SYSTEMATIC REVIEW

The accuracy of plasma natriuretic peptide levels for diagnosis of cardiac dysfunction and chronic heart failure in community-dwelling elderly: a systematic review

BERT VAES1, WOUTER DE RUIJTER2, JACOBIJN GUSSEKLOO2, JAN DEGRYSE3

1Department of General Practice, Université Catholique de Louvain, 1200 Woluwe-Saint-Lambert, Belgium
2Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, The Netherlands
3Department of General Practice, Université Catholique de Louvain, 1200 Woluwe-Saint-Lambert, Belgium

Address correspondence to: B. Vaes. Tel: (+32) 2 7645344; Fax: (+32) 2 7645327. Email: Bert.Vaes@uclouvain.be

Abstract

Background: measurement of plasma natriuretic peptide levels has been proposed as a simple, accessible test to assist the diagnosis of cardiac dysfunction and heart failure. Most studies have been hospital based and have investigated the relationship between natriuretic peptides and cardiac dysfunction or heart failure in younger populations.

Objective: we performed a systematic review to evaluate the diagnostic accuracy of plasma natriuretic peptide measurement in elderly patients from the general population.

Methods: electronic searches of MEDLINE and EMBASE from January 1985 to May 2008 were performed. Diagnostic cohort and cross-sectional studies on the accuracy of natriuretic peptides for diagnosis of cardiac dysfunction or chronic heart failure in people aged 75 and over in the community were included. The quality of the selected studies was assessed with the modified QUADAS tool and the data extracted by two independent reviewers.

Results: five studies were included. The general quality of the studies was moderate. The extracted data could not be pooled. Negative likelihood ratios for cardiac dysfunction ranged from 0.09 to 0.29.

Conclusion: we found limited evidence supporting the use of plasma natriuretic peptide measurement for diagnosis of cardiac dysfunction or heart failure in the elderly of 75 years and over in the general population. Important questions about the implementation of plasma natriuretic peptide measurement in daily practice remain unresolved.

Keywords: aged, natriuretic peptides, heart failure, cardiac dysfunction, diagnosis, community, elderly

Introduction

In our ageing society, the burden of chronic heart failure is rising. The prevalence increases with age from 0.7% in people aged 55–64 years to 2.7% in those aged 65–74 years and 13.0% in those aged 75–84 years [1]. Heart failure not only has negative consequences for functional status and well-being but also leads to increased mortality [1]. However, diagnosing chronic heart failure is notoriously difficult, especially in the elderly who often have multiple co-morbidities and may present with many other possible causes for dyspnoea, fatigue or peripheral oedema. With increasing average patient age, primary care physicians will become increasingly important as the principal diagnosticians and treating physicians. In this setting, poor availability of routine echocardiography leads to considerable over- and under-diagnosis of heart failure [2, 3]. This emphasises the need for a simple test, easily applicable in primary care settings, to identify the elderly patients at risk and to initiate timely treatment to reduce mortality and improve quality of life.

Over the last decade, brain natriuretic peptide (BNP) and its amino-terminal portion N-terminal pro-BNP (NT-proBNP) have been extensively studied. Not only their accuracy as diagnostic markers for cardiac dysfunction and heart
failure [4–9] but also their possible applications as prognostic markers [10] and therapeutic agents [11] have been investigated. However, most studies are hospital based and have investigated the relationship between natriuretic peptide levels and cardiac dysfunction or heart failure in younger populations. Because age has been shown to be an important confounder for the plasma level of natriuretic peptides [12–13], cut-off values used in younger populations cannot be applied in old age. In addition, because of a different prevalence (pre-test probability) of cardiac dysfunction and heart failure in inpatient and outpatient settings, results from hospital-based studies cannot be extrapolated to primary care settings.

Therefore, we performed a systematic review to investigate the accuracy of plasma natriuretic peptide levels for diagnosis of cardiac dysfunction and heart failure in community-dwelling people aged 75 and over.

Methods

Search question and search strategy

A ‘PIRT’ (Patient-Index test-Reference test-Target condition), analogous to the ‘PICO’ [14] (Patient-Intervention-Comparison-Outcome) for systematic reviews of interventional studies, was created for systematic review of diagnostic studies. To create a homogeneous study population of elderly patients, we searched for population-based studies that included only patients aged 75 years and over, or studies with subgroup analysis of patients aged 75 years and over. Only studies that were based in the general population or studies with participants referred for further investigation by a general practitioner were included. Natriuretic peptides represented the index test of the search and no restriction for a specific type of test was used. Because of the absence of a gold-standard test for cardiac dysfunction and heart failure, all possible reference tests were included. The target condition was cardiac dysfunction, systolic and/or diastolic and chronic heart failure. Based on the results of the search, subgroup analyses for index and reference test and target condition were performed to guarantee homogeneity.

MEDLINE (PubMed) and EMBASE data from January 1985 to 31 May 2008 were searched for all studies of the accuracy of natriuretic peptides for diagnosis of cardiac dysfunction or chronic heart failure. To have a large safety net, a broad search strategy based only on the index test and target condition was constructed. In MEDLINE, the MeSH term ‘Aged’ was added. The search strategy for MEDLINE and EMBASE is available from the authors. Additional searching of reference lists of relevant articles and reviews was performed.

Selection of studies

Both population-based diagnostic cohort and cross-sectional studies were included, with no language restrictions. Case-control studies, studies investigating acute dyspnoea or acute heart failure, studies in specific patient populations (e.g. patients with chronic obstructive pulmonary disease, diabetics) and studies examining the use of natriuretic peptides as prognostic markers or therapeutic agents were all excluded. The first reviewer (B.V.) divided the resulting articles into three categories (definitely excluded, included and in doubt) based on title and abstract. All studies in the last two categories, plus a random selection of the excluded articles, were checked by the second reviewer (W.R.). Disagreements were resolved by consensus or by a third reviewer (J.D.). A log of the excluded articles, with the reasons for exclusion, was kept.

Quality assessment

Two reviewers (B.V. and W.R.) independently assessed the quality of the selected studies with the modified QUADAS tool [15]. Five extra items, relevant for this systematic review, were added (Table 1).

Data extraction

Each reviewer independently extracted the data from the selected studies. If multiple cut-off values for the natriuretic peptide were presented, the cut-off value giving the greatest sum of sensitivity and specificity was used. For each cut-off value, sensitivity, specificity and positive and negative likelihood ratio were reported. The negative likelihood ratio [(1 - sensitivity)/specificity] and the positive likelihood ratio (sensitivity/(1 - specificity)) for all the included studies were calculated, if it was not already available. The likelihood ratio
The diagnostic accuracy of natriuretic peptides in community-dwelling elderly

Total $n = 4704$

<table>
<thead>
<tr>
<th>Database</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE</td>
<td>$n = 2656$</td>
</tr>
<tr>
<td>EMBASE</td>
<td>$n = 2048$</td>
</tr>
</tbody>
</table>

Duplicates $n = 1508$

Screening abstract and title $n = 3196$

Eligible for further review $n = 80$

Eligible for full text evaluation $n = 80$

Excluded, $n = 75$
- No (sub)population $\geq 75$ y, $n = 62$
- Systematic review, $n = 2$
- Review, $n = 1$
- No diagnostic study, $n = 7$
- No population-based study, $n = 2$
- Acute heart failure, $n = 1$

Included in review $n = 5$

General population of (unselected) elderly ($\geq 75$ y), asymptomatic AND symptomatic, $n = 4$

General population of (selected) elderly ($\geq 75$ y), symptomatic, $n = 1$

Figure 1. Retrieval of eligible studies: flowchart.

was used to convert the pre-test probability (prevalence) of cardiac dysfunction or heart failure into a post-test probability, which takes the result into account [16]. The post-test probability is the probability that the patient has cardiac dysfunction or heart failure, given a negative or positive test result. The post-test probability for a negative and positive test result (PTP$^-$ and PTP$^+$, respectively) was calculated by dividing the post-test odds \( \left\{ \text{likelihood ratio} \times \frac{\text{pre-test probability}}{1 - \text{pre-test probability}} \right\} \) by (post-test odds $+ 1$). The negative predictive value (NPV) equals 1 minus the post-test probability of a negative test. The positive predictive value (PPV) equals the post-test probability of a positive test.

The decision whether to pool the data was based on the number of studies found and the observed differences in the index test and target condition used.

Results

The search strategy and results are presented in Figure 1. Five studies that met the inclusion criteria were identified [17–21]. Four population-based studies [22–25] that included participants, or a subgroup of participants, with a mean or median age of 75 years and older, were excluded because they also included patients younger than 75 in their analysis. We
differentiated between studies of non-selected populations [17–20], including symptomatic and asymptomatic participants, and a selected population, including only symptomatic participants [21]. Two articles concerned the same study population, but one article analysed the accuracy of BNP [17] and the other NT-proBNP [19] for the diagnosis of cardiac dysfunction.

Table 1 shows the description of the individual items from the ‘extended’ modified QUADAS tool and Table 2 reports the evaluation of each item in the selected studies. The general quality of the studies was moderate, with no study scoring <9/16 and a maximum score of 11/16. Item (b) was scored by a specialist in the field. The BNP test from Shionogi [17, 18] was considered to be outdated. Several important items were unclear or missing. The study by Sivakumar et al. [21] was carried out in an outpatient echocardiography service in a district general hospital, but it was unclear whether the participants were referred only by general practitioners or also by hospital physicians. Most studies did not mention the time period between the index test and the reference test, so it was impossible to evaluate whether a progression or recovery bias or a treatment paradox was present in these studies. Item 7 was scored ‘unclear’ for all included studies, because it was not mentioned whether the laboratory doctor who interpreted the level of natriuretic peptide was blinded for the results of the reference test. No study used a pre-specified cut-off value. Most studies used discriminatory values derived from receiver operating characteristic analysis.

Table 3 shows the results for studies from an unselected study population. Hedberg et al. [18] included only participants of 75 years old, whereas the other three studies included a subgroup analysis of participants aged 75 years and over. Costello-Boerrigter et al. [19] and Abhayaratna et al. [20] used NT-proBNP as the index test, while Redfield et al. [17] and Hedberg et al. [18] used BNP. All four used echocardiography as the reference test. An ejection fraction (EF) ≤40% was mostly used as the cut-off value. Hedberg et al. [18] used the wall motion index that corresponded best to EF <40% as the target condition. Abhayaratna et al. [20] used a combination of systolic (EF ≤40% or EF ≤50%) and diastolic dysfunction as target condition. This explains the higher prevalence of the target condition found by Abhayaratna et al. [20]. It was not possible to determine the prevalence of the target condition for the study of Redfield et al. [17], although this study used the same study population as did Costello-Boerrigter et al. [19], because the number of participants receiving one particular index test was different in each study. No statistical analysis for heterogeneity was performed because of the low number of articles found.

We found negative likelihood ratios ranging from 0.13 to 0.29. The diagnostic impact of natriuretic peptide measurement as expressed by the likelihood ratios depends on the prevalence of the target condition in the population involved (pre-test probability), giving a reduction of a pre-test probability of 21% to a post-test probability of 5.7% for the study of Abhayaratna et al. [20] and a reduction of a pre-test probability of 4.7% to a post-test probability of 0.64% for the study of Costello-Boerrigter et al. [19]. The post-test probability of a positive test was less acceptable, ranging from 21.0 to 59.7%, depending on the pre-test probability.

Because of a different target condition, the diagnostic accuracy for NT-proBNP as determined by Costello-Boerrigter et al. [19] and Abhayaratna et al. [20] is difficult to compare. Costello-Boerrigter et al. [19] found acceptable gender-specific sensitivities and specificities for systolic dysfunction. Abhayaratna et al. [20] found gender-specific negative likelihood ratios and PTP− in the same range as Costello-Boerrigter et al. [19], for a combined target condition of systolic and diastolic dysfunction with lower cut-off values. Because a 2 × 2 table could not be extracted from the subgroup analysis by Redfield et al. [17], the data could not be pooled with that of the study by Hedberg et al. [18]. Hedberg et al. [18] found a negative likelihood ratio for systolic dysfunction in the same range as Redfield et al. [17] with a much lower cut-off value. With this low cut-off value, they found a low PTP− of 1.7% and a low PTP+ of 34.8%.

The one study concerning a selected population investigated the utility of NT-proBNP in a cohort of 100 very elderly (range 75–94 years) patients with suspected cardiac disorders referred for echocardiography [21]. For the diagnosis of systolic dysfunction (EF <50%, prevalence 25%) the area under the curve (AUC) was 0.71 (95% CI 0.69–0.89). An NT-proBNP level of 424 pg/mL had a sensitivity of 96%, a specificity of 45% and a positive and negative likelihood ratio of 1.75 and 0.09, respectively. The post-test probability for a positive test was 36.8%, and the post-test probability for a negative test was 2.9%. The authors found that patients with diastolic dysfunction or failure had lower plasma concentrations of NT-proBNP than patients without cardiac dysfunction.

Table 2. Qualification of the articles with the modified QUADAS tool and extra items

<table>
<thead>
<tr>
<th>Item (a)</th>
<th>Item (b)</th>
<th>Item (c)</th>
<th>Item (d)</th>
<th>Item (e)</th>
<th>Item (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ present</td>
<td>− not present</td>
<td>? unclear</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unselected population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abhayaratna, 2006 [20]</td>
</tr>
<tr>
<td>+ + ? + + + + + + ? + + − − − − + + +</td>
</tr>
<tr>
<td>Costello-Boerrigter 2006 [19]</td>
</tr>
<tr>
<td>+ + ? + + + + ? + + + + + + − − +</td>
</tr>
<tr>
<td>Hedberg, 2004 [18]</td>
</tr>
<tr>
<td>+ + ? + + + + ? + + + + + + ? + +</td>
</tr>
<tr>
<td>Redfield, 2002 [17]</td>
</tr>
<tr>
<td>+ + + + + + + + + + + + + + + + +</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Selected population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sivakumar, 2006 [21]</td>
</tr>
<tr>
<td>? + + ? + + + + + + ? + + + + + + + ? − + + + +</td>
</tr>
</tbody>
</table>

+ = present; − = not present; ? = unclear.
Table 3. Search results for an unselected population ≥75 years old

<table>
<thead>
<tr>
<th>Sample size</th>
<th>≥75 years</th>
<th>% Men ≥75 years</th>
<th>Target condition</th>
<th>Prevalence</th>
<th>n/tot (%)</th>
<th>Index test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abhayaratna, 2006 [20]</td>
<td>301</td>
<td>51.5</td>
<td>EF ≤40% or advanced DD-NEF</td>
<td>Men 20/155 (13)</td>
<td>452 pg/mL</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Women 23/146 (16)</td>
<td>375 pg/mL</td>
<td>83</td>
</tr>
<tr>
<td>Costello-Boerrigter, 2006 [19]</td>
<td>275</td>
<td>?</td>
<td>EF ≤50% or advanced DD-NEF</td>
<td>Men 32/155 (21)</td>
<td>426 pg/mL</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Women 28/146 (19)</td>
<td>375 pg/mL</td>
<td>80</td>
</tr>
<tr>
<td>Hedberg, 2004 [18]</td>
<td>407</td>
<td>49.6</td>
<td>EF ≤40%</td>
<td>Men 13/275 (4.7)</td>
<td>1,024 pg/mL</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Women 28/407 (6.9)</td>
<td>28 pg/mL</td>
<td>75</td>
</tr>
<tr>
<td>Redfield, 2002 [17]</td>
<td>?</td>
<td>?</td>
<td>LVSD, WMI that best corresponded to EF &lt;40%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Results**

<table>
<thead>
<tr>
<th>Cut-off value</th>
<th>AUC</th>
<th>Sens (%)</th>
<th>Spec (%)</th>
<th>LR+</th>
<th>LR−</th>
<th>PTP+ (%)</th>
<th>PTP− (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>452 pg/mL</td>
<td>0.86</td>
<td>85</td>
<td>81</td>
<td>4.1</td>
<td>0.19</td>
<td>38.0</td>
<td>2.7</td>
</tr>
<tr>
<td>375 pg/mL</td>
<td>0.89</td>
<td>83</td>
<td>80</td>
<td>4.8</td>
<td>0.21</td>
<td>47.8</td>
<td>3.8</td>
</tr>
<tr>
<td>426 pg/mL</td>
<td>0.84</td>
<td>81</td>
<td>86</td>
<td>4.2</td>
<td>0.23</td>
<td>52.8</td>
<td>5.7</td>
</tr>
<tr>
<td>375 pg/mL</td>
<td>0.91</td>
<td>86</td>
<td>86</td>
<td>6.3</td>
<td>0.17</td>
<td>59.7</td>
<td>3.8</td>
</tr>
<tr>
<td>1,024 pg/mL</td>
<td>0.89</td>
<td>86</td>
<td>84</td>
<td>5.6</td>
<td>0.13</td>
<td>59.7</td>
<td>3.8</td>
</tr>
<tr>
<td>879 pg/mL</td>
<td>0.86</td>
<td>89</td>
<td>89</td>
<td>5.4</td>
<td>0.29</td>
<td>21.6</td>
<td>0.64</td>
</tr>
<tr>
<td>&gt;73 pg/mL</td>
<td>0.88</td>
<td>79</td>
<td>89</td>
<td>5.4</td>
<td>0.29</td>
<td>21.6</td>
<td>0.64</td>
</tr>
</tbody>
</table>

EF, ejection fraction; DD-NEF, diastolic dysfunction with normal ejection fraction; LVSD, left ventricular systolic dysfunction; WMI, wall motion index; AUC, area under curve; Sens, sensitivity; Spec, specificity; LR+, positive likelihood ratio; LR−, negative likelihood ratio; PTP+, post-test probability of positive test; PTP−, post-test probability of negative test.
**Discussion**

In this systematic review, we investigated the accuracy of plasma natriuretic peptide levels for diagnosis of cardiac dysfunction and chronic heart failure in community-dwelling elderly patients aged 75 and over. Limited data were retrieved from a broad search strategy that included some 3,000 articles. The five studies that we found delivered limited information about the use of natriuretic peptide analysis as a tool for diagnosis of cardiac dysfunction and chronic heart failure in our target population. In general, the studies we found confirmed earlier observations in other populations that natriuretic peptide levels can best be used to exclude cardiac dysfunction or chronic heart failure [8].

Age is an important confounder for the plasma level of natriuretic peptides [12–13]. The mechanisms underlying the age-related increase have not been fully elucidated, although renal impairment, myocardial fibrosis and subtle diastolic dysfunction, not detectable by current techniques, have been suggested [17, 26]. Systematic reviews published in this field have not used age as an inclusion criterion, nor has a subgroup analysis based on age been performed [4–9]. Only Ewald et al. [9] examined the effect of age on test performance, but they pooled results of studies in which the mean age of participants was <80 years. Our systematic review is the first to use age as a selection criterion.

Several unclear issues regarding the diagnostic abilities of natriuretic peptides in the elderly remain. First, most studies that investigated the diagnosis of chronic heart failure have ignored prognostically significant diastolic dysfunction [27]. Heart failure and preserved left ventricular ejection fraction (HFPEF) is relatively uncommon in younger patients but increases in importance in the elderly. It has, however, been debated whether diastolic and systolic heart failure are separate pathophysiological entities, as diastolic impairment at rest is a common if not universal accompaniment of left ventricular systolic dysfunction [28]. To date, there is no reference standard to indicate diastolic dysfunction in the way that the left ventricular EF is generally accepted as the reference standard for systolic dysfunction. Furthermore, no specific treatment has yet been shown, convincingly, to reduce morbidity and mortality in patients with HFPEF, and management of HFPEF should be aimed at causes of diastolic dysfunction such as myocardial ischaemia, hypertension and diabetes and precipitating factors like atrial fibrillation [28].

Second, multiple determinants are known to influence circulating levels of natriuretic peptides, potentially limiting their use as diagnostics, especially in the very elderly, who often present with multiple co-morbidities. An understanding of the determinants affecting natriuretic peptide levels is a prerequisite for their optimal use as a tool for diagnosis of cardiac dysfunction or chronic heart failure in the community [29]. When ordering a natriuretic peptide test in patients with atrial fibrillation or renal impairment, physicians should realize that plasma levels can be raised without accompanying echocardiographic evidence of abnormal cardiac structure or function [20].

Third, as for every day clinical practice, several considerations should be taken into account when interpreting the results from our systematic review:

- In most studies retrieved in this systematic review, natriuretic peptide levels were used as a screening test, not as a diagnostic tool for heart failure or high-risk patients. So far, cost effectiveness studies into such screening are lacking.
- The place of the natriuretic peptide test in the diagnostic algorithm is still unclear. The results of our analysis indicate that natriuretic peptide levels are most efficient in excluding cardiac dysfunction or chronic heart failure. Studies estimating the added value of natriuretic peptides beyond history taking and clinical examination or ECG, representing everyday clinical practice, remain extremely scarce.
- Regarding the choice of natriuretic peptide test, it has been shown that NT-proBNP is more stable than BNP [30], and may have lower intra-individual and inter-individual variation [31].
- Cut-off values vary considerably between studies, with important effects on test characteristics. Rutten et al. even suggest the use of ‘double’ cut-off values, one exclusionary (‘rule out’) and one confirmatory (‘rule in’) [32].
- Gender is a known confounder for natriuretic peptides, which are found to be higher in women than in men [33]. However, in elderly patients, gender-specific cut-off values appeared to be higher in men [17, 19, 20].

In conclusion, we found limited evidence for the usefulness of natriuretic peptide measurement for the diagnosis of cardiac dysfunction or heart failure in elderly patients aged 75 and over from the general population. Important questions about the implementation of the natriuretic peptide test in daily practice remain unsolved. This systematic review emphasises the importance of future research on the diagnostic accuracy of natriuretic peptide levels in elderly people in the community, not only because of the high accessibility of measuring natriuretic peptides (as opposed to echocardiography) in primary care, but also because the forthcoming epidemic of heart failure will predominantly affect elderly people.

**Key points**

- The presence of multiple co-morbidities often compromises the diagnosis of chronic heart failure in elderly patients.
- Natriuretic peptides have been proposed as a simple, accessible test to assist the diagnosis of cardiac dysfunction and heart failure.
- Most studies, however, were hospital-based and investigated natriuretic peptides in younger populations.
- We found limited evidence supporting the use of plasma natriuretic peptide measurement for diagnosis of cardiac...
dysfunction or heart failure in the elderly of 75 years and over in the general population.

- Important questions about the implementation of plasma natriuretic peptide measurement in daily practice remain unresolved.

Conflicts of interest

None.

References

B. Vaes et al.


29. Jaffe AS, Apple FS, Babuin L. Why we don’t know the answer may be more important than the specific question. Clin Chem 2004; 50: 1495–7.


Received 12 November 2008; accepted in revised form 8 July 2009