Clinical research

Using BNP to develop a risk score for heart failure in primary care

David Adlam1*, Paul Silcocks2, and Nigel Sparrow3

1 Department of Cardiovascular Medicine, Queen’s Medical Centre, Nottingham, UK
2 Trent Institute for Health Services Research, University of Nottingham Medical School, Queen’s Medical Centre, Nottingham, UK
3 Newthorpe Medical Practice, Nottingham, UK

Received 25 February 2004; revised 4 December 2004; accepted 27 January 2005; online publish-ahead-of-print 15 March 2005

See page 1052 for the editorial comment on this article (doi:10.1093/eurheartj/ehi244)

Aims Chronic heart failure is a common condition with high mortality. Accurate diagnosis in primary care is difficult. Elevated B-type natriuretic peptide (BNP) is associated with left ventricular systolic dysfunction and increased mortality. Prognostic scoring systems using BNP may help to stratify risk in primary care patients. The aim of this research was to establish the independent variables which predict mortality in a primary care population-prescribed loop diuretics and to generate and validate a scoring system for heart failure in general practice.

Methods and results Five hundred and thirty-two patients were followed up for a mean of 6.4 years after attending a research clinic for clinical assessment, electrocardiogram (ECG), echocardiography, and BNP. Multivariate analysis was used to establish independent prognostic variables and to generate a prognostic scoring system.

The score generated was: 0.50/2 BNP + 5/2 age + 50/2 (CVA + sex + diabetes + ECG).
The cut-off scores for risk groups were; 25th percentile, 411; 50th percentile, 475; 75th percentile, 524; Harrell’s c = 0.75.

Conclusion Developing prognostic scoring systems provides a means of risk stratifying patients without relying on a single cut-off diagnostic value for BNP. Further validation of such scoring systems may improve future management of community heart failure patients.

KEYWORDS
Primary care; Heart failure; Brain natriuretic peptide; Echocardiography; Prognosis

Introduction

Chronic heart failure is a common condition, especially in the elderly. In those 75 years of age, the prevalence may exceed 80 cases per 1000.1 It is associated with a high mortality and considerable morbidity, in particular, from breathlessness and reduced exercise tolerance. The identification of patients with chronic heart failure is known to benefit from treatments

*Corresponding author. Tel: +44 1865 760 177.
E-mail address: davidadlam@doctors.org.uk

© The European Society of Cardiology 2005. All rights reserved. For Permissions, please e-mail: journals.permissions@oupjournals.org
degree of technical expertise, and availability to primary
care physicians may be limited. Even where open access
echocardiography services are available, the referral for
this investigation is much lower than the population
prevalence of the disease.9

B-type natriuretic peptide (BNP) is closely correlated
with left ventricular systolic dysfunction in most13–15
but not all studies.16,17 It is also an independent predictor
of mortality.16,18 The development of bedside test kits
for BNP has the potential to allow rapid accurate
assessment of BNP in a primary care setting.10,21 The practical
application of BNP measurement to primary care practice
has been limited by difficulties in establishing a cut-off
threshold for diagnosis.22,23 This has led some to
suggest that there is no role for BNP measurement in
current clinical practice.22 The development of scoring
systems in primary care, utilizing BNP and other readily
available clinical information, may allow patients with
a worse prognosis to be identified without relying
purely on a single cut-off threshold value for BNP.

In this study, we followed up a population of patients
prescribed loop diuretics in primary care, who attended
a research clinic for assessment. The aim of this study
was to use the independent variables predictive of
mortality to develop and validate a scoring system in
primary care, using BNP and other readily available
clinical information.

Methods

Seven general practices in Nottingham, UK, with a total list size
of 60,728 were surveyed between January 1995 and December
1998. The study area included both urban and rural practices,
with list sizes ranging from 4000 to 12,000. This practice
population was under the care of 27 full-time equivalent general
practitioners. In total, 1366 patients were found to be taking
loop diuretics. These patients were invited by letter to attend
a research clinic at the local hospital where they underwent
full clinical assessment, electrocardiography, BNP measurement,
and echocardiography. Each of these elements was carried out
blinded to the results of other parts of the assessment.
Electrocardiograms (ECGs) were analysed by a primary care
physician (N.S.) and classified as normal or abnormal. To best
mimic conditions in primary care practice, the primary care
physician had no prior specialized ECG training and no specific
instruction on what constituted abnormality. This simple
approach to ECG analysis has previously been demonstrated to
be both accurate and reproducible in primary care.24 Echocar-
diography was performed by an experienced technician. Left
ventricular ejection fractions were measured using a phased
array sector scanner (Vingmed CFM 700, Oslo, Norway). The
images were analysed using a computer-assisted, video overlay,
echocardiographic analysis system (Thoracenter, Erasmus
University, Rotterdam, The Netherlands). An apical four-
chamber view was used for imaging, and a modified Simpson’s
single plane disc method was used for analysis. The BNP assay
used has been described elsewhere.16 In brief, plasma samples
were acidified and extracted using pre-activated Sep Pak C18
cartridge (Waters Corporation, Milford, MA, USA). The eluates
were dried under vacuum using a centrifugal evaporator and
stored. The precipitates were resuspended in assay buffer and
assayed by radioimmunoassay (Peninsula Laboratories, St Helens,
UK).

Of 737 patients who attended the research clinic, ejection
fraction and BNP measurements were obtained in 570 patients.
Ethical approval for this study was obtained from the Queen’s
Medical Centre Ethics Committee, Nottingham. The investi-
gation conforms to the principles outlined in the Declaration of
Helsinki (BMJ 1964;i:177).

The vital status of patients was subsequently established a
mean of 6.4 years (4.0–7.1) and median 6.3 years after assess-
ment. Five hundred and thirty-two patients were traced using
local authority records. Thirty-eight patients were lost to
follow up. The derivation of the study population is shown in
Figure 1.

Statistics

Variables significantly associated with a worse prognosis were
identified for descriptive purposes by an initial univariate

60,728 – Total GP practice population studied from seven practices in Nottinghamshire,
England

1366 – Patients taking loop diuretics according to primary care prescribing records. All
were invited to attend the research clinic

737 patients attended research clinic for clinical assessment, electrocardiography,
echocardiography, and plasma brain natriuretic peptide measurement

570 patients had measurable ejection fraction and plasma brain natriuretic peptide and
were included in this analysis

Vital status determined on 532 patients mean 6.4 years after assessment using local
authority records of which 528 had evaluable follow-up durations 38 patients lost to follow-up

466 patients with no missing values (apart from smoking status)

Figure 1 Diagram demonstrating how the study population was derived.
analysis. Because some of these variables might be redundant for predictive purposes, a Cox regression analysis was then performed to establish which factors retained prognostic importance when considered jointly. Cigarette consumption data were found to be incomplete in the primary care records of the patients studied, and this variable was omitted from further consideration.

A predictive model was developed using the whole data set as follows. Having excluded incomplete observations, a 100% simple random sample with replacement was drawn and a stepwise Cox regression was performed, the list of variables selected being stored. This process was repeated 100 times, and those variables which were selected in ≥60% of the repetitions (i.e. for which the lower 95% confidence interval for the probability of selection was ≥50%) were then fitted simultaneously to generate a prognostic score. This admittedly ad hoc procedure ensured that only variables consistently selected would be used for the final model, which would therefore not suffer from over-fitting and is similar to that described by Austin and Tu. The number of events per parameter was 11.38.

The performance of the prognostic score was then evaluated on the data set by means of Harrell’s c-statistic. The c-statistic is a rank-based measure, a value of 1 indicating perfect concordance of the score and outcome, whereas a value of 0.5 indicates only chance agreement. The evaluation incorporated a correction for optimism by means of a 0.632 bootstrap. A simplification of the score and outcome, whereas a value of 0.5 indicates only chance agreement. The evaluation incorporated a correction for optimism by means of a 0.632 bootstrap. A simplification of the score and outcome, whereas a value of 0.5 indicates only chance agreement. The evaluation incorporated a correction for optimism by means of a 0.632 bootstrap. A simplification of the score and outcome, whereas a value of 0.5 indicates only chance agreement. The evaluation incorporated a correction for optimism by means of a 0.632 bootstrap.

### Results

The characteristics of patients in this study are shown in Table 1. Using the Kaplan–Meier method, all-cause mortality in this population was calculated as 37.14% at 5 years (190/528 evaluable). The death certification data over the whole follow-up period show LVF as the primary cause of death in 21% (44/209) of patients. A further 18 patient characteristics (9%) listed LVF as a contributing factor in the cause of death. Cardiovascular disease was given as the cause of death in 56% (117/209) of patients, cancer in 10% (21/209), and lung disease in 18% (38/209). Other causes of death accounted for 15% (32/209) of deaths. No information on cause of death was available in one patient.

For comparability of univariate and multiple variable analyses, further results apply to the data set without missing values (except when otherwise specified). Mortality was higher in patients with lower ejection fractions (Table 2). For those patients with ejection fraction measurements in the lowest quartile (median 25%, IQR 22–28%), 56% were alive at 5 years. This compares with the patients in the highest quartile of ejection fractions (median 54%, IQR 51–59%), where 69% survived 5 years. BNP level at the time of assessment was also significantly correlated with mortality (Table 3). Of patients with BNP measurements in the highest quartile (median 135 pg/mL, IQR 103–165), 34% survived 5 years, whereas in those with BNP in the lowest quartile (median 28 pg/mL, IQR 22–32), 81% survived 5 years.

Table 4 shows the univariate analysis. The variables significantly associated with mortality are in bold. The effect of BNP only appears small because it is expressed per increment of pg/mL, but it is statistically highly significant. When repeated stepwise, Cox regressions were performed to identify the independent variables predictive of mortality; the variables consistently selected were age at assessment, BNP, sex, past history of stroke, past history of diabetes, and the ECG assessment being normal or abnormal. Table 5 shows pairwise $r^2$ values: those corresponding to variables consistently selected are in bold and it is clear that these variables were largely uncorrelated with each other, whether this was in the full data set or the one without missing values.

The variables consistently identified were combined in a prognostic score, on the basis of coefficients in a Cox regression. The optimal score was found to be: $\{0.0069 \times \text{BNP (pg/mL)}\} + 0.068 \times \text{age at assessment} + 0.549 \times \text{past history of stroke (0 if no past history, 1 if past history)} \} + 0.486 \times \text{sex (1 if male, 0 if female)} + 0.587 \times \text{past history of diabetes (0 if no past history, 1 if past history)} + 0.637 \times \text{ECG abnormality (0 absent, 1 present)}$.

The proportional hazards assumption was tested for this model, using the Stata function `phstest` (Table 6). Overall, there is no evidence contrary to the assumption, although there is evidence that the assumption is violated for diabetes. However, in this context, the prime interest is in the performance of the overall score so the result for diabetes was discounted. The overall performance as measured by Harrell’s c was 0.756 and...
correction for optimism by a 0.632 bootstrap only reduced this to a value of 0.748. This compares with Harrell’s c for BNP in the univariate analysis of 0.695. Thus, the combined score gives a useful improvement over absolute BNP alone. To generate a more user-friendly prognostic score, a simplified version of this score was also tested. The simple score was 0.50/C2BNP + 5/C2age + 5/C2(CVA + sex + diabetes + ECG) for which Harrell’s c was 0.752 [and for which the test for non-proportional hazards gave a \( \chi^2 \) of 1.49 (1 df), \( P = 0.2216 \)]. There was therefore no practical difference in the discrimination provided by the optimal and simplified scoring systems as estimated by Harrell’s c, although this is a rank-based measure, it may miss small differences in discrimination.

The Kaplan–Meier survival estimates for the groupings of the simplified score are shown in Figure 2. The scores are grouped by quartiles, with the lowest score correlating with the best survival. A crude sensitivity analysis was also performed by imputing missing values first

<table>
<thead>
<tr>
<th>Ejection fraction (split at quartiles) median (IQR)</th>
<th>Mean ejection fraction (%)</th>
<th>Alive at 5 years (Kaplan–Meier estimate) (%)</th>
<th>SE</th>
<th>Lower 95% CL</th>
<th>Upper 95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 (22–28)</td>
<td>23.93</td>
<td>56.11</td>
<td>4.64</td>
<td>46.54</td>
<td>64.62</td>
</tr>
<tr>
<td>36 (34–38)</td>
<td>35.87</td>
<td>63.92</td>
<td>4.64</td>
<td>54.07</td>
<td>72.20</td>
</tr>
<tr>
<td>45 (43–46)</td>
<td>44.64</td>
<td>62.84</td>
<td>4.56</td>
<td>53.21</td>
<td>71.03</td>
</tr>
<tr>
<td>54 (51–59)</td>
<td>55.38</td>
<td>68.68</td>
<td>4.53</td>
<td>58.86</td>
<td>76.62</td>
</tr>
</tbody>
</table>

IQR, interquartile range; CL, confidence limit.

<table>
<thead>
<tr>
<th>BNP (pg/mL) (split at quartiles) median (IQR)</th>
<th>Mean BNP</th>
<th>Alive at 5 years (Kaplan–Meier estimate) (%)</th>
<th>SE</th>
<th>Lower 95% CL</th>
<th>Upper 95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 (22–32)</td>
<td>26.2</td>
<td>81.13</td>
<td>3.56</td>
<td>72.94</td>
<td>87.06</td>
</tr>
<tr>
<td>44 (40–44)</td>
<td>44.66</td>
<td>72.64</td>
<td>4.42</td>
<td>62.86</td>
<td>80.25</td>
</tr>
<tr>
<td>72 (62–81)</td>
<td>71.40</td>
<td>61.89</td>
<td>4.59</td>
<td>52.22</td>
<td>70.16</td>
</tr>
<tr>
<td>135 (103–165)</td>
<td>146.90</td>
<td>33.74</td>
<td>4.77</td>
<td>24.61</td>
<td>43.08</td>
</tr>
</tbody>
</table>

IQR, interquartile range; CL, confidence limit.
with the highest and next with the lowest values observed for each variable. This resulted in the same variables being identified.

Figure 3 shows the relative survival in each of the quartile population groups identified by the simple scoring system (compared with that expected among the general population matched for age and sex). This shows that even those patients in this study population with the best scores have a worse prognosis than age- and sex-matched patients in the region in which this study was performed (relative survival at 5 years = 63.28%, 95% CI 58.54–67.64), and in effect, general mortality can be ignored for prognostic purposes in these patients.

Discussion

The results of this study confirm that BNP is an independent predictor of mortality in the primary care population studied. This is consistent with the findings of other community-based studies. The relationship between BNP and prognosis may be directly related to the association between elevated BNP and left ventricular systolic dysfunction. In our study population, however, the correlation between BNP and left ventricular systolic dysfunction at the time of assessment was not as close as that found in other studies. It is possible
Interpretation of the scoring system in this study is limited by the population studied. We studied a population of patients prescribed loop diuretics in primary care who were willing and able to attend a research clinic for assessment (i.e. prevalent cases). This population was selected to include most patients in primary care with significant left ventricular systolic dysfunction, but will also include some patients with normal left ventricular function-prescribed loop diuretics for other reasons.\(^\text{9,29}\) The scoring system may not apply equally to other populations and would need to be more widely validated. Even those patients with the lowest scores had a higher mortality relative to age- and sex-matched controls in the region. However, a BNP scoring system would probably not be used as a population screening tool\(^\text{20}\) but rather to risk stratify patients identified for assessment by their primary care physician. The clinical judgement, which leads general practitioners to prescribe loop diuretics, may by itself identify a population who would benefit from further investigation.\(^\text{9}\) Figure 1 demonstrates how the study population was derived from the primary care population as a whole. We have previously shown that our study population is, in general, younger, more likely to be male, and more likely to have a history of angina than the community population as a whole.\(^\text{9}\) This is the result of the inevitable degree of selection bias inherent to any voluntary attendance study and underlines the importance of broader future validation of this prognostic score in different patient populations.

Another key question is how to use a scoring system such as that proposed in this study to modify patient management. Clearly, it would be useful if patients with the lowest scores could be managed in primary care, whereas those with the highest scores could be referred for more urgent out-patient hospital management. This scoring system is devised from prevalent cases in a primary care population-prescribed loop diuretics. Diuretic therapy has been shown to reduce levels of circulating natriuretic peptides in patients treated for class IV heart failure.\(^\text{31}\) Untreated, newly presenting, or incident cases may behave differently to our study patients. Any such scoring system would therefore have to be prospectively assessed before an algorithm of alternative approaches to patient management could be proposed. The separation of the Kaplan–Meier survival estimates with our simple score suggests that it may well be possible to develop a scoring system to risk stratify patients and prioritize patients for hospital management. However, it remains to be determined if those patients with the best prognostic scores can be safely managed without echocardiography. A prospective study of this type may also be able to establish whether more intense medical therapy or surveillance in patients with higher scores improves their subsequent prognosis. There is, at present, little evidence to support the use of BNP to monitor response to treatment in patients with chronic heart failure, although this is the subject of ongoing research.\(^\text{22}\)

Acknowledged limitations in this study include the fact that interpretation of ECG and echocardiography was performed by a single observer. Therefore, there is no
measure of inter-observer variability. We have previously reported that the simple classification of the ECG into 'normal' vs. 'abnormal' by a primary care physician is reproducible and shows little variation from that of a cardiologist. Thirty-eight patients in this study were lost to follow-up. Most had moved from the study area. We do not feel that this unduly influenced the results presented. Smoking data was not included in our scoring system owing to a large proportion of missing data from patient primary care records. Inclusion of smoking as a variable may further refine and strengthen the score. It is possible that inclusion of other parameters not included in our analysis, such as presenting symptoms or New York Heart Association class, could further improve the prognostic accuracy of this scoring system.

The results of this study conform to those of a previous study by Nielson et al. that followed up 126 patients for a mean of 4.3 years from a primary care population with a history of symptoms of heart disease. The authors found that combining elevated N-terminal atrial natriuretic peptide (using a cut-off value of 0.8 nmol/L) with an abnormal ECG improved the ability to predict a poor prognosis. This study also found that the predictive value of these factors varies with age. Our study population is more than four times that of Nielson and Hilden and had a longer follow-up (mean 6.7 years). As a result, there was almost 10 times the number of deaths during follow-up in our analysis. We also identify the prognostic predictive value of natriuretic peptides and an abnormal ECG in a similar population and show that age and sex affect this predictive power. We continue to use this data to derive a risk score which uses absolute values of BNP rather than a cut-off and includes age and sex and important aspects of past medical history.

There is growing data on the correlation between BNP and left ventricular systolic dysfunction and mortality. If BNP is to become a clinically useful test in primary care, it must become quick and easy to perform and interpret. The development of rapid bedside test kits should allow rapid assessment of BNP. Interpretation of the prognostic implication of the results may be facilitated by the development of simple scoring systems in primary care, using BNP and other information readily obtained.

Acknowledgements

The study was funded by a grant from the Trent Regional Health Authority and the Nottinghamshire Multidisciplinary Audit Advisory Group. The authors wish to thank Dr A. Cowley and Dr Hetmanski for their work and Professor J.R. Hampton for his advice during the preparation of this manuscript.

References


