-discharge B-type natriuretic peptide predicts early recurrence of decompensated heart failure in patients admitted to a general medical unit', *European Journal of Heart Failure,* 7, 566–571.

Vitarelli, A., Tiukinhoy, S., Di Luzio, S., Zampino, M. & Gheorghiade, M. (2003). 'The role of echocardiography in the diagnosis and management of heart failure', *Heart Failure Reviews,* 8, 181–189.

Vrtovec, B., Delgado, R., Zewail, A., Thomas, C.D., Richartz, B.M. & Radovancevic, B. (2003). 'Prolonged QTc interval and high B-type natriuretic peptide levels together predict mortality in patients with advanced heart failure', *Circulation,* 107, 1764–1769.

Watanabe, J., Shiba, N., Shinozaki, T., Koseki, Y., Karibe, A., Komaru, T., Miura, M., Fukuchi, M., Fukahori, K., Sakuma, M., Kagaya, Y. & Shirato, K. (2005). 'Prognostic value of plasma brain natriuretic peptide combined with left ventricular dimensions in predicting sudden death of patients with chronic heart failure', *Journal of Cardiac Failure,* 11, 50–55.

Weinberg, E.O., Shimpo, M., Hurwitz, S., Tominaga, S., Rouleau, J.L. & Lee, R.T. (2003). 'Identification of serum soluble ST2 receptor as a novel heart failure biomarker', *Circulation,* 107, 721–726.

Whiting P.R.A., Reitsma Jb, Bossuyt Pm, Kleijnen J. (2003). 'The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews', *BMC Medical Research Methodology,* 3, 25.

Wijeysundera, H.C., Hansen, M.S., Stanston, E., Cropp, A.S., Hall, C. & Dhalla, N.S.E.A. (2003). 'Neurohormones and oxidative stress in nonischemic cardiomyopathy: relationship to survival and the effect of treatment with amlodipine', *American Heart Journal,* 146, 291–297.

Wright, S.P., Doughty, R.N., Pearl, A., Gamble, G.D., Whalley, G.A., Walsh, H.J., Gordon, G., Bagg, W., Oxenham, H., Yandle, T., Richards, M. & Sharpe, N. (2003). 'Plasma amino-terminal pro-brain natriuretic peptide and accuracy of heart-failure diagnosis in primary care - A randomized, controlled trial', *Journal of the American College of Cardiology,* 42, 1793–1800.

Yu, C.M. & Sanderson, J.E. (1999). 'Plasma brain natriuretic peptide - an independent predictor of cardiovascular mortality in acute heart failure', *European Journal of Heart Failure,* 1, 59–65.

Zugck, C., Haunstetter, A., Kruger, C., Kell, R., Schellberg, D., Kubler, W. & Haass, M. (2002). 'Impact of beta-blocker treatment on the prognostic value of currently used risk

predictors in congestive heart failure', *Journal of the American College of Cardiology,* 39, 1615–1622.

B-type natriuretic peptide assays in the diagnosis of heart failure in the non-hospital setting

---

May 2007

MSAC Application 1087

Assessment report (Part B)

Executive summary

Part A of this report assesses the use of two B-type natriuretic peptide assays (BNP and NT-proBNP) in three key areas (diagnosis, monitoring and prognosis) for suspected and diagnosed heart failure (HF) patients, with the diagnostic use occurring in the hospital setting. Part B of this report assesses the diagnostic use of the two B-type natriuretic peptide assays to rule out HF in patients presenting in a non-hospital setting.

The procedure

B-type natriuretic peptide testing involves a blood test to determine the level of cardiac neurohormone circulating in the blood of a patient suspected or diagnosed with HF. Levels of two types of cardiac neurohormone can be tested—brain natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP).

These B-type natriuretic peptides act as counter-regulatory hormones to stabilise circulatory function. In an attempt to maintain cardiac output from a failing heart, the renin-angiotensin-aldosterone system is activated to enhance blood volume retention, circulatory vasoconstriction and ventricular remodelling, in order to maintain ventricular pre-load. This physiological response to the failing heart actually increases the workload of the heart because of an increase in vascular resistance and after-load. The circulatory volume overload stretches cardiac myocytes which then release the B-type natriuretic peptides to stabilise circulatory function.

Both peptides have been implicated as diagnostic biomarkers for suspected HF in clinical practice. In this context it is suggested that assays or tests of these peptides may complement conventional diagnostic strategies and thus assist with the identification of suspected HF in symptomatic patients. Patients with low levels of the cardiac neurohormones are '**ruled out**' for HF through these tests and are investigated for differential diagnoses; those **not excluded** from HF may go on to other confirmatory testing such as an echocardiogram.

Medical Services Advisory Committee – role and approach

The Medical Services Advisory Committee (MSAC) was established by the Australian Government to strengthen the role of evidence in health financing decisions in Australia. The MSAC advises the Minister for Health and Ageing on the evidence relating to the safety, effectiveness and cost-effectiveness of new and existing medical technologies and procedures, and under what circumstances public funding should be supported.

A rigorous assessment of evidence is thus the basis of decision-making when funding is sought under Medicare. A team from Adelaide Health Technology Assessment (AHTA), Discipline of Public Health, School of Population Health and Clinical Practice, University of Adelaide was engaged to conduct a systematic review of the literature (Part A and Part B of this report) on B-type natriuretic peptide assays in the diagnosis, monitoring and prognosis of HF. An advisory panel with expertise in this area then evaluated the evidence and provided advice to the MSAC.

MSAC's assessment of B-type natriuretic peptide assays in a non-hospital setting

Clinical need

Heart failure is commonly cited to afflict 300,000 Australians, with approximately 30,000 new cases occurring each year. However, these figures underestimate the number of patients *suspected* of having HF each year and thus who would potentially receive a B-type natriuretic peptide test. Patients presenting with symptoms like dyspnoea (breathlessness) may have HF or, alternatively, chronic obstructive pulmonary disease, pneumonia, emphysema or other lung diseases.

The BEACH data on Australian General Practice Activity show that 0.83 per cent (95%CI 0.6%, 1.0%) of presenting problems in 2004–05 were for 'shortness of breath, dyspnoea' (AIHW 2005). Given that there were almost 89 million Medicare Benefits Schedule claims for general practitioner (GP) consultations in 2005 and BEACH data suggest that 54.5 per cent of GP attendances were for patients aged 45 years and older, then approximately 398,893 GP attendances in patients aged 45+ years were for symptoms of dyspnoea in 2005.

Only patients with *new* symptoms of HF are likely to receive a B-type natriuretic peptide test because for those patients with uncontrolled previously diagnosed HF, management and monitoring are required, rather than a diagnosis. New symptoms of suspected HF were reported in 24.6 per cent of patients in the CASE study (Krum et al 2001). This estimate is not ideal as it relates to patients aged 60 years and older and includes a much wider definition of suspected HF than just dyspnoea symptoms. However, the study sample included patients who did or did not have a prior diagnosis of HF, which is similar to the target population in Australian general practice. Wright et al (2003) reported that of patients aged 40+ years *without* a previous history of HF and presenting to GPs with symptoms of dyspnoea and/or oedema of recent onset, 70 per cent were suspected of having HF. These two rates were applied to indicate the likely range in clinical need for the B-type natriuretic peptide tests in patients aged 45+ years with symptoms of dyspnoea and/or oedema of recent onset. Approximately 97,956 to 279,225 patients could require a B-type natriuretic peptide test according to the information given in Krum et al (2001) and Wright et al (2003), respectively. It is unclear at what rate patients presenting with *new* symptoms of suspected HF are referred to hospital or cardiologists/physicians. However, BEACH data (AIHW 2001) indicate that 19.7 per cent of patients in general practice with confirmed *or* suspected HF will be referred to hospital or a cardiologist/physician. Assuming these patients do *not* receive a B-type natriuretic peptide test prior to referral, then between 78,658 and 224,218 patients per year, aged 45+ years and with dyspnoea and/or oedema of recent onset and suspected of HF, would be eligible to receive a B-type natriuretic peptide test.

Safety

The likelihood of adverse events occurring during B-type natriuretic peptide testing is small and similar to that of other venepuncture blood tests. False positive test results may theoretically cause harm to the patient through sequelae such as inappropriate treatment for HF. False negative test results are rare as the tests have very high negative predictive value. The fact that the population in question is not acutely ill means that even those

very few false negative test results (which would be associated with inappropriately delayed treatment for HF) are unlikely to be harmful to these patients. There were no studies in the available evidence base that actually reported physical or psychological adverse events as a result of B-type natriuretic peptide testing in a non-hospital setting.

Effectiveness

Diagnosis

BNP assays

The effectiveness of supplementing conventional diagnostic assessment in the non-hospital setting with BNP testing was evaluated by a small volume of evidence, with the highest quality data obtained from one high quality level II diagnostic accuracy study. Because there was no direct evidence available evaluating the effect of the BNP test on patient health outcomes in this setting, a linked evidence approach was undertaken. Evidence concerning the diagnostic accuracy of the test was to be linked to evidence of the impact of the test on patient management, and this then linked to the effect of this type of patient management on patient health outcomes. Five studies provided evidence of the diagnostic accuracy of the test. Overall, this body of evidence was relatively consistent in its findings that the BNP blood test is sensitive, with a high negative predictive value. The specificity of the test is variable. Its main role, therefore, appears to be as a 'first line' test, as a negative result on the test 'rules out' the diagnosis of HF, so that differential diagnoses can be investigated.

The impact of the introduction of a BNP test on the management of patients by GPs could not be directly determined. There were no change-in-management studies, associated with the use of BNP tests in a non-hospital setting, that were available for inclusion in this review at the time of searching.

There were also no studies available to assess the direct impact of BNP testing on health outcomes. A systematic review of the impact of early treatment on health outcomes for patients with and without HF was beyond the scope of this report. High level evidence does suggest that early treatment for HF is beneficial for the patient. It is unlikely, however, that use of the BNP test would result in an *earlier* identification of HF, than currently, for these symptomatic suspected HF patients. Use of the test *is* likely to assist in earlier identification of alternative diagnoses for those patients 'ruled out' from HF—and most of these alternative diagnoses (pulmonary diseases, asthma, anaemia) have established treatments. Should this alternative pathology be severe enough, early identification and treatment is likely to be beneficial to the patient.

The populations studied in the included diagnostic studies are reasonably applicable to the target population in Australia, that is patients presenting to general practice with symptoms (eg dyspnoea) suggestive of HF. As a group, however, they may have had slightly more severe symptoms than is usual in general practice, as most were selected on the basis of *referral* from a GP on suspicion of HF. The results of the studies are largely generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in diagnosing and managing symptomatic suspected HF.

In conclusion, on the basis of the evidence presented, BNP testing appears to be a valuable 'first line' diagnostic test that, when added to conventional diagnostic

assessment, assists the GP to determine that patient symptoms are *not* caused by HF. The clinical impact of the test is, however, currently unknown in the non-hospital setting.

NT-proBNP assays

The effectiveness of NT-proBNP testing added to conventional diagnostic assessment was evaluated by a small volume of evidence (5 studies), with the most reliable data being obtained from one high quality level II intervention study and one high quality level II diagnostic accuracy study. Overall, the diagnostic accuracy evidence was relatively consistent in its findings that NT-proBNP assays are sensitive with (in general) high negative predictive values (>90%), indicating that the test effectively 'rules out' HF in patients with a negative result. The specificity of the test is low and variable.

One good quality randomised controlled trial reported the impact of NT-proBNP testing on clinical diagnoses formulated in general practice. A 13 per cent improvement [95%CI 5.5, 21.0, p=0.002] was observed in correct diagnoses in the NT-proBNP trial arm compared to the control trial arm, with the main impact occurring by enabling GPs to correctly 'rule out' HF.

There were no studies available to assess the direct impact of NT-proBNP testing on health outcomes. A systematic review of the impact of early treatment on health outcomes for patients with and without HF was beyond the scope of this report. As mentioned above, high level evidence suggests that early treatment for HF is beneficial for the patient, but it is unlikely that use of the NT-proBNP test would result in an earlier identification of HF, than currently, for these patients. Use of the test would, however, identify earlier alternative diagnoses for those patients 'ruled out' from HF, most of which have established treatments. Should this alternative pathology be severe enough, then early identification and treatment is likely to be beneficial to the patient.

The populations studied in the available evidence base are similar to the target population in Australia, that is patients presenting to general practice with symptoms—primarily dyspnoea and/or oedema of recent onset—suggestive of HF. The results of the studies are generalisable to the Australian healthcare context, with most being conducted in developed countries (including New Zealand) with similar standards of practice in diagnosing and managing suspected HF.

In conclusion, on the basis of the evidence presented, NT-proBNP assays appear to be sensitive 'first line' diagnostic tests that, when added to conventional diagnostic assessment, appear to change the practice of GPs. Their main role is to assist the GP to correctly 'rule out' HF more frequently in those patients presenting with dyspnoea and/or oedema of recent onset and suspected of HF. The clinical impact of the test (ie on patient health) is, however, currently unknown in the non-hospital setting.

Economic implications

B-type natriuretic peptide assays in a non-hospital setting are 'first line' tests. The extent of their effectiveness, (ie in terms of life-years or quality-adjusted life-years (QALYs) gained) depends on the extent to which they hasten the establishment of the correct definitive diagnosis, as well as on the influence of a correct diagnosis on the outcome of the disease. In turn, the extent of their *cost-effectiveness* hinges on the value of the additional information made available by the tests in terms of health gain and resource savings compared to the cost of providing the tests.

This analysis has largely been limited to symptomatic patients (in line with the evidence base), that is those presenting with dyspnoea and/or oedema of recent onset and suspected of HF. Appraisal of the economic implications of using B-type natriuretic peptide tests in general practice was hindered by the absence of any randomised controlled trial in that setting with a *health outcome* as an endpoint. Thus, it has not been possible to estimate an incremental cost-effectiveness ratio based on life-years saved or QALYs saved in the non-hospital setting.

Decision analytic modelling involves asking 'What if ...' questions. In this instance, as there were insufficient, adequate data to populate a model, it involved asking what circumstances might substantially reduce the cost-effectiveness of a B-type natriuretic peptide test. The costs and outcomes associated with B-type natriuretic peptide testing will usually depend on the GP's referral propensity. The extent of immediate cost offsets depends on whether or not the GP decides to order an echocardiogram and initially undertake self-management of the patient; or to refer the test positive patient to a cardiologist who may or may not order an echocardiogram. The extent of downstream costs (or savings) also depends on the same referral propensity, and on the proportions of patients correctly diagnosed.

Three scenarios have been presented illustrating different types of possible referral patterns that have increasing levels of resource use. In all three scenarios the use of B-type natriuretic peptide testing is cost saving (from $50 to $86 per patient tested), due primarily to the reduced need for echocardiograms and/or cardiologist referral in patients who test negative for HF. These scenarios reflect current clinical practice guideline recommendations and assume that in Australia *all* patients over 45 years of age presenting with dyspnoea and/or oedema of recent onset and suspected of HF would receive an echocardiogram. However, the data available suggests that actual echocardiogram referral may range from 3.8 per cent to 17.7 per cent for patients with new symptoms suggestive of HF. Unfortunately there are no robust data available on the referral patterns of GPs presented with patients reporting dyspnoea and/or oedema of recent onset and suspected of HF. It is likely that referral patterns will vary widely between GPs and, despite B-type natriuretic peptide tests reducing the number of 'ruled out' patients going on to an echocardiogram, it is unknown what impact the introduction of these tests will have on the current echocardiogram referral rate for those who test positive. Results of a one-way sensitivity analysis suggest that B-type natriuretic peptide testing may not be cost saving if GPs currently refer this patient group to echocardiography at a rate of 60 per cent or lower.

Despite this, when testing is confined to general practice patients with dyspnoea and/or oedema of recent onset and suspected of HF, most of those patients who test negative on the B-type natriuretic peptide assay are still likely to have a clinically important pathology. In this situation, the potential for improvement in health outcome may warrant the amount of resources used in testing and follow-up.

One-way sensitivity analysis has suggested that diminishing marginal returns could arise should the testing extend to general practice populations where there is a high probability that the patient has HF. They may also arise if testing occurs on increasing proportions of patients with minor levels of symptoms but lacking any clinically important pathology. The incremental cost of testing this much larger population would not be counterbalanced by the increment in health gains from testing. It is therefore of critical importance that B-type natriuretic peptide testing is only ordered for those patients with

dyspnoea and/or oedema of recent onset, for whom there is real uncertainty as to whether the symptoms are caused by HF or an alternative pathology. In practice, the Australian Government can employ several strategies for ensuring that this test is ordered appropriately, including: (1) facilitating intensive education programs for health professionals (particularly important given that this test is most effective when used to 'rule out' HF); (2) restricting payment of benefits for the item to specific indications; and (3) limiting augmentation of the 'pathology cap' to a level that is consistent with the restricted use of the item.

Finally, the results of the key trial in the emergency department (ED) setting (see Part A of this report), perhaps with some adjustment for the severity of clinical presentation, is relevant for the rural and remote setting where the GP decides to admit the patient to the local hospital and continue inpatient management.

The additional Australian Government expenditure due to the introduction of B-type natriuretic peptide assays into a non-hospital setting for patients presenting with dyspnoea and/or oedema of recent onset is estimated to range between $4.0 million and $11.3 million per year. This expenditure is likely to be offset by savings on fewer echocardiograms and cardiologist referrals and earlier management of non-HF diagnoses, but the extent of these offsets is presently unknown.

Recommendation

'MSAC has considered the safety, effectiveness and cost effectiveness of the use of assays of B-type natriuretic peptides in the diagnosis of heart failure in patients presenting with dyspnoea in the non-hospital setting when compared with current clinical practice +/- echocardiography.

MSAC finds that there is sufficient evidence that B-type natriuretic peptide assays, when used in the diagnosis of heart failure in patients presenting with dyspnoea, are safe and effective (diagnostically accurate).

MSAC finds that there is major uncertainty around the cost effectiveness in the non - hospital setting.

MSAC recommends that public funding is <u>not supported</u> for the use of assays of B-type natriuretic peptides in the diagnosis of heart failure in patients presenting with dyspnoea in the non-hospital setting at this time.

MSAC further recommends research on the use of BNP in the general practice setting to identify appropriate usage and the patient group most likely to benefit in the non hospital setting.'

The Minister for Health and Ageing accepted this recommendation on 29 August 2007.

Glossary

ACE	Angiotensin-converting enzyme
Area under the curve	Calculated as the area under a receiver operator characteristic curve, the area under the curve (AUC) provides a numerical description of the accuracy of a diagnostic test. A test with no diagnostic value has an AUC of 0.5, while a perfect test has an AUC of 1.0.
BNP	Brain (or B-type) natriuretic peptide
Dyspnoea	A distressful sensation of uncomfortable breathing or breathlessness
ECG	Electrocardiogram
GP	General practitioner
HF	Heart failure
ITT	Intention-to-treat
LVEF	Left ventricular ejection fraction
LVSD	Left ventricular systolic dysfunction
MBS	Medicare Benefits Schedule
Meta-analysis	Meta-analysis refers to the statistical analysis of a number of individual study results for the purpose of integrating findings.
MSAC	Medical Services Advisory Committee
NT-proBNP	N-terminal proBNP (nucleotides 1–76)
NYHA	New York Heart Association
OECD	Organisation for Economic Co-operation and Development
Power	Power refers to the ability of a statistical test to reject a false null hypothesis.
QALY	Quality-adjusted life-year
Receiver operator characteristic curve	A receiver operator characteristic curve (ROC) is a plot of sensitivity against 1 minus specificity for different values of a diagnostic test. It highlights the trade-off between sensitivity and specificity, and gives an overall indication of the diagnostic accuracy of a test.
Sensitivity	Sensitivity refers to the proportion of people with a disease who report a positive test result.
Specificity	Specificity refers to the proportion of people without a disease who report a negative test result.

Introduction

The Medical Services Advisory Committee (MSAC) has reviewed the use of B-type natriuretic peptide assays in the diagnosis of suspected heart failure (HF) in the hospital setting, along with the prognosis and monitoring of patients with HF (see Part A of this report).

The MSAC has also reviewed the use of B-type natriuretic peptide tests in the diagnosis of suspected HF in the non-hospital setting in Part B of this report. The MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Scheme in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. The MSAC adopts an evidence-based approach to its assessments, based on reviews of the scientific literature and other information sources, including clinical expertise.

The MSAC's terms of reference and membership are at Part A, . The MSAC is a multidisciplinary expert body, comprising members drawn from such disciplines as diagnostic imaging, pathology, surgery, internal medicine and general practice, clinical epidemiology, health economics, consumer health and health administration.

Rationale for assessment

Roche Diagnostics Australia and Abbott Diagnostics Australia have made separate applications to the Medical Services Advisory Committee (MSAC) to have the Elecsys® proBNP and the AxSYM® BNP assays placed on the Medicare Benefits Schedule for the diagnosis, monitoring and prognosis of HF. These assays are performed on patient blood extracted through a simple blood test. They measure brain natriuretic peptide (BNP) or the by-product of the cleavage from the precursor of BNP to BNP (NT-proBNP) and would be performed by clinical laboratories, either in a public hospital or a private pathology laboratory. It is suggested that measurement of BNP and/or NT-proBNP would not replace traditional clinical investigations but may assist in the selection of patients who would benefit most from receiving further investigations. These assays are not designed to screen patients without risk factors but to act as a 'first line' test for individuals who are suspected of having HF due to various signs or symptoms (such as new symptoms of dyspnoea and oedema).

BNP and NT-proBNP testing are considered new medical services requiring a new Medicare item number.

For background information on suspected HF, the B-type natriuretic peptides and the assays used to assess B-type natriuretic peptide levels, see Part A of this report.

Intended purpose

In the terminology coined by Sackett and colleagues, the proposed value of BNP and NT-proBNP assays is as 'SnOut' tests. These tests are considered highly sensitive; therefore, a negative test result 'rules *out*' the diagnosis (Sackett et al 1991). It is proposed, therefore, that B-type natriuretic peptide tests would act as '**first line**' **diagnostic tools** to identify patients who should or should not be referred for echocardiography to

confirm a clinical diagnosis of HF or to explore alternative diagnoses. As such they could be used *either* as supplemental *or* replacement tests in the diagnostic workup. For those patients 'ruled out' from HF they would *replace* the usual confirmatory HF tests. For those patients not excluded from HF, they would *supplement* the usual confirmatory HF tests. The role of the tests is not to act as a 'reference standard' as their specificity is generally not high.

Comparators

Diagnosis

Comparators

The most common method of diagnosing HF is by clinical examination. This involves a review of the patient's medical history and a physical examination, including observation, palpation and auscultation. The World Health Organization criteria (Table 43) may be used to diagnose HF. The subjective nature of a diagnosis made on clinical features alone is a weakness of this method. Clinical evaluation may be used in conjunction with objective tests, including electrocardiograms, chest X-rays, and echocardiography when available (NHF & CSANZ 2002). Laboratory investigations (eg blood count, creatinine and urinalysis) are also part of the routine diagnostic evaluation for HF.

Table 43 Modified World Health Organization criteria for assessment of possible chronic heart failure, 1995

Symptoms	Dyspnoea, chronic fatigue, oedema, exercise intolerance
Signs	Third or fourth heart sounds, heart murmur, cardiomegaly, pulmonary crackles, raised jugular venous pressure, dependent oedema
Causative factors	Angina, previous myocardial infarction, hypertension, valvular heart disease/rheumatic fever, cardiomyopathy
Possible HF is considered if patients have:	• 2 symptoms • ≥ 2 signs • ≥ 1 symptom and ≥ 1 sign • ≥1 symptom and ≥ 1 causative factor

HF = heart failure
Source: Krum 2001

Heart failure cannot be diagnosed or excluded reliably on the basis of clinical examination alone (see Table 44). Although incorrect treatment due to misdiagnosis may alleviate the patient's symptoms, it may obscure an underlying problem that worsens over time. This is particularly relevant in the elderly population, who are most at risk of HF and in whom multiple diseases are common (Remme & Swedberg 2001). Furthermore, when elderly patients experience symptoms upon exertion, they are likely to restrict their activity levels to reduce the symptoms, which can lead to deconditioning. Diagnosis within this population is difficult as onset may be slow, and HF may be asymptomatic at lower levels of exertion (Shamsham & Mitchell 2000).

Table 44 European Society of Cardiology definition of heart failure.

I	Symptoms of heart failure (at rest or during exercise)
	and
II	Objective evidence (preferably by echocardiography) of cardiac dysfunction (systolic and/or diastolic) (at rest)
	and (in cases where the diagnosis is in doubt)
III	Response to treatment directed towards heart failure

Source: Swedberg et al 2005

Despite the need for accurate diagnosis, many physicians, particularly in general practice, rely on clinical grounds alone to diagnose HF since the availability of echocardiograms may be limited. It also requires the services of an experienced cardiologist for interpretation, and patients with dyspnoea may find it difficult to lie down long enough for an echocardiogram (Hobbs 2002).

The reference standard

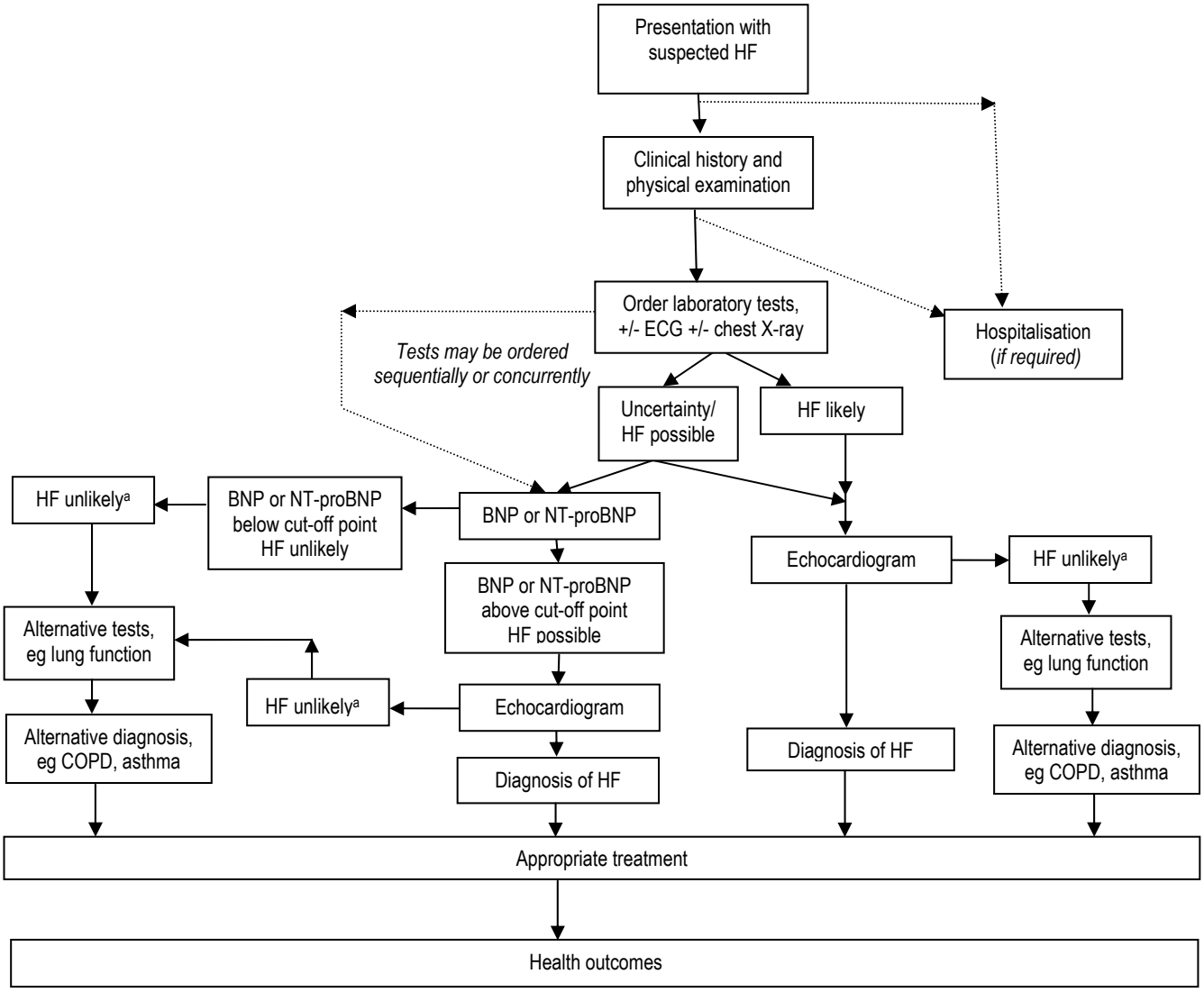
The diagnostic accuracy of B-type natriuretic peptide assays can be assessed against the objective measure of ventricular function provided by the transthoracic echocardiogram (de Denus et al 2004). Echocardiography uses ultrasound to image the heart and surrounding tissues, providing structural and functional information. Left ventricular ejection fraction is the key parameter for distinguishing patients with cardiac systolic dysfunction from those with preserved systolic function (Remme & Swedberg 2001). Measurements of left ventricular relaxation (left ventricular end diastolic diameter) and filling pressures (via catheterisation or Doppler echocardiography) are considered the best objective measures to assess diastolic HF, but echocardiography may also be used for this purpose by assessing mitral inflow and pulmonary venous flow (Dhir et al 2004).

The transthoracic echocardiogram is a painless, non-invasive procedure that takes between 15 and 30 minutes, and involves the patient lying still on their back on the examination table with their chest exposed. The radiologist or technician applies gel onto the skin to allow the transducer to slide against the skin, emitting ultrasound waves that bounce back, or 'echo' off the structures of the heart (Penn State College of Medicine 2004).

However echocardiography is an imperfect reference standard. The 'reference standard' should be the echocardiogram result taken in conjunction with a clinical diagnosis of HF (based on all information including signs, symptoms and other tests, eg chest X-ray). The 'gold standard' for diagnosing HF is usually *consensus* cardiologist opinion integrating clinical (signs and symptoms of HF) and objective tests, including echocardiography.

To assess the effectiveness of the BNP and NT-proBNP assays as 'first line' diagnostic tests, the effect on patient relevant outcomes of the addition of either of these assays to existing diagnostic strategies (ie clinical examination/diagnostic workup in conjunction with laboratory tests, chest X-ray, ECG and/or echocardiogram) would need to be compared to the effect on patient relevant outcomes of the existing diagnostic strategies alone (see Figure 11).

Figure 11 Generic clinical pathway for use of B-type natriuretic peptide assays in the diagnosis of heart failure in a non-hospital setting



HF = heart failure; ECG = electrocardiogram; COPD = chronic obstructive pulmonary disease; [a] systolic heart failure unlikely; diastolic heart failure may still be a possibility

Clinical need and burden of disease

The clinical need for an additional tool to assist in the diagnosis of HF is dependent on the number of patients *suspected* of having HF due to symptoms, signs and causative factors (eg diabetes mellitus, kidney problems, obesity, hypertension, coronary artery disease). In the Australian Cardiac Awareness Survey and Evaluation (CASE) Study, it was found that 4,807 (21.8%) from 22,060 consecutive patients over 60 years of age who presented to their general practitioner (GP) were suspected of having HF (Krum et al 2001). Of these suspected HF patients, only 420 had objective clinical evidence of HF, resulting in a confirmed:suspected HF ratio of 1:11. In contrast, Hobbs et al (2000) reported a much lower confirmed:suspected HF ratio of 1:3, also in the non-hospital setting.

Heart failure in Australia is commonly cited to afflict 300,000 individuals, with approximately 30,000 new cases occurring per annum. However, as is evidenced by these confirmed:suspected HF ratios, these figures could substantially underestimate the number of *suspected* HF patients that would potentially receive B-type natriuretic peptide testing. Patients can present with similar symptoms to HF but receive alternative diagnoses. The most common acute symptom associated with suspected HF is dyspnoea (breathlessness). Alternative diagnoses can include chronic obstructive pulmonary disease, pneumonia, emphysema or other lung diseases, anaemia and asthma.

In the 2005 calendar year, 88,680,935 GP attendances were claimed in Australia overall on the Medicare Benefits Schedule. The BEACH data on Australian General Practice Activity show that 0.83 per cent (95%CI 0.6%, 1.0%) of presenting problems in general practice in 2004–05 were for 'shortness of breath, dyspnoea' (AIHW 2005). Given that BEACH data also suggest that 54.5 per cent of GP attendances were for patients aged 45 years and older, approximately 398,893 GP consults in 2005 were for patients aged 45+ years with symptoms of dyspnoea.

A robust estimate could not be identified for the number of patients presenting to GPs with oedema of recent onset (another symptom commonly associated with a diagnosis of suspected HF). However, given that there is a likely overestimate of BEACH dyspnoea cases as being suspected of HF (ie the dyspnoea category would include cases of known asthma), it is likely that this overestimate would in some part correct for the missing data on cases of oedema of recent onset.

Only patients with *new* symptoms of HF are likely to receive a B-type natriuretic peptide test because for those patients with uncontrolled previously diagnosed HF, management and monitoring is required rather than a diagnosis. New symptoms of suspected HF were reported in 24.6 per cent of patients in the CASE study (Krum et al 2001). This estimate is not ideal as it relates to patients aged 60 years and older and includes a much wider definition of suspected HF than just dyspnoea and/or oedema of recent onset. However, the study sample included patients who did or did not have a prior diagnosis of HF, which is similar to the target population in Australian general practice. Wright et al (2003) reported that of patients aged 40+ years *without* a previous history of HF and presenting to GPs with symptoms of dyspnoea and/or oedema of recent onset, 70 per cent were suspected of HF. These two rates were applied to indicate the likely range in clinical need for the B-type natriuretic peptide tests in patients aged 45 years and over with new symptoms of dyspnoea and/or oedema. Approximately 97,956 to 279,225 patients could require a B-type natriuretic peptide test using the data provided by Krum

et al (2001) and Wright et al (2003), respectively. It is unclear at what rate patients presenting with *new* symptoms of suspected HF are referred to hospital or cardiologists/physicians. However, BEACH data (AIHW 2001) indicate that 19.7 per cent of patients in general practice with confirmed *or* suspected HF are referred to hospital or a cardiologist/physician. Assuming that these patients do *not* receive a B-type natriuretic peptide test prior to referral, then between 78,658 and 224,218 patients per year, aged 45+ years and with dyspnoea and/or oedema of recent onset and suspected of HF, would be eligible to receive a B-type natriuretic peptide test.

Current treatments

Heart failure

Early treatment of HF is important for preventing or retarding progression of the disease (Hammerer-Lercher et al 2001). Treatment strategies will depend on the cause and severity of the disease.

When a specific cause of HF is able to be identified, it should be addressed and if possible corrected (Leibovitch 2005). This could involve withdrawal of drugs which dampen cardiac function, or treating potentially reversible diseases. For instance, if HF is due to hypertension, thyroid dysfunction, sleep apnoea or renal failure, then these causes should be addressed (Leibovitch 2005). Alternatively, if HF is due to an abnormal heart valve, the valve could be surgically replaced (American Heart Association 2005).

Most patients who experience HF will be advised to consider a number of non-pharmacological measures such as taking regular physical exercise, reducing their intake of dietary sodium to below 2,000 mg/day, and limiting fluid intake (1.5 L/day for mild to moderate HF and 1 L/day in severe HF). Smoking and alcohol intake are strongly discouraged. Patients' weight gain is monitored and they may be vaccinated against influenza and pneumococcal disease (NHF & CSANZ 2002).

A number of pharmacological agents are available for the treatment of systolic HF (left ventricular ejection fraction <40%), depending on the classification of the patient's presenting symptoms (eg New York Heart Association (NYHA) classes I-IV). Pharmacological agents include diuretics with or without angiotensin-converting enzyme (ACE) inhibitors. Depending on the progress of the patient and the reduction of symptoms, these agents may be supplemented by the use of beta-blockers. In cases of persistent oedema, spironolactone with or without digoxin may be prescribed. If these pharmacological measures are ineffective or cannot be tolerated by the patient, a heart transplant may be considered for patients <65 years of age with no other major comorbidity (NHF & CSANZ 2002).

Strong evidence of pharmacological effectiveness in treating HF is reported in two systematic literature reviews on beta-blockers (Shibata et al 2001) and ACE inhibitors (Garg & Yusuf 1995). The former review on beta-blockers analysed 22 single or double-blinded randomised controlled trials that assessed five different beta-blocker agents. The pooled analysis included 10,480 patients (5,507 with active treatment; 4,973 as a placebo group) who were followed up for a mean of 11 months with a completeness of follow-up of 85 per cent. Most patients were categorised at baseline as NYHA functional class III (63.3%). The pooled effect (odds ratio, OR) measures due to beta-blocker therapy for all-

cause mortality and hospitalisation were 0.65 (95%CI 0.57, 0.74; p<0.00001) and 0.63 (95%CI 0.56, 0.71; p<0.00001), respectively. Beta-blocker therapy therefore conferred a 35 per cent reduction in the chance of dying and a 37 per cent reduction in the probability of hospitalisation in HF patients relative to placebo treatment. A systematic review of ACE inhibitor effectiveness (Garg & Yusuf 1995), which included 32 randomised trials, resulted in a pooled analysis of 7,105 patients each randomised to a placebo or one of eight ACE inhibitors (predominately Enalapril). The majority of HF patients included in the meta-analysis were classified as NYHA class II or III. A risk reduction of 23 per cent (OR = 0.77; 95%CI 0.67, 0.88) was reported for all-cause mortality and 35 per cent for hospitalisation (OR = 0.65; 95%CI 0.57, 0.74) relative to placebo treatment. Taken together, these systematic reviews suggest that beta-blockers and ACE inhibitors are clearly effective treatment options for HF.

Cardiac assist devices such as implantable cardioverter-defibrillators have been found to reduce mortality rates, but are associated with very high costs. Heart transplantation is also very effective, but a scarcity of resources (ie human hearts) limits availability of the technique (Leibovitch 2005).

In the Heart Outcomes Prevention Evaluation (HOPE) study, it was found that patients at risk of cardiovascular events—such as those with coronary artery disease, stroke, peripheral vascular disease or diabetes and one other risk factor such as hypertension, without any evidence of HF or left ventricular dysfunction—benefited from receiving treatment. In a randomised controlled trial it was found that patients who received ramipril (ACE inhibitor) were less likely to develop heart disease or experience cardiovascular events than patients who received a placebo (Aurbach et al 2004).

Differential diagnoses

Given the wide variety of alternative diagnoses possible for patients suspected of HF who present primarily with dyspnoea, including chronic obstructive pulmonary disease, pneumonia, emphysema or other lung diseases, anaemia and asthma, it is difficult to determine the effectiveness and availability of treatments for these pathologies without conducting another systematic literature review.

It is, however, probable that correct and early identification of the alternative diagnosis and prompt treatment would be beneficial for the patient, particularly in cases where there is a severe or acute presentation.

Potential impact of the test

Should B-type natriuretic peptide assays be publicly funded in the diagnosis of HF in the non-hospital setting, the potential impact is likely to be extensive. There is a large population that requires testing for new HF-like symptoms (diagnosis). These patients currently represent the largest diseased population within Australia. Given the ageing of the Australian population, prevalence of risk factors (eg obesity, physical inactivity) and the fact that more individuals are surviving acute coronary events (ie myocardial infarction), HF incidence and prevalence will continue to rise.

The technology for testing BNP and NT-proBNP levels is already established. Pathology services currently offer testing via a fee-for-service arrangement. Should these tests

receive public funding, a substantial increase in use of current pathology services for B-type natriuretic peptides would be driven by GPs using the test to rule out HF in symptomatic patients suspected of HF. This cost would be borne by the Commonwealth.

The unit cost of the B-type natriuretic peptide tests is estimated to be $50.59 per test on the basis of laboratory benchmarking data (see economic considerations section and Appendix G of Part A of this report).

Marketing status of the technology

At the time of writing, all therapeutic products marketed in Australia require listing on the Australian Register of Therapeutic Goods (ARTG) unless they have an exemption.

According to the applications submitted to the MSAC, both the Abbott AxSYM BNP and Roche NT-proBNP analytic systems are exempt from the *Therapeutic Goods Administration Act 1998* because the proposed diagnostic tests are not used for blood screening, are not used by consumers, do not contain material of human/animal origin, are not listed on the Pharmaceutical Benefits Schedule and are not used for human immunodeficiency virus or hepatitic C testing.

Current reimbursement arrangement

Although B-type natriuretic peptide assays are being used in Australia, there are currently no items on the Medicare Benefits Schedule that cover these products. This assessment is being conducted to determine whether these tests should receive public funding for use in the non-hospital setting.

Approach to assessment

Objectives

The objective of this assessment is to determine whether there is sufficient evidence, in relation to clinical need, safety, effectiveness and cost-effectiveness, to use B-type natriuretic peptide (BNP or NT-proBNP) assays in the diagnosis of heart failure (HF) in the non-hospital setting.

Research questions

Safety

- Is the use of the BNP assay as a 'first line' diagnostic test in the non-hospital setting, in conjunction with standard clinical assessment[13] ± echocardiography, as safe as, or safer than, standard clinical assessment ± echocardiography alone in the diagnosis of heart failure?

- Is the use of the NT-proBNP assay as a 'first line' diagnostic test in the non-hospital setting, in conjunction with standard clinical assessment ± echocardiography, as safe as, or safer than, standard clinical assessment ± echocardiography alone in the diagnosis of heart failure?

Diagnostic effectiveness

Direct evidence

- Is the use of the BNP assay as a 'first line' diagnostic test in the non-hospital setting, in conjunction with standard clinical assessment ± echocardiography, as, or more, effective at improving the health outcomes associated with suspected heart failure than standard clinical assessment ± echocardiography alone?

- Is the use of the NT-proBNP assay as a 'first line' diagnostic test in the non-hospital setting, in conjunction with standard clinical assessment ± echocardiography, as, or more, effective at improving the health outcomes associated with suspected heart failure than standard clinical assessment ± echocardiography alone?

Linked evidence[14]

- What is the diagnostic accuracy of the BNP assay when used to diagnose heart failure in the non-hospital setting compared to clinical diagnosis and/or echocardiography?

---

[13] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs

[14] Used in situations where direct evidence of diagnostic effectiveness is not available or where it is limited

- What is the diagnostic accuracy of the NT-proBNP assay when used to diagnose heart failure in the non-hospital setting compared to clinical diagnosis and/or echocardiography?

- Does the BNP assay affect the clinical management or treatment options available to patients suspected of heart failure in the non-hospital setting?

- Does the NT-proBNP assay affect the clinical management or treatment options available to patients suspected of heart failure in the non-hospital setting?

- Does the BNP assay and possible alterations in clinical management in the non-hospital setting impact on the health outcomes associated with suspected heart failure?

- Does the NT-proBNP assay and possible alterations in clinical management in the non-hospital setting impact on the health outcomes associated with suspected heart failure?

Diagnostic cost effectiveness[15]

- Is the BNP assay cost-effective as a 'first line' test in the non-hospital setting, in conjunction with standard clinical assessment ± echocardiography, in the diagnosis of heart failure compared to standard clinical assessment ± echocardiography alone?

- Is the NT-proBNP assay cost-effective as a 'first line' test in the non-hospital setting, in conjunction with standard clinical assessment ± echocardiography, in the diagnosis of heart failure compared to standard clinical assessment ± echocardiography alone?

Expert advice

An advisory panel with expertise in pathology, clinical biochemistry, general practice, and consumer issues was established to evaluate the evidence from this Assessment Report and to provide advice to the MSAC from a clinical or consumer perspective. In selecting members for advisory panels, the MSAC's practice is to approach the appropriate medical colleges, specialist societies and associations and consumer bodies for nominees. Membership of the advisory panel associated with this MSAC assessment is provided in Part A, Appendix B).

Review of the literature

Literature sources and search strategies

The medical literature was searched to identify relevant studies concerning B-type natriuretic peptides for the period between 1988 and August 2005. B-type natriuretic

[15] Only investigated if there was evidence of clinical effectiveness.

peptide assays were first reported in 1988. Part A, Appendix C, describes the electronic databases that were used for this search and other sources of evidence that were investigated. Grey literature was included in the search strategy. Unpublished literature, however, was not canvassed as it is difficult to search for this literature exhaustively and systematically and trials that are difficult to locate are often smaller and of lower methodological quality (Egger et al 2003). It is, however, possible that these unpublished data could impact on the results of this assessment.

The search terms, presented in Part A, Appendix C, were used to identify literature in electronic bibliographic databases on the safety, effectiveness and cost-effectiveness of using B-type natriuretic peptide assays in the diagnosis of HF.

Inclusion/exclusion criteria

In general, studies were excluded if they:

- did not address the research question;
- did not provide information on the pre-specified target population;
- did not include one of the pre-specified interventions;
- did not compare results to the pre-specified comparator;
- did not address one of the pre-specified outcomes and/or provided inadequate data on these outcomes (in some instances, a study was included to assess one or more outcomes but had to be excluded for other outcomes due to data inadequacies); or
- did not have the appropriate study design.

Where two (or more) papers reported on different aspects of the same study, such as the methodology in one and the findings in the other, they were treated as one study. Similarly, if the same data were duplicated in multiple articles, results from the most comprehensive, or most recent article only were included.

The criteria for including studies relevant to determining the *safety* of B-type natriuretic peptide assays can be found in Box 7.

Box 7 Study selection criteria for assessing safety

Selection criteria	Inclusion criteria
Population	Symptomatic patients with suspected heart failure in the non-hospital setting
Intervention	BNP or NT-proBNP assays in conjunction with standard clinical assessment[a] ± echocardiography
Comparator	Standard clinical assessment[a] ± echocardiography
Outcomes	Adverse events—physical, psychological due to testing (anxiety due to a true positive or a false positive diagnosis), delay in management associated with a false negative diagnosis
Study design	Randomised or non-randomised controlled trials, cohort studies, registers, case series, case reports or systematic reviews of these study designs
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 04/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base.

[a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs

Diagnostic assessment framework

This assessment of the diagnostic use of the BNP and NT-proBNP assays follows the methodology outlined in the MSAC *Guidelines for the assessment of diagnostic technologies* handbook (MSAC 2005).

In order to assess the effectiveness of the BNP and NT-proBNP tests in the diagnosis of HF in the non-hospital setting, there needed to be a consideration of their diagnostic accuracy (in comparison to a reference standard), their impact on the clinical management of the patient with suspected HF, and their ultimate impact on patient health outcomes. The first goal of this assessment was to find *direct evidence* of the effectiveness of the BNP and NT-proBNP tests on health outcomes. No direct evidence was available, so a *linked evidence* approach was undertaken. This is an approach where studies which assess, individually, the diagnostic accuracy of the tests, the impact on patient management and the impact on health outcomes are linked together through a narrative.

The criteria for including studies on *diagnostic effectiveness* are presented in Part B, Appendix L (for the direct evidence approach), and Box 8 and Box 9 (for the linked evidence approach).

Box 8 Study selection criteria for assessing diagnostic accuracy in the non-hospital setting (linked evidence approach)

Selection criteria	Inclusion criteria
Population	Symptomatic patients with suspected heart failure in the non-hospital setting
Intervention	NT-proBNP or BNP diagnostic assays
Comparator	Clinical diagnosis or echocardiography
Reference standard	Clinical diagnosis using all data, including echocardiogram
Outcomes	Sensitivity and specificity (and therefore rates of false positives and negatives), likelihood ratios and diagnostic odds ratios, negative predictive values, diagnostic yield
Study design	Cross-sectional studies where patients are cross-classified on the test and reference standard. Case-control diagnostic studies were only acceptable if cross-sectional studies were not available, or were limited. Systematic reviews of these study designs were also acceptable
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

Box 9 Study selection criteria for assessing effectiveness of diagnosis in the non-hospital setting (linked evidence approach)

	Change in management	Change in health outcomes
Selection criteria	**Inclusion criteria**	**Inclusion criteria**
Population	Symptomatic patients with suspected HF in the non-hospital setting	Patients with HF or alternative diagnosis[c] in the non-hospital setting
Intervention	NT-proBNP or BNP diagnostic assays as 'first line' tests in conjunction with standard clinical assessment[a] ± echocardiography[b]	Treatment for HF (eg ACE inhibitors, beta-blockers, surgery) or for alternative diagnosis
Comparator(s)	Standard clinical assessment ± echocardiography	No (or delayed) treatment for HF or alternative diagnosis
Outcomes	Primary: treatment rates, method of treatment, time to diagnosis, rate of referral to specialist	Primary: rate of survival/ death, symptom resolution (dyspnoea, oedema), quality of life, functional status
	Secondary: rates of echocardiogram/ supportive diagnostic testing	Secondary: confirmation of HF by left ventricular ejection fraction (LVEF) <50%[d], hospital length of stay, rate of readmission
Study design	Randomised or non-randomised controlled trials or cohort studies, uncontrolled before-and-after case series (with 20 or more participants) or systematic reviews of these study designs.	
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005	
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base	

HF = heart failure; [a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs; [b] Echocardiogram is likely to be used in the diagnostic pathway if, on the basis of the 'first line' tests (eg BNP, NT-proBNP, physical examination, chest X-ray, laboratory tests, ECG), the patient is still suspected of HF. Those patients 'ruled out' for HF on the basis of these 'first line' tests, however, will not receive an echocardiogram; [c] Given the multitude of alternative diagnoses for patients presenting with HF-like symptoms, it was not possible to assess treatment effectiveness systematically in this patient group; [d] It is acknowledged that LVEF is only effective at detecting systolic dysfunction, and may not detect diastolic dysfunction.

Search results

The process of study selection for this report went through six phases:

1. All reference citations from all literature sources were collated into an Endnote 8.0 database;

2. Duplicate references were removed;

3. Studies were excluded, on the basis of the complete citation information, if it was obvious that they did not meet the inclusion criteria. All other studies were retrieved for full-text assessment;

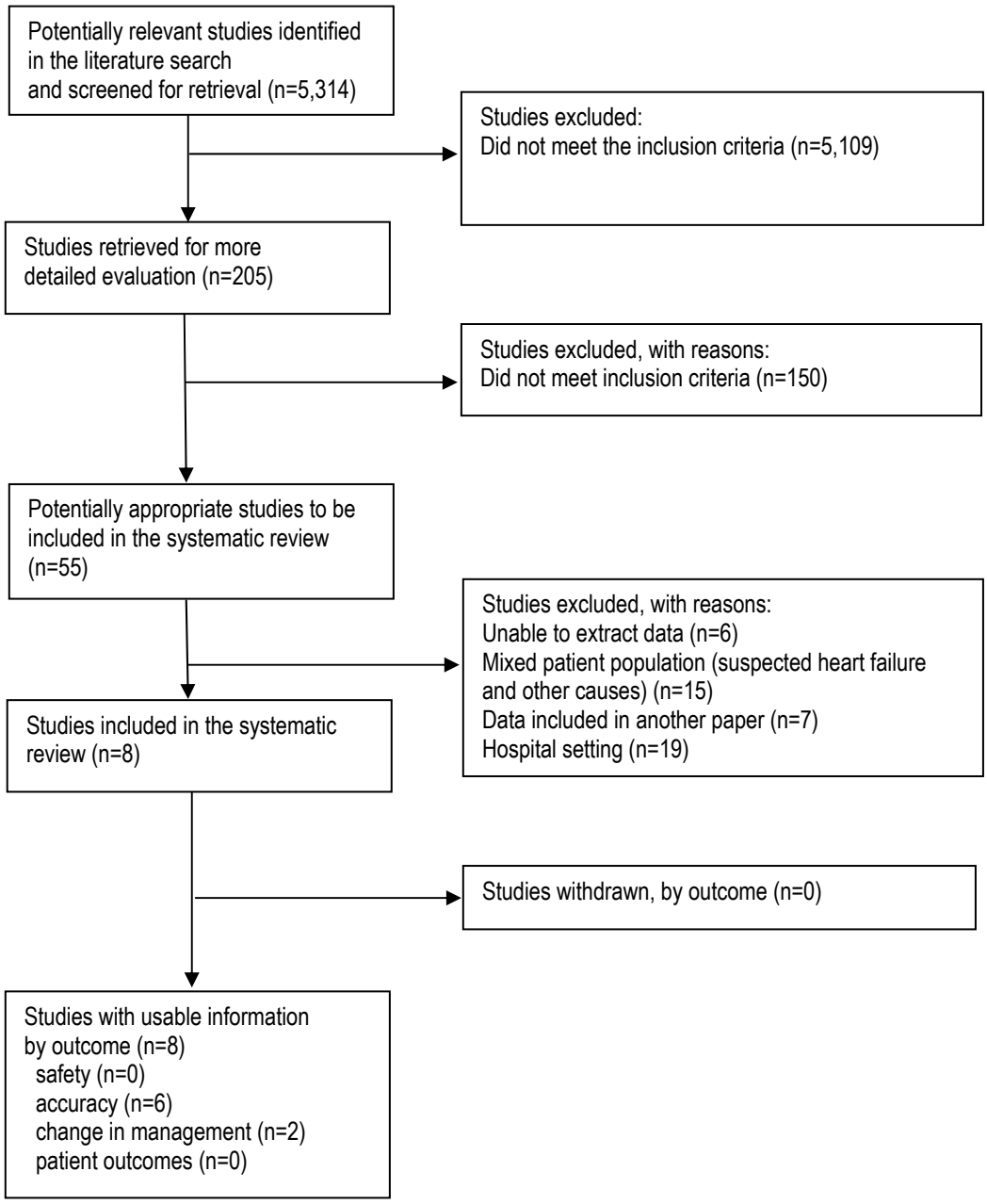
4. Inclusion criteria were independently applied to the full-text articles by two or more researchers. Those articles meeting the criteria formed part of the evidence base. The remainder provided background information;

5. The reference lists of the included articles were pearled for additional relevant studies. These were retrieved and assessed according to phase 4; and

6. The evidence base consisted of articles from phases 4 and 5 that met the inclusion criteria.

Any doubt concerning inclusions at Phase 4 was resolved by group consensus. The results of the process of study selection—to collate the evidence base for assessing diagnostic effectiveness in the non-hospital setting—are provided in Figure 12.

Figure 12 Summary of the process used to identify and select studies for the assessment of diagnostic effectiveness in the non-hospital setting



Adapted from Moher et al (1999)

Data extraction and analysis

A profile of key characteristics was developed for each included diagnostic study (Part B, Appendix M). Studies that were unable to be retrieved or that met the inclusion criteria but contained insufficient or inadequate data for inclusion are provided in Part B, Appendix N. Definitions of all technical terms and abbreviations are provided in the Glossary.

Diagnostic accuracy

The appropriate population for diagnostic accuracy studies (in linked evidence) included in this assessment consisted of symptomatic patients with suspected HF in a non-hospital setting. Studies were excluded that recruited patients based on referral for echocardiography without indicating whether the referral was for clinically suspected HF. In studies reporting diagnostic accuracy for HF as well as diagnostic accuracy for left ventricular dysfunction, only data referring to the former were extracted. However, data were extracted on diagnostic accuracy for left ventricular dysfunction when that was all that was presented.

Data from the *linked* evidence on the diagnostic accuracy of the B-type natriuretic peptide tests was extracted using the classic 2 x 2 table, whereby the results of the diagnostic test were cross-classified against the results of the reference standard (Armitage et al 2002; Deeks 2001), and Bayes' Theorem was applied:

		Cardiac status (based on reference standard—clinical diagnosis using all clinical data, including echocardiogram)		
		Heart failure (HF)	*Normal*	
Index test *(BNP or NT-proBNP)*	*Test +*	True positive	False positive	Total positive
	Test -	False negative	True negative	Total negative
		Total with HF	Total without HF	

The sensitivity of the index test (BNP or NT-proBNP) was calculated as the proportion of people with HF who have positive diagnostic test results:

Sensitivity (true positive rate) = Number of true positives / total with HF * 100

The specificity of the index test (BNP or NT-proBNP) was calculated as the proportion of people without HF who have normal diagnostic test results:

Specificity (true negative rate) = Number of true negatives / total without HF * 100

When a 95% confidence interval was not provided in the relevant study, it was calculated using exact binomial methods.

Due to the very small number (2) of diagnostic accuracy studies with raw data available in the non-hospital setting, summary measures of test accuracy were not calculated and meta-analysis was not undertaken.

For *linked* evidence intervention studies, that is those studies assessing change in management/treatment through use of the test or change in health outcomes from that treatment, descriptive statistics (eg means, standard deviations) were extracted or

calculated from the individual studies for all the safety and effectiveness pre-specified outcomes. A statistically significant difference in outcomes was assumed at $p<0.05$.

All statistical calculations and testing were undertaken using the statistical computer package Stata version 8.2 (Stata Corporation 2004). All data regarding B-type natriuretic peptide assay levels were presented as pg/mL. Data presented as pmol/L in the original studies were converted to pg/mL by multiplying by 8.457 and 3.456 for NT-proBNP and BNP, respectively (molecular weights of NT-proBNP and BNP are 8,457 and 3,456 respectively) (Januzzi & Maisel 2004). Further, to convert these data to Standard International Units, as used in Australia, the pg/mL should be converted to their original pmol/L values (by dividing by 8.457 and 3.456 for NT-proBNP and BNP, respectively) and then converting the pmol/L to mmol/L.[16]

Appraisal of the evidence

The evidence presented in the selected studies was assessed and classified using the dimensions of evidence defined by the National Health and Medical Research Council (NHMRC 2000).

These dimensions (Table 45) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from the literature identified as informing a particular intervention. The last two require expert clinical input as part of their determination.

Table 45 Evidence dimensions

Type of evidence	Definition
Strength of the evidence	
Level	The study design used, as an indicator of the degree to which bias has been eliminated by design.[a]
Quality	The methods used by investigators to minimise bias within a study design.
Statistical precision	The *p*-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect.
Size of effect	The distance of the study estimate from the 'null' value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

[a]See Table 5

Strength of the evidence in individual studies

The three subdomains (level, quality and statistical precision) are collectively a measure of the strength of the evidence.

[16] 1 mmol is 10^{-3} of a mole and 1 pmol is 10^{-12} of a mole; therefore. the conversion factor from pmol/L to mmol/L can be achieved by dividing by 10^{8}

Level

The 'level of evidence' reflects the effectiveness of a study design to answer a particular research question. Effectiveness is based on the probability that the design of the study has reduced or eliminated the impact of bias on the results.

The new version of the NHMRC evidence hierarchy provides a ranking of various study designs ('levels of evidence') by the type of research question being addressed (NHMRC 2005). Table 46 is an abbreviated version of this evidence hierarchy and includes the research questions relevant to a *linked* assessment of diagnostic effectiveness. The *Diagnosis* column in this evidence hierarchy was used for ranking the diagnostic accuracy studies. The *Intervention* column was used for ranking the studies on change in management due to the test and change in health outcomes due to treatment.

Quality

Study quality was presented in this Assessment Report both in terms of the components of quality (eg selection bias, misclassification bias, reviewer bias) and as an overall quality score.

The appraisal of studies pertaining to the diagnostic accuracy of the NT-proBNP and BNP assays was conducted using the QUADAS tool, a checklist developed by the Centre for Reviews and Dissemination, York, United Kingdom (Whiting et al 2003). Studies assessing change in management and change in patient health outcomes for this *linked evidence* approach were critically appraised using the Downs and Black instrument (Downs & Black 1998).

Table 46 Designation of intervention and diagnostic levels of evidence

Level	Intervention §	Diagnosis **
I *	A systematic review of level II studies	A systematic review of level II studies
II	A randomised controlled trial	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, §§ among consecutive patients with a defined clinical presentation ††
III-1	A pseudorandomised controlled trial (i.e. alternate allocation or some other method)	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, §§ among non-consecutive patients with a defined clinical presentation††
III-2	A comparative study with concurrent controls: Non-randomised, experimental trial † Cohort study Case-control study Interrupted time series with a control group	A comparison with reference standard that does not meet the criteria required for levels II and III-1 evidence
III-3	A comparative study without concurrent controls: Historical control study Two or more single arm study ‡ Interrupted time series without a parallel control group	Diagnostic case-control study ††
IV	Case series with either post-test or pre-test/post-test outcomes	Study of diagnostic yield (no reference standard) ‡‡

* A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence; § Definitions of these study designs are provided on pages 7–8 in *How to use the evidence: assessment and application of scientific evidence* (NHMRC 2000); † This also includes controlled before-and-after (pre-test/post-test) studies, as well as indirect comparisons (ie using A vs B and B vs C, to determine A vs C); ‡ Comparing single arm studies ie. case series from two studies; ** The dimensions of evidence apply only to studies of diagnostic accuracy. To assess the effectiveness of a diagnostic test there also needs to be a consideration of the impact of the test on patient management and health outcomes. See MSAC (2004) *Guidelines for the assessment of diagnostic technologies.* Available at: <www.msac.gov.au>; §§ The validity of the reference standard should be determined in the context of the disease under review. Criteria for determining the validity of the reference standard should be pre-specified. This can include the choice of reference standard(s) and its/their timing in relation to the index test. The validity of the reference standard can be determined through quality appraisal of the study. See Whiting P et al 2003; †† Well-designed population-based case-control studies (eg population-based screening studies where test accuracy is assessed on all cases, with a random sample of controls) do capture a population with a representative spectrum of disease and thus fulfil the requirements for a valid assembly of patients. However, in some cases the population assembled is not representative of the use of the test in practice. In diagnostic case-control studies a selected sample of patients already known to have the disease are compared with a separate group of normal/healthy people known to be free of the disease. In this situation patients with borderline or mild expressions of the disease, and conditions mimicking the disease, are excluded, which can lead to exaggeration of both sensitivity and specificity. This is called spectrum bias because the spectrum of study participants will not be representative of patients seen in practice; ‡‡ Studies of diagnostic yield provide the yield of diagnosed patients, as determined by an index test, without confirmation of the accuracy of this diagnosis by a reference standard. These may be the only alternatives when there is no reliable reference standard.

Note 1: Assessment of comparative harms/safety should occur according to the hierarchy presented for each of the research questions, with the proviso that this assessment occurs within the context of the topic being assessed. Some harms are rare and cannot feasibly be captured within randomised controlled trials; physical harms and psychological harms may need to be addressed by different study designs; harms from diagnostic testing include the likelihood of false positive and false negative results; harms from screening include the likelihood of false alarm and false reassurance results.

Note 2: When a level of evidence is attributed in the text of a document, it should also be framed according to its corresponding research question, eg level II intervention evidence, level IV diagnostic evidence, level III-2 prognostic evidence etc.

Source: NHMRC (2005)

Summary appraisal of strength of the diagnostic evidence

Individual studies assessing diagnostic effectiveness were graded according to the pre-specified quality and applicability criteria (MSAC 2005) as shown in Table 47.

Table 47 Grading system used to rank included diagnostic studies

Validity criteria	Description	Grading system
Appropriate comparison	Did the study evaluate a direct comparison of the index test strategy versus the comparator test strategy?	C1 direct comparison CX other comparison
Applicable population	Did the study evaluate the index test in a population that is representative of the subject characteristics (age and sex) and clinical setting (disease prevalence, disease severity, referral filter and sequence of tests) for the clinical indication of interest?	P1 applicable P2 limited P3 different population
Quality of study	Was the study designed to avoid bias?	Study design: NHMRC level of evidence
	High quality = no potential for bias based on pre-defined key quality criteria	Study quality: Q1 high quality
	Medium quality = some potential for bias in areas other than those pre-specified as key criteria	Q2 medium
	Poor quality = poor reference standard and/or potential for bias based on key pre-specified criteria	Q3 poor reference standard poor quality or insufficient information

Statistical precision

Statistical precision was determined using standard statistical principles. Small confidence intervals and p-values give an indication as to the probability that the reported effect is real (NHMRC 2000).

Size of effect in individual studies

It is important to establish whether statistically significant differences are also clinically important in terms of health outcomes. Where appropriate, the size of the effect needs to be determined, as well as whether the 95% confidence interval includes only clinically important effects (NHMRC 2000). This evaluation of clinical importance is necessarily limited by the evidence available to document the impact on health outcomes by the diagnostic tests.

Relevance of evidence in individual studies

The outcome being measured in the studies should be appropriate and clinically relevant. Inadequately validated (predictive) surrogate measures of a clinically relevant health outcome should be avoided (NHMRC 2000). Once again, this evaluation of outcomes is limited by the evidence available to assess the impact of the B-type natriuretic peptide tests on health outcomes.

The body of evidence

Appraisal of the body of evidence was conducted along the lines suggested by the NHMRC in their guidance on clinical practice guideline development (NHMRC 2005). Five components are considered essential by the NHMRC when judging the body of evidence:

- the volume of evidence—which includes the number of studies sorted by their methodological quality and relevance to patients;
- the consistency of the study results—whether the better quality studies had results of a similar magnitude and in the same direction, that is homogenous or heterogenous findings;
- the potential clinical impact—appraisal of the precision, size and clinical importance or relevance of the primary health outcomes used to determine the safety and effectiveness of the test;
- the generalisability of the evidence to the target population; and
- the applicability of the evidence—integration of this evidence for conclusions about the net clinical benefit of the index test in the context of Australian clinical practice.

A matrix for assessing the body of evidence for each research question, according to the components above, was adapted for this assessment (see Table 48) (NHMRC 2005).

Table 48 Body of evidence assessment matrix

Component	A Excellent	B Good	C Satisfactory	D Poor
Volume of evidence	Several level I or II studies with low risk of bias	One or two level II studies with low risk of bias or a SR/multiple level III study with low risk of bias	Level III studies with low risk of bias, or level I or II studies with moderate risk of bias	Level IV studies, or level I to III studies with high risk of bias
Consistency	All studies consistent	Most studies consistent and inconsistency may be explained	Some inconsistency reflecting genuine uncertainty around clinical question	Evidence is inconsistent
Clinical impact	Very large	Substantial	Moderate	Slight or restricted
Generalisability	Population(s) studied in body of evidence is/are the same as the target population	Population(s) studied in the body of evidence is/are similar to the target population	Population(s) studied in body of evidence is/are different to target population but it is clinically sensible to apply this evidence to target population	Population(s) studied in body of evidence is/are different to target population and it is hard to judge whether it is sensible to generalise to target population
Applicability	Directly applicable to Australian healthcare context	Applicable to Australian healthcare context with few caveats	Probably applicable to Australian healthcare context with some caveats	Not applicable to Australian healthcare context

Results of assessment

Are B-type natriuretic peptide assays safe in the diagnosis of heart failure in the non-hospital setting?

Summary – Safety of B-type natriuretic peptide assays

None of the studies that met the inclusion criteria for this assessment reported physical harms ensuing from the B-type natriuretic peptide testing procedure. Similarly, none of the diagnostic studies available for this assessment of B-type natriuretic peptide testing investigated the impact of the diagnosis on the patients' psychological wellbeing.

B-type natriuretic peptide testing involves a simple blood test. Blood is extracted using a standard venepuncture technique and is collected in tubes containing ethylenediaminetetraacetic acid. Common after-care is to apply manual pressure and/or a dressing to the wound to assist with haemostasis. B-type natriuretic peptide testing is therefore a minimally invasive procedure and in some cases would be one of several blood tests that the patient might undergo during the diagnostic process. Like all blood tests, harms can occur if the venepuncture procedure is done incorrectly by the health practitioner. Similarly, patients with blood clotting disorders or receiving blood thinners require careful observation to ensure bleeding from the wound is controlled.

None of the studies included in Part B of this Assessment Report mentioned physical harms occurring as a result of B-type natriuretic peptide testing in the non-hospital setting.

Psychological harms are a theoretical risk for patients undergoing this test. False positive test results could mean that the patient undergoes the stress of receiving an initial diagnosis of heart failure (HF), along with a battery of generally more invasive diagnostic tests and, in some cases, treatment or medications that prove, eventually, to be completely unnecessary. False negative test results provide false reassurance to the patient that he/she is well, potentially resulting in poor health outcomes due to inappropriately delayed treatment.

The impact of B-type natriuretic peptide testing on patients' psychological wellbeing was not evaluated in any of the diagnostic studies that met the criteria for inclusion in Part B of this Assessment Report.

Are B-type natriuretic peptide assays effective in the diagnosis of heart failure in the non-hospital setting?

BNP assays (direct evidence of effectiveness)

There was no direct evidence available concerning the effectiveness or impact of BNP assays on patient health outcomes in the non-hospital setting.

BNP assays (linked evidence of effectiveness)

Summary – Linked evidence of diagnostic effectiveness of BNP assays

The small number of diagnostic accuracy studies in the non-hospital setting were relatively consistent in their findings that BNP tests are sensitive, with a high negative predictive value, meaning a negative test result effectively 'rules out' HF in a patient. The specificity of the test is, however, variable.

There were no studies available that assessed the impact of the BNP test on the management of patients by general practitioners. A systematic review of treatment effectiveness for patients with and without HF was beyond the scope of Part B of this Assessment Report.

Five studies met the inclusion criteria for providing linked evidence of the effectiveness of BNP testing in the non-hospital setting (see Table 49). The rates of HF in the non-hospital setting were low, in the range 20%–34%. In general, the patient populations were adult, with symptoms of suspected HF, who were referred by their general practitioners (GPs) to a clinic or imaging service for further assessment.

Are BNP assays accurate?

All five studies—including two that reported on suspected HF specifically caused by left ventricular systolic dysfunction—measured BNP levels using a radioimmunoassay (Cowie et al 1997; Hobbs et al 2002, 2004; Landray et al 2000; Sim et al 2003). Sensitivity of the tests was in the range 80%–100% and specificity in the range 18%–88%. Negative predictive values at an optimised cut-off point were in the range 98%–100% (see Table 50).

Do BNP assays change patient management?

There were no studies that met the inclusion criteria that reported on the impact of BNP test results on the decision-making or patient management practices of GPs.

Does treatment on the basis of a BNP assay change health outcomes?

Linked evidence was not systematically assessed to address the impact of treatment on patient health outcomes.

A section in the Introduction discusses, in a non-systematic manner, the high level evidence supporting the well-known beneficial effects of the early treatment of HF. The BNP test is, however, unlikely to result in earlier diagnosis and treatment of HF as it has a high false positive rate in the non-hospital setting (when prevalence of HF is in the range 20%–34%) and an echocardiogram is required to confirm the diagnosis. It is, however, possible that use of a BNP test would result in earlier identification of *alternative* conditions in those patients 'ruled out' from HF, resulting in appropriate treatment and benefiting the health of these patients. The benefit is likely to be higher as the severity of the alternative condition increases.

Table 49 Summary of included BNP diagnostic accuracy studies in the non-hospital setting—characteristics and quality appraisal

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests[a]	Outcomes assessed	Study quality[b]	Applicability[b]
			N	Selection criteria				
(Cowie et al 1997)	Prospective cohort – cross-classified	Rapid access clinic (primary care) Hillingdon district, west London, UK	106 [ITT: 122]	Suspected cases of new HF referred by GPs	Medical hx Clinical exam ECG CXR	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Hobbs et al 2004)	Prospective cohort – cross-classified	Primary care Four practices, England, UK	103	Patients with suspected HF (but unvalidated) randomly sampled from GP practices	Medical hx Clinical exam ECG Spirometry	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P1
(Landray et al 2000)	Prospective cohort – cross-classified	Primary care Clinic, Banbury, Oxford, UK	126	Patients referred by GPs with suspected HF	ECG CXR	Sensitivity Specificity	Level III-2 diagnostic evidence Q3	P1
(Sim et al 2003)	Prospective cohort – cross-classified	Open-access echo-cardiography service Newport, South Wales, UK	83	Patients with symptoms of dyspnoea referred by GPs	Not stated	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P1
(Zaphiriou et al 2005)	Prospective cohort – cross-classified	Rapid access HF clinic Five hospitals, UK	306	Consecutive patients with new symptoms suggestive of HF, referred by their GPs	Medical hx Clinical exam ECG Blood tests CXR	Sensitivity Specificity Negative predictive value	Level II diagnostic evidence Q1	P1

HF = heart failure; hx = history; ECG = electrocardiogram; CXR = chest X-ray; GP = general practitioner; ITT = intention-to-treat; [a] Only tests that were mentioned or inferred from the study are included – there may be other prior tests that were not specifically reported; [b] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

Table 50 Summary of included BNP diagnostic accuracy studies in the non-hospital setting—results and precision estimates

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
(Cowie et al 1997)	Level III-1 diagnostic evidence Q2	106 [ITT: 122]	Age: 24–87 yrs M/F: 59/63 Excluded if HF hx	HF: 29/106 (27%)	Consensus clinical diagnosis – all data (cardiologists)	BNP Radio-immunoassay Peninsula Laboratories		77	97 [82,100]	84 [74,92]
(Hobbs et al 2004)	Level III-2 diagnostic evidence Q3	103	Age: n/a for subgroup M/F: n/a for subgroup Unclear if HF hx	HF: n/a HF (caused by LVSD): 21/103 = 20%	Consensus clinical diagnosis – all data (study investigators – and cardiologists in equivocal cases)	BNP Immuno-radiometric assay Shinogi, Japan		279	80 [28,100]	88 [84,92]
(Landray et al 2000)	Level III-2 diagnostic evidence Q3	126	Age: 74±9 yrs M/F: 68/58 Unclear if HF hx	HF (caused by LVSD): (32%)		BNP Immuno-radiometric assay Shionoria assay, Shionogi	Echocardiography	10 17.9 76	92 [85,96] 88 [80,94] 66 [56,75]	18 [11,27] 34 [25,44] 87 [79,93]
(Sim et al 2003)	Level III-2 diagnostic evidence Q3	83	Age: 72 (37–87) yrs M/F: 40/43 Unclear if HF hx	HF (caused by LVSD): 26/83 (31%)		BNP Radio-immunoassay (in-house) Bachem Ltd	Echocardiography – two independent echo-cardiographers	19 20	100 [87,100] 96 [80,100]	49 [36,63] 58 [44,71]
(Zaphiriou et al 2005)	Level II diagnostic evidence Q1	306	Age: median 74 (52–87) yrs M/F: 130/176 Patients with HF hx excluded	HF: 104/306 (34%)	Clinical diagnosis – according to pre-determined criteria (cardiologist)	BNP Immuno-fluorescence assay Biosite Diagnostics		30 65 100	95 [89,98] 87 [79,93] 79 [70,87]	35 [26,45] 57 [47,67] 72 [62,81]

HF = heart failure; hx = history; n/a = not available; ITT = intention-to-treat; LVSD = left ventricular systolic dysfunction; [a] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

NT-proBNP assays (direct evidence of effectiveness)

There was no direct evidence available concerning the effectiveness or impact of NT-proBNP assays on patient health outcomes.

NT-proBNP assays (linked evidence)

Summary – Linked evidence of diagnostic effectiveness of NT-proBNP

The three diagnostic accuracy studies in the non-hospital setting were relatively consistent in their findings that NT-proBNP tests are sensitive. In general, the assays have a high negative predictive value. The clinical impact of the test was most obvious in terms of patient management. Good quality level II intervention evidence reported a clinically important improvement in the proportion of correct diagnoses for those general practitioners receiving NT-proBNP test results. The impact of the NT-proBNP test on patient health outcomes was not tested directly and a systematic review of treatment effectiveness for patients with and without HF was beyond the scope of Part B of this Assessment Report.

Five studies met the inclusion criteria for providing linked evidence of the effectiveness of NT-proBNP testing in the non-hospital setting (see Table 51). Three of the studies provided evidence of the diagnostic accuracy of the Elecsys chemiluminescent sandwich immunoassay developed by Roche Diagnostics. Two further studies reported on the impact of NT-proBNP testing on patient management by the GP.

The patient populations in the diagnostic accuracy studies were primarily adults presenting with suspected HF. In the change-in-management studies the patients were adults who presented with dyspnoea and were suspected of HF. The prevalence of HF in the diagnostic accuracy studies was in the range 9%–34%. Two of the three studies reported on HF caused by left ventricular systolic dysfunction, with prevalence rates of 9%–20%.

Are NT-proBNP assays accurate?

Three studies assessed the accuracy of the NT-proBNP tests at correctly identifying or ruling out HF in patients suspected of having the condition. Sensitivity rates were uniformly high but specificity rates varied considerably despite the fact that the same assay was being tested in the studies (see Table 52).

Sensitivity of the NT-proBNP tests was in the range 96%–100%, specificity in the range 18%–56%, and negative predictive values for an optimal cut-off point in the range 97%–100%. Therefore, like BNP testing, it would appear that in general NT-proBNP assays have high sensitivity at detecting HF but lower specificity. The variability in rates is quite wide, although the negative predictive values are uniformly high (>90%).

Do NT-proBNP assays change patient management?

In a Danish study of 345 patients referred to a hospital-based clinic by their GPs for dyspnoea, echocardiographic confirmation of HF occurred for 81 patients. Nielsen et al

(2004) determined that for 68 of these 81 patients, either no or inadequate treatment was being administered at the time of examination—thereby suggesting that earlier diagnosis (possibly through NT-proBNP testing) might enable earlier and more effective treatment of HF (level IV intervention evidence). Nielsen et al (2004) also determined that 51 per cent of 287 patients could have safely forgone echocardiography on the basis of the 'rule out' results from the NT-proBNP test.

Wright et al (2003) conducted a single-blind randomised controlled trial (level II intervention evidence) in Auckland, New Zealand, to assess the impact of NT-proBNP testing on the diagnoses by GPs evaluating patients with recent symptoms of dyspnoea and/or oedema. This good quality trial compared the GP diagnoses prior to and after the receipt of the NT-proBNP results versus customary clinical assessment at the same time points in the control group. Each patient was assessed separately by a cardiologist, and the results of that testing formed the basis for a reference standard diagnosis of HF or non-HF, according to pre-defined criteria, by a panel of cardiologists and a GP who reviewed all clinical data. The primary outcome of this trial was an increase in GP diagnostic accuracy. NT-proBNP testing resulted in 32/152 (21%) accurate corrections to the initial diagnosis by the GP, whereas in the control group there were 12/153 (8%) accurate corrections to the initial diagnosis. Therefore, the overall improvement in correct diagnosis for the NT-proBNP trial arm compared to the control trial arm was 13 per cent [95%CI 5.5, 21.0; p=0.002]. This is a clinically important outcome, although the confidence interval does include clinically unimportant effects. The main impact of the correction of diagnosis in the NT-proBNP trial arm was to enable GPs to correctly 'rule out' HF. This reinforces the notion that B-type natriuretic peptide testing is an effective 'SnOut' test—high sensitivity, so that a negative test rules *out* HF. Patients with a negative test are effectively screened out from the need to undergo further testing for HF and can instead be tested for alternative pathologies.

Does treatment on the basis of a NT-proBNP assay change health outcomes?

The review of HF treatment provided in the Introduction points to very high level evidence of treatment effectiveness for some therapies. Therefore, a systematic assessment of all the various HF treatments was considered unnecessary and unproductive in the context of determining the impact of early treatment on patients with HF. The NT-proBNP test is, however, unlikely to result in *earlier* diagnosis and treatment of HF, as it has a high false positive rate in the non-hospital setting (prevalence of HF between 9% and 34%) and an echocardiogram is required to confirm the diagnosis. It is, however, probable that patients 'ruled out' from HF earlier, as a consequence of a NT-proBNP test, would receive earlier and more accurate treatment than would otherwise be received when conventional diagnostic strategies are used. The impact of earlier and more accurate treatment on the patient would depend on the nature and severity of the large number of alternative pathologies that present with HF-like symptoms, eg asthma, chronic obstructive pulmonary disease and pneumonia.

Table 51 Summary of included NT-proBNP diagnostic accuracy studies in the non-hospital setting—characteristics and quality appraisal

Study Author(s) (Year)	Study design	Setting Region, site	Study population		Prior tests	Outcomes assessed	Study quality[a]	Applicability[a]
			N	Selection criteria				
(Gustafsson et al 2003)	Prospective cohort – cross-classified	Primary care Copenhagen General Practitioners' Laboratory, Denmark	367	Patients referred by GPs for echocardiography to confirm / rule out suspected HF	Medical hx Medication use ECG	Sensitivity Specificity Negative predictive value	Level III-1 diagnostic evidence Q2	P1
(Hobbs et al 2002)	Prospective cohort – cross-classified	Primary care Four practices, England, UK	103	Patients with suspected HF (unvalidated) randomly sampled from GP practices	Medical hx Clinical exam ECG Spirometry	Sensitivity Specificity Negative predictive value	Level III-2 diagnostic evidence Q3	P1
(Zaphiriou et al 2005)	Prospective cohort – cross-classified	Rapid access HF clinic Five hospitals, UK	306	Consecutive patients with new symptoms suggestive of HF, referred by their GPs	Medical hx Clinical exam ECG Blood tests CXR	Sensitivity Specificity Negative predictive value	Level II diagnostic evidence Q1	P1

HF = heart failure; hx = history; CXR = chest X-ray; ECG = electrocardiogram; GP = general practitioner; [a] The assessment of study quality and applicability followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

Table 52 Summary of included NT-proBNP diagnostic accuracy studies in the non-hospital setting—results and precision estimates

Study Author(s) (Year)	Study quality[a]	Study population			Reference standard	Index test specifications	Comparator specifications	Results		
		N	Characteristics	Disease prevalence				Cut-off point (pg/mL)	Sensitivity [95%CI]	Specificity [95%CI]
(Gustafsson et al 2003)	Level III-1 diagnostic evidence Q2	367	Age: 69 (39–84) yrs M/F: 169/198 Unclear if HF hx	HF (caused by LVSD): 33/367 (9%)		NT-proBNP Chemiluminescent sandwich immuno-assay Elecsys, Roche Diagnostics	Echocardiography LVEF ≤0.40 LVEF ≤0.30 *(LVEF level ruling in/out LVSD causing HF)*	125	97 [91,99] 100 [96,100]	46 [36,56] 56 [46,66]
(Hobbs et al 2002) (Hobbs et al 2004)	Level III-2 diagnostic evidence Q3	103	Age: n/a for subgroup M/F: n/a for subgroup Unclear if HF hx	HF: n/a HF (caused by LVSD): 21/103 = 20%	Consensus clinical diagnosis – all data (study investigators – and cardiologists in equivocal cases)	NT-proBNP Enzyme linked immunosorbent assay (ELISA) Roche Diagnostics		304	100 [92,100]	18 [10,29]
(Zaphiriou et al 2005)	Level II diagnostic evidence Q1	306	Age: median 74 (52–87) yrs M/F: 130/176 Patients with HF hx excluded	HF: 104/306 (34%)	Clinical diagnosis – according to pre-determined criteria (cardiologist)	NT-proBNP ELISA assay Elecsys system, Roche Diagnostics		125 166	98 [93,100] 96 [90,99]	35 [26,45] 43 [33,53]

HF = heart failure; n/a = not available; hx = history; LVEF = left ventricular ejection fraction; LVSD = left ventricular systolic dysfunction; [a] The assessment of study quality followed the approach outlined in the 'Approach to assessment' chapter, specifically the section on 'Strength of the evidence in individual studies'.

Discussion

In diagnostic studies of either BNP or NT-ProBNP assays, rates of heart failure (HF) were in the range 9%–34% (median 29%) in non-hospital settings. This is as expected, and probably applicable to that portion of the Australian population presenting to the non-hospital setting with symptoms suggestive of HF.

The populations studied in the included diagnostic studies are reasonably applicable to the target population in Australia, that is patients presenting to general practice with symptoms (eg dyspnoea) suggestive of HF. As a group, however, they may have had slightly more severe symptoms than is usual in general practice, as most were selected on the basis of *referral* from a general practitioner (GP) on suspicion of HF.

Safety of NT-proBNP and BNP assays

Studies included in Part B of this Assessment Report did not mention physical or psychological harms occurring as a consequence of B-type natriuretic peptide testing. The likelihood of physical harms is low and similar to that of any other blood test. However, it is possible that false positive results could engender inappropriate treatment leading to adverse events for the patient. False negative results could also lead to delayed treatment but the impact is likely to be minimal in this patient population as the false negative rate is very small for B-type natriuretic peptide assays, and these patients are not acutely ill. Psychological harms (eg anxiety) can be associated with the delivery of a diagnosis (whether correct or incorrect); however, this is the case with all diagnostic tests.

Effectiveness of BNP assays in the diagnosis of heart failure in the non-hospital setting

Diagnostic accuracy

The value of BNP tests is as a highly sensitive test such that a negative result 'rules out' the diagnosis (Sackett et al 1991). This is based on the notion that sensitivity and negative predictive value share the 2 x 2 diagnostic table cell for false negatives (FN) in their denominator. As sensitivity increases toward 100 per cent, FN decreases toward zero. As FN decreases toward zero, negative predictive value increases toward 100 per cent. The negative predictive value at optimal cut-off points for the BNP test was predominantly greater than 90 per cent in the studies included in this Assessment Report. BNP testing therefore appears to act as a 'first line' diagnostic tool to identify patients that should or should not be referred to echocardiography or other diagnostic tests to confirm a clinical diagnosis of HF. Its role, therefore, is not to act as a 'reference standard' test for HF as its specificity is generally not high.

Impact on patient management

There were no studies that reported the impact of the BNP test on patient management in a non-hospital setting. It is likely, however, that this impact would reflect the results

seen with the use of the NT-proBNP test in New Zealand general practice (Wright et al 2003). That is, the main impact on patient management is to assist GPs to correctly 'rule out' HF in patients presenting with dyspnoea and/or oedema of recent onset.

Impact on health outcomes

The direct impact of BNP testing in the non-hospital setting on the health outcomes of patients was not evaluated in any of the studies available for this review.

To definitively determine the direct effect of BNP testing on patient health outcomes, a trial (probably across several clinical practices) would need to be conducted with sufficient sample size to detect a statistically significant difference in health outcomes (that was also clinically important) for patients suspected of HF who receive BNP-assisted diagnostic assessment compared to conventional diagnostic assessment in the non-hospital setting.

In the absence of such data, linked high-level evidence of treatment effectiveness (see Introduction) suggests that early treatment for HF is beneficial for the patient. However, a BNP test is unlikely to result in an *earlier* definitive diagnosis of HF, as it has a high false positive rate in the non-hospital setting and an echocardiogram is still required to confirm the diagnosis. Assuming a HF prevalence of 30 per cent in Australian general practice, treating on the basis of this test alone would result in about one-quarter of the patients being inappropriately treated for HF (see Economic Considerations section). Therefore, the BNP test may not benefit the health of some patients still suspected of HF after the test. It is, however, possible that BNP testing could benefit the health of those patients 'ruled out' from HF, as they would have an earlier assessment and diagnosis of the correct alternative pathology, and treatment could be instituted earlier.

An overall evaluation of the body of evidence supporting BNP testing is provided in Table 53.

Table 53 Assessment of body of diagnostic evidence for BNP assay in the non-hospital setting

Component	A Excellent	B Good	C Satisfactory	D Poor
Volume of evidence		One or two level II studies with low risk of bias or a SR/multiple level III study with low risk of bias		
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact				Unknown[a]
Generalisability			Population(s) studied in body of evidence different to target population but it is clinically sensible to apply this evidence to target population	
Applicability		Applicable to Australian healthcare context with few caveats		

SR = systematic review; [a] Relates to the impact of BNP testing on patient management and health outcomes. No *direct* evidence of the impact of the test was found. *Linked* evidence is suggestive that the clinical impact would primarily be for those symptomatic patients 'ruled out' from heart failure.

Effectiveness of NT-proBNP assays in the diagnosis of heart failure in the non-hospital setting

Diagnostic accuracy

The evidence base for assessing the diagnostic accuracy of NT-proBNP testing was not as extensive as that available for BNP testing. However, like BNP testing it would appear that in general NT-proBNP assays have high sensitivity at detecting HF but lower specificity. The variability in specificity rates is quite wide, although the negative predictive values are uniformly high (>90%).

NT-proBNP assays, as is the case for BNP assays, appear to be of value as a highly sensitive test such that a negative test result 'rules out' the diagnosis of HF.

Impact on patient management

Given that NT-proBNP testing appears to have a high negative predictive value for ruling out HF, it is likely that many patients could safely forgo further diagnostic testing for HF on the basis of this one blood test. Nielsen et al (2004) determined that 51 per cent of 287 patients could safely forgo echocardiography on the basis of a NT-proBNP test (level IV intervention evidence).

Nielsen et al (2004) also suggested that GPs are undertreating patients despite having a clinical suspicion that they may have HF, and therefore the use of NT-proBNP testing in general practice would allow earlier and more effective treatment of HF prior to echocardiographic confirmation. However, this suggestion would need to be considered in the context of the high false positive rate associated with B-type natriuretic peptide tests, and therefore the likelihood of patients without HF being inappropriately treated for HF on the basis of the test alone (ie without a confirmatory echocardiogram or cardiology consult).

The good quality level II intervention evidence provided by Wright et al (2003) was conducted in a New Zealand adult population presenting in general practice with symptoms of dyspnoea and/or oedema of recent onset and suspected of HF. This population is applicable to the Australian situation. This trial very clearly reported the impact of NT-proBNP testing on clinical diagnoses formulated by GPs. A 13 per cent improvement [95%CI 5.5, 21.0, p=0.002] was observed in correct diagnoses in the NT-proBNP trial arm as compared to the control trial arm. The main effect of NT-proBNP testing on clinical decision-making was to assist GPs to correctly 'rule out' HF.

Impact on health outcomes

The direct impact of NT-proBNP testing on health outcomes was not assessed in any of the included studies. High level evidence of the effect of early treatment for HF (from linked evidence) on patient health outcomes indicates that early treatment for HF is beneficial, although this was not investigated systematically in this report. However, in symptomatic patients suspected of HF a NT-proBNP test is unlikely to result in the *earlier* definitive diagnosis and treatment of HF, as an echocardiogram is still required to

confirm the diagnosis (see comments earlier). It is, however, possible that NT-proBNP testing could benefit the health of those patients 'ruled out' from HF, as they would have an earlier assessment and diagnosis of the correct alternative pathology, and treatment could be instituted earlier.

An overall assessment of the body of evidence supporting NT-proBNP testing is provided in Table 54.

Table 54 Assessment of body of diagnostic evidence for NT-proBNP assay in the non-hospital setting

Component	A Excellent	B Good	C Satisfactory	D Poor
Volume of evidence		One or two level II studies with low risk of bias or a SR/multiple level III study with low risk of bias		
Consistency		Most studies consistent and inconsistency may be explained		
Clinical impact			Moderate[a]	Unknown[b]
Generalisability		Population(s) studied in the body of evidence are similar to the target population		
Applicability		Applicable to Australian healthcare context with few caveats		

SR = systematic review; [a] Relates to the impact of NT-proBNP testing on patient management; [b] Relates to the impact of NT-proBNP, as measured *directly*, on patient health outcomes. *Linked* evidence is suggestive that the clinical impact would primarily be for those symptomatic patients 'ruled out' from heart failure.

Current clinical guidance

Recent evidence-based clinical practice *Guidelines for the prevention, detection and management of chronic heart failure in Australia (2006)* (National Heart Foundation of Australia and The Cardiac Society of Australia and New Zealand 2006) do not recommend the *routine* measurement of BNP or NT-proBNP levels in the diagnosis of HF.

The guidelines suggest that the use of these tests is context dependent. Patients who have a very high likelihood of HF, on the basis of initial clinical assessment, should be treated as having HF and an echocardiogram ordered. This is primarily because the negative predictive value of the B-type natriuretic peptide tests, in this context, would be reduced, and there is no evidence that the tests provide any additional diagnostic information to that provided by an echocardiogram.

However, in cases of diagnostic or clinical uncertainty, where an echocardiogram is not easily accessible or available, it is suggested that the measurement of BNP or NT-proBNP levels may be considered so that *alternative* diagnoses can then be investigated in a timely manner. Raised B-type natriuretic peptide levels would still warrant further investigation, however, including echocardiography (National Heart Foundation of Australia and The Cardiac Society of Australia and New Zealand 2006).

What are the economic considerations?

Summary

This economic analysis has been largely limited to symptomatic patients (in line with the evidence base), that is those presenting with dyspnoea and/or oedema of recent onset and suspected of heart failure (HF).

The costs and outcomes associated with B-type natriuretic peptide testing will usually depend on the referral propensity of the general practitioner (GP). The extent of immediate cost offsets depends on whether or not the GP decides to order an echocardiogram and initially undertake self-management of the patient; or decides to refer the test positive patient to a cardiologist who may or may not order an echocardiogram. The extent of downstream costs (or savings) also depends on the same referral propensity and on the proportions of patients correctly diagnosed. No randomised controlled trial with health outcome data could be located in the non-hospital setting.

Lacking valid evidence regarding changes in GP referral behaviour and patient health outcomes, it was not possible to draw conclusions regarding the incremental cost per life-year (or per QALY) saved. Nevertheless, some deductions could be made.

Three scenarios are presented illustrating different types of possible referral patterns that have increasing levels of resource use. In all three scenarios the use of B-type natriuretic peptide testing is cost saving (from $50 to $86 per patient tested). However, the scenarios reflect current clinical practice guideline recommendations and assume that currently *all* patients over 45 years of age presenting with dyspnoea and/or oedema of recent onset and suspected of HF would receive an echocardiogram. The data available suggests that actual echocardiogram referral for these patients may be considerably lower. Results of a one-way sensitivity analysis suggest that B-type natriuretic peptide testing may not be cost saving if GPs currently refer this patient group to echocardiography at a rate of 60% or lower. Despite this, when testing is confined to general practice patients with dyspnoea and/or oedema of recent onset and suspected of HF, most of those patients who test negative on the B-type natriuretic peptide assay are still likely to have a clinically important pathology. In this situation the potential for improvement in health outcome may warrant the amount of resources used in testing and follow-up.

It has been determined that diminishing marginal returns could arise should the testing extend to general practice populations where there is a high probability of HF or when there are increasing proportions of patients with minor levels of symptoms but lacking any clinically important pathology.

The additional Australian Government expenditure due to the introduction of B-type natriuretic peptide assays into the non-hospital setting for patients presenting with dyspnoea and/or oedema of recent onset is estimated to be in the range $4.0–$11.3 million per year. This expenditure is likely to be offset by savings on fewer echocardiograms and cardiologist referrals and earlier management of non-HF diagnoses, but the extent of these offsets is presently unknown.

The purpose of an economic evaluation is to assist decision-makers in ensuring that society's ultimately scarce resources are allocated to those activities from which we will get the most value. That is, it seeks to enhance economic efficiency.

Economic evaluation under the MSAC process focuses on the scarce resources available within the Australian health system. It asks whether these scarce resources would be better spent on producing the amount of health gain obtainable through the intervention in question or through the identified comparator intervention.

An expert advisory panel considered that there was sufficient diagnostic effectiveness evidence for B-type natriuretic peptide testing in the non-hospital setting to warrant an economic analysis.

Objective

The aim of the present economic evaluation is to review the economic considerations associated with adding B-type natriuretic peptide testing [17] to current Australian protocols in the diagnosis of (ie ruling out) HF in a non-hospital setting, and to provide an indication of the extent of uncertainty entailed. The perspective of the analysis is that of society. Direct costs of informal care and indirect costs (ie productivity costs) are not considered.

As often occurs in economic evaluation, a paucity of data has necessitated several crucial assumptions. It is important that the results be interpreted in the light of the likely validity of these assumptions.

Introduction of B-type natriuretic peptide testing in a non-hospital setting

Appraising the potential economic impact of introducing B-type natriuretic peptide testing into Australian practice is difficult because no randomised controlled trials assessing patient health outcomes in the non-hospital setting could be located. There are also some serious concerns about the number of patients whom GPs might consider for the test. Therefore, in line with the best evidence available, the analysis in a non-hospital setting has been limited to patients presenting with dyspnoea and/or oedema of recent onset who are suspected of HF.

B-type natriuretic peptide tests are triage tests. Their cost-effective use in the non-hospital setting would require that the value of the additional information made available by the tests in terms of health gain and resource savings would be sufficient to justify the cost of providing the tests. Current laboratory benchmarking data suggest that B-type natriuretic peptide tests would cost $50.59 per test. Because bulk-billing occurs in the vast majority of such cases, it is appropriate to regard these unit costs as representing the opportunity cost of the test. In contrast, the Medicare Benefits Schedule (MBS) fee for an echocardiogram is $231.

[17] Due to insufficient evidence for costs and outcomes for each peptide assay, it was assumed that the cost and effectiveness of BNP and NT-proBNP testing were equivalent.

In the non-hospital setting, clinical practice and referral patterns vary widely, but three typical scenarios have been identified under which a patient presenting to a GP with dyspnoea and/or oedema of recent onset may require a B-type natriuretic peptide test for diagnostic purposes. These scenarios differ in the extent to which extrapolation of the health outcome results of the key trial by Mueller et al (2004b) in the hospital setting (the only trial where health outcome data are available—see Part A of this report) may be applicable.

In the first typical scenario the GP would customarily order an echocardiogram and, on the basis of the result, either manage the patient directly or refer the patient to a cardiologist. With the availability of B-type natriuretic peptide testing, and where the GP has real uncertainty as to the diagnosis of HF, either a blood sample or the patient would be sent to a pathology laboratory where batch testing is performed, and those patients with a positive test would then be referred for an echocardiogram. In doing so, the GP is accepting that a delay of perhaps 1 day or so will not substantially influence the patient's prognosis. The evidence for patient outcome in this scenario is limited. Since the Mueller et al (2004b) trial was performed on an emergency department (ED) population most of whom were subsequently admitted as inpatients, it is not directly applicable here. Where the GP has high suspicion that the patient has HF after an initial diagnostic assessment, B-type natriuretic peptide tests would not be required and the patient would be referred directly for an echocardiogram and/or to a cardiologist.

In the second typical scenario, the GP decides to refer the patient to a cardiologist for assessment either because the patient has a history of (uncontrolled) HF or because the GP feels unable to manage the patient appropriately without additional advice. Once again, in doing so the GP is accepting that a delay in definitive assessment will not substantially influence the patient's prognosis. The Mueller at al (2004b) trial is again not applicable to this situation.

The third typical scenario may arise in the rural and remote setting where the GP decides to admit the patient to the local hospital. Inpatient management would be continued by the admitting GP. The Mueller et al (2004b) trial may be relevant for this scenario, with some adjustment for the degree of severity of the presenting condition. In the absence of better evidence, this scenario will not be considered further because Part A of this Assessment Report has considered the Mueller et al trial in depth. For the sake of completeness, it should be noted that if and when a point-of-care test becomes available in Australia, this may well find a place in rural and remote practice.

Alternatives to these scenarios are possible, including sharing of care between the GP and a cardiologist with or without admission to hospital, but the three identified scenarios provide a basis for consideration of the main health outcome and resource use issues in the non-hospital setting.

The evaluation has not considered the (probably uncommon) situation where, after an initial assessment, the GP decides that the patient should immediately be referred to hospital without conducting a B-type natriuretic peptide test or awaiting its result. This situation thus leads to an ED presentation under essentially similar circumstances to those considered in the economic analysis based on the key trial by Mueller et al (2004b), which has been discussed in Part A of this Assessment Report.

Diagnostic accuracy and impact on patient management of B-type natriuretic peptide testing in the non-hospital setting

During the evaluation the highest levels of evidence identified were in one randomised controlled trial (Wright et al 2003) with level II intervention evidence and two diagnostic accuracy studies (Cowie et al 1997; Zaphiriou et al 2005) with level II diagnostic evidence, which reported on the use of B-type natriuretic peptide testing in the non-hospital setting against the correct reference standard. However, these studies report neither health outcomes nor the propensity of GPs to order cardiologist assessment and echocardiography in routine practice. Unless ambulatory care patients are assumed to have the same prognosis as patients in an ED setting (ie the Mueller et al (2004b) trial), which is doubtful, there is no evidence-based data on which to determine the magnitude of the difference in health outcome in the non-hospital setting without conducting a randomised controlled trial specifically for that setting.

Wright et al (2003) conducted a prospective, randomised controlled trial (level II intervention evidence) of the effect of NT-proBNP on the accuracy of HF diagnosis in New Zealand general practice. Patients had presented to their GP with symptoms of dyspnoea and/or peripheral oedema of recent onset. The primary end point was the accuracy of the GPs' diagnoses both with and without use of a NT-proBNP test result compared to a full cardiologic assessment including echocardiography (as assessed by a panel of three cardiologists and a GP). Diagnostic accuracy improved 21 per cent in the NT-proBNP group and 8 per cent in the control group (p=0.002). The main impact was in enabling GPs to correctly rule out HF. There was a difference of 10.6 per cent in patients 'ruled out' for a HF diagnosis in the NT-proBNP group compared to the control group. Whereas initially the GPs suspected that 70 per cent of the patients had HF, only 25 per cent of patients had a confirmed diagnosis of HF following the full cardiologic assessment. The alternative diagnoses reached by the GPs were not reported in the trial, and neither were patient health outcomes. The design of the study also does not allow for an assessment of the percentage of patients whom the GP, with and without the availability of the NT-proBNP test, would have referred for echocardiography, with or without follow-up cardiologist assessment, in routine practice.

This study also reports the diagnostic accuracy of NT-proBNP testing versus the panel diagnosis (using all clinical investigations other than NT-proBNP) of HF (see Table 55).

Table 55 Reconstruction of NT-proBNP test results from the intervention arm of Wright et al (2003) [a]

	HF +	HF –	Total
Test +	39 (25.7%)	42 (27.6%)	81 (53.3%)
Test –	4 (2.6%)	67 (44.1%)	71 (46.7%)
Total	43 (28.3%)	109 (71.7%)	152 (100%)

[a] Based on the NT-proBNP arm of Wright et al (2003). Reference standard was panel diagnosis by three cardiologists and a general practitioner, and included echocardiography.

In the diagnostic accuracy study (level III-1 diagnostic evidence) by Cowie et al (1997), GPs in a London district referred all suspected cases of HF to a rapid-access clinic for diagnosis by a panel of three cardiologists using all clinical information, including echocardiography. This study was part of the Hillingdon Heart Failure Study. Of the 122 patients referred, 35 (29%) were judged to satisfy the case definition for HF. Of the 87 with other diagnoses, 19 (22%) had chronic obstructive pulmonary disease, 17 (20%) obesity, 12 (14%) angina, 7 (8%) venous insufficiency, 7 (8%) anxiety, 5 (6%)

palpitation/arrhythmias, 2 (2%) pulmonary fibrosis, 2 (2%) malignant disorders and 9 (10%) other diagnoses; the diagnosis was not known in the remaining 7 patients. The area under the receiver-operating-characteristic curve, indicating the overall accuracy of BNP testing, was 0.96. With a negative predictive value of 98 per cent, the positive predictive value of BNP testing was 70 per cent. No health outcomes were reported. The design of this study also does not allow for an assessment of the proportion of patients whom the GP, with and without the availability of the BNP test, would have referred for echocardiography, with or without follow-up cardiologist assessment, in routine practice.

In the diagnostic accuracy study (level II diagnostic evidence) by Zaphiriou et al (2005), GPs referred patients with new symptoms suggestive of HF to rapid-access HF clinics in major cities in the United Kingdom. After full assessment, the diagnosis of HF was confirmed in 34 per cent of patients. The area under the receiver-operating-characteristic curve was 0.84 (95%CI 0.79-0.89) for BNP and 0.85 (0.81-0.90) for NT-proBNP. At the manufacturers' recommended decision cut-off points, NT-proBNP had a higher negative predictive value (0.97) than BNP (0.87), but a lower positive predictive value (0.44 versus 0.59). Again, the design of the study does not allow for an assessment of the proportion of patients whom the GP, with and without the availability of the BNP or NT-proBNP test, would have referred for echocardiography, with or without follow-up cardiologist assessment, in routine practice.

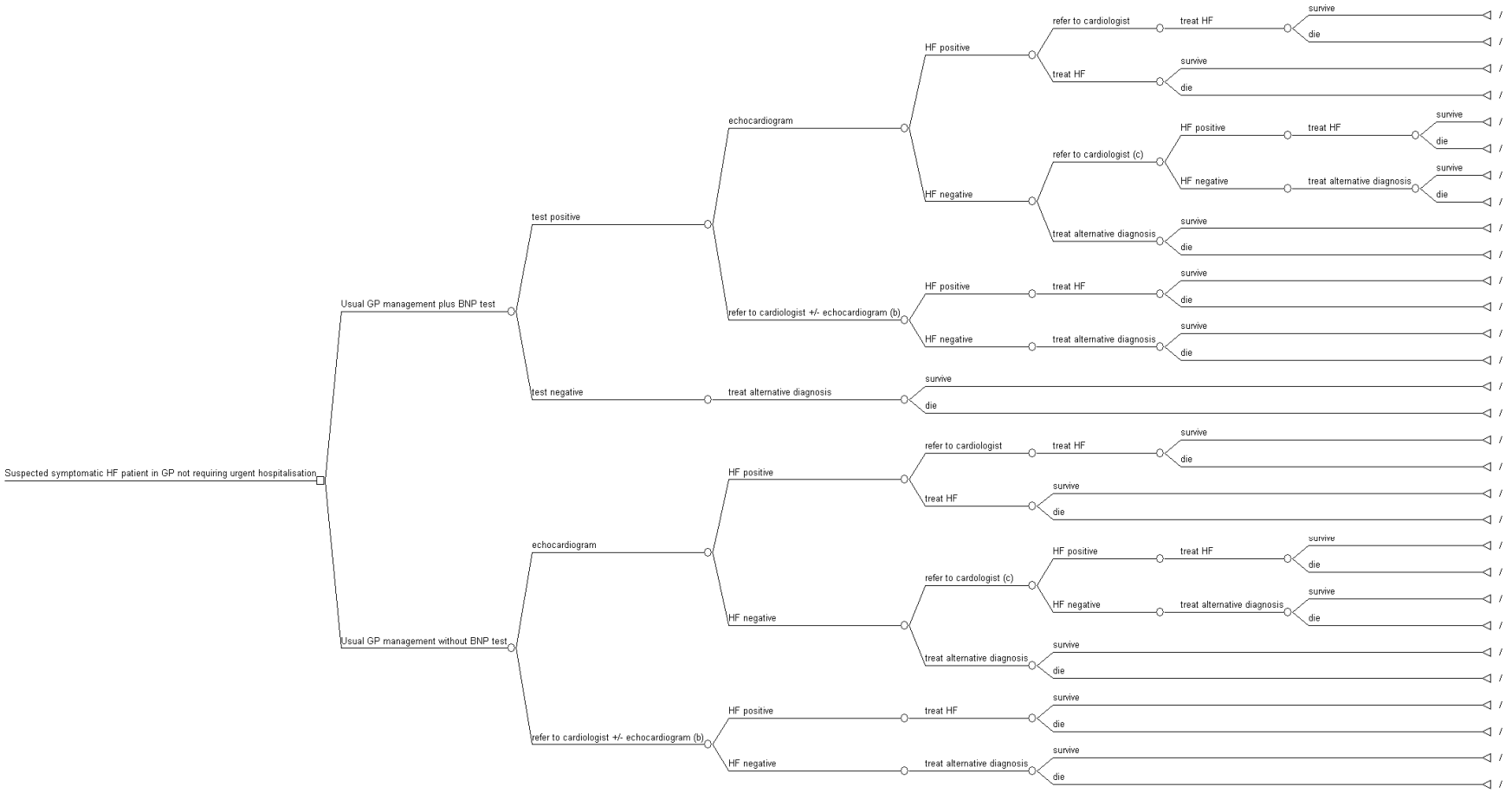
These latter two studies were attempting to assess the diagnostic accuracy of B-type natriuretic peptide tests in the non-hospital setting. GPs recruiting patients for these studies were requested to refer all incident cases of patients with suspected HF to a referral centre where the reference test (clinical diagnosis after all investigations, including echocardiography) could be applied. However, there is the possibility that practitioners with a special interest in HF may have been more likely to respond to an invitation to join the research project. Further, it is unclear whether or not the studies by Zaphiriou et al (2005) and Cowie et al (1997): (1) report selective referral of the more challenging cases presenting to a GP, (2) reflect an unrealistic propensity of GPs to refer numbers of patients apparently meeting the criteria, or (3) reflect referral behaviour that remained unchanged despite participation in a trial. These studies may not be indicative of the behaviour of a typical GP.

Towards a decision analytic model for ambulatory management in the non-hospital setting

Since a careful search of the literature has not revealed a randomised controlled trial in the non-hospital setting with final health outcome indicators, the possibility of modelling has been considered.

The population in this model would consist of primary care patients presenting with dyspnoea and/or oedema of recent onset and suspected of HF. The decision whether or not to order a B-type natriuretic peptide test would be made after an appropriate clinical history and physical examination. Usual management in general practice plus B-type natriuretic peptide testing would be compared with usual management in general practice without B-type natriuretic peptide testing (see Figure 13).

Figure 13 Decision tree for clinical diagnosis strategy with and without B-type natriuretic peptide testing in the non-hospital setting for patients presenting with dyspnoea and/or oedema of recent onset, not requiring urgent hospitalisation, and suspected of heart failure (a)



HF = heart failure; (a) Patients with initial *probable* heart failure diagnosis should not receive BNP test but directly undergo echocardiography (National Heart Foundation of Australia and The Cardiac Society of Australia and New Zealand 2006); (b) Some patients would have received an echocardiogram prior to presenting to the GP; (c) As echocardiography is an imperfect reference standard

Unit costs

The unit costs applicable to the economic model are shown in Table 56.

The value of the resources used in healthcare interventions was assumed to be 100 per cent of the MBS fee (ie over-schedule charges were ignored).

Table 56 Unit costs of resources used for the management of patients presenting with dyspnoea and/or oedema of recent onset in the non-hospital setting

MBS item no.	Schedule fee	Short definition MBS item no. description
23	$30.85	**Level B practitioner consult** Professional attendance involving taking a selective history, examination of the patient with implementation of a management plan in relation to 1 or more problems, OR a professional attendance of less than 20 minutes duration involving components of a service to which item 36, 37, 38, 40, 43, 44, 47, 48, 50 or 51 applies.
36	$58.55	**Level C practitioner consult** Professional attendance involving taking a detailed history, an examination of multiple systems, arranging any necessary investigations and implementing a management plan in relation to 1 or more problems, and lasting at least 20 minutes, OR a professional attendance of less than 40 minutes duration involving components of a service to which item 44, 47, 48, 50 or 51 applies
110	$128.05	**Professional attendance** Consultant physician (other than in psychiatry), referred consultation – surgery or hospital (professional attendance at consulting rooms or hospital by a consultant physician in the practice of his or her specialty (other than in psychiatry) where the patient is referred to him or her by a medical practitioner)
116	$64.10	**Subsequent professional attendances** Each attendance subsequent to the first in a single course of treatment
73910	$17.50	**Patient episode initiation** Initiation of a patient episode by collection of a specimen for a service (other than a service described in item 73901, 73903 or 73905 or in Group P9) if the specimen is collected by an approved pathology practitioner or an employee of an approved pathology authority from a person in a residential aged care home or institution
55113	$230.65	**Echocardiography** M-MODE and 2 DIMENSIONAL REAL TIME ECHOCARDIOGRAPHIC EXAMINATION of the heart from at least 2 acoustic windows, with measurement of blood flow velocities across the cardiac valves using pulsed wave and continuous wave Doppler techniques, and real time colour flow mapping from at least 2 acoustic windows, with recordings on video tape or digital medium, not being a service associated with a service to which an item in Subgroups 1 (with the exception of item 55054) or 3, or another item in this Subgroup (with the exception of items 55118 and 55130), applies, for the investigation of symptoms or signs of cardiac failure, or suspected or known ventricular hypertrophy or dysfunction, or chest pain (R)
11509	$30.85	**Lung function test** MEASUREMENT OF RESPIRATORY FUNCTION involving a permanently recorded tracing and written report, performed before and after inhalation of bronchodilator, with continuous technician attendance in a laboratory equipped to perform complex respiratory function tests (the tests being performed under the supervision of a specialist or consultant physician or in the respiratory laboratory of a hospital) - each occasion at which 1 or more such tests are performed

A decision analytic model for ambulatory management in the non-hospital setting would have to take into account:

- test performance
- GP referral behaviour (for echocardiography and/or to a cardiologist)
- costs of the various treatment pathways

- patient health outcomes.

In the absence of evidence about the differences in GP referral behaviour and patient health outcomes, it would be difficult to complete the model calculations and thus draw valid conclusions regarding the incremental cost per life-year (or per QALY) saved associated with the use of B-type natriuretic peptide testing.

Nevertheless, some deductions can be made by inspection of the structure of the model and by a comparison of the immediate costs of diagnosis through a reconstruction of the results of Wright et al (2003). Three scenarios have been presented illustrating different types of possible referral patterns that have increasing levels of resource use. Drug costs have not been included in either arm as it is unknown what proportions and type of alternative diagnoses are present in patients with new symptoms of dyspnoea and oedema, suspected of HF, and treatment is wholly dependent on the presenting pathology.

Scenario 1: Where the GP would always order an echocardiogram, unless the B-type natriuretic peptide test is negative, and would also manage the patient.

In this situation the initial cost of the B-type natriuretic peptide test is offset by the cost savings from not performing an echocardiogram in those patients who are negative on the B-type natriuretic peptide test. There is no good evidence available regarding downstream costs or cost savings for those patients ruled out from HF and having earlier and correct identification of an alternative diagnosis. However, for this group of patients there may be health benefits (depending on the severity of the alternative pathology) due to this earlier identification. For those patients with a positive B-type natriuretic peptide test there is likely to be no increase in downstream costs or health benefits as *no more* patients will receive a diagnosis of HF, nor will the diagnosis occur earlier than currently. Delays associated with instituting a B-type natriuretic peptide test prior to echocardiographic confirmation are unlikely to have an impact as the patient group is not acutely ill. Although without a cardiology referral it is likely that some cases of HF will be missed, this is a function of echocardiography as an imperfect reference standard and is unrelated to the impact of B-type natriuretic peptide testing. Any harms associated with a false negative B-type natriuretic peptide test are likely to be minimal as the patient population is not acutely ill and the test's false negative rate is very small.

Table 55 (see above) indicates that, using the test parameters and prevalence from the NT-proBNP arm of Wright et al (2003), GP assessment plus the NT-proBNP test correctly rules out 44.1 per cent of patients and incorrectly rules out (ie false negative) 2.6 per cent of patients. The 'gold standard' rule out following cardiologist assessment and echocardiography is 71.7 per cent. Sensitivity of the test is 90.7 per cent and specificity is 61.5 per cent.

In Scenario 1, and using the data from Table 55, there will be immediate cost savings of $50 per patient (Table 57).

Table 57 Illustrative comparison of the immediate costs of diagnosis with and without the availability of a B-type natriuretic peptide test for ambulatory management of patients presenting with dyspnoea and/or oedema of recent onset in general practice: where the GP would *always* order an echocardiogram, unless the B-type natriuretic peptide test is negative, and self-manage the patient

	NT-proBNP test performed on all patients (n=100)	NT-proBNP test not available (n=100)
Level C Practitioner consult @ $58.55	$5,855	$5,855
NT-proBNP test @ $50.59 [a]	$5,059	-
Level B Practitioner consult @ $30.85	$3,085 [b]	-
Echocardiogram @ $230.65	$12,294 [c]	$23,065
Level B Practitioner consult @ $30.85	-	$3,085 [d]
Lung function test [e] @ $30.85	$1,441 [f]	$2,360 [g]
Level B Practitioner consult @ $30.85	$3,085 [h]	$2,360 [i]
Lung function test @ $30.85	$919 [j]	-
Total cost	**$31,738**	**$36,725**

[a] Batch testing and bulk billing assumed; [b] Patient receives NT-proBNP test results and, on basis of result, GP orders echocardiography or a lung function test; [c] Proportion of patients testing positive on NT-proBNP assay = 53.3 per cent (calculated from Wright et al 2003; see Table 55); [d] Patient receives echocardiography results, and either treatment for HF is initiated or GP orders lung function test for patient identified as HF-negative on basis of echocardiogram; [e] Lung function test used as a representative of tests the GP would conduct in HF-negative patients to establish an alternative diagnosis; [f] Proportion of patients testing negative on NT-proBNP assay = 46.7 per cent (Wright et al 2003); [g] Nielsen et al 2004 found that 23.5 per cent of patients referred by GP received echocardiographic confirmation of HF. Therefore, assumption is that 76.5 per cent of echocardiograms in a non-NT-proBNP test population will be HF-negative —concordant with 71.7 per cent of patients being HF-negative after full cardiologic assessment in Wright et al 2003 and notion that echocardiography is an imperfect reference standard and fails to identify diastolic HF; [h] Patient receives results of echocardiography or lung function test and treatment initiated (where warranted) in those with 'definitive' diagnosis, that is patients with negative echocardiogram referred for lung function test; [i] Patient receives results of lung function test and treatment initiated (where warranted); [j] Proportion of NT-proBNP test positive patients who are in turn echocardiogram negative is unknown. Assume, given small false negative rate on NT-proBNP test, that it equates to HF-negative rate applicable to unfiltered (ie no NT-proBNP test) population (from tablenote g) minus NT-proBNP test negative (HF-negative) rate = 76.5–46.7 = 29.8 per cent of total population would receive lung function test at this juncture.

Scenario 2: Where the GP would always order an echocardiogram, unless the B-type natriuretic peptide test is negative, and refer to a cardiologist if the echocardiogram is positive.

In this situation the initial cost of the B-type natriuretic peptide test is offset by the cost savings from not performing an echocardiogram in those patients who test negative on the B-type natriuretic peptide assay. False negatives on the B-type natriuretic peptide test may or may not have been picked up through use of an echocardiogram and thus it is unclear whether there would be any savings on cardiology referrals. These would be small in any event, given the low false negative rate.

The evidence is not available regarding downstream costs or cost savings for those patients ruled out from HF and having earlier and correct identification of alternative diagnoses. However, as mentioned above, for this group of patients there may be health benefits due to the earlier correct diagnosis. For those patients with a positive B-type natriuretic peptide test there is likely to be no increase in downstream costs or health benefits as *no more* patients will receive a diagnosis of HF, nor will the diagnosis occur earlier than currently. Delays associated with instituting a B-type natriuretic peptide test prior to echocardiographic confirmation are unlikely to impact on these patients as they are not acutely ill. Although without a cardiology referral for patients with a negative echocardiogram it is likely that some cases of HF will be missed, this is a function of echocardiography as an imperfect reference standard and is unrelated to the impact of B-

type natriuretic peptide testing. Any harms associated with a false negative B-type natriuretic peptide test are likely to be minimal as the patient population is not acutely ill and the test's false negative rate is very small.

In Scenario 2, and using the data from Table 55, there will be immediate cost savings of $64 per patient (Table 58).

Table 58 Illustrative comparison of the immediate costs of diagnosis with and without the availability of a B-type natriuretic peptide test for ambulatory management of patients presenting with dyspnoea and/or oedema of recent onset in general practice: where the GP would *always* order an echocardiogram, unless the B-type natriuretic peptide test is negative, and always refer echocardiogram positive patients to a cardiologist

	NT-proBNP test performed on all patients (n=100)	NT-proBNP test not available (n=100)
Level C Practitioner consult @ $58.55	$5,855	$5,855
NT-proBNP test @ $50.59 [a]	$5,059	-
Level B Practitioner consult @ $30.85	$3,085 [b]	-
Echocardiogram @ $230.65	$12,294 [c]	$23,065
Level B Practitioner consult @ $30.85	-	$3,085 [d]
Lung function test [e] @ $30.85	$1,441 [f]	$2,360 [g]
Level B Practitioner consult @ $30.85	$3,085 [h]	$2,360 [i]
Lung function test @ $30.85	$919 [j]	-
Cardiology consult @ $128.05	$1,604 [k]	$3,009 [k]
Total cost	**$33,342**	**$39,734**

[a] Batch testing and bulk billing assumed; [b] Patient receives NT-proBNP test results and, on basis of result, GP orders echocardiography or a lung function test; [c] Proportion of patients testing positive on NT-proBNP assay = 53.3 per cent (calculated from Wright et al 2003; see Table 55); [d] Patient receives echocardiography results and, if positive, is referred to a cardiologist or, if negative, is referred for a lung function test; [e] Lung function test used as a representative of tests the GP would conduct in HF-negative patients to establish an alternative diagnosis; [f] Proportion of patients testing negative on NT-proBNP assay = 46.7 per cent (Wright et al 2003); [g] Nielsen et al 2004 found that 23.5 per cent of patients referred by GP received echocardiographic confirmation of HF. Therefore, assumption is that 76.5 per cent of echocardiograms in a non-NT-proBNP test population will be HF-negative—concordant with 71.7 per cent of patients being HF-negative after full cardiologic assessment in Wright et al 2003 and notion that echocardiography is an imperfect reference standard and fails to identify diastolic HF; [h] Patient receives results of echocardiography and, if positive, is referred to cardiologist or, if negative, receives lung function test OR patient receives results of lung function test and treatment is initiated (if warranted); [i] Patient receives results of lung function test and treatment initiated (where warranted); [j] Proportion of NT-proBNP test positive patients who are in turn echocardiogram negative is unknown. Assume, given small false negative rate on NT-proBNP test, that it equates to HF-negative rate applicable to unfiltered (ie no NT-proBNP test) population (from tablenote g) minus NT-proBNP test negative (HF-negative) rate = 76.5–46.7 = 29.8 per cent of total population would receive lung function test at this juncture; [k] See tablenotes g and j—assume 23.5 per cent of echocardiograms ordered are test positive and thus require a cardiology consult.

Scenario 3: Where the GP would always order an echocardiogram *and* refer to a cardiologist, unless the B-type natriuretic peptide test is negative.

In this situation the initial cost of the B-type natriuretic peptide test is offset by the cost savings from not referring to a cardiologist and from not performing an echocardiogram in those patients who test negative on the B-type natriuretic peptide assay.

The available evidence is limited regarding downstream costs or cost savings for those patients ruled out from HF and having earlier and correct identification of alternative diagnoses. For this group of patients, however, health benefits are possible due to the earlier correct diagnosis. For those patients with a positive B-type natriuretic peptide test there is likely to be no increase in downstream costs or health benefits as *no more* patients will receive a diagnosis of HF, nor will the diagnosis occur earlier than currently. Delays associated with instituting a B-type natriuretic peptide test prior to echocardiographic and cardiologic confirmation are unlikely to impact on these patients as they are not

acutely ill. With a cardiology referral for all patients with a positive B-type natriuretic peptide test, regardless of echocardiogram result, it is likely that all cases of HF will be captured. Any harms associated with a false negative B-type natriuretic peptide test are likely to be minimal as the patient population is not acutely ill and the test's false negative rate is very small.

In Scenario 3, and using the data from Table 55, there will be immediate cost savings of $86 per patient (see Table 59).

Table 59 Illustrative comparison of the immediate costs of diagnosis with and without the availability of a B-type natriuretic peptide test for ambulatory management of patients presenting with dyspnoea and/or oedema of recent onset in general practice: where the GP would *always* refer to a cardiologist and order an echocardiogram, unless the B-type natriuretic peptide test is negative

	NT-proBNP test performed on all patients (n=100)	NT-proBNP test not available (n=100)
Level C Practitioner consult @ $58.55	$5,855	$5,855
NT-proBNP test @ $50.59 [a]	$5,059	-
Level B Practitioner consult @ $30.85	$3,085 [b]	-
Cardiology consult @ $128.05	$6,825 [c]	$12,805
Echocardiogram @ $230.65	$12,294 [c]	$23,065
Lung function test [d] @ $30.85	$1,441 [e]	$2,212 [g]
	+ $771 [f]	
	= $2,212	
Level B Practitioner consult @ $30.85	$2,212 [h]	$2,212 [h]
Total cost	**$37,542**	**$46,149**

[a] Batch testing and bulk billing assumed; [b] Patient receives NT-proBNP test results and, on basis of result, GP orders echocardiography or a lung function test; [c] Proportion of patients testing positive on NT-proBNP assay = 53.3 per cent (calculated from Wright et al 2003; see Table 55); [d] Lung function test used as a representative of tests the GP would conduct in HF-negative patients to establish an alternative diagnosis; [e] Proportion of patients testing negative on NT-proBNP assay = 46.7 per cent (Wright et al 2003); [f] Proportion of NT-proBNP test positive patients who are in turn HF-negative after full cardiology assessment is unknown. Assume, given small false negative rate on NT-proBNP test, that it equates to HF-negative rate applicable to unfiltered (ie no NT-proBNP test) population receiving full cardiology assessment (71.7%, Wright et al 2003) minus NT-proBNP test negative rate (46.7%) = 25 per cent of total population would receive lung function test at this juncture; [g] Wright et al 2003 reported 71.7 per cent of patients referred by GPs for full cardiologic assessment were HF-negative. These then would require a lung function test; [h] Patients receiving lung function test would receive result from GP and treatment initiated (where warranted).

It is likely that actual clinical practice in the non-hospital setting would be a hybrid of all three scenarios given above. In all scenarios the initial cost of the B-type natriuretic peptide test is offset by:

- fewer patients (those who are test negative) in the B-type natriuretic peptide test arm who go on to a cardiologist referral and/or an echocardiogram; and

- some patients who were previously incorrectly diagnosed as having HF being now managed with their correct (non-HF) diagnosis sooner than would otherwise have happened.

The amount of the savings associated with the use of a B-type natriuretic peptide test depends on the propensity of the GP to refer or to order tests. Evidence for this propensity has not been located. The trial by Wright et al (2003) does not address this issue. The extent of downstream costs and cost savings depends on the same propensity to refer test positive patients, and on the proportions of patients correctly diagnosed.

Health outcomes must also be considered. This can be done by taking the cardiologist assessment plus echocardiogram as the 'gold standard' for the diagnosis of HF and, in addition, assuming that patients thereby ruled out for HF will have their true diagnosis established.

Under these conditions, any patient who has an echocardiogram and cardiologist assessment will experience the health outcome they would have had whether or not they had a prior B-type natriuretic peptide test, although their definitive management may be delayed while waiting for the full cardiology assessment.

Patients ruled out from HF by a B-type natriuretic peptide test will receive the same management they would have received following rule out with full cardiology assessment, but without the delay in waiting for these referrals or of receiving inappropriate treatment. A very minor proportion of these patients will be false negatives.

Thus, any patient benefit from use of the B-type natriuretic peptide test is basically is a trade-off between earlier and delayed management. That is, the delay associated with use of B-type natriuretic peptide test *prior* to an echocardiogram in patients *with* HF vs the early management associated with a B-type natriuretic peptide test *instead of* an echocardiogram in patients *without* HF. The extent of this trade-off can only be determined from a study of GP referral behaviour concerning patients with dyspnoea and/or oedema of recent onset who are suspected of HF.

In conclusion, in all three scenarios the use of B-type natriuretic peptide testing is cost saving (from $50 to $86 per patient tested). However, it must be noted that each scenario reflects current clinical practice guideline recommendations for the diagnosis of HF (National Heart Foundation of Australia and The Cardiac Society of Australia and New Zealand 2006) and assumes that currently *all* patients presenting with dyspnoea and/or oedema of recent onset and suspected of HF (limited to 45+ years of age in this analysis) would receive an echocardiogram.

The data available suggests that actual echocardiogram referral may range from 3.8 per cent (BEACH data, AIHW 2001) to 17.7 per cent (CASE study, Krum et al 2001) for patients with new symptoms suggestive of HF. It has been suggested that GPs often rely on clinical grounds alone to diagnose and treat HF as access to echocardiograms may be limited (Hobbs 2002). Should this be the case, the cost savings associated with reducing echocardiogram and cardiologist referrals may no longer apply, although there may be some savings on current inappropriate HF treatment for those patients who are B-type natriuretic peptide test negative. As all the scenarios presented assume that current practice is 'best practice', they *may substantially overestimate the cost savings* associated with these tests.

Table 60 illustrates the impact on costs/cost savings per patient of varying the usual rate of GP referral of symptomatic patients to echocardiography for scenario 1. An assumption underlying this one-way sensitivity analysis is that, when the NT-proBNP test is available, those patients suspected of HF and with a positive B-type natriuretic peptide test—who are not referred for echocardiography—would instead be diagnosed and treated for HF. An equivalent assumption has been applied for when the NT-proBNP test is not available. In this case, those patients suspected of HF—but not referred for echocardiography—would instead be diagnosed and treated for HF on the basis of clinical symptoms alone. Further assumptions applied to both arms are that the treatment would be reviewed at a follow-up consultation, and that the cost of treatment

(ie drug treatment) over this period would be trivial (and thus was not included). Given these assumptions, the results outlined in Table 60 suggest that B-type natriuretic peptide testing may not be cost saving if GPs currently refer to echocardiography, at a rate of 60 per cent or lower, those patients presenting with dyspnoea and/or oedema of recent onset who are suspected of HF.

Table 60 One-way sensitivity analysis for Scenario 1: impact on cost savings per patient of varying proportion of symptomatic patients referred by GP for echocardiography

Echocardiogram referrals	NT-proBNP performed on all patients (n=100)	NT-proBNP test not available (n=100)	Cost savings per patient
100% (base case/CPGs[a])	$31,738	$36,725	$50
90%	$30,272	$33,947	$37
80%	$29,043	$31,168	$21
70%	$27,342	$28,390	$10
60%	**$25,901**	**$25,611**	**-$3**
50%	$24,672	$22,833	-$18
40%	$23,442	$20,054	-$34
30%	$22,213	$17,276	-$49
20%	$20,983	$14,497	-$65
10%	$19,754	$11,719	-$80

[a] CPGs = according to clinical practice guidelines. **Bold** indicates the threshold where the test is no longer cost saving.

Unfortunately there are no robust data available on the referral patterns of GPs who are presented with patients reporting dyspnoea and/or oedema of recent onset and suspected of HF. It is likely that referral patterns will vary widely between GPs and, despite B-type natriuretic peptide tests reducing the number of 'ruled out' patients going on to an echocardiogram, it is also unknown what impact the introduction of these tests will have on the current echocardiogram referral rate for those who test positive.

Despite this, when testing is confined to general practice patients with dyspnoea and/or oedema of recent onset and suspected of HF, most of those patients who test negative on the B-type natriuretic peptide assay are still likely to have a clinically important pathology. In this situation the potential for improvement in health outcome may warrant the amount of resources used in testing and follow-up.

Are there diminishing marginal returns associated with B-type natriuretic peptide testing?

Diminishing marginal returns is a descriptor for the typical situation where, keeping other factors constant, using more and more of a variable input (viz more B-type natriuretic peptide tests) leads to less and less additional output (viz better health). Hence, there may be a level of testing intensity beyond which it is not cost-effective to proceed.

Two situations are considered:

- where the level of symptoms is the same but the distribution of underlying pathology between HF and other diseases may vary
- where the level of symptoms is less.

Same level of symptoms

Suppose the level of patient symptoms (of dyspnoea and/or peripheral oedema) is similar, but there are in fact less true HF patients in the population, perhaps due to improved health promotion and prevention strategies. This would mean that the introduction of B-type natriuretic peptide testing would lead to the ruling out of more patients so that their true diagnosis could be established. This would not lead to a situation of diminishing marginal returns. That is, B-type natriuretic peptide testing in the primary care of patient populations with a lower prevalence of HF and a correspondingly higher prevalence of other clinically important pathologies (ie underlying the presenting symptoms) should still lead to a better diagnostic process compared to the counterfactual of not testing.

The following discussion relies on recognition that B-type natriuretic peptide assays are a 'rule out' test. These tests have characteristics that help the GP to exclude a diagnosis of HF and so focus on an alternative diagnosis, which will often be respiratory in nature. It is important to recognise that, while the test can rule out HF with a high degree of accuracy, it cannot be relied upon to confirm a diagnosis of HF. This is illustrated in Table 61, in which the calculations are based on NT-proBNP sensitivity of 90.7 per cent and specificity of 61.5 per cent from the intervention arm of Wright et al (2003).

As noted in the systematic review section of this report, the prevalence of HF reported in the included studies was in the range 9%–34% (median 29%) in the non-hospital setting. The proportions of patients with a final diagnosis of HF were 25 per cent in Wright et al (2003), 29 per cent in Cowie et al (1997) and 34 per cent in Zaphiriou et al (2005).

Table 61 Illustration of the characteristics of B-type natriuretic peptide assays as a 'rule out' test (using test accuracy from Wright et al 2003)

Prevalence of true HF (HF+) in the tested population	Probability that patient does *not* have HF	PPV	NPV	False positives per 100 tests	False negatives per 100 tests
90%	10%	95.5%	42.4%	3.85	8.37
70%	30%	84.6%	73.9%	11.55	6.51
50%	50%	70.2%	86.9%	19.25	4.65
30%	70%	50.2%	93.9%	26.95	2.79
10%	90%	20.7%	98.3%	34.65	0.93

HF = heart failure; PPV = positive predictive value; NPV = negative predictive value

Arguably, an informed pre-test probability for an eventual definitive diagnosis of HF in a patient presenting currently in Australia, with a similar history of dyspnoea and/or oedema of recent onset, is approximately 30 per cent. Applying Bayes' theorem with a pre-test probability of HF of 30 per cent and with the sensitivity and specificity from the intervention arm of the Wright et al (2003) trial conducted in New Zealand, the post-test probability would rise only to 50 per cent. That is, there is only a 1 in 2 chance that the B-type natriuretic peptide test would correctly identify HF in patients with a positive test. While the false negative rate will be low, the false positive rate can approach 40 per cent. In this situation, however, most patients who are 'false positive' still have some form of clinically important pathology that would need further investigation. Therefore, if there is a high false positive rate and a high rate of echocardiogram referral, this may not be unnecessary referral if it leads to the alternative diagnosis being correctly established in a timely manner, either by the cardiologist or on referral back to the GP.

Diminishing marginal returns could arise if the GP has a very high index of suspicion for HF but uses the BNP test anyway. In this case the cost of the additional B-type natriuretic peptide test may be unnecessary as the patient is likely to have HF and the echocardiogram will probably be performed in any event. The cost implications are explored in Table 62.

Table 62 Illustration of the characteristics of B-type natriuretic peptide assays as a 'rule out' test in 100 patients in general practice with acute dyspnoea and/or oedema of recent onset suggestive of heart failure (using Scenario 1 diagnostic pathway)

Number of patients who do *not* have HF	Prevalence or likelihood of true HF (HF+) in the tested population	Number of HF– patients who will be correctly ruled out (true negatives)	Number of patients who will be referred for echocardiogram[a]	Total cost of testing plus echocardiogram[b]
10	90%	6.15	85.48	$39,160
30	70%	18.45	75.04	$36,752
50	50%	30.75	64.60	$34,344
70	30%	43.05	54.16	$31,936
90	10%	55.35	43.72	$29,528

HF = heart failure; [a] Assuming all test positive patients are referred, ie the extreme situation; [b] See Table 57, Scenario 1, for cost items; assuming sensitivity = 90.7 per cent and specificity = 61.5 per cent (Wright et al 2003).

When compared with the cost of echocardiography alone (see Scenario 1, Table 57), Table 62 illustrates that it is *unlikely* that diminishing marginal returns will emerge if the test is used on patients with dyspnoea/or oedema of recent onset. However, it is also likely that the test will *not* be cost saving if it is *probable* that the patient has HF (>70% prevalence of true HF in patient population). (See Table 57 for cost of the comparator in Scenario 1).

If the B-type natriuretic peptide test is negative, it is reasonable to consider an alternative diagnosis. However, if the B-type natriuretic peptide test is positive, confirmatory testing with echocardiography, with or without cardiologist referral, is highly desirable. As pointed out in the Discussion section of this report, prevalence rates of HF in non-hospital settings in the included studies were in the range 9%–34% (median 29%). Given this prevalence rate and the data in Table 61, the percentage of false positives expected in a non-hospital setting would be approximately 27 per cent. Treatment for HF on the basis of a positive B-type natriuretic peptide test *alone* should therefore not be undertaken because approximately one-quarter of patients would be inappropriately treated.

Lesser level of symptoms

With B-type natriuretic peptide testing in general practice, diminishing marginal returns could perhaps arise in the testing of patient populations with less severe symptoms (eg fatigue).

This can be envisaged as the inclusion in the population tested of an increasing proportion of patients who do not have any *clinically important* form of organic pathology.

The three studies relied on in this economic analysis in the non-hospital setting—those by Wright et al (2003), Cowie et al (1997) and Zaphiriou et al (2005)—do not appear to have been designed to test for this possibility. The test parameters calculated in these studies are derived in symptomatic populations. It is important to recognise that the economic properties of B-type natriuretic testing discussed in previous sections of this

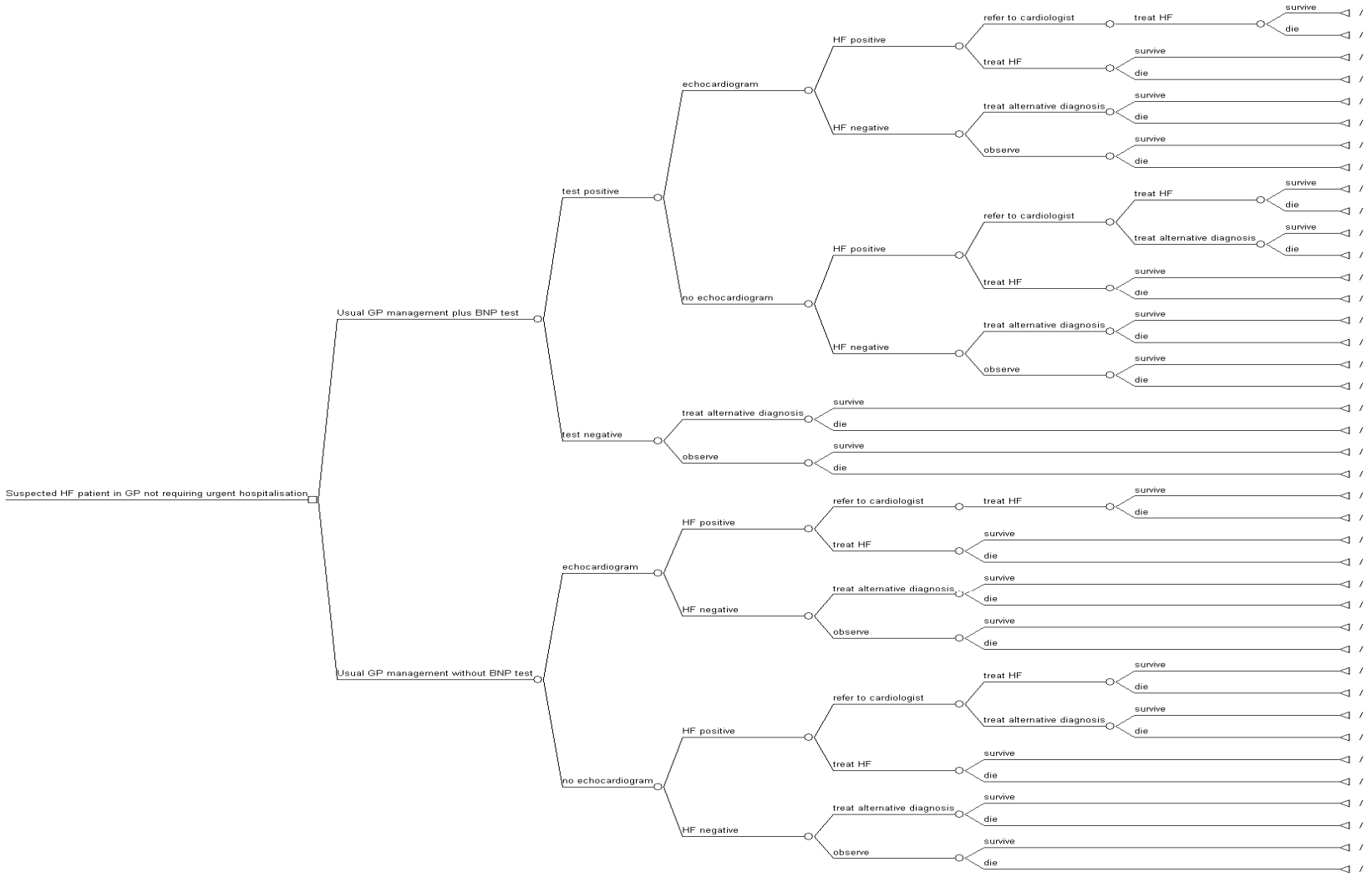
report are based on patients presenting with a level of symptoms reflecting these three key studies. It is unclear whether the three studies are likely to be representative of primary care in Australia, and whether such tight selection of symptomatic and suspected HF patients would be followed in routine general practice.

The structure of the decision analytic model needs to be changed to reflect this situation (see Figure 14). In particular, the population of patients in the model is now not only those who present with actual dyspnoea and/or oedema of recent onset (symptomatic), but also contains those merely suspected to be in HF. There are options for the GP to observe rather than treat patients who are not positively diagnosed as having HF, that is those patients who may be diagnosed as not having any clinically important pathology.

This decision tree thus illustrates how, in testing patient populations with minor levels of symptoms, diminishing marginal returns may emerge. As there is a possibility of a broader spread of use than strictly 'dyspnoea and/or oedema of recent onset in the diagnosis of HF', it is anticipated that, should B-type natriuretic peptide testing be publicly funded under Medicare, the restriction for use will need to be carefully audited.

In conclusion, it has been determined that diminishing marginal returns could arise should the testing extend to general practice populations where: (1) there is a high probability that the patient has HF or (2) there are increasing proportions of patients with minor levels of symptoms but lacking any clinically important pathology. Basically, the incremental cost of testing this latter much larger population would not be counterbalanced by the increment in health gains from testing.

Figure 14 Decision tree for clinical diagnosis strategy with and without B-type natriuretic peptide testing in the non-hospital setting for patients with suspected (but not necessarily symptomatic) heart failure



Financial incidence analysis

The financial impact on the Australian Government of a new pathology test depends upon the size of the population to whom this test applies and the net expenditure generated by the new test (ie the expenditure on the new test minus potential downstream savings).

The evidence contained in this report indicates that the most useful role of the B-type natriuretic peptide test in primary care is to assist GPs when they are uncertain as to whether patients presenting with dyspnoea and/or oedema recent onset have suspected HF or alternative (probably respiratory) pathology. From a practical point of view the Advisory Panel believes that it would be possible to restrict the payment of Medicare Benefits for this test to those patients presenting with these indications. With this restriction in place, two estimates of the likely cost to Medicare can be made using the following data sources: Medicare Benefits Schedule reclaim data, the BEACH study on Australian General Practice Activity (AIHW 2005), and Wright et al (2003).

Estimation of expenditure based on presentation of dyspnoea and/or oedema of recent onset and suspected of heart failure

Clinical need for the B-type natriuretic peptide tests was estimated in the 'Clinical need' section of this report. This detail is reproduced below and costs have been applied in order to estimate the financial implications to the Commonwealth of introducing B-type natriuretic peptide testing.

In the 2005 calendar year, 88,680,935 GP attendances were claimed overall. The BEACH study on Australian General Practice Activity (AIHW 2005) reports that between April 2004 and March 2005 'shortness of breath, dyspnoea' was responsible for 779 (0.83%; 95%CI 0.6%, 1.0%) problems of the total 94,386 GP problems assessed in the BEACH cohort. As BEACH data suggest that 54.5 per cent of GP attendances were for patients aged 45 years and older, approximately 398,893 GP attendances in patients aged 45+ years were for symptoms of dyspnoea.

A robust estimate could not be identified for the number of patients presenting to GPs with oedema of recent onset. However, given that there is a likely overestimate of BEACH dyspnoea cases as being suspected of HF (ie the dyspnoea category would include cases of known asthma), it is likely that this overestimate would in some part correct for the missing data on cases of oedema of recent onset.

Only patients with *new* symptoms of HF are likely to receive a B-type natriuretic peptide test because, for those patients with uncontrolled previously diagnosed HF, management and monitoring is required rather than a diagnosis. New symptoms of possible HF were reported in 24.6 per cent of patients in the CASE study (Krum et al 2001). This estimate is not ideal as it relates to patients aged 60 years and older and includes a much wider definition of suspected HF than just dyspnoea and/or oedema of recent onset. However, the study sample included patients who did or did not have a prior diagnosis of HF similar to the target population in Australian general practice. Wright et al (2003) reported that, of patients aged 40 years and over *without* a previous history of HF and presenting to GPs with symptoms of dyspnoea and/or oedema of recent onset, 70 per cent were suspected of HF. These two rates were applied to indicate the likely range in clinical need for the B-type natriuretic peptide tests in patients aged 45 years and older

with new symptoms of dyspnoea and/or oedema. Approximately 97,956 and 279,225 patients could require a B-type natriuretic peptide test according to the data provided by Krum et al (2001) and Wright et al (2003), respectively. It is unclear at what rate patients presenting in the non-hospital setting with *new* symptoms of suspected HF are referred to hospital or cardiologists/physicians. However, BEACH data (AIHW 2001) indicate that 19.7 per cent of patients with confirmed *or* suspected HF are referred to hospital or a cardiologist/physician. Assuming these patients do *not* receive a B-type natriuretic peptide test prior to referral, then between 78,658 and 224,218 patients aged 45 years and over, with dyspnoea of recent onset and suspected of HF, would be eligible to receive a B-type natriuretic peptide test annually.

The benchmark price of a B-type natriuretic peptide test is $50.59. Thus, the additional Australian Government expenditure due to the introduction of B-type natriuretic peptide assays into the non-hospital setting for patients presenting with dyspnoea and/or oedema of recent onset is estimated to be between $3,979,323 and $11,343,186 per year. This expenditure is likely to be offset by savings on fewer echocardiograms and cardiologist referrals and earlier management of non-HF diagnoses, but the extent of these offsets is presently unknown.

The most important assumption is that only suspected *symptomatic* HF patients (eg those presenting with dyspnoea and/or oedema of recent onset) will be tested. There is the potential for patients presenting with minor symptoms or indications to be tested; should this become widespread practice, there is a concern that the expenditure implications could be large. In practice, however, the Australian Government can employ several strategies to ensure that a test is ordered appropriately, including: (1) facilitating intensive education programs for health professionals (particularly important given that this test is most effective when there is considerable uncertainty regarding the differential diagnosis of a patient and the test is then used to 'rule out' HF); (2) restricting payment of benefits for the item to specific indications; and (3) limiting augmentation of the 'pathology cap' to a level that is consistent with the restricted use of the item.

The initial impact of a 'leakage' scenario on Government expenditure may in fact have been absorbed to some extent in the above financial estimates (~$4.0–11.3 million per year) through the use of the two studies in Australian and New Zealand general practice (Krum et al 2001; Wright et al 2003) to estimate the prevalence of suspected HF. It is likely that both studies overestimate the prevalence of suspected HF as the participating GPs in these studies would have been sensitive to the study aims and been more primed to active identification and assessment of suspected HF than is usual in normal general practice.

The use of B-type natriuretic peptide testing in specialist settings has been ignored because there are no itemised MBS data for cardiac specialists to estimate the potential annual reclaim for symptomatic suspected HF patients. Furthermore, it is unclear whether these specialists, who are already experts in detecting cases of HF, will use B-type natriuretic peptide testing as frequently as GPs. The effect of adding specialist reclaim for B-type natriuretic peptide testing will increase the overall expenditure estimate, but this is difficult to quantify.

Conclusions

Safety

The likelihood of adverse events as a consequence of the B-type natriuretic peptide testing procedure is low and similar to that of any blood test. Psychological or physical harms are a possibility due to the inevitability of false positive results being associated with the tests, and sequelae such as inappropriate treatment. The false negative rate from B-type natriuretic peptide testing is very low and, as the population is not acutely ill, delayed treatment associated with a false negative test result is unlikely to be harmful. None of the included studies in this report, however, reported on patient physical or psychological harms as a consequence of B-type natriuretic peptide testing in the non-hospital setting.

Diagnostic effectiveness

BNP testing

The effectiveness of supplementing conventional diagnostic assessment in the non-hospital setting with BNP testing was evaluated by a small volume of evidence, with the most reliable evidence obtained from one high quality level II diagnostic accuracy study. There was no direct evidence available evaluating the effect of the BNP test on patient health outcomes in the non-hospital setting. A linked evidence approach was therefore attempted. Evidence concerning the diagnostic accuracy of the test was to be linked through a narrative to evidence of the impact of the test on patient management, and this then linked to the effect of this type of patient management on patient health outcomes.

Only five studies met the criteria for inclusion in the review and all provided evidence on the diagnostic accuracy of the test. Overall, this body of evidence was relatively consistent in its findings that the BNP blood test is sensitive with a high negative predictive value and variable specificity. Its main role, therefore, appears to be as a 'first line' test as a negative result on the test 'rules out' the diagnosis of heart failure (HF), so that differential diagnoses for symptomatic patients can be investigated.

The impact of the introduction of a BNP test on the management of patients by GPs could not be directly determined. There were no change-in-management studies associated with the BNP test in the non-hospital setting that were available for inclusion in this review at the time of searching.

There were no studies available to assess the direct impact of BNP testing on patient health outcomes. A systematic review of the impact of early treatment on health outcomes for patients with and without HF was beyond the scope of this report. High level evidence (see Introduction) suggests that early treatment for HF is beneficial for the patient. It is unlikely, however, that use of the BNP test would result in an *earlier* identification of HF than currently for these symptomatic suspected HF patients. Patients still require an echocardiogram and/or a cardiology consultation to receive a definitive diagnosis of HF. Use of the test *is* likely to assist in earlier identification of alternative diagnoses for those patients 'ruled out' from HF—and most of these

alternative diagnoses (pulmonary diseases, asthma, anaemia) have established treatments. Should this alternative pathology be severe enough, early identification and treatment of the condition is likely to be beneficial to the patient.

The populations studied in the included diagnostic studies are reasonably applicable to the target population in Australia, that is patients presenting in the non-hospital setting with symptoms (eg dyspnoea) suggestive of HF. As a group, however, they may have had slightly more severe symptoms than is usual in general practice, as most were selected on the basis of *referral* from a GP on suspicion of HF. The results of the studies are largely generalisable to the Australian healthcare context, with most being conducted in developed countries with similar standards of practice in diagnosing and managing symptomatic suspected HF.

In conclusion, on the basis of the evidence presented, BNP testing appears to be a valuable 'first line' diagnostic test that, when added to conventional diagnostic assessment, assists the GP to determine which HF-like symptoms exhibited by patients are *not* caused by HF. The clinical impact of the test on patient health is, however, currently unknown in the non-hospital setting.

NT-proBNP testing

The effectiveness of NT-proBNP testing added to conventional diagnostic assessment was evaluated by a small volume of evidence (5 studies), with the best evidence being obtained from one high quality level II intervention study and one high quality level II diagnostic accuracy study. Overall, the diagnostic accuracy evidence was relatively consistent in its findings that NT-proBNP assays are sensitive with, in general, high negative predictive values (>90%), indicating that the test effectively 'rules out' HF in patients with a negative result. The specificity of the test is low and variable.

The impact of the test was obvious in terms of patient management by the GP. Good quality level II intervention evidence (randomised controlled trial) reported a clinically important improvement in the proportion of correct diagnoses of HF in those GPs receiving NT-proBNP test results. A 13 per cent improvement [95%CI 5.5, 21.0, p=0.002] was observed in correct diagnoses in the NT-proBNP trial arm compared to the control trial arm, with the main impact occurring by enabling GPs to correctly 'rule out' HF.

There were no studies available to assess the direct impact of NT-proBNP testing on patient health outcomes. A systematic review of the impact of early treatment on health outcomes for patients with and without HF was beyond the scope of this report. High level evidence (presented in the Introduction) suggests that early treatment for HF is beneficial for the patient. It is, however, unlikely that use of the NT-proBNP test would result in an earlier identification of HF than currently for these patients due to the low positive predictive value of the test, and thus the need for further confirmatory testing. Use of the test would, however, identify earlier alternative diagnoses—most of which have established treatments—for those patients 'ruled out' from HF. Should this alternative pathology be severe enough, early identification and treatment is likely to benefit the health of the patient.

The populations studied in the available evidence base are similar to the target population in Australia, that is patients presenting in the non-hospital setting with symptoms—

primarily dyspnoea and/or oedema of recent onset—suggestive of HF. The results of the studies are generalisable to the Australian healthcare context, with most being conducted in developed countries (including New Zealand) with similar standards of practice in diagnosing and managing suspected HF.

In conclusion, on the basis of the evidence presented, NT-proBNP assays appear to be sensitive 'first line' diagnostic tests that, when added to conventional diagnostic assessment, contribute to a change in the clinical practice of GPs. Their main role is to assist the GP to correctly 'rule out' HF more frequently in those patients presenting with dyspnoea and/or oedema of recent onset and suspected of HF. The clinical impact of the test (ie on patient health) is, however, currently unknown in the non-hospital setting.

Economic considerations

The extent of the effectiveness of B-type natriuretic peptide assays in the non-hospital setting (ie in terms of life-years or quality-adjusted life-years (QALYs) gained) depends on the extent to which they hasten the establishment of the correct definitive diagnosis, as well as on the influence of a correct diagnosis on the outcome of HF. In turn, the extent of their *cost-effectiveness* hinges on the value of the additional information made available by the tests in terms of health gain and resource savings compared to the cost of providing the tests.

This economic analysis has been primarily limited to a population of symptomatic patients (in line with the evidence base), that is those presenting with dyspnoea and/or oedema of recent onset and suspected of HF. Appraisal of the economic implications of using B-type natriuretic peptide assays in the non-hospital setting was hampered by the absence of any randomised controlled trial in that setting with a *health outcome* as an endpoint. Thus, it has not been possible to estimate an incremental cost-effectiveness ratio based on life-years saved or QALYs saved in the non-hospital setting.

There were insufficient data to populate a decision analytic model to assess the cost-effectiveness of these assays in the non-hospital setting. Therefore, this economic analysis is based on identifying those circumstances which might substantially reduce the cost-effectiveness of a B-type natriuretic peptide test.

It was determined that the costs and outcomes associated with B-type natriuretic peptide testing will usually depend on the GP's referral propensity. The extent of immediate cost offsets depends on whether or not the GP decides to order an echocardiogram and initially undertake self-management of the patient; or to refer the test positive patient to a cardiologist who may or may not order an echocardiogram. The extent of downstream costs (or savings) also depends on the same referral propensity, and on the proportions of patients correctly diagnosed.

Three scenarios were presented illustrating different types of possible GP referral patterns that have increasing levels of resource use. In all three scenarios the use of B-type natriuretic peptide testing is cost saving (from $50 to $86 per patient tested), primarily through savings on echocardiogram and cardiologist referrals in test negative patients. These scenarios reflect current clinical practice guideline recommendations and assume that in Australia *all* patients (over 45 years of age) presenting with dyspnoea and/or oedema of recent onset, and suspected of HF, would receive an echocardiogram. The data available, however, suggests that actual echocardiogram referral may range from

3.8 per cent to 17.7 per cent for patients with new symptoms suggestive of HF. Unfortunately there are no robust data available on the likely referral patterns of GPs presented with patients reporting dyspnoea and/or oedema of recent onset and suspected of HF. It is also probable that referral patterns will vary widely between GPs, due in part to accessibility to echocardiograms and cardiologists. Results of a one-way sensitivity analysis suggest that B-type natriuretic peptide testing may not be cost saving if GPs currently refer this patient group to echocardiography at a rate of 60 per cent or lower.

Despite it being shown that B-type natriuretic peptide tests reduce the number of 'ruled out' patients going on to an echocardiogram, it is unknown what impact the introduction of these tests will have on the current echocardiogram referral rate for those who test positive.

Given this, it is still likely that when testing is confined to general practice patients with dyspnoea and/or oedema of recent onset and suspected of HF, most of those patients who test negative on the B-type natriuretic peptide assay will have a clinically important pathology. In this situation the potential for improvement in health outcome may warrant the amount of resources used in testing and follow-up.

It has been determined that diminishing marginal returns could arise should the testing extend to general practice populations where: (1) there is a high likelihood that the patient has HF or (2) there are increasing proportions of patients with minor levels of symptoms but lacking any clinically important pathology. It is therefore of critical importance that B-type natriuretic peptide testing is only ordered for those patients with dyspnoea and/or oedema of recent onset, for whom there is real uncertainty as to whether the symptoms are caused by HF or an alternative pathology. Finally, the results of the key trial in the ED setting (see Part A of this report), perhaps with some adjustment for the severity of clinical presentation, are relevant for the rural and remote setting where the GP decides to admit the patient to the local hospital and continue inpatient management.

The additional Australian Government expenditure due to the introduction of B-type natriuretic peptide assays into the non-hospital setting for patients presenting with dyspnoea and/or oedema of recent onset is estimated to range between $4.0 million and $11.3 million annually. This expenditure is likely to be offset by savings on fewer echocardiograms and cardiologist referrals and earlier management of alternative (non-HF) diagnoses, but the extent of these offsets cannot currently be determined.

Recommendation

'MSAC has considered the safety, effectiveness and cost effectiveness of the use of assays of B-type natriuretic peptides in the diagnosis of heart failure in patients presenting with dyspnoea in the non-hospital setting when compared with current clinical practice +/- echocardiography.

MSAC finds that there is sufficient evidence that B-type natriuretic peptide assays, when used in the diagnosis of heart failure in patients presenting with dyspnoea, are safe and effective (diagnostically accurate).

MSAC finds that there is major uncertainty around the cost effectiveness in the non - hospital setting.

MSAC recommends that public funding is <u>not supported</u> for the use of assays of B-type natriuretic peptides in the diagnosis of heart failure in patients presenting with dyspnoea in the non-hospital setting at this time.

MSAC further recommends research on the use of BNP in the general practice setting to identify appropriate usage and the patient group most likely to benefit in the non hospital setting.'

The Minister for Health and Ageing accepted this recommendation on 29 August 2007.

Appendix L Inclusion criteria

Box 10 Study selection criteria for assessing diagnostic effectiveness in the non-hospital setting (direct evidence approach)

Selection criteria	Inclusion criteria
Population	Symptomatic patients with suspected heart failure in the non-hospital setting
Intervention	NT-proBNP or BNP diagnostic assays as 'first line' tests in conjunction with standard clinical assessment[a] ± echocardiography[b]
Comparator	Standard clinical assessment ± echocardiography
Outcomes	Health outcomes: Primary: rate of survival/death, symptom resolution (dyspnoea, oedema), quality of life, functional status Secondary: hospital / intensive care unit length of stay, discharge diagnosis, rates of echocardiogram usage
Study design	Randomised or non-randomised controlled trials or cohort studies or case-control studies or systematic reviews of these study designs.
Search period	Because BNP was first described in the literature in 1988, the search period was restricted to 1988 – 08/2005
Language	Studies in languages other than English were only translated and included if they represented a higher level of evidence than that available in the English language evidence base

[a] Clinical assessment of signs, symptoms, laboratory tests, chest X-rays, ECGs; [b] Echocardiogram is likely to be used in the diagnostic pathway if, on the basis of the 'first line' tests (eg BNP, NT-proBNP, physical examination, chest X-ray, laboratory tests, ECG), the patient is still suspected of heart failure. Those patients 'ruled out' for heart failure on the basis of these 'first line' tests, however, will not receive an echocardiogram.

Appendix M Studies included in this review

Diagnosis in the the non-hospital setting

Studies on BNP assays

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Cowie et al 1997)	Prospective cohort – cross-classified	Rapid access clinic (primary care) Hillingdon district, west London, UK	106 [ITT: 122]	Suspected cases of new HF referred by GPs	Age: 24–87 yrs M/F: 59/63 HF: 29/106 (27%) Excluded if HF hx	Consensus clinical diagnosis Three cardiologists using all data[b] to determine if met case definition recommended by European Society of Cardiology	BNP Radio-immunoassay Peninsula Laboratories		CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value
(Hobbs et al 2004)	Prospective cohort – cross-classified	Primary care Four practices, England, UK [substudy of ECHOES study]	103	Patients with suspected HF (but unvalidated) randomly sampled from GP practices 607 in total but only 103 with suspected HF	Age: n/a for subgroup M/F: n/a for subgroup HF: n/a HF (caused by LVSD): 21/103 = 20% Unclear if HF hx	Consensus clinical diagnosis Study investigators (+ 3 cardio-vascular clinicians in equivocal cases) using all data including	BNP Immuno-radiometric assay (IRMA) Shionogi, Japan		CX P1 Level III-2 diagnostic evidence Q3 [Quadas = 9/14]	Sensitivity Specificity Negative predictive value *As it relates to HF caused by LVSD, not HF overall*

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
						ECG, echo-cardiogram, clinical exam, spirometry. Used HF criteria developed by European Society of Cardiology				
(Landray et al 2000)	Prospective cohort – cross-classified	Primary care Clinic, Banbury, Oxford, UK	126	Patients referred by GPs with suspected HF	Age: 74±9 yrs M/F: 68/58 HF (caused by LVSD): (32%) Unclear if HF hx		BNP Immuno-radiometric assay Shionoria assay, Shionogi	Echo-cardiography	CX P1 Level III-2 diagnostic evidence Q3 [Quadas = 8/14]	Sensitivity Specificity
(Sim et al 2003)	Prospective cohort – cross-classified	Open-access echo-cardiography service Newport, South Wales, UK	83	Patients with symptoms of dyspnoea referred by GPs Excluded: patients with both dyspnoea and heart murmur	Age: 72 (37–87) yrs M/F: 40/43 HF (caused by LVSD): 26/83 (31%) Unclear if HF hx		BNP Radioimmuno-assay (in-house) Bachem Ltd	Echo-cardiography – (two independent echo-cardiographers)	CX P1 Level III-2 diagnostic evidence Q3 [Quadas = 10/14]	Sensitivity Specificity Negative predictive value False positive rate False negative rate
(Zaphiriou et al 2005)	Prospective cohort – cross-classified	Rapid access HF clinic	306	Consecutive patients with new	Age: median 74 (52–87) yrs	Clinical diagnosis –	BNP		CX P1	Sensitivity Specificity

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
		Aberdeen Royal Infirmary; Western General and Royal Infirmary, Glasgow; Whittington Hospital, London; Charing Cross Hospital, London, UK		symptoms suggestive of HF, referred by their GPs to a rapid access HF clinic Excluded: patients with previous documented hx of HF	M/F: 130/176 HF: 104/306 (34%) Patients with HF hx excluded	A cardiologist defined HF if at least one symptom present (shortness of breath, fatigue, leg oedema) at rest or on exertion and objective evidence of cardiac dysfunction at rest (including echo-cardiography), as recommended by the European Society of Cardiology.	Immuno-fluorescence assay Biosite Diagnostics		Level II diagnostic evidence Q1 [Quadas = 12/14]	Negative predictive value

M/F = male/female; hx = history; HF = heart failure; LVSD = left ventricular systolic dysfunction; n/a = not available; ECG = electrocardiogram; [a] For an explanation of the appraisal system used in this assessment, refer to section on 'Strength of the evidence in individual studies' in chapter on 'Approach to assessment'; [b] Excluding BNP data.

Studies on NT-proBNP assays

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
(Gustafsson et al 2003)	Prospective cohort – cross-classified	Primary care Copenhagen General Practitioners' Laboratory, Copenhagen, Denmark	367	Patients referred by GPs for echo-cardiography to confirm / rule out suspected HF	Age: 69 (39–84) yrs M/F: 169/198 HF (caused by LVSD): 33/367 (9%) Unclear if HF hx		NT-proBNP Chemi-luminescent sandwich immuno-assay Elecsys, Roche Diagnostics	Echo-cardiography Full echo-cardiographic and Doppler echo-cardiographic study; LVEF level indicating severity of LVSD causing HF	CX P1 Level III-1 diagnostic evidence Q2 [Quadas = 11/14]	Sensitivity Specificity Negative predictive value *HF caused by LVSD*
(Hobbs et al 2002) (Hobbs et al 2004)	Prospective cohort – cross-classified	Primary care Four practices, England, UK [substudy of ECHOES study]	103	Patients with suspected HF (but unvalidated) randomly sampled from GP practices 607 in total but only 103 with suspected HF	Age: n/a for subgroup M/F: n/a for subgroup HF: n/a HF (caused by LVSD): 21/103 = 20% Unclear if HF hx	Consensus clinical diagnosis Study investigators (+ 3 cardio-vascular clinicians in equivocal cases) using all data including ECG, echo-cardiogram, clinical exam, spirometry. Used HF	NT-proBNP Enzyme linked immuno-sorbent assay (ELISA) Roche Diagnostics		CX P1 Level III-2 diagnostic evidence Q3 [Quadas = 9/14]	Sensitivity Specificity Negative predictive value *As it relates to HF, not HF caused by LVSD*

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a] Appropriate comparison Applicable population Study quality	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)					
						criteria developed by European Society of Cardiology				
(Nielsen et al 2004)	Post-test case series	Primary care Hospital-based clinic, Denmark	345	Consecutive patients referred by 74 GPs with dyspnoea symptoms Excluded: missing blood sample or other diagnostic data.	Age: median 65 (1889) yrs M/F: 51%/49% HF: 81/345 (24%) Patients with HF hx included		NT-proBNP Chemi-luminescent sandwich immuno-assay Roche Diagnostics	Echo-cardiogram	C1 P1 Level IV intervention evidence Q3	*(potential)* Change in management
(Wright et al 2003)	Randomised, single-blind controlled trial	Primary care Participating general practices Auckland, New Zealand	305 [ITT: 307] BNP arm: n=153 Control arm: n=152	Patients presenting to GP with symptoms of dyspnoea and/or oedema of recent onset Excluded: patients <40 years; requiring urgent hospital admission; unable to	BNP arm: Age: 69±11 yrs M/F: 52/101 HF: 43/153 (28%) Patients with HF hx not included Control arm: Age: 72±11 yrs M/F: 55/97 HF: 34/152 (22%) Patients with HF	Consensus clinical diagnosis Three cardiologists and one GP using all clinical data[b], including ECG, CXR, echocardiogram. Applied criteria of European Society of Cardiology Working Group re HF definition. If	Clinical diagnosis (patient hx, clinical examination, ECG, CXR, echo-cardiography, blood tests) + NT-proBNP NT-proBNP - Radio-immunoassay	Clinical diagnosis (patient hx, clinical examination, ECG, CXR, echo-cardiography, blood tests)	C1 P1 Level II intervention evidence Q1 [Downs & Black = 21/27 + adequately powered for primary outcome (increase in diagnostic	Change in management

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a] Appropriate comparison Applicable population Study quality	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)					
				provide informed consent; hospital admission with diagnosis of HF; echo-cardiography for assessment of LVD; outpatient cardiology assessment for dyspnoea or oedema in previous 12 months	hx not included	doubt remained, beneficial response to Rx was considered	(In-house)		accuracy)]	
(Zaphiriou et al 2005)	Prospective cohort – cross-classified	Rapid access HF clinic Aberdeen Royal Infirmary; Western General and Royal Infirmary, Glasgow; Whittington Hospital, London; Charing Cross Hospital, London, UK	306	Consecutive patients with new symptoms suggestive of HF, referred by their GPs to a rapid access HF clinic Excluded: patients with previous documented hx of HF	Age: median 74 (52–87) yrs M/F: 130/176 HF: 104/306 (34%) Patients with HF hx excluded	Clinical diagnosis A cardiologist defined HF if at least one symptom present (shortness of breath, fatigue, leg oedema) at rest or on exertion and objective evidence of cardiac dysfunction at rest (including	NT-proBNP ELISA assay Elecsys system, Roche Diagnostics		CX P1 Level II diagnostic evidence Q1 [Quadas = 12/14]	Sensitivity Specificity Negative predictive value

Study Authors (Year)	Study design	Setting Region, site	Study population			Reference standard	Index test	Comparator	Appraisal[a]	Outcomes
			N	Selection criteria	Characteristics (eg age, gender, disease prevalence)				Appropriate comparison Applicable population Study quality	
						echo-cardiography), as recommended by the European Society of Cardiology.				

M/F = male/female; hx = history; HF = heart failure; ECG = electrocardiography; CXR = chest X-ray; LVEF = left ventricular ejection fraction; LVSD = left ventricular systolic dysfunction; n/a = not available; Rx = treatment; GP = general practitioner; [a] For an explanation of the appraisal system used in this assessment, refer to section on 'Strength of the evidence in individual studies' in 'Approach to assessment' chapter; [b] Excluding NT-proBNP data.

Appendix N Excluded studies

Diagnosis

Unable to extract relevant data

Alehagen, U., Lindstedt, G. et al (2003). 'Utility of the amino-terminal fragment of pro-brain natriuretic peptide in plasma for the evaluation of cardiac dysfunction in elderly patients in primary health care', *Clinical Chemistry*, 49 (8), 1337–1346.

Bettencourt, P., Ferreira, A. et al (2000). 'Evaluation of brain natriuretic peptide in the diagnosis of heart failure', *Cardiology*, 93 (1–2), 19–25.

Reasons for referral for echocardiography included suspected HF and other reasons (therefore unclear if LVD rates relate solely to HF)

Atisha, D., Bhalla, M. A. et al (2004). 'A prospective study in search of an optimal B-natriuretic peptide level to screen patients for cardiac dysfunction', *American Heart Journal*, 148 (3), 518–523.

Bhalla, V., Isakson, S. et al (2005). 'Diagnostic ability of B-type natriuretic peptide and impedance cardiography: Testing to identify left ventricular dysfunction in hypertensive patients', *American Journal of Hypertension*, 18 (2), 73S–81S.

Epshteyn, V., Morrison, K. et al (2003). 'Utility of B-type natriuretic peptide (BNP) as a screen for left ventricular dysfunction in patients with diabetes', *Diabetes Care*, 26 (7), 2081–2087.

Hutcheon, S.D., Gillespie, N.D. et al (2002). 'B-type natriuretic peptide in the diagnosis of cardiac disease in elderly day hospital patients', *Age Ageing*, 31 (4), 295–301.

Lubien, E., DeMaria, A. et al (2002). 'Utility of B-natriuretic peptide in detecting diastolic dysfunction - Comparison with Doppler velocity recordings', *Circulation*, 105 (5), 595–601.

Maisel, A.S., Koon, J. et al (2001). 'Utility of B-natriuretic peptide as a rapid, point-of-care test for screening patients undergoing echocardiography to determine left ventricular dysfunction', *American Heart Journal*, 141 (3), 367–374.

McLean, A.S., Tang, B. et al (2003). 'Increased B-type natriuretic peptide (BNP) level is a strong predictor for cardiac dysfunction in intensive care unit patients', *Anaesthesia and Intensive Care*, 31 (1), 21–27.

Mueller, T., Gegenhuber, A. et al (2004). 'Head-to-head comparison of the diagnostic utility of BNP and NT-proBNP in symptomatic and asymptomatic structural heart disease', *Clinica Chimica Acta*, 341 (1–2), 41–48.

Talwar, S., Squire, I.B. et al (1999). 'Plasma N-terminal pro-brain natriuretic peptide and the ECG in the assessment of left-ventricular systolic dysfunction in a high risk population', *European Heart Journal*, 20 (23), 1736–1744.

Vourvouri, E.C., Schinkel, A.F. et al (2003). 'Screening for left ventricular dysfunction using a hand-carried cardiac ultrasound device', *European Journal of Heart Failure*, 5 (6), 767–774.

Yamamoto, K., Burnett, J.C., Jr. et al (1996). 'Superiority of brain natriuretic peptide as a hormonal marker of ventricular systolic and diastolic dysfunction and ventricular hypertrophy', *Hypertension*, 28 (6), 988–994.

Yamamoto, K., Burnett J.C, Jr. et al (2000). 'Clinical criteria and biochemical markers for the detection of systolic dysfunction', *Journal of Cardiac Failure*, 6 (3), 194–200.

Data included in another paper

Knudsen, C.W., Riis, J.S. et al (2004). 'Diagnostic value of a rapid test for B-type natriuretic peptide in patients presenting with acute dyspnoea: effect of age and gender', *European Journal of Heart Failure*, 6 (1), 55–62.

Maisel, A.S., Clopton, P. et al (2004). 'Impact of age, race, and sex on the ability of B-type natriuretic peptide to aid in the emergency diagnosis of heart failure: results from the Breathing Not Properly (BNP) multinational study', *American Heart Journal*, 147 (6), 1078–1084.

Maisel, A.S., McCord, J. et al (2003). 'Bedside B-type natriuretic peptide in the emergency diagnosis of heart failure with reduced or preserved ejection fraction - Results from the breathing not properly multinational study', *Journal of the American College of Cardiology*, 41 (11), 2010–2017.

McCullough, P.A., Hollander, J.E. et al (2003). 'Uncovering heart failure in patients with a history of pulmonary disease: rationale for the early use of B-type natriuretic peptide in the emergency department', *Academy of Emergency Medicine*, 10 (3), 198–204.

Morrison, L.K., Harrison, A. et al (2002). 'Utility of a rapid B-natriuretic peptide assay in differentiating congestive heart failure from lung disease in patients presenting with dyspnea', *Journal of the American College of Cardiology*, 39 (2), 202–209.

Steg, P.G., Joubin, L. et al (2005). 'B-type natriuretic peptide and echocardiographic determination of ejection fraction in the diagnosis of congestive heart failure in patients with acute dyspnea', *Chest*, 128 (1), 21–29.

Wu, A.H.B., Omland, T. et al (2004). 'The effect of diabetes on B-type natriuretic peptide concentrations in patients with acute dyspnea - An analysis from the breathing not properly multinational study', *Diabetes Care*, 27 (10), 2398–2404.

Abstracts – where a higher level of evidence was available in full text

Hobbs, F.D.R., Davis, R.C. et al. (2000). 'Plasma N-terminal pro-brain natriuretic peptide has similar predictive value to brain natriuretic peptide in diagnosis of heart failure in the community', *European Heart Journal,* 21, 133.

Hobbs, F.D.R., Davis, R.C. et al. (2001). 'Performance characteristics of N terminal pro brain natriuretic peptide (NT-ProBNP) and BNP assays in the diagnosis of heart failure in community settings', *Journal of the American College of Cardiology,* 37(2),147A.

Inomata, T., Nakano, H. et al. (2004). 'Exceptions of the diagnostic value for heart failure of plasma brain-type natriuretic peptide levels/disproportion between brain natriuretic peptide and norepinephrine levels implies constrictive pericarditis', *European Heart Journal,* 25, 627.

Jourdain, R., Logeart, D. et al. (2003). 'Brain natriuretic peptide usefulness for diagnosing heart failure in elderly patients (BUD study): about 300 patients over 75', *European Heart Journal,* 24, 116.

Luchner, A., Hengstenberg, C. et al. (2002). 'Automated measurement of N-terminal proBNP for biochemical detection of left ventricular dysfunction', *European Heart Journal,* 23, 572.

Luong, M.V., Auziere, L. et al. (2003). 'Echocardiographic determinants of brain natriuretic peptide levels in patients suspected of heart failure', *European Heart Journal,* 24, 412.

Maisel, A., Harrison, A. et al. (2001). 'B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnoea', *European Heart Journal,* 22, 377.

Passino, C., Bramanti, F. et al. (2003). 'Eighteen-minute N terminal pro-brain natriuretic peptide versus 24-hour brain natriuretic peptide assay as diagnostic markers in congestive heart failure', *European Heart Journal,* 24, 360.

Triepels, R.H., Busscher, S. et al. (2003). 'N-terminal pro-brain natriuretic peptide (NT-proBNP) as screening test for early stage heart failure', *Clinical Chemistry,* 49(6), A37–A38.

Trochu, J. N., Pattier, S. et al. (2003). 'Accuracy of N terminal pro-brain natriuretic peptide for the diagnosis of congestive heart failure in patients admitted with dyspnoea in the emergency department', *European Heart Journal,* 24, 260.

Vinereanu, D., Lim, P. et al. (2003). 'Echocardiographic screening for heart failure must analyse left-ventricular longitudinal function. Comparison versus brain natriuretic peptide.' *European Heart Journal,* 24, 351.

Wieczorek, S.J., Ferrier, A. et al. (2000). 'B-type natriuretic peptide for the evaluation of congestive heart failure as determined from the New York Heart Association Classification system and the six minute walk test', *American Journal of Clinical Pathology,* 114(2), 312.

Zaninotto, M., Mion, M. et al. (2003). 'N-terminal pro-brain natriuretic peptide in the differential diagnosis of dyspnea in an emergency department', *Clinical Chemistry,* 49(6), A41.

Zaphiriou, A., Robb, S. et al. (2003). 'Using brain natriuretic peptide and N terminal pro-brain natriuretic peptide to rule out heart failure: does it work in clinical practice? Results of the UK natriuretic peptide study', *European Heart Journal,* 24, 260.

Hospital setting

See studies included in Part A of this Assessment Report (Appendix E).

References

AIHW (2001). *A general practice view of cardiovascular disease and diabetes in Australia.* AIHW cat. no. CVD 17 (Cardiovascular Disease Series No. 18), Canberra, AIHW.

AIHW (2005). *General practice activity in Australia 2004–05*. Australian Institute of Health and Welfare (General Practice Series No. 18), AIHW Cat. No. GEP 18. Canberra.

American Heart Association (2005). *Congestive heart failure* [Internet]. Available from: http://www.americanheart.org/presenter.jhtml?identifier=4585 [Accessed 14 April].

Armitage, P., Berry, G. & Matthews, J.N.S. (2002). *Statistical methods in medical research,* Oxford, Blackwell Science.

Aurbach, A., Russ, W., Battegay, E., Bucher, H.C., Brecht, J.G., Schadlich, P.K. & Sendi, P. (2004). 'Cost-effectiveness of ramipril in patients at high risk for cardiovascular events: a Swiss perspective', *Swiss Medical Weekly,* 134, 399–405.

Cowie, M.R., Struthers, A.D., Wood, D.A., Coats, A.J., Thompson, S.G., Poole-Wilson, P.A. & Sutton, G.C. (1997). 'Value of natriuretic peptides in assessment of patients with possible new heart failure in primary care', *Lancet,* 350, 1349–1353.

De Denus, S., Pharand, C. & Williamson, D.R. (2004). 'Brain natriuretic peptide in the management of heart failure: the versatile neurohormone', *Chest,* 125, 652–668.

Deeks, J.J. (2001). 'Systematic reviews of evaluations of diagnostic and screening tests', *British Medical Journal,* 323, 157–162.

Dhir, M., Arora, U. & Nagueh, S.F. (2004). 'The role of echocardiography in the diagnosis and prognosis of patients with heart failure', *Expert Reviews in Cardiovascular Therapy,* 2, 141–144.

Downs, S.H. & Black, N. (1998). 'The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions', *Journal of Epidemiology and Community Health,* 52, 377–384.

Egger, M., Juni, P., Bartlett, C., Holenstein, F. & Sterne, J. (2003). 'How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study', *Health Technology Assessment,* 7, 1–76.

Garg, R. & Yusuf, S. (1995). Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. *Journal of the American Medical Association,* 273, 1450–1456.

Gustafsson, F., Badskjaer, J., Hansen, F.S., Poulsen, A.H. & Hildebrandt, P. (2003). 'Value of N-terminal proBNP in the diagnosis of left ventricular systolic dysfunction in primary care patients referred for echocardiography', *HeartDrug,* 3, 141–146.

Hammerer-Lercher, A., Neubauer, E., Muller, S., Pachinger, O., Puschendorf, B. & Mair, J. (2001). 'Head-to-head comparison of N-terminal pro-brain natriuretic peptide, brain natriuretic peptide and N-terminal pro-atrial natriuretic peptide in diagnosing left ventricular dysfunction', *Clinica Chimica Acta,* 310, 193–197.Hobbs, F.D. (2002). 'Unmet need for diagnosis of heart

failure: the view from primary care', *Heart,* 88 Suppl 2, 9–11.Hobbs, F.D., Davis, R.C. & Lip, G.Y. (2000). 'ABC of heart failure: Heart failure in general practice', *British Medical Journal,* 320, 626–629.

Hobbs, F.D., Davis, R.C., Roalfe, A.K., Hare, R. & Davies, M.K. (2004). 'Reliability of N-terminal proBNP assay in diagnosis of left ventricular systolic dysfunction within representative and high risk populations', *Heart,* 90, 866–870.

Hobbs, F.D., Davis, R.C., Roalfe, A.K., Hare, R., Davies, M.K. & Kenkre, J.E. (2002). 'Reliability of N-terminal pro-brain natriuretic peptide assay in diagnosis of heart failure: cohort study in representative and high risk community populations', *British Medical Journal,* 324, 1498.

Januzzi, J.L. & Maisel, A.S. (2004). 'Routine measurement of natriuretic peptide to guide the diagnosis and management of chronic heart failure', *Circulation,* 109, e325–e326; author reply e325–e326.

Krum, H. (2001). 'Guidelines for management of patients with chronic heart failure in Australia', *Medical Journal of Australia,* 174, 459–466.

Krum, H., Tonkin, A.M., Currie, R., Djundjek, R. & Johnston, C. I. (2001). 'Chronic heart failure in Australian general practice. The Cardiac Awareness Survey and Evaluation (CASE) Study', *Medical Journal of Australia,* 174, 439–444.

Landray, M.J., Lehman, R. & Arnold, I. (2000). 'Measuring brain natriuretic peptide in suspected left ventricular systolic dysfunction in general practice: Cross-sectional study', *British Medical Journal,* 320, 985–986.

Leibovitch, E.R. (2005). 'Once upon a time, heart failure was easy', *Geriatrics,* 60, 31–33.

Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D. & Stroup, D.F. (1999). 'Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses', *Lancet,* 354, 1896–1900.

MSAC (2005). *Guidelines for the assessment of diagnostic technologies*, Canberra, ACT, Commonwealth of Australia.

Mueller, C., Scholer, A., Laule-Kilian, K., Martina, B., Schindler, C., Buser, P., Pfisterer, M. & Perruchoud, A. P. (2004b). 'Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea', *New England Journal of Medicine,* 350, 647–654.

National Heart Foundation of Australia & The Cardiac Society of Australia and New Zealand (Chronic Heart Failure Guidelines Expert Writing Panel) (2006). *Guidelines for the prevention, detection and management of chronic heart failure in Australia*, 2006. Australia, National Heart Foundation of Australia. [Internet] Available at: www.heartfoundation.com.au. [Accessed 3 January 2007]

NHF & CSANZ (2002). *Guidelines on the Contemporary Management of the Patient with Chronic Heart Failure in Australia*, National Heart Foundation of Australia and Cardiac Society of Australia & New Zealand.

NHMRC (2000). *How to use the evidence: assessment and application of scientific evidence*, Canberra, National Health and Medical Research Council.

NHMRC (2005). *NHMRC additional levels of evidence and grades for recommendations for developers of guidelines. Pilot Program 2005*, National Health and Medical Research Council, Australian Government. [Internet]. Available from: www.nhmrc.gov.au/consult/index.htm [Accessed 2 June 2005].

Nielsen, L.S., Svanegaard, J., Klitgaard, N.A. & Egeblad, H. (2004). 'N-terminal pro-brain natriuretic peptide for discriminating between cardiac and non-cardiac dyspnoea', *European Journal of Heart Failure,* 6, 63–70.

Penn State College of Medicine (2004). *Transthoracic echocardiography.* [Internet]. Available from: http://www.hmc.psu.edu/healthinfo/t/transthoracicecho.htm [Accessed 3 June 2004].

Remme, W.J. & Swedberg, K. (2001). 'Guidelines for the diagnosis and treatment of chronic heart failure', *European Heart Journal,* 22, 1527–1560.

Sackett, D.L., Hayes, R.B., Guyatt, G.H. & Tugwell, P. (1991). *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Boston, MA: Little, Brown and Company.

Shamsham, F. & Mitchell, J. (2000). Essentials of the diagnosis of heart failure, *American Family Physician,* 61, 1319–1328.

Shibata, M.C., Flather, M.D. & Wang, D. (2001). 'Systematic review of the impact of beta blockers on mortality and hospital admissions in heart failure', *European Journal of Heart Failure,* 3, 351–357.

Sim, V., Hampton, D., Phillips, C., Lo, S.N., Vasishta, S., Davies, J. & Penney, M. (2003). 'The use of brain natriuretic peptide as a screening test for left ventricular systolic dysfunction - cost-effectiveness in relation to open access echocardiography', *Family Practice,* 20, 570–574.

Stata Corporation (2004). *Intercooled Stata 8.2 for Windows*, College City, Texas, Stata Corporation.

Swedberg, K., Cleland, J., Dargie, H., Drexler, H., Follath, F., Komajda, M., Tavazzi, L., Smiseth, O.A., Gavazzi, A., Haverich, A., Hoes, A., Jaarsma, T., Korewicki, J., Levy, S., Linde, C., Lopez-Sendon, J.L., Nieminen, M.S., Pierard, L. & Remme, W.J. (2005). 'Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology', *European Heart Journal,* 26, 1115–1140.

Whiting P.R.A., Reitsma Jb, Bossuyt Pm, Kleijnen J. (2003). 'The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews', *BMC Medical Research Methodology,* 3, 25.

Wright, S.P., Doughty, R.N., Pearl, A., Gamble, G.D., Whalley, G.A., Walsh, H.J., Gordon, G., Bagg, W., Oxenham, H., Yandle, T., Richards, M. & Sharpe, N. (2003). 'Plasma amino-terminal pro-brain natriuretic peptide and accuracy of heart-failure diagnosis in primary care - A randomized, controlled trial', *Journal of the American College of Cardiology,* 42, 1793–1800.

Zaphiriou, A., Robb, S., Murray-Thomas, T., Mendez, G., Fox, K., Mcdonagh, T., Hardman, S.M.C., Dargie, H.J. & Cowie, M.R. (2005). 'The diagnostic accuracy of plasma BNP and NTproBNP in patients referred from primary care with suspected heart failure: Results of the UK natriuretic peptide study', *European Journal of Heart Failure,* 7, 537–541.