European Heart Journal (2006) **27**, 330–337
doi:10.1093/eurheartj/ehi631

NT-proBNP testing for diagnosis and short-term prognosis in acute destabilized heart failure: an international pooled analysis of 1256 patients

The International Collaborative of NT-proBNP Study

James L. Januzzi[1*†], Roland van Kimmenade[2†], John Lainchbury[3], Antoni Bayes-Genis[4], Jordi Ordonez-Llanos[5], Miguel Santalo-Bel[6], Yigal M. Pinto[2], and Mark Richards[3]

[1] *Cardiology Division, Massachusetts General Hospital, Yawkey 5984, 55 Fruit Street, Boston, MA 02114, USA;* [2] *Cardiology Department, University Hospital, Maastricht, The Netherlands;* [3] *Christchurch Cardioendocrine Research Group, Department of Medicine, Christchurch School of Medicine and Health Sciences, Christchurch, New Zealand;* [4] *Cardiology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain;* [5] *Biochemistry Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; and* [6] *Emergency Medicine, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain*

Received 5 July 2005; revised 14 September 2005; accepted 13 October 2005; online publish-ahead-of-print 17 November 2005

KEYWORDS
Natriuretic peptides;
Diagnosis;
Prognosis

Aims Experience with amino-terminal pro-brain natriuretic peptide (NT-proBNP) testing for evaluation of dyspnoeic patients with suspected acute heart failure (HF) is limited to single-centre studies. We wished to establish broader standards for NT-proBNP testing in a study involving four sites in three continents.
Methods and results Differences in NT-proBNP levels among 1256 patients with and without acute HF and the relationship between NT-proBNP levels and HF symptoms were examined. Optimal cut-points for diagnosis and prognosis were identified and verified using bootstrapping and multi-variable logistic regression techniques. Seven hundred and twenty subjects (57.3%) had acute HF, whose median NT-proBNP was considerably higher than those without (4639 vs. 108 pg/mL, $P < 0.001$), and levels of NT-proBNP correlated with HF symptom severity ($P = 0.008$). An optimal strategy to identify acute HF was to use age-related cut-points of 450, 900, and 1800 pg/mL for ages <50, 50–75, and >75, which yielded 90% sensitivity and 84% specificity for acute HF. An age-independent cut-point of 300 pg/mL had 98% negative predictive value to exclude acute HF. Among those with acute HF, a presenting NT-proBNP concentration >5180 pg/mL was strongly predictive of death by 76 days [odds ratio = 5.2, 95% confidence interval (CI) = 2.2–8.1, $P < 0.001$].
Conclusion In this multi-centre, international study, NT-proBNP testing was valuable for diagnostic evaluation and short-term prognosis estimation in dyspnoeic subjects with suspected or confirmed acute HF and should establish broader standards for use of the NT-proBNP in dyspnoeic patients.

Introduction

Recently, use of both B-type natriuretic peptide (BNP) and its amino-terminal fragment, N-terminal pro-brain natriuretic peptide (NT-proBNP) has been found to be useful as an adjunct to standard clinical evaluation for the diagnosis and triage of dyspnoeic patients,[1–3] because these markers are considerably higher in patients with acute destabilized heart failure (HF). As such, the utility of serum testing for the natriuretic peptides has been recognized and incorporated in consensus documents and guidelines for the diagnosis and management of HF.[4–6]

Large-scale experience with BNP was reported previously, in a study of 1586 patients,[2] providing useful information concerning this marker in a heterogeneous, multi-national patient population. Although several similarly designed studies have been published supporting the use of NT-proBNP for the evaluation of dyspnoeic patients,[1,7,8] each was restricted to single centres and on occasion had limited numbers of patients. As such, optimal NT-proBNP concentrations for confirming or excluding the presence of acute destabilized HF in acutely dyspnoeic patients have yet to be definitively established. Further, the prognostic utility of NT-proBNP at presentation in the setting of acute destabilized HF, where very high levels of natriuretic peptides at presentation might obscure the prognostic value of the marker, remains undefined.

In order to clarify these issues, we performed this international multi-centre analysis of 1256 subjects to establish optimal NT-proBNP cut-points for diagnosis or exclusion of acute HF and to evaluate the prognostic

* Corresponding author. Tel: +1 617 726 3443; fax: +1 617 643 1620.
E-mail address: jjanuzzi@partners.org
† The first two authors contributed equally to this manuscript.

significance of elevated NT-proBNP in the setting of acute HF.

Methods

Component studies

The study population consisted of patients from three previously reported prospective clinical trials of NT-proBNP testing from Christchurch, New Zealand, Barcelona, Spain, and Boston, MA, USA, each performed to explore the use of NT-proBNP testing in dyspnoeic Emergency Department (ED) patients.[1,7,8] All three prospective trials had compatible inclusion/exclusion criteria and obtained similar clinical information. In addition, prospectively gathered data from 367 patients in a previously unpublished registry of patients with acute HF enrolled at the University Hospital of Maastricht, The Netherlands were included to complete the International Collaborative of NT-proBNP (ICON) data set.

The Christchurch study[8] comprised 205 patients presenting with dyspnoea to the ED. In this trial, the results of blinded NT-proBNP concentration were compared with a final adjudicated diagnosis, rendered utilizing the European Society of Cardiology Guidelines.[6] For the purposes of the present study, 195 subjects had complete data and were included for analysis. The Barcelona study[7] comprised 100 dyspnoeic patients presenting to the ED, and blinded NT-proBNP results were subsequently compared with the final diagnosis, which was assigned by a panel of physicians utilizing all available clinical data pertaining to each subject. Of the original 100 patients, 95 had complete data and were included in this study. Similar to the earlier trials, the ProBNP Investigation of Dyspnoea in the Emergency Department (PRIDE) study[1] was a prospective, blinded study of NT-proBNP testing performed in Boston, MA, USA, which examined 599 dyspnoeic subjects in the ED. Similar to the New Zealand and Barcelona studies, the final diagnosis for each patient was assigned by study cardiologists blinded to NT-proBNP levels, using available clinical data from presentation to 76-day follow-up. All patients from PRIDE were eligible for the present analysis.

The final source of data for the present analysis was the Second Maastricht Registry of Congestive Heart Failure (MARCH II). This prospective registry consists of patients with acute HF consecutively admitted to the University Hospital of Maastricht, between 1 January 2003 and 15 March 2004. During this period, 419 patients with acute HF and dyspnoea of varying severities were recruited. Of those enrolled, 367 had presenting NT-proBNP data available for analysis and were included for the present analysis.

Pooled methodology

All four component data sets had comparable information available, including standard demographics, past medical history and drug therapy, presenting symptoms and signs [including severity of breathlessness by the New York Heart Association (NYHA) classification], physical examination, results of serum chemistry tests, radiographic studies (typically plain chest radiographs), electrocardiography results, and finally, the results of NT-proBNP testing. Glomerular filtration rate (GFR) was estimated using the modified diet in renal disease (MDRD) equation.[9]

Follow-up for vital status among HF subjects was complete in 100% of subjects with acute HF from ICON through 76 days from presentation. In subjects from Christchurch, vital status was ascertained via review of hospital records as well as contact with caregivers or patients, when appropriate. In each case, data were corroborated utilizing the New Zealand Hospital Information Service and the Canterbury District Health Board Patient Management System database. In subjects from Barcelona, vital status was ascertained via review of medical records and phone follow-up with each subject. In subjects from Boston, follow-up was achieved utilizing review of hospital records and phone follow-up with each patient and/or their physicians. Deaths were corroborated using the Social Security Death Index. Finally, in subjects from Maastricht, follow-up was achieved utilizing hospital records as well as city hall reports for all deaths in Maastricht.

NT-proBNP testing

For each trial, blood was collected into EDTA tubes and NT-proBNP was measured using a validated, commercially available immunoassay (Elecsys® ProBNP, Roche Diagnostics, Indianapolis, IN, USA), using established methodology. This assay has been reported to have $<0.001\%$ cross-reactivity with bioactive BNP, and in the constituent studies in this report, this assay had inter-run coefficients of variation ranging from 0.9 to 5.5%. For the purposes of this report, NT-proBNP levels are expressed in pg/mL (to convert pg/mL to pmol/L, multiply by 0.118).

Measurement of troponin T (cTnT, Troponin T Stat®, Roche Diagnostics) was performed at each institution using standard methodology.

Statistical analysis

Comparisons of clinical characteristics between patients across all four studies were performed utilizing χ^2 tests for categorical data and the Wilcoxon rank sum test for continuous data. Similar methods were subsequently used to compare clinical characteristics between patients with and without acute HF in the final pooled analysis. Correlations between left ventricular ejection fraction and log-transformed NT-proBNP concentrations were performed using Spearman correlation. Comparisons of NT-proBNP levels between groups categorized by NYHA symptom severity and by final diagnostic categories were performed using the Kruskal–Wallis test. For these analyses, SPSS software (SPSS Inc., Chicago, IL, USA) was used.

Cut-points for diagnosis or exclusion of acute HF

Candidate NT-proBNP diagnostic cut-points were determined with the use of logistic regression analyses, with resulting receiver operating characteristic (ROC) curves. In each case, log-transformed NT-proBNP concentrations were used as the single independent variable in the logistic regressions, and to ensure that there were no spurious significant regression results due to multiple comparisons, the Bonferroni inequality was used to set the alpha threshold for declaring the log(NT-proBNP) regression coefficient significantly different from zero (i.e. the critical α value for a significant tail probability was set at 0.01 for each of the determinations, so that all cut-points would enjoy at most a joint 0.05 probability of a type I error).

For each logistic regression, log-transformed NT-proBNP results were entered in a single forward step, with the tail probability to enter set at $P = 0.01$ and to remove the effect from the regression at $P = 0.02$. These analyses were carried out with the STATA statistics package (STATA Corp., College Station, TX, USA).

Each of the estimation procedures was coded in the STATA programming language, then subjected to the STATA bootstrap prefix command for 100 bootstrap repeated random samples with replacement to estimate 95% confidence limits for each of the estimated cut-points in NT-proBNP. These limits are expressed as lower 95% confidence limit and upper 95% confidence limit. The bootstrap sample size was 1256 (the size of the entire data set). Each of the estimation procedures determined the threshold in estimated probability of acute HF that maximized predictive accuracy subject to different constraints on sensitivity or specificity for each of the cut-points. This threshold was then transformed with the use of the estimated logistic equation to a log(NT-proBNP) threshold, which was then, in turn, transformed to a diagnostic cut-point in NT-proBNP expressed at pg/mL.

The constraints on cut-point optimization were dictated by clinical considerations. The study population was divided into three clinically distinct sub-populations on the basis of age. For patients under 50 years, a minimum sensitivity of the NT-proBNP cut-point for diagnosing acute HF was set at 95% because of the recognition

of lower expected values in the setting of HF in younger subjects, together with the large potential adverse outcome of missing a young patient with acute HF. At the other extreme, patients older than 75 years required the emphasis on optimal specificity because of the recognition of the known age-related increase in basal NT-proBNP and the undesirability of incorrectly diagnosing and/or treating older patients for acute HF. In the intermediate group of patients, those 50–75 years old, a balance in sensitivity and specificity was required (while maximizing predictive accuracy within these constraints). Because gender has known independent effects on NT-proBNP levels,[10] we also examined further categorization by gender (and age) to improve accuracy of NT-proBNP for diagnosis.

For determining the cut-point for ruling out acute HF, a similar procedure as for diagnosis was used to identify an NT-proBNP concentration yielding optimal negative predictive value (NPV). As negative predictive value is subjected to a constraint based on specificity, a minimal, reasonable degree of specificity was required, and this constraint was set at 60% in bootstrap replications. In addition, we similarly examined the manufacturer-recommended cut-points for excluding HF.

Cut-points for short-term prognosis in acute HF

To determine a single prognostic cut-point for the prediction of mortality at 76 days, again, using log(NT-proBNP) as the sole predictor variable, predictive accuracy was maximized subject to a minimal specificity constraint, using bootstrap replications, as was done for identifying optimal cut-points for diagnosis. Once identified, the optimal cut-point was entered into multi-variable logistic regression analysis.

The STATA validation bootstrapping programme, SWBOOT, was used to perform 100 bootstrapped repeats of multi-variable stepwise logistic regression on candidate predictors of 76-day mortality. Factors entered into the analysis included source of data, elements from past and present medical history, symptoms and signs, medication use, and results of diagnostic studies including radiography, electrocardiography, haematology, and blood chemistries (including NT-proBNP results).

After the SWBOOT run, the resulting validated candidate independent variables were again entered stepwise into a logistic model in the same fashion described for the diagnostic cut-points. In this validation run, the aforementioned logistic regression procedure was used except that the alpha tail probability for an independent variable to enter the model was set at the conventional screening level of 0.05 and to remove of 0.1. Again, using the Bonferroni inequality to protect against type I errors, the tail probability of the test statistic was set at 0.01 to enter and 0.02 to remove the effect from the model. STATA bootstrap estimation was used to estimate 95% confidence intervals (CIs) for the odds ratios computed for the remaining independent predictors of 76-day mortality.

Linearity of the logit of proportion of patients with adjudicated acute HF or with 76 day mortality in log(NT-proBNP) was judged not to be an important issue, because each logistic regression fit was confirmed to be adequate by application of the Hosmer–Lemeshow goodness of fit test.

The Kaplan–Meier survival curves were constructed to compare 76-day mortality rates in groups divided as a function of NT-proBNP concentrations, using the log-rank test to compare the significance of the rates of mortality.

For all statistical analyses, all *P*-values are two-sided, with composite results less than 0.05 considered significant.

Results

The baseline demographics, past medical history, medication use, symptoms/signs, and lab results of the four study cohorts were reflective of the prevalence of acute HF in each study (*Table 1*). Notably, study subjects from the Boston trial were younger than the Christchurch, Barcelona and Maastricht patients ($P < 0.001$). *Table 1* details the demographics, past medical histories, medication use, symptoms/signs, results of physical examination, and outcomes of laboratory testing as a function of the source of the data.

Diagnoses, symptom severity, and NT-proBNP concentration

Of the 1256 dyspnoeic subjects in this pooled analysis, 720 (57.3%) had acute HF, whereas 536 (42.7%) did not. Of those without acute HF exacerbation at the time of enrolment, 55 subjects (4.4% overall) had a prior diagnosis of HF. The results of NT-proBNP testing between those with and without acute HF are detailed in *Figure 1*. The median NT-proBNP concentration of those patients with acute HF exacerbation (4639 pg/mL) was significantly higher than those with neither acute nor prior HF (108 pg/mL, $P < 0.001$ for difference). Notably, within the subjects without acute HF, 55 subjects were present with prior HF without acute exacerbation at the time of study enrolment. These 55 subjects had higher NT-proBNP concentrations [949 pg/mL, inter-quartile range (IQR) = 269–2590 pg/mL] than those without prior HF; however, the NT-proBNP concentrations among those with acute destabilized HF were significantly higher than in the 55 non-acute HF subjects with prior HF ($P < 0.001$ for difference).

Of those patients with acute HF at the time of enrolment ($n = 720$), 55 (7.6%) had NYHA Class II symptoms, whereas 348 (48.3%) had Class III symptoms and 317 (44.0%) had Class IV symptoms. *Figure 2* demonstrates the significant relationship between NYHA symptom severity and NT-proBNP levels; as symptom severity rose, a significant increase in median NT-proBNP levels was observed ($P = 0.008$), although significant overlap existed between the three groups.

Effect of left ventricular ejection fraction

Of the 720 subjects in ICON with acute HF, 655 (91%) had available left ventricular ejection fraction data. The relationship between left ventricular ejection fraction and NT-proBNP among subjects with acute HF in ICON is demonstrated in *Figure 3*; among these subjects a modest, but significant, relationship existed between ventricular function and natriuretic peptide concentrations ($r = -0.289$, $P < 0.001$).

Of those with available ejection fraction data, when categorized using a left ventricular ejection fraction of 50%, 293 (45%) subjects had preserved systolic function, whereas the balance had acute HF with impaired systolic function. Those subjects with acute HF but preserved left systolic function had lower NT-proBNP concentrations (3070 pg/mL, IQR = 1344–7974 pg/mL) when compared with those with impaired systolic function (6536 pg/mL, IQR = 2777–13407 pg/mL, $P < 0.001$ for difference). The overall sensitivity of NT-proBNP in subjects with preserved systolic function was 84% when compared with 92% in those with impaired systolic function. Of the subjects with non-systolic HF below the threshold for diagnosis, seven (2.4%) had NT-proBNP concentrations <300 pg/mL.

Table 1 Baseline demographics, past medical history, symptoms and signs, results of physical examination, and laboratory testing of the 1256 study subjects, categorized with respect to the centre of origin

Characteristics	Boston ($n = 599$)	Barcelona ($n = 95$)	Christchurch ($n = 195$)	Maastricht ($n = 367$)	Overall ($n = 1256$)
Acute destabilized heart failure (%)	35.0	84.2	33.8	100	57
Age (mean ± SD)	62.5 ± 17.1	70.9 ± 11.3	70.8 ± 14.3	76.0 ± 11.0	68.3 ± 15.9
Male gender (%)	50.8	60.0	48.7	51	51
African ethnicity (%)	7.3	0	0	0.5	3.7
Prior history					
Hypertension (%)	48.7	64.2	48.7	58.0	53
Coronary artery disease (%)	27.9	33.7	42.6	60.5	40
Prior myocardial infarction (%)	13.0	25.3	28.7	42.4	25
Prior HF (%)	25.0	44.2	26.2	50.7	34
COPD or asthma (%)	36.1	51.6	55.4	23.8	37
Tobacco use (past or present) (%)	62.4	56.8	69.7	43.1	56
Loop diuretic use (%)	29.5	48.4	39.5	62.3	42
Symptoms/signs					
PND (%)	12.4	61.1	13.8	28.1	21
Orthopnoea (%)	17.4	81.1	39.5	51.4	36
Lower extremity oedema (%)	17.2	54.7	32.8	49.3	32
Chest pain (%)	42.7	29.5	51.3	30.9	40
Cough (%)	36.9	55.8	46.7	24.8	36
Fever (%)	9.3	6.3	7.2	5.2	8
Sputum production (%)	9.5	55.8	41.5	13.6	19
Physical examination					
Pulse rate, mean ± SD (range)	87.6 ± 22.8	97.8 ± 26.4	90.7 ± 25.7	96.4 ± 26.8	91.4 ± 25.0
Jugular venous distension (%)	24.5	53.7	30.8	66.2	33
Rales on lung exam (%)	33.0	83.2	48.2	74.5	49
Wheezing on lung exam (%)	14.2	38.9	31.8	10.1	23
S3 gallop (%)	0.8	11.7	12.3	4.9	5
Lower extremity oedema (%)	24.5	54.7	32.8	63.2	39
ECG findings					
Sinus rhythm (%)	70.5	57.9	70.8	62.4	67
Atrial fibrillation or flutter (%)	12.5	33.7	18.5	37.3	22
Left ventricular hypertrophy (%)	4.0	9.5	4.1	12.8	7
Left bundle branch block (%)	5.7	14.7	10.3	18.3	11
Chest radiography findings					
Interstitial oedema (%)	16.9	66.3	15.5	23.4	22
Infiltrate/pneumonia (%)	14.2	15.8	5.6	8.4	11
Pleural effusion (%)	17.2	16.8	17.9	19.1	18
Cephalization of vessels (%)	8.0	66.3	17.4	33.0	31
Cardiomegaly (%)	18.3	69.5	33.3	33.0	24
GFR (mL/min/1.73 m^2) mean ± SD	74.0 ± 30.3	80.7 ± 22.7	96.7 ± 30.0	55.1 ± 23.5	72.5 ± 31.8
Troponin T (ng/mL) mean ± SD	0.04 ± 0.26	0.09 ± 0.27	0.25 ± 1.3	0.20 ± 0.64	0.12 ± 0.62

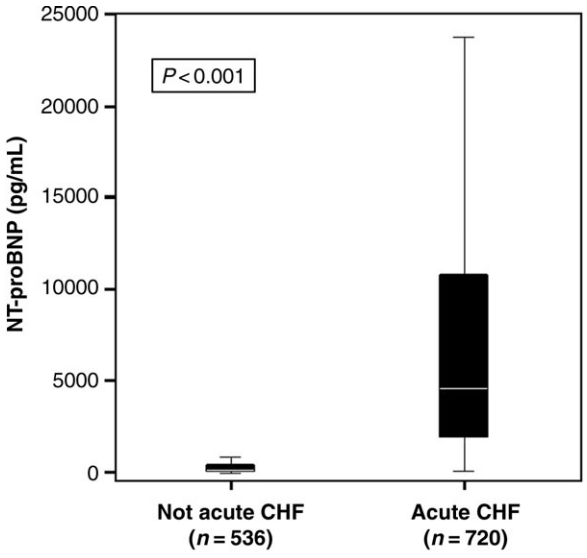
Cut-point analysis: identification of acute HF

Following bootstrapping methodology with a goal of optimal diagnostic accuracy across the wide age range in ICON, we selected cut-points emphasizing strong sensitivity for younger subjects, strong specificity for older subjects, and a balance of both for subjects at intermediate age ranges. In order to optimally confirm the diagnosis of HF, we found that dividing patients into three age groups (of <50, 50–75, and >75 years) yielded the optimal diagnostic accuracy, balanced with ease of use. Bootstrapping methods demonstrated that this 'triple cut-point' strategy was superior to using a single, age-independent cut-point (of 1243 pg/mL).

Notably, although female gender was associated with a trend towards higher median NT-proBNP levels among patients without acute HF (190 vs. 160 pg/mL, $P = 0.12$), the median and IQR of NT-proBNP concentration in female patients with acute HF ($n = 350$; 5801 pg/mL, IQR = 2300–12 149 pg/mL) were not significantly higher than those in male patients with acute HF ($n = 370$; 5645 pg/mL, IQR = 2411–11 215 pg/mL; $P = 0.66$), and there was no further advantage in sensitivity by adding gender stratification to age.

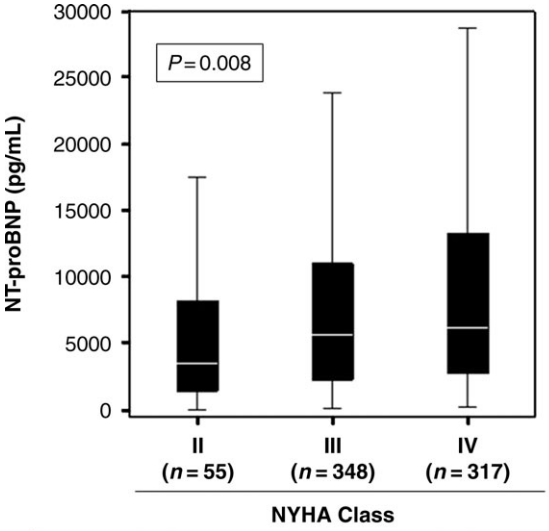
Example ROC curves derived from the proposed 'age-alone' diagnostic strategy are depicted in *Figure 4*. Among those patients <50 ($n = 184$), 50–75 years ($n = 537$), and >75 years of age ($n = 535$), NT-proBNP had an area under the ROC curve of 0.99, 0.93, and 0.86 for the diagnosis of acute HF (all $P < 0.0001$).

The bootstrap-validated optimal cut-points (with 95% CI) were estimated to be 450 (145, 1463 pg/mL), 900 (676, 1244 pg/mL), and 1800 pg/mL (1281, 2641 pg/mL) for the identification of acute HF in subjects aged <50, 50–75,



Diagnostic category	Median NT-proBNP	IQR
Not acute CHF	108 pg/mL	37–381 pg/mL
Acute CHF	4639 pg/mL	1882–10 818 pg/mL

Figure 1 NT-proBNP values between diagnostic groups in the ICON study.



NYHA Class	Median NT-proBNP	IQR
II	3512 pg/mL	1395–8588 pg/mL
III	5610 pg/mL	2260–11 001 pg/mL
IV	6196 pg/mL	2757–13 295 pg/mL

Figure 2 Median NT-proBNP concentrations in subjects with acute HF, expressed as a function of NYHA symptom severity. Boxes represent IQRs, while whiskers the 5th and 95th percentiles in each category.

and >75 years, respectively. The diagnostic sensitivities, specificities, positive predictive value (PPV) and NPV, and accuracies for these age-alone confirmatory cut-points are depicted in *Table 2*. Overall, this confirmatory strategy was associated with a sensitivity of 90%, a specificity of 84%, and a PPV of 88%. The accuracy for confirming acute HF using this strategy was 86%.

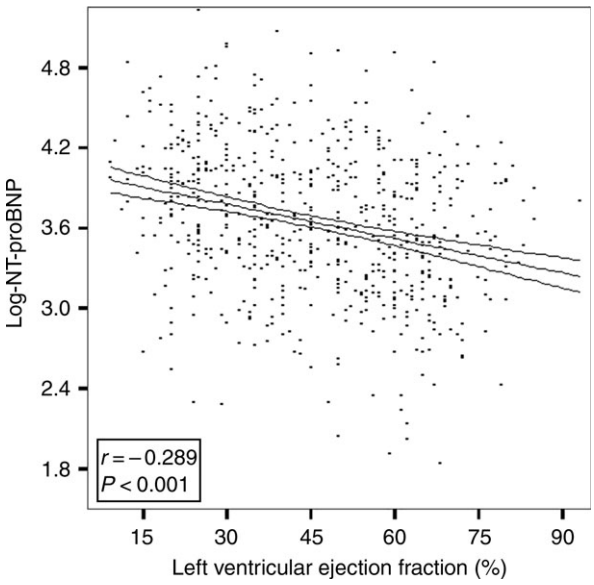


Figure 3 Linear regression relationship between systolic and NT-proBNP concentrations in the 655 subjects with acute HF and available left ventricular ejection fraction data.

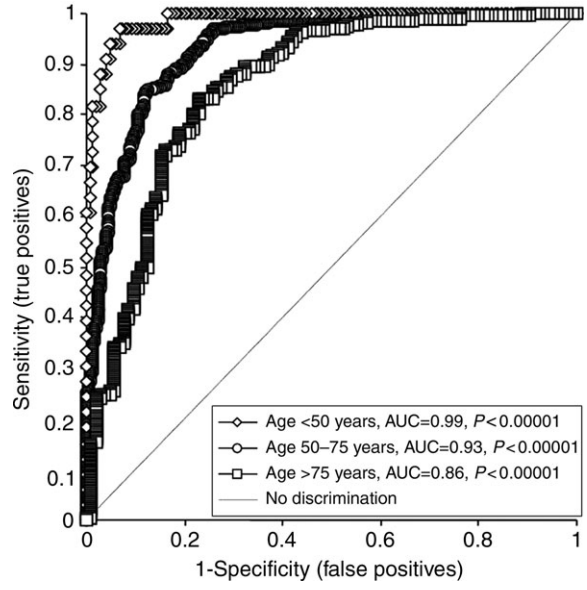


Figure 4 ROC curves for NT-proBNP-based diagnosis of acute HF across three age groups. As depicted, NT-proBNP had high AUC in each age group.

Cut-point analysis: exclusion of acute HF

In recognition of the value of NT-proBNP for the exclusion of acute HF, we examined manufacturer's recommended cut-points of 125 pg/mL for patients <75 years of age and 450 pg/mL for those ≥75 years of age and found them to have an overall NPV of 95%. After examining numerous age-related cut-point strategies, we found that using an age-independent approach for 'ruling out' acute HF in our cohort was superior, with a single cut-point of 300 pg/mL (95% CI = 241–369) demonstrating a sensitivity of 99%, a specificity of 60%, and an NPV of 98% (*Table 2*).

Table 2 Optimal NT-proBNP cut-points for the diagnosis or exclusion of acute HF among dyspnoeic patients

Category	Optimal cut-point	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Confirmatory ('rule in') cut-points						
<50 years ($n = 184$)	450 pg/mL	97	93	76	99	94
50–75 years ($n = 537$)	900 pg/mL	90	82	83	88	85
>75 years ($n = 535$)	1800 pg/mL	85	73	92	55	83
Rule in, overall		90	84	88	66	85
Exclusionary ('rule out') cut-point						
All patients ($n = 1256$)	300 pg/mL	99	60	77	98	83

Prognostic implications of NT-proBNP

Among 1256 subjects in the present analysis, 109 (8.6%) died within 76 days of presentation. Of these, 89 (81.7%) had acute HF. Among these patients with acute HF, the median NT-proBNP levels were higher in those patients dying by 76 days of follow-up (10 426 pg/mL, IQR = 5611–23 818 pg/mL) when compared with those surviving (4873 pg/mL, IQR = 2204–10 897 pg/mL; $P < 0.001$ for difference).

ROC analyses were performed in an effort to identify optimal NT-proBNP cut-points for prediction of short-term mortality in acute HF. Among those patients with acute HF ($n = 720$), ROC analysis of NT-proBNP to predict short-term mortality had an area under the curve (AUC) of 0.76 ($P < 0.0001$); the optimal NT-proBNP cut-point for predicting 76-day mortality in patients with acute HF was 5180 pg/mL, which was 68% sensitive and 72% specific for predicting short-term death, with a PPV of 19% and an NPV of 96%.

Multi-variable analysis: mortality

Top candidates predictive of mortality in acute HF based on bootstrap replications are detailed in *Table 3* and included results from testing of cTnT, NT-proBNP, haemoglobin, and renal function, as well as age. These candidates were then entered into multi-variable logistic regression analyses to determine their significance for predicting mortality in acute HF. Although significant in multi-variable analyses, both age (OR = 1.03, 95% CI = 1.0–1.05, $P = 0.02$) and serum creatinine (OR = 1.35, 95% CI = 1.02–1.81, $P = 0.036$) were rejected on the basis of the strict Bonferroni criterion of tail $P \leq 0.01$. Accordingly, the final independent predictors of 76-day mortality in acute HF are detailed in *Table 4*; for patients with acute HF ($n = 720$), the independent predictors of death by 76 days were haemoglobin (OR = 0.92, 95% CI = 0.87–0.97), cTnT > 0.03 ng/mL (OR = 3.4, 95% CI = 1.6–5.2), and NT-proBNP > 5180 pg/mL (OR = 5.2, 95% CI = 2.2–8.1).

The Kaplan–Meier curves depicting the differences in 76-day mortality rates between those patients with acute HF above and below the threshold of 5180 pg/mL are demonstrated in *Figure 5*. By 76 days, a significant difference in survival existed between the two groups (log-rank test, $P < 0.00001$), the curves for which diverged early and continued to separate by the end of the 76-day period.

Table 3 Top five candidate predictors of mortality in subjects with acute HF, based on 100 bootstrap replications

Factor	Number of replications out of 100 in which factor selected
A	
NT-proBNP >5180 pg/mL	100
Age	89
Creatinine	73
Haemoglobin	72
Paroxysmal nocturnal dyspnoea	64
B	
cTnT > 0.3 ng/mL	99
NT-proBNP > 5180 pg/mL	91
Haemoglobin	86
Age	72
Creatinine	69

As cTnT values were missing in 48 (6.7%) subjects, bootstrap replications were first performed without cTnT results (and are displayed, A), followed by the subsequent analysis in the presence of cTnT values (B).

Table 4 Independent predictors of 76-day mortality among those with acute HF

Predictor	Odds ratio	95% CI	*P*-value
NT-proBNP >5180 pg/mL	5.2	2.2–8.1	<0.001
Troponin T > 0.03 ng/mL[a]	3.4	1.6–5.2	<0.001
Haemoglobin	0.92	0.87–0.97	0.006

Neither age nor NYHA classification was independent predictor of short-term death in the presence of NT-proBNP results.
[a]Troponin T results missing in 48 (6.7%) subjects.

Discussion

Our results demonstrate the importance of NT-proBNP for the diagnosis and exclusion of acute HF in dyspnoeic patients and offer guidance concerning the importance of age stratification for improving diagnostic sensitivity for younger patients, while preserving the specificity of the test for older patients with dyspnoea and suspected acute HF. In addition, in our data set, the presenting NT-proBNP was not only useful for diagnosis, but also strongly predicted likelihood for short-term mortality in subjects with acute HF, with a more than five-fold increase in risk for death by 76 days among those with marked elevation in NT-proBNP concentrations.

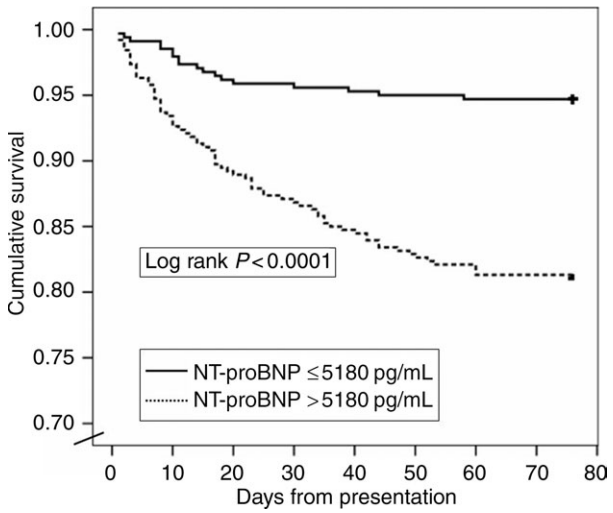


Figure 5 The Kaplan–Meier curves demonstrating survival rates of patients with acute HF ($n = 720$) during the first 76 days following presentation, expressed as a function of NT-proBNP concentration (log-rank test, $P < 0.001$).

The importance of age stratification to improve the accuracy of natriuretic peptide testing to confirm the diagnosis of acute HF is supported by the observation of a direct relationship between age and levels of both NT-proBNP and BNP.[11–13] This relationship between age and natriuretic peptide levels is likely consequent to age-related changes in left ventricular compliance,[14] as well as decreasing GFR.[15] Notably, although age stratification improves the ability of NT-proBNP to identify a high likelihood for acute HF, we found no need for age stratification to exclude the diagnosis. The confirmation and exclusion cut-points for NT-proBNP will help clinicians more confidently utilize the marker in the evaluation of the dyspnoeic patient, preserving sensitivity for younger patients with suspected HF, while optimizing specificity for elderly patients. Although age stratification for identifying acute HF might be expected to potentially increase the risk for a higher proportion of older subjects falling below the cut-points identifying acute HF, but above the cut-point excluding the diagnosis (a phenomenon referred to as the 'grey zone' for natriuretic peptides), our data demonstrate that in fact the overall incidence of a 'grey zone' diagnosis was only 16%, actually less than the 26% incidence reported for BNP, using a single cut-point.[16] In the situation of a 'grey zone' diagnosis, clinical judgement is often necessary to ascertain the correct diagnosis,[1,3] including consideration of situations in acute HF with lower than expected NT-proBNP concentrations, such as non-systolic HF[17] and obesity.[18] Indeed, in our analysis, those subjects with non-systolic HF had significantly lower NT-proBNP concentrations, with reduced sensitivity for confirming acute HF. Notably, among those with non-systolic HF below the threshold for acute HF, >97% had an NT-proBNP value above the 'rule out' cut-point of 300 pg/mL.

The large number of patients in our study both with and without acute HF afforded us the opportunity to further examine the relationship between gender and NT-proBNP levels. It is well known that among healthy subjects, NT-proBNP levels are higher in older female subjects when compared with age-matched male subjects,[12,19] possibly due to a higher prevalence of diastolic abnormalities or more significant age-related reductions in GFR in women. No significant gender-related differences in median or inter-quartile NT-proBNP levels were noted among our patients with or without acute HF, and the addition of gender stratification to age added no further diagnostic value for evaluating subjects.

The admission NT-proBNP concentration was so strongly predictive of short-term mortality among our patients with acute HF that its presence in multi-variable models overwhelmed the prognostic impact of other traditional risk factors for mortality, such as age or NYHA classification. The importance of NT-proBNP testing for prognosis among outpatients with chronic HF was recently described;[20–22] however, the role of NT-proBNP testing for risk stratification for patients with an acute exacerbation of HF remained incompletely understood. Most prior analyses examined the risk associated with persistent post-treatment elevation of NT-proBNP[23,24] and were restricted to a relatively small number of subjects. In one study, baseline NT-proBNP levels did not appear to predict short-term hazard.[24] In contrast, our study is more powered for such an analysis and demonstrates the importance of NT-proBNP concentration at presentation: although NT-proBNP was nearly universally elevated among patients with acute HF, among all factors analysed, marked elevation of NT-proBNP (essentially just above the median NT-proBNP concentration in our HF patients) was the single strongest predictor of death by just slightly longer than 2 months from presentation. Our data thus establish the importance of NT-proBNP at presentation not only for diagnosis, but also for simultaneous short-term risk assessment in acute HF.

With data now supportive of the utility of NT-proBNP in a wide variety of cardiovascular states, the importance of this marker is established. Although comparative studies are limited in number,[8,25,26] both NT-proBNP and BNP appear to deliver important diagnostic and prognostic information in a wide variety of patient types; the choice of which marker to use should be based on the differences in analytical performance, the needs of the institution utilizing the assays, and individual clinician comfort with the results from the assays.

Limitations of our study include the fact as a pooled analysis it lacks pre-defined endpoints, despite the similar designs and goals of the respective data sources. In this setting, the possibility of chance observations may be increased.[27] In addition, nearly half of the subjects were derived from a single study,[1] with attendant effects on demographics and results of the pooled analysis. Although every patient in the ICON study was dyspnoeic, further stratification of subjects using the NYHA classification represents a subjective measure, influenced by numerous factors, including the opinion of the caregiver evaluating the subject. Indeed, the inadequacy of physician-estimated measures of disease severity in acute HF is well described[28] and may explain, in part, the significant overlap of NT-proBNP concentrations between NYHA classifications in ICON. Finally, the cut-points proposed by the present study for identifying or excluding acute HF are similar to those from the PRIDE study,[1] with the exception of the addition of 1800 pg/mL for those >75 years. Notably, however, the addition of this cut-point is relevant, as the age-related effects on NT-proBNP results are significant, and the average age of patients with acute HF is rising.

In conclusion, in the first large-scale international analysis of NT-proBNP testing for evaluating patients with suspected acute HF, we demonstrate the utility of NT-proBNP testing for both diagnosis and exclusion of acute HF. As well, we demonstrate the value of NT-proBNP for the short-term estimation of risk for mortality in acute HF.

Acknowledgements

The authors wish to thank John Newell for expert statistical assistance, as well as the co-investigators and staff at their respective institutions. The ICON Study was supported partly by an unrestricted grant from Roche Diagnostics, Inc.

Conflict of interest: J.L.J. has received speaking, consulting and research grants from Roche Diagnostics, Dade-Behring, and Ortho Diagnostics; R.v.K., speaking and research grants from Roche Diagnostics; J.L., none declared; A.B.-G., speaking and research grants from Roche Diagnostics; J.O.-L., speaking, research, and consulting grants from Roche Diagnostics; M.S.-B., none declared; Y.M.P., speaking and research grants from Roche Diagnostics; M.R., speaking, research, and consulting grants from Roche Diagnostics.

References

1. Januzzi JL, Camargo CA, Anwaruddin S, Baggish AL, Chen A, Krauser D, Tung R, Cameron R, Nagurney JT, Lloyd-Jones DM, Chae CU, Melanson SF, Sluss P, Lewandrowski EL, Lewandrowski KB. The N-terminal ProBNP investigation of dyspnea in the Emergency Department (PRIDE) Study. *Am J Cardiol* 2005;**95**:948–954.
2. Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P, Omland T, Storrow AB, Abraham WT, Wu AH, Clopton P, Steg PG, Westheim A, Knudsen CW, Perez A, Kazanegra R, Herrmann HC, McCullough PA. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med* 2002;**347**:161–167.
3. McCullough PA, Nowak RM, McCord J, Hollander JE, Herrmann HC, Steg PG, Duc P, Westheim A, Omland T, Knudsen CW, Storrow AB, Abraham WT, Lamba S, Wu AH, Perez A, Clopton P, Krishnaswamy P, Kazanegra R, Maisel AS. B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from breathing not properly (BNP) Multinational Study. *Circulation* 2002;**106**:416–422.
4. Christenson R, Apple F, Cannon C, Francis G, Jesse R, Morrow D, Newby LKR, Storrow JA, Tang HW, Wu A. National Assocation of Clinical Biochemistry Laboratory Medicine Practice Guidelines: Biomarkers of Acute Coronary Syndrome and Heart Failure. http://www.nacb.org/lmpg/card_biomarkers_lmpg_draft_pdf.stm (2004).
5. Liu P, Arnold JM, Belenkie I, Demers C, Dorian P, Gianetti N, Haddad H, Howlett J, Ignazewski A, Jong P, McKelvie R, Moe G, Parker JD, Rao V, Rouleau JL, Teo K, Tsuyuki R, White M, Huckel V, Issac D, Johnstone D, LeBlanc MH, Lee H, Newton G, Niznick J, Ross H, Roth S, Roy D, Smith S, Sussex B, Yusuf S. The 2002/3 Canadian Cardiovascular Society consensus guideline update for the diagnosis and management of heart failure. *Can J Cardiol* 2003;**19**:347–356.
6. Remme WJ, Swedberg K. Guidelines for the diagnosis and treatment of chronic heart failure. *Eur Heart J* 2001;**22**:1527–1560.
7. Bayes-Genis A, Santalo-Bel M, Zapico-Muniz E, Lopez L, Cotes C, Bellido J, Leta R, Casan P, Ordonez-Llanos J. N-terminal probrain natriuretic peptide (NT-proBNP) in the emergency diagnosis and in-hospital monitoring of patients with dyspnoea and ventricular dysfunction. *Eur J Heart Fail* 2004;**6**:301–308.
8. Lainchbury JG, Campbell E, Frampton CM, Yandle TG, Nicholls MG, Richards AM. Brain natriuretic peptide and n-terminal brain natriuretic peptide in the diagnosis of heart failure in patients with acute shortness of breath. *J Am Coll Cardiol* 2003;**42**:728–735.
9. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999;**130**:461–470.
10. Emdin M, Passino C, Del Ry S, Prontera C, Galetta F, Clerico A. Influence of gender on circulating cardiac natriuretic hormones in patients with heart failure. *Clin Chem Lab Med* 2003;**41**:686–692.
11. Raymond I, Groenning BA, Hildebrandt PR, Nilsson JC, Baumann M, Trawinski J, Pedersen F. The influence of age, sex and other variables on the plasma level of N-terminal pro brain natriuretic peptide in a large sample of the general population. *Heart* 2003;**89**:745–751.
12. Redfield MM, Rodeheffer RJ, Jacobsen SJ, Mahoney DW, Bailey KR, Burnett JC Jr. Plasma brain natriuretic peptide concentration: impact of age and gender. *J Am Coll Cardiol* 2002;**40**:976–982.
13. Ray P, Arthaud M, Lefort Y, Birolleau S, Beigelman C, Riou B. Usefulness of B-type natriuretic peptide in elderly patients with acute dyspnea. *Intensive Care Med* 2004;**30**:2230–2236.
14. Arbab-Zadeh A, Dijk E, Prasad A, Fu Q, Torres P, Zhang R, Thomas JD, Palmer D, Levine BD. Effect of aging and physical activity on left ventricular compliance. *Circulation* 2004;**110**:1799–1805.
15. McCullough PA, Duc P, Omland T, McCord J, Nowak RM, Hollander JE, Herrmann HC, Steg PG, Westheim A, Knudsen CW, Storrow AB, Abraham WT, Lamba S, Wu AH, Perez A, Clopton P, Krishnaswamy P, Kazanegra R, Maisel AS. B-type natriuretic peptide and renal function in the diagnosis of heart failure: an analysis from the Breathing Not Properly Multinational Study. *Am J Kidney Dis* 2003;**41**:571–579.
16. Knudsen CW, Clopton P, Westheim A, Klemsdal TO, Wu AH, Duc P, McCord J, Nowak RM, Hollander JE, Storrow AB, Abraham WT, McCullough PA, Maisel AS, Omland T. Predictors of elevated B-type natriuretic peptide concentrations in dyspneic patients without heart failure: an analysis from the breathing not properly multinational study. *Ann Emerg Med* 2005;**45**:573–580.
17. O'Donoghue M, Chen A, Baggish AL, Anwaruddin S, Krauser DG, Tung R, Januzzi JL. The effects of ejection fraction on N-terminal ProBNP and BNP levels in patients with acute CHF: analysis from the ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) Study. *J Card Fail* 2005;**11**(Suppl. 5):9–14.
18. Krauser DG, Lloyd-Jones DM, Chae CU, Cameron R, Anwaruddin S, Baggish AL, Chen A, Tung R, Januzzi JL. Effect of body mass index on natriuretic peptide levels in patients with acute congestive heart failure: a ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) substudy. *Am Heart J* 2005;**149**:744–750.
19. McDonagh TA, Holmer S, Raymond I, Luchner A, Hildebrant P, Dargie HJ. NT-proBNP and the diagnosis of heart failure: a pooled analysis of three European epidemiological studies. *Eur J Heart Fail* 2004;**6**:269–273.
20. Hartmann F, Packer M, Coats AJ, Fowler MB, Krum H, Mohacsi P, Rouleau JL, Tendera M, Castaigne A, Anker SD, Amann-Zalan I, Hoersch S, Katus HA. 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* 2004;**110**:1780–1786.
21. Groenning BA, Raymond I, Hildebrandt PR, Nilsson JC, Baumann M, Pedersen F. Diagnostic and prognostic evaluation of left ventricular systolic heart failure by plasma N-terminal pro-brain natriuretic peptide concentrations in a large sample of the general population. *Heart* 2004;**90**:297–303.
22. Gardner RS, Ozalp F, Murday AJ, Robb SD, McDonagh TA. N-terminal pro-brain natriuretic peptide. A new gold standard in predicting mortality in patients with advanced heart failure. *Eur HeartJ* 2003;**24**:1735–1743.
23. Bettencourt P, Azevedo A, Pimenta J, Frioes F, Ferreira S, Ferreira A. N-terminal-pro-brain natriuretic peptide predicts outcome after hospital discharge in heart failure patients. *Circulation* 2004;**110**:2168–2174.
24. O'Brien RJ, Squire IB, Demme B, Davies JE, Ng LL. Pre-discharge, but not admission, levels of NT-proBNP predict adverse prognosis following acute LVF. *Eur J Heart Fail* 2003;**5**:499–506.
25. Mueller T, Gegenhuber A, Poelz W, Haltmayer M. Head-to-head comparison of the diagnostic utility of BNP and NT-proBNP in symptomatic and asymptomatic structural heart disease. *Clin Chim Acta* 2004;**341**:41–48.
26. Yeo KT, Wu AH, Apple FS, Kroll MH, Christenson RH, Lewandrowski KB, Sedor FA, Butch AW. Multicenter evaluation of the Roche NT-proBNP assay and comparison to the Biosite Triage BNP assay. *Clin Chim Acta* 2003;**338**:107–115.
27. Brophy JM, Dagenais GR, McSherry F, Williford W, Yusuf S. A multivariate model for predicting mortality in patients with heart failure and systolic dysfunction. *Am J Med* 2004;**116**:300–304.
28. Maisel A, Hollander JE, Guss D, McCullough P, Nowak R, Green G, Saltzberg M, Ellison SR, Bhalla MA, Bhalla V, Clopton P, Jesse R. Primary results of the Rapid Emergency Department Heart Failure Outpatient Trial (REDHOT). A multicenter study of B-type natriuretic peptide levels, emergency department decision making, and outcomes in patients presenting with shortness of breath. *J Am Coll Cardiol* 2004; **44**:1328–1333.