Accuracy of B-Type Natriuretic Peptide Tests to Exclude Congestive Heart Failure

Systematic Review of Test Accuracy Studies

Markus Battaglia, MD, MPH; Daniel Pewsner, MD; Peter Juni, MD; Matthias Egger, MD, MSc, MFPHM; Heiner C. Bucher, MD, MPH; Lucas M. Bachmann, MD, PhD

Background: Congestive heart failure (CHF) is a major public health problem. The use of B-type natriuretic peptide (BNP) tests shows promising diagnostic accuracy. Herein, we summarize the evidence on the accuracy of BNP tests in the diagnosis of CHF and compare the performance of rapid enzyme-linked immunosorbent assay (ELISA) and standard radioimmunosorbent assay (RIA) tests.

Methods: We searched electronic databases and the reference lists of included studies, and we contacted experts. Data were extracted on the study population, the type of test used, and methods. Receiver operating characteristic (ROC) plots and summary ROC curves were produced and negative likelihood ratios pooled. Random-effect meta-analysis and metaregression were used to combine data and explore sources of between-study heterogeneity.

Results: Nineteen studies describing 22 patient populations (9 ELISA and 13 RIA) and 9093 patients were included. The diagnosis of CHF was verified by echocardiography, radionuclide scan, or echocardiography combined with clinical criteria. The pooled negative likelihood ratio overall from random-effect meta-analysis was 0.18 (95% confidence interval [CI], 0.13-0.23). It was lower for the ELISA test (0.12; 95% CI, 0.09-0.16) than for the RIA test (0.23; 95% CI, 0.16-0.32). For a pretest probability of 20%, which is typical for patients with suspected CHF in primary care, a negative result of the ELISA test would produce a posttest probability of 2.9%; a negative RIA test, a posttest probability of 5.4%.

Conclusions: The use of BNP tests to rule out CHF in primary care settings could reduce demand for echocardiography. The advantages of rapid ELISA tests need to be balanced against their higher cost.

Arch Intern Med. 2006;166:1073-1080

CME course available at www.archinternmed.com

Ruling out CHF early in the diagnostic process is important but difficult. Clinical symptoms and signs such as edema, shortness of breath, or persistent coughing or wheezing are unspecific and can be absent. 2,5 Particularly in elderly patients and patients with comorbid disorders that can mimic CHF (eg, obese patients in primary care or patients with acute pulmonary conditions attending emergency departments), diagnostic uncertainty may lead to delayed therapy and unnecessary echocardiograms.

See also pages 1063 and 1081

B-type (brain) natriuretic peptide (BNP) is a neurohormone that is secreted in response to volume expansion and pressure overload of cardiac ventricles. 6 In recent years, 2 tests measuring BNP in plasma have been developed, an enzyme-linked immunosorbent assay (ELISA) and a radioimmunosorbent assay (RIA). The ELISA test is a bedside test and would be particularly suitable to assist rapid diagnosis on site in primary and emergency care settings. The comparative accuracy of...
the 2 tests is, however, unclear. We performed a systematic review of the literature and meta-analysis to compare the diagnostic accuracy of ELISA and RIA assays.

We used methods recommended by the Cochrane methods group on systematic reviews of screening and diagnostic tests.7

IDENTIFICATION OF STUDIES

We searched Medline and EMBASE (January 1990 to March 2004), the Cochrane Library (2004, issue 1) and MEDION (a database of diagnostic test reviews set up by Belgian and Dutch colleagues) (December 1971 to March 2004) to identify diagnostic studies evaluating the accuracy of BNP in the diagnosis of CHF. Search strategies combined free text terms and Medical Subject Headings (MeSH) relating to heart failure and diagnostic accuracy studies. The detailed search strategy is available on request. We considered studies in any language. We supplemented electronic searches by hand searching reference lists of relevant articles and reviews and by contacting experts and manufacturers of BNP tests.

STUDY SELECTION AND DATA EXTRACTION PROCEDURES

Studies were eligible if they compared any type of BNP assay in asymptomatic patients or patients with suspected acute CHF with echocardiographic findings or findings from radionuclide scans, with or without additional clinical criteria. Besides this, the minimum requirement for inclusion was enough information to fill the 2×2 table. Two reviewers independently examined titles and abstracts of all potentially relevant articles and obtained full articles of all citations meeting the selection criteria. When necessary, we contacted authors to clarify issues, for example, when data from the same patients may have been included in several studies. We excluded studies that did not report the number of true-positive, false-positive, true-negative, and false-negative findings, and studies that were restricted to patients with diastolic dysfunction or examined the use of BNP as a prognostic marker. Final decisions on eligibility were made by consensus.

We abstracted study characteristics, methodologic quality, and results from each selected article. Study characteristics included the setting and type of population examined, publication year, test type and cutoff, reference standards, and the source of funding (main support from industry vs other). For comparison in statistical calculations we converted BNP levels from picogram per milliliter to picomole per liter where necessary, using a factor of 0.289 (1/3.463), based on the molecular weight of the BNP. Data were abstracted in duplicate using an electronic data entry form developed for this purpose. Discrepancies were resolved by consensus.

ASSESSMENT OF STUDY QUALITY

We assessed all articles that met the selection criteria for study quality. We selected items based on theoretical considerations and empirical data, including study design (case-control design or other), type of recruitment (prospective or retrospective, consecutive, or other), use of different reference tests (yes or no), application of reference tests (applied to all or only a fraction of study participants), and blind interpretation of test results (yes or no).

DATA SYNTHESIS

We calculated sensitivities, specificities, likelihood ratios, and their standard errors. We examined individual study results and between-study heterogeneity by plotting sensitivity and specificity in the receiver operating characteristics (ROC) space and used the regression model proposed by Moses et al10 to calculate summary ROC curves for both assays separately. We calculated the F statistic, which describes the percentage of total variation across studies that is due to heterogeneity rather than chance.11 Mild heterogeneity will account for less than 30% of the variation, and pronounced heterogeneity will account for more than 30%.

Because the BNP test is mainly used to rule out CHF, we were particularly interested in the likelihood ratio of a negative result. The likelihood ratio is a measure of a test result’s ability to modify pretest probabilities and is used to convert the estimated probability of the suspected diagnosis before the test result was known (pretest probability) into a posttest probability, which takes the result into account.12

METAREGRESSION ANALYSIS

We used random-effects metaregression13 to investigate sources of variation in the negative likelihood ratios. The following variables were considered: type of test (ELISA or RIA), mean age of study population, proportion of male participants, presence of typical symptoms of heart failure, type of setting, and year of publication. We also examined the following items on study quality: patient recruitment (consecutive or other), blinding of test results (yes or no), study design (prospective or retrospective), and funding source (industry or other). Finally, we explored whether the type of reference test (echocardiography, radionuclide scan, or echocardiography combined with clinical criteria) influenced results. Results are presented as sensitivities, specificities, negative likelihood ratios, and summary ROC curves, with 95% confidence intervals (CIs). All analyses were
performed using Stata Statistical software (version 8.2; Stata Corp, College Station, Tex).

**RESULTS**

**IDENTIFICATION OF STUDIES AND STUDY QUALITY**

*Figure 1* summarizes the process of identifying and selecting studies. Nineteen studies, which described 22 study populations, met our inclusion criteria. Study characteristics are listed in Table 1. Among the 19 studies, 10 had a prospective design,"8 enrolled patients consecutively,17,19,22,24-28 and 11 reported blind interpretation of test results.14,17,19,22-24,26-31 There were no diagnostic case-control studies, and reference tests were applied equally to all study participants. Thirteen studies examined the accuracy of the RIA assay,14,22,23,27,32,5, the ELISA test,23,24,28,29,31, and the ELISA N-terminal pro-BNP (NT-pro-BNP) test.26 Studies used either the ejection fraction determined by echocardiography (cutoff, 30%-50%) or r-

![Table 1. Characteristics of Studies on Test Accuracy of BNP Tests for the Diagnosis of Congestive Heart Failure](https://archinte.jamanetwork.com/)

(continued)
dionuclide scans (cutoff, 35%-40%) as the reference standard; 8 studies used a combination of echocardiography and clinical symptoms. The median sample size of the 19 studies was 139 patients (range, 52-3177); the analysis was based on 9093 patients. Five studies were sponsored by industry.23,25,27,29,31

CHARACTERISTICS OF STUDY POPULATIONS AND PRETEST PROBABILITIES

Study populations and settings were heterogeneous (Table 1). The mean age of patients ranged from 51 to 79 years; the percentage of men, from 35% to 95%. Five studies were conducted in patients with acute dyspnea in tertiary care settings, with the prevalence (pretest probability) of heart failure ranging from 39% to 72%.14,15,19,28,29,31 Similarly high prevalences were observed in patients examined after a myocardial infarction and in patients with an existing diagnosis of heart failure.14,15,19,20,23,29,31 Among outpatients who were referred by general practitioners, prevalence of heart failure ranged from 10% to 31%.16,17,20,22,27 In screening studies of patients with risk factors for coronary heart disease and CHF, the prevalence was below 10%.18,21,32

**Table 1. Characteristics of Studies on Test Accuracy of BNP Tests for the Diagnosis of Congestive Heart Failure (cont)**

<table>
<thead>
<tr>
<th>Source</th>
<th>Quality of Study*</th>
<th>Individuals Tested, No.</th>
<th>Mean Age, y</th>
<th>Proportion Male, %</th>
<th>Type of Test</th>
<th>Cutoff, pg/mL (pmol/L)</th>
<th>Reference Standard</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hobbs et al,26</td>
<td>Blind, consecutive, test described</td>
<td>297</td>
<td>62</td>
<td>52</td>
<td>ELISA</td>
<td>NR (36)</td>
<td>ECHO</td>
<td>Randomly sampled patients, general population aged over 45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>76</td>
<td>74</td>
<td>50</td>
<td>NR (36)</td>
<td>Patients with existing heart failure diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>66</td>
<td>69</td>
<td>55</td>
<td>NR (36)</td>
<td>Patients taking diuretics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>114</td>
<td>66</td>
<td>55</td>
<td>NR (36)</td>
<td>Patients at high risk for heart failure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hutcheon et al,27</td>
<td>Consecutive, prospective, test described</td>
<td>299</td>
<td>NR</td>
<td>35</td>
<td>RIA</td>
<td>170 (49)</td>
<td>ECHO</td>
<td>Referrals for suspected cardiovascular disease</td>
</tr>
<tr>
<td>Jourdain et al,28</td>
<td>Blind, consecutive, prospective, test described</td>
<td>125</td>
<td>74</td>
<td>48</td>
<td>ELISA</td>
<td>300 (87)</td>
<td>ECHO and symptoms</td>
<td>Referrals for acute dyspnea by general practice or cardiologist to emergency department</td>
</tr>
<tr>
<td>Maisel et al,29</td>
<td>Blind, prospective, test described</td>
<td>1514</td>
<td>64</td>
<td>56</td>
<td>ELISA</td>
<td>100 (29)</td>
<td>ECHO</td>
<td>Acute dyspnea patients in emergency setting</td>
</tr>
<tr>
<td>McGeoch et al,30</td>
<td>Blind, test described</td>
<td>100</td>
<td>79</td>
<td>48</td>
<td>RIA</td>
<td>122.5 (35)</td>
<td>ECHO</td>
<td>Selected heart failure patients from 2 primary care practices</td>
</tr>
<tr>
<td>Morrison et al,31</td>
<td>Blind, test described</td>
<td>321</td>
<td>NR</td>
<td>95</td>
<td>ELISA</td>
<td>94 (27)</td>
<td>ECHO and symptoms</td>
<td>Veterans, convenience sample, all dyspnea on exertion, 75% at rest</td>
</tr>
<tr>
<td>Vasan et al,32</td>
<td>prospective, test described</td>
<td>3177</td>
<td>58</td>
<td>46</td>
<td>RIA</td>
<td>NR</td>
<td>ECHO</td>
<td>General population, Framingham offspring</td>
</tr>
</tbody>
</table>

Abbreviations: BNP, B-type natriuretic peptide; ECHO, echocardiography; ELISA, enzyme-linked immunosorbent assay; NR, not reported; NT-pro-BNP, N-terminal pro-BNP assay (not converted into picograms per milliliter); RIA, radioimmunosorbent assay; RN, radionuclide.

*Only quality items that were met by the study are listed. The following items could be met: prospective recruitment of study participants, consecutive recruitment, adequate description of test, blind interpretation of test results. There were no diagnostic case-control studies; in all studies, reference tests were applied to all study participants.

†Numbers indicate different subgroups of patients as defined by setting.

TEST PERFORMANCE CHARACTERISTICS

Table 2 lists for each study population the number of true-positive, false-positive, false-negative, and true-negative test findings, as well as sensitivity, specificity, and negative likelihood ratio. The median sensitivity was 87% (range, 68%-98%). Specificity tended to be lower and more heterogeneous (median,
72%; range, 19%-98%). The combined negative likelihood ratio overall from random-effect meta-
analysis was 0.18 (95% CI, 0.13-
0.23). It was lower for the ELISA test (0.12; 95% CI, 0.09-0.16) than for the RIA test (0.23; 95% CI, 0.16-
0.32) (Figure 2). This difference was unlikely to be a chance finding (P from test of interaction=.009). The type of test explained a substan-
tial proportion of between study het-
erogeneity in the negative likeli-
hood ratio: the I² statistic was reduced from 66% to 45% when in-
cluding type of test in the metaregression model. The differ-
ence in the performance of the 2 test systems is also evident from the summary ROC curves shown in 

**Figure 3.**

The other variables considered in metaregression analyses were not strongly associated with the nega-
tive likelihood ratio, with the ex-
ception of presence of symptoms and type of test. The presence of symp-
toms was associated with a lower negative likelihood ratio (P=.07); however, the association was much attenuated and became nonsignif-
ificant (P=.62) when the type of test system was entered into the model. 

**Figure 4** shows the estimated posttest probabilities for different pretest probabilities assuming con-
stant of the negative likelihood ra-
tio. For example, for a pretest prob-
ability of 20% (which is typical for pa-
ients with suspected heart fail-
ure in primary care), a negative re-
sult of the ELISA test would pro-
duce a posttest probability of 2.9%,
and a negative RIA test would pro-
duce a posttest probability of 5.4%. Con-
versely, the corresponding post-
test probabilities among patients with acute symptoms in an em-
ergency setting, assuming a pretest prob-
ability of 60%, would be 15% and
26%, respectively.

We summarized the accuracy of the 2 available BNP screening tests for ex-
cluding CHF in different study popu-
lations, ranging from asymptomatic patients in the community to pa-
tients presenting with acute dyspnea in emergency departments. We found that negative results of both tests ac-
curately rule out the diagnosis if pa-
tients are at relatively low risk of CHF. The use of BNP tests in low-risk pa-
tients in primary care settings could reduce demand for echocardiogra-
phy and referrals of patients to spe-
cialists. The ELISA tests, which al-
low bedside testing and provide 
results in a few minutes, performed somewhat better than the RIA tests, which must be sent to a laboratory. However, the ELISA advantages must be weighed against its higher cost compared with RIA ($55 vs $11).

**STRENGTHS AND LIMITATIONS**

This systematic review and meta-
analysis is based on 9093 partici-
pants from 22 separate study popu-
lations, 1854 (20%) of whom were diagnosed as having CHF. Studies were generally of good methodo-

### Table 2. Results of Studies on Test Accuracy of BNP Tests for the Diagnosis of Congestive Heart Failure*

<table>
<thead>
<tr>
<th>Study</th>
<th>Test Type</th>
<th>True Positive</th>
<th>False Positive</th>
<th>False Negative†</th>
<th>True Negative</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Negative Likelihood Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choy et al, 1994</td>
<td>RIA</td>
<td>24</td>
<td>11</td>
<td>4</td>
<td>18</td>
<td>85.7</td>
<td>62.1</td>
<td>0.23 (0.09-0.60)</td>
</tr>
<tr>
<td>Davis et al, 1994</td>
<td>RIA</td>
<td>30</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>93.8</td>
<td>90.0</td>
<td>0.07 (0.02-0.27)</td>
</tr>
<tr>
<td>Davidson et al, 1996</td>
<td>RIA</td>
<td>20</td>
<td>28</td>
<td>0</td>
<td>39</td>
<td>97.6</td>
<td>58.2</td>
<td>0.04 (0.00-0.64)</td>
</tr>
<tr>
<td>Cowie et al, 1997</td>
<td>RIA</td>
<td>28</td>
<td>12</td>
<td>1</td>
<td>65</td>
<td>96.6</td>
<td>84.4</td>
<td>0.04 (0.01-0.28)</td>
</tr>
<tr>
<td>McDonagh et al, 1998</td>
<td>RIA</td>
<td>28</td>
<td>158</td>
<td>9</td>
<td>1057</td>
<td>75.7</td>
<td>87.0</td>
<td>0.28 (0.16-0.49)</td>
</tr>
<tr>
<td>Bettencourt et al, 2000</td>
<td>RIA</td>
<td>55</td>
<td>8</td>
<td>12</td>
<td>26</td>
<td>82.1</td>
<td>76.5</td>
<td>0.23 (0.14-0.40)</td>
</tr>
<tr>
<td>Landray et al, 2000</td>
<td>RIA</td>
<td>34</td>
<td>57</td>
<td>5</td>
<td>29</td>
<td>87.2</td>
<td>33.7</td>
<td>0.38 (0.16-0.91)</td>
</tr>
<tr>
<td>Smith et al, 2000</td>
<td>RIA</td>
<td>11</td>
<td>50</td>
<td>1</td>
<td>93</td>
<td>91.7</td>
<td>65.0</td>
<td>0.13 (0.02-0.84)</td>
</tr>
<tr>
<td>Yamamoto et al, 2000</td>
<td>RIA</td>
<td>40</td>
<td>149</td>
<td>11</td>
<td>266</td>
<td>76.4</td>
<td>64.1</td>
<td>0.34 (0.20-0.57)</td>
</tr>
<tr>
<td>Dao et al, 2001</td>
<td>ELISA</td>
<td>95</td>
<td>12</td>
<td>2</td>
<td>141</td>
<td>97.9</td>
<td>92.2</td>
<td>0.02 (0.01-0.09)</td>
</tr>
<tr>
<td>Maisel et al, 2001</td>
<td>ELISA</td>
<td>82</td>
<td>2</td>
<td>13</td>
<td>103</td>
<td>86.3</td>
<td>98.1</td>
<td>0.14 (0.08-0.23)</td>
</tr>
<tr>
<td>Valli et al, 2001</td>
<td>RIA</td>
<td>49</td>
<td>17</td>
<td>9</td>
<td>78</td>
<td>84.5</td>
<td>82.1</td>
<td>0.19 (0.10-0.35)</td>
</tr>
<tr>
<td>Hobbs et al, 2002†</td>
<td>ELISA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>NT-pro-BNP</td>
<td>6</td>
<td>85</td>
<td>0</td>
<td>206</td>
<td>92.9</td>
<td>70.8</td>
<td>0.10 (0.01-1.46)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>24</td>
<td>42</td>
<td>0</td>
<td>10</td>
<td>98.0</td>
<td>19.2</td>
<td>0.10 (0.01-1.66)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>5</td>
<td>34</td>
<td>1</td>
<td>26</td>
<td>83.3</td>
<td>43.3</td>
<td>0.38 (0.08-2.36)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>3</td>
<td>59</td>
<td>0</td>
<td>52</td>
<td>87.5</td>
<td>46.8</td>
<td>0.27 (0.02-3.59)</td>
</tr>
<tr>
<td>Hutchison et al, 2002</td>
<td>RIA</td>
<td>27</td>
<td>123</td>
<td>4</td>
<td>145</td>
<td>87.1</td>
<td>54.1</td>
<td>0.24 (0.09-0.60)</td>
</tr>
<tr>
<td>Jourdain et al, 2002</td>
<td>ELISA</td>
<td>85</td>
<td>5</td>
<td>5</td>
<td>30</td>
<td>94.4</td>
<td>85.7</td>
<td>0.06 (0.03-0.15)</td>
</tr>
<tr>
<td>Maisel et al, 2002</td>
<td>ELISA</td>
<td>670</td>
<td>185</td>
<td>74</td>
<td>585</td>
<td>90.1</td>
<td>76.0</td>
<td>0.13 (0.11-0.16)</td>
</tr>
<tr>
<td>McGech et al, 2002</td>
<td>RIA</td>
<td>41</td>
<td>13</td>
<td>19</td>
<td>27</td>
<td>68.3</td>
<td>67.5</td>
<td>0.47 (0.31-0.72)</td>
</tr>
<tr>
<td>Morrison et al, 2002</td>
<td>ELISA</td>
<td>116</td>
<td>4</td>
<td>19</td>
<td>182</td>
<td>85.9</td>
<td>97.8</td>
<td>0.14 (0.09-0.22)</td>
</tr>
<tr>
<td>Vasan et al, 2002</td>
<td>RIA</td>
<td>168</td>
<td>805</td>
<td>22</td>
<td>2182</td>
<td>88.4</td>
<td>73.0</td>
<td>0.16 (0.11-0.23)</td>
</tr>
</tbody>
</table>

Abbreviations: BNP, B-type natriuretic peptide; CI, confidence interval; ELISA, enzyme-linked immunosorbent assay; NT-pro-BNP, N-terminal pro-BNP assay (not converted into picograms per milliliter); RIA, radioimmunosorbent assay.

*To 0 cells, 0.5 was added to allow calculation of sensitivity, specificity, and other statistics. Numbers indicate different subgroups of patients as defined by setting.
†In cases of 0 false negatives, 0.5 was added to 0 to allow calculation of sensitivity, specificity, and other statistics.
logic quality, although reporting was sometimes incomplete. Both tests had been examined in a range of populations, but there was little evidence that test performance depended on patient characteristics or that the difference in performance observed between the ELISA and RIA test was explained by differences in study populations. Studies used different reference tests and criteria for the diagnosis of CHF. Different reference standards could have introduced heterogeneity in test accuracy, but this was not confirmed by our analyses.

We acknowledge that we could only adjust for information that was aggregated at the study level. Individual patient data would have been preferable but were not available. Individual patient data would also have allowed us to examine the effect of the choice of different BNP cutoffs. Despite a comprehensive literature search, we could not identify any studies that directly compared the 2 test systems. We therefore believe that our study makes an important contribution to the best available evidence on the accuracy of BNP tests for the diagnosis of CHF.

RESULTS IN CONTEXT WITH OTHER STUDIES

Doust and colleagues recently did a systematic review of the diagnostic accuracy of natriuretic peptides in the diagnosis of heart failure. This review differs in several respects: we excluded 6 studies that were in-

Figure 2. Forest plots showing the negative and positive likelihood ratios (LRs) for enzyme-linked immunosorbent assay (ELISA) and radioimmunosorbent assay (RIA) testing along with the corresponding pooled estimates using a random-effects model. CI indicates confidence interval; numbers following Hobbs et al indicate different subgroups of patients in that study, as defined by setting (Table 1).

Figure 3. Receiver operating characteristic (ROC) plot along with the summary ROC (SROC) curves stratified for studies of the enzyme-linked immunosorbent assay (ELISA) and radioimmunosorbent assay (RIA).
IMPLICATIONS AND FURTHER RESEARCH

It is widely acknowledged that CHF is a major public health problem. Recent data from the Framingham Heart Study indicate that the lifetime risk of CHF is 1 in 5, for both men and women. In the United Kingdom and elsewhere, CHF is a leading cause of hospitalization among people older than 65 years. Although prognosis has improved in recent decades, mortality remains high: the 5-year mortality from 1990 to 1999 was 59% for men and 45% for women aged 65 to 74 years in the Framingham study. With further improvements in the prognosis of CHF, improved survival after acute myocardial infarction, and an aging population, the burden of CHF will continue to increase in the years to come. B-type natriuretic peptide tests have the potential to guide clinical decisions, particularly in patients at lower risk in primary care and emergency departments. Applied early in the diagnostic process in patients with suspected cardiac failure, negative BNP test findings can help rule out CHF and thus avoid unnecessary referral to echocardiography. If the test result is positive, confirmation by echocardiography will generally be required. Early diagnosis of left ventricular dysfunction or CHF may be important: treatment with angiotensin-converting enzyme inhibitors can improve the prognosis of patients with left ventricular dysfunction without overt heart failure and patients with symptomatic CHF.

Future diagnostic accuracy studies should model the probability of CHF given the BNP test result and also take into account relevant clinical data (such as the presence of symptoms and age). Such studies should be adequately powered so that the influence of clinical indicators on test performance can be examined. The resulting prediction rules should then be validated in other populations, and their cost-effectiveness should ideally be investigated in randomized trials. Indeed, although our review indicates that BNP tests are useful to rule out CHF in patients at lower risk, the effect of the introduction of such tests on clinical outcomes or costs remains unclear.