Establishing the discharge diagnosis of patients admitted for suspected heart failure: development of a simple score including BNP or NT-proBNP

Jochem Hogenhuis, Adriaan A. Voors, Tiny Jaarsma, Arno W. Hoes, Mike J. de Jongste, Hans L. Hillege, Dirk J. van Veldhuisen
Abstract

**Background:** Previous studies showed that 18% - 33% of patients admitted for suspected heart failure (HF) receive a false-positive discharge diagnosis of heart failure. We aimed to identify the most important independent predictors of a (false) HF discharge diagnosis to construct a score to improve the discharge diagnosis of HF.

**Methods:** Using univariable and multivariable logistic regression and Receiver Operating Characteristics (ROC) analysis, the value of potentially important diagnostic determinants including B-type Natriuretic Peptide (BNP) and N-terminal proBNP (NT-proBNP) was assessed and combined in a scoring rule.

**Results:** In the present study population (n = 540, 71 ± 11 years of age, 62% male), an outcome panel identified 6% with a discharge diagnosis other than HF. Univariable predictors of the discharge diagnosis HF were dyspnoea at rest, absence of COPD/asthma, absence of anaemia, renal dysfunction, BNP > 100 pg/ml and NT-proBNP > 300 pg/ml. In multivariable logistic analyses, all of these, except BNP > 100 pg/ml, remained independent predictors of a HF discharge diagnosis. The ROC area of a model combining these predictors was 0.78.

**Conclusion:** In our population, a discharge diagnosis other than HF occurred in only 6% of the patients. Four readily available patient characteristics together with NT-proBNP are independent predictors of a discharge diagnosis of patients admitted for suspected HF, and these were combined in a scoring rule to optimize the discharge diagnosis of these patients. This score should be validated in other populations before its use in clinical practice can be recommended.

Introduction

Discharge diagnoses have been instrumental in providing estimates of and time trends in the incidence of heart failure (HF). However, studies in the Netherlands, the United States and in Sweden showed that respectively 20%, 33% and 18% of the patients hospitalized for HF receive a false-positive discharge diagnosis of HF,¹-³ which may suggest that up to one third of the patients in these studies may be discharged without optimal medical treatment, which may lead to inappropriate care and unnecessary high expenses.

Although B-type Natriuretic Peptide (BNP)⁴ and N-terminal pro BNP (NT-proBNP)⁵ are known to be useful in the emergency diagnosis of HF, little evidence is available on their diagnostic value in establishing the discharge diagnosis in patients admitted for suspected HF. When BNP or NT-proBNP is an independent predictor of a HF discharge diagnosis in these patients, it could be combined with routinely available patient characteristics to develop a scoring rule to be applied in clinical practice. Therefore, we aimed to: 1) investigate whether NT-proBNP, BNP and routinely available patient characteristics, including co-morbidity, are independent predictors of the discharge diagnosis of HF in patient admitted with suspected HF and 2) develop a score combining the most important predictors of a HF discharge diagnosis.
Methods

Study population
All 540 patients in this study were recently admitted for suspected HF (NYHA II-IV), and were included in a multicenter HF trial conducted in the Netherlands (COACH). Data for the present study were collected before randomization. All participating sites (n = 17) were centres with vast experience in HF care. Patients were at least 18 years of age and admitted with a clear suspicion of heart failure. The design, inclusion and exclusion criteria of the trial have been described in detail elsewhere. In short, COACH is a randomized controlled trial investigating the effect of education and counselling on readmission for HF and mortality.

Availability and measurement of BNP and NT-proBNP levels
Of the 1049 patients included in the COACH, 541 patients had both BNP and NT-proBNP plasma levels available. Temporary, usually in the starting phase of the study, unavailability of the necessary laboratory facilities was the main reason for missing BNP and NT-proBNP data. Blood was collected shortly before discharge between 8:00 AM and 4:00 PM, after patients had been clinically stabilised and recovered enough to go home. Ten millilitres of whole blood was taken from an antecubital vein and collected into tubes containing potassium ethylenediaminetetraacetic acid (EDTA; 1 mg/ml blood) when patients were in a supine position. The tubes were centrifuged for 10 minutes (2500 x g) and the plasma was separated and stored in polyepropene tubes at -70°C to -80°C.

BNP measurement: In 364 out of the 541 patients, BNP levels were determined on site in whole blood within 4 hours after blood collection. In 177 out of the 541 patients BNP levels were determined in plasma at the Core Laboratory at the University Medical Center Groningen. All measurements were performed using a fluorescence immunoassay kit (Triage®, Biosite Incorporated, San Diego, CA). Details on the system provided by the manufacturer indicated the analytical sensitivity of the assay is less than 5.0 pg/ml. The system has been extensively validated. The measurable range of the BNP assays was 5.0-5000.0 pg/ml.

NT-proBNP measurement: All measurements of NT-proBNP levels were performed in plasma at the Core laboratory on an Elecsys™ 2010 analyser, a commercially available electrochemiluminescent sandwich immunoassay (Elecsys proBNP, Roche Diagnostics, Mannheim, Germany). The intra-assay precision (coefficient of variation) is 1.2 – 1.5%, and the inter-assay precision (coefficient of variation) is 4.4 – 5.0%, with an analytical range of 5 – 35000 pg/ml.

Renal function
Serum creatinine was determined from the blood drawn shortly before discharge, in the local laboratory at each centre. Estimated Glomerular Filtration Rates (eGFR’s) were calculated using the Levey – modified Modification of Diet in Renal Disease formula:

eGFR (ml/min/1.73m²) = 186 x SCR\(^{1.154}\) x age\(^{-0.203}\) x (0.742 if female) x (1.21 if black)

Renal dysfunction was defined as an eGFR < 60 ml/min/1.73m².
Anaemia
The definition of anaemia according to the World Health Organisation was used; haemoglobin levels < 7.5 mmol/l (12 g/dl) for women and < 8.1 for men (13 g/dl).\textsuperscript{11}

Determination of discharge diagnosis i.e. reference (or ‘gold’) standard
Two cardiologists that were blinded to BNP and NT-proBNP plasma levels reviewed the discharge reports of the 540 patients that were admitted for suspected HF. These reports included information on physical examination, medical history and the results of ECG, radiological examinations, laboratory blood tests, echocardiography and effects of treatment. To determine the actual discharge diagnoses of these patients, the cardiologists independently classified the discharge diagnosis as ‘HF’ or ‘other discharge diagnosis’. In case of disagreement about the discharge diagnosis between the two cardiologists, they discussed the discharge diagnosis until agreement was reached. The cardiologists could not draw any conclusion about the discharge diagnosis of one patient because the discharge report did not contain sufficient information. This patient was excluded from the analysis.

Statistical analyses
To prevent chance findings we a priori decided to study a limited number of potential diagnostic determinants. Selection of the 11 variables was based on the literature and availability in every day practice. Univariable logistic analyses were performed to identify which variables were associated with the discharge diagnosis HF. Because BNP and NT-proBNP levels had a skewed distribution, the levels were dichotomized (BNP below or above or equal to 100 pg/ml\textsuperscript{4} and NT-proBNP below or above or equal to 300 pg/ml\textsuperscript{5} was used. A p-value <0.15 was required to enter a variable into the multivariable logistic analyses. Either BNP or NT-proBNP was included in the multivariable regression analysis (and not both of these) because of the danger of multicollinearity and, first and foremost, because in clinical practice usually only one of the two is available. A Lemeshow and Hosmer goodness of fit statistic was calculated to evaluate the fit of the multivariable logistic regression model. The diagnostic performance of the combination of independent predictors (p-value < 0.05 in the multivariable analysis) was quantified by calculating the area under the receiver operating characteristic (ROC) curve. An ROC area of 0.5 signifies no discriminatory value (like a flip of a coin), while an area of 1.0 means perfect discrimination between those with and without a HF discharge diagnosis.

Since any model is too optimistic when the model is used on the dataset from which it was developed (over fitting), we used bootstrapping techniques in order to internally validate the final multivariate model and to adjust the estimated ROC area under the curve and regression coefficients. The performance of a model after bootstrapping is more in concordance with the performance that can be expected in future patients.\textsuperscript{12,13}

Eighteen patients had 27 missing values. Since missing values usually do not occur at random, deletion of subjects with a missing value may lead to biased results and loss of power. Hence, we imputed any missing values by using a regression method with the addition of a random error term (SPSS version 11.0). The imputation was based on the correlations between each variable with missing values and all other variables from the 522 (97%) complete datasets.

In order to create an easy applicable diagnostic rule or score, we transformed the original
regression coefficients (after adjustment for over fitting) of the variables in the final model to points according to their relative contribution to the estimation of the discharge diagnosis. We calculated the score for each patient and estimated the absolute percentages of correctly diagnosed patients across score categories. To allow the construction of a diagnostic points score haemoglobin and eGFR were dichotomised into presence of anaemia (according to WHO definition) and renal dysfunction ($< 60 \text{ ml/min/1.73m}^2$) respectively.

The analyses were performed with S-Plus version 6.1 except for the imputation procedure, as stated above.

**Results**

Demographic and clinical characteristics of the 540 patients are presented in table 1. Mean age of the 540 patients was $71 (\pm 11)$ years, and more than half of the population was male (62%). The characteristics were not significantly different between patients with available BNP/NT-proBNP levels (n=540) and the total patient group included in COACH (n=1049).

The consensus panel classified 506 patients (94%) with a discharge diagnosis of heart failure, and in 34 patients (6%) another discharge diagnosis was established. These included: acute myocardial infarction, unstable angina pectoris, atrial fibrillation, chronic obstructive pulmonary disease (COPD)/asthma and pulmonary embolism.
The median BNP value of the patients with HF as discharge diagnosis was 459 pg/ml (interquartile range 212 – 938 pg/ml), and the median BNP value of the patients with another discharge diagnosis was 297 pg/ml (interquartile range 135 – 632 pg/ml). The proportion of patients with BNP > 100 pg/ml differed between the two groups (91% vs. 79%; P=0.03). The median NT-proBNP value of the patients with a HF discharge diagnosis was 2718 pg/ml (interquartile range 1379 – 6168 pg/ml), while the value in patients with another discharge diagnosis was 1501 pg/ml (interquartile range 739 – 4920 pg/ml). The proportion of patients with NT-proBNP > 300 pg/ml was 97% and 82%, respectively (p<0.01) (table 1).

The following variables were associated with the discharge diagnosis 'HF' (p<0.15) in the univariable analysis: BNP > 100 pg/ml, NT-proBNP > 300 pg/ml, dyspnoea at rest (NYHA class IV) at hospital admission, absence of COPD / asthma, higher haemoglobin and lower eGFR (table 1). After bootstrapping, exactly the same variables showed p-values < 0.15, although the individual odds ratios slightly changed.

All of these, except BNP, were independent predictors of the discharge diagnosis HF the multivariable analyses, and the Lemeshow and Hosmer goodness of fit statistic of the model was 5.0 (p=0.76). The ROC area of the combination of these 5 items was 0.78 (95% confidence interval 0.70-0.86; figure 1). Internal validation of the multivariable model by means of bootstrapping techniques yielded an ROC area of 0.67. Regression coefficients were rounded tot whole points according to their relative contribution to the discharge diagnosis (table 2). The score ranged from 0 to 14 point. For each score category, sensitivity, specificity and predictive values were calculated (table 3).


**Discussion**

Dyspnoea at rest (NYA class IV), absence of COPD/Asthma, renal dysfunction, absence of anaemia and NT-proBNP > 300 pg/ml are independent predictors of a discharge diagnosis of HF in patients admitted to hospital with suspected HF.

In contrast to previous studies, where 18%, 20% and 33% of HF discharge diagnoses were false-positive,\(^1\),\(^2\) we found that only 6% of the patients that were admitted with suspected HF did not have HF as the discharge diagnosis after close examination of all available information. This is at least partly attributable to the patient selection typically involved in randomized trials and more the rigorous evaluation (more intensive monitoring) of patients in such studies. In addition, it also reflects the experience of the participating clinics and cardiologists. Because of the low number of non HF discharge diagnoses we limited the number of diagnostic determinants to be evaluated.

---

**Table 2:** Independent predictors (and their points assigned in the score) of a discharge diagnosis of heart failure in patients admitted with suspected heart failure. Results from multivariable logistic regression analyses.

<table>
<thead>
<tr>
<th>Variables</th>
<th>OR (95% CI)</th>
<th>P-value</th>
<th>Points for rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of anaemia</td>
<td>2.10 (0.97 - 4.54)</td>
<td>0.02</td>
<td>2</td>
</tr>
<tr>
<td>NYHA class IV</td>
<td>2.51 (0.98 - 6.43)</td>
<td>0.01</td>
<td>3</td>
</tr>
<tr>
<td>Absence of COPD/Asthma</td>
<td>2.01 (0.96 - 4.19)</td>
<td>0.01</td>
<td>2</td>
</tr>
<tr>
<td>Renal dysfunction (eGFR&lt;60ml/min/1.73m²)</td>
<td>2.07 (0.97 - 4.40)</td>
<td>&lt;0.01</td>
<td>2</td>
</tr>
<tr>
<td>NT-proBNP &gt;300pg/ml</td>
<td>4.95 (1.62 - 15.15)</td>
<td>&lt;0.001</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 3:** Sensitivity, specificity and predictive values of different categories of the score.

<table>
<thead>
<tr>
<th>Risk category of simplified model (points)</th>
<th>Heart failure as discharge diagnosis (n=506)</th>
<th>Other discharge diagnosis (n=34)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low (0) n=0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low (2-5) n=24</td>
<td>16 (67%)</td>
<td>8 (33%)</td>
<td>96%</td>
<td>6%</td>
<td>94%</td>
<td>9%</td>
</tr>
<tr>
<td>Medium (6-9) n=222</td>
<td>203 (91%)</td>
<td>19 (9%)</td>
<td>93%</td>
<td>26%</td>
<td>95%</td>
<td>21%</td>
</tr>
<tr>
<td>High (10-14) n=276</td>
<td>270 (98%)</td>
<td>6 (2%)</td>
<td>53%</td>
<td>18%</td>
<td>91%</td>
<td>2%</td>
</tr>
</tbody>
</table>
Our study identified other determinants of a diagnosis of HF than studies in other settings. Previous studies in an emergency department setting showed that age, history of HF, previous myocardial infarction, hypertension, rales, pulmonal congestion and orthopnea were independent predictors of the diagnosis of HF, while in our study NYHA class and several co-morbidities (COPD/asthma, anaemia, renal dysfunction) were the most powerful predictors (table 2). This difference could be explained by the difference in clinical setting. All patients in the present study were admitted with suspected HF and therefore almost all patients had evidence of cardiac disease. In contrast, about half of the patients in the other studies did not have HF at all. A remarkable finding of the present study is that the haemoglobin plasma levels were lower in the patients with other discharge diagnoses. The fact that anaemia causes symptoms similar to HF, notably dyspnoea and fatigue may have led to a suspected diagnosis of HF at admission. A similar phenomenon is known from patients with COPD and also shown in our data.

The well established emergency department BNP rule out value of 100 pg/ml and the NT-proBNP rule out value of 300 pg/ml were univariable predictors of the discharge diagnosis HF, but in multivariable analysis only NT-proBNP > 300 pg/ml remained a statistically significant determinant. Our results for NT-proBNP confirm the findings from a recent study by Baggish et al. who developed a score for the diagnosis of HF in an acute HF setting that included elevated NT-proBNP.

The reference ('gold') standard that was used in the present study to classify the discharge diagnosis is not perfect due to subjectivity. However, an outcome panel is a common way to classify the definite diagnosis and it has been used in many other diagnostic studies, including many in the field of heart failure. Furthermore, according to the Standards for Reporting of Diagnostic Accuracy (STARD) initiative, a consensus panel is the best proxy reference in the absence of an ideal standard. Although by bootstrapping the final model we adjusted the model for over-fitting, this final model and the points score needs to be validated externally in a new sample of patients that is about to be discharged from the hospital after admission for suspected HF.

We observed high positive predictive values and low negative predictive values for our score categories, which indicates that this score is more a test to confirm the discharge diagnosis HF than to reject it. In case of local availability of BNP instead of NT-proBNP, the score can be adjusted, although, as mentioned above, the performance of the model including BNP (ROC area 0.76) is worse than for the model including NT-proBNP (ROC area 0.78). The corresponding score for BNP is: absence of anaemia 2 points, NYHA class IV 4 points, absence of COPD/asthma 2 points, renal dysfunction 3 points and BNP > 100 pg/ml 3 points.

We conclude that a limited number of readily obtainable data from physical examination or medical history (dyspnoea at rest and absence of COPD/asthma) together with renal dysfunction, and absence of anaemia and the NT-proBNP value > 300 pg/ml may improve the assessment of the discharge diagnosis of patients admitted with suspected HF, and thus optimize therapeutic interventions in these patients after discharge. Although we internally validated the score by means of bootstrapping, it needs to be validated externally in a new sample of patients that is about to be discharged from the hospital after admission for suspected HF, before it can be recommended for clinical practice.
Acknowledgments

The NHF-COACH study is financially supported by the Netherlands Heart Foundation (Grant 2000Z003). Prof. Van Veldhuisen is an Established Investigator of the Netherlands Heart Foundation (Grant D97.017). We are indebted to Roche Diagnostics (Mannheim, Germany) for providing NT-proBNP assay kits. We are grateful to Biosite Incorporated (San Diego, CA) for providing BNP assay kits, and to Novartis (Arnhem, the Netherlands) for an unrestricted grant to invest in BNP Triage® meters. We gratefully acknowledge Peter Zuithoff of the Julius Center for Health Sciences and Primary Care for his statistical advice.
References


