

A validated clinical and biochemical score for the diagnosis of acute heart failure: The ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) Acute Heart Failure Score

Aaron L. Baggish, MD,^a Uwe Siebert, MD, MSc, MPH,^b John G. Lainchbury, MD,^d Renee Cameron, MS,^a Saif Anwaruddin, MD,^a Annabel Chen, MD,^a Daniel G. Krauser, MD,^a Roderick Tung, MD,^a David F. Brown, MD,^c A. Mark Richards, MD, PhD,^d and James L. Januzzi Jr, MD^a *Boston, MA; and Christchurch, New Zealand*

Background No method integrating amino-terminal pro-brain natriuretic peptide (NT-proBNP) testing with clinical assessment for the evaluation of patients with suspected acute heart failure (HF) has been described.

Methods Amino-terminal pro-brain natriuretic peptide results and clinical factors from 599 patients with dyspnea were analyzed. The β coefficients of the 8 independent predictors of HF were used to assign a weighted integer score for predictor. The sum of these integers provided a diagnostic HF "score" for each patient. Receiver operating characteristic curve analysis determined the optimal cut point for the diagnosis of acute HF. The performance of the score was evaluated in the development cohort and subsequently in a patient population from a separate clinical trial of patients with dyspnea conducted in Christchurch, New Zealand.

Results Eight factors comprised the score: elevated NT-proBNP (4 points), interstitial edema on chest x-ray (2 points), orthopnea (2 points), absence of fever (2 points), loop diuretic use, age >75 years, rales, and absence of cough (all 1 point). Median scores in patients with acute HF were higher than those without acute HF (9 vs 3 points, $P < .001$). At a cut point of ≥ 6 points, the score had a sensitivity of 96% and a specificity of 84% for the diagnosis of acute HF ($P < .001$). The score improved diagnostic accuracy over NT-proBNP testing alone and retained discriminative capacity in patients in whom clinical uncertainty was present. Lastly, the accuracy of the score was validated in the external data set of patients with suspected acute HF.

Conclusion We report a simple and accurate scoring system combining NT-proBNP testing and clinical assessment for the diagnosis or exclusion of acute HF in patients with dyspnea. (*Am Heart J* 2006;151:48-54.)

The signs and symptoms of acute heart failure (HF) are frequently nonspecific, highly variable, and may also be observer dependent, thereby rendering accurate diagnosis a significant clinical challenge.^{1,2} Testing for amino-terminal pro-brain natriuretic peptide (NT-proBNP) is a useful adjunct to routine assessment for differentiating acute HF from other etiologies of dyspnea.³⁻⁵ However, factors such as comorbid illnesses, age, renal failure, body mass, and baseline systolic

dysfunction may affect NT-proBNP levels in manners that obscure the diagnosis of acute HF when this marker is used in isolation.⁶

A combination of NT-proBNP or brain-type natriuretic peptide testing and standard clinical assessment has been suggested to be superior to either tool used in isolation^{4,7}; however, the optimal method to combine natriuretic peptide testing and clinical factors of proven value for predicting a diagnosis of HF has not been accomplished. One such method might be to combine natriuretic peptide testing and clinical evaluation in a scoring system. Clinical scoring systems have been demonstrated to be useful in numerous clinical situations including acute aortic dissection, contrast-induced nephropathy, and acute coronary syndromes.⁸⁻¹⁰

We hypothesized that a diagnostic scoring system integrating NT-ProBNP testing and routine clinical assessment would optimize diagnostic accuracy for detecting acute HF among patients presenting to an

From the ^aDepartment of Medicine, ^bInstitute for Technology Assessment and ^cDepartment of Emergency Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, and ^dDepartment of Medicine, Christchurch Hospital, Christchurch, New Zealand.

Supported by a grant from Roche Diagnostics, Indianapolis, IN.

Submitted November 28, 2004; accepted February 15, 2005.

Reprint requests: James L. Januzzi Jr, MD, Department of Medicine, Massachusetts General Hospital, Yawkey 5800, 55 Fruit Street, Boston, MA 02114.

E-mail: jjanuzzi@partners.org

0002-8703/\$ - see front matter

© 2005, Mosby, Inc. All rights reserved.

doi:10.1016/j.ahj.2005.02.031

emergency department (ED) with acute dyspnea. We report a simple, accurate, and novel tool incorporating NT-proBNP testing and routine clinical assessment for this purpose.

Methods

Study design

The patients studied for the present analysis were derived from a prior study of NT-proBNP testing for the evaluation of patients with dyspnea, the PRIDE study.⁷ The design, results, and conclusions of this study have been described.⁷ Briefly, 599 consecutive patients with dyspnea were enrolled in a trial examining the value of NT-proBNP testing compared with clinical judgment for the identification of acute HF. Exclusion criteria for the study included age <21 years, severe renal insufficiency (defined as a serum creatinine >2.5 mg/dL), dyspnea after chest trauma, dyspnea secondary to severe coronary ischemia (identified with >0.1 mV ST-segment elevation or ST-segment depression on 12-lead electrocardiogram, if performed on presentation), >2-hour time delay after urgent intravenous loop diuretic administration (above any baseline maintenance dose), and unblinded natriuretic peptide level measurement. Clinical data and a blinded NT-proBNP level were obtained for each study participant. At the end of standard clinical assessment, the managing physician in the ED was asked to provide a diagnosis and the percent likelihood of acute HF for each patient (range 0%-100%). These results were compared with NT-proBNP test results.

Statistical analysis of the PRIDE database was used to identify independent predictors of acute HF. These variables were used in the present study to generate a weighted scoring system for the diagnosis and exclusion of acute HF.

Methods of measurement

After enrollment, clinical characteristics obtained in the ED were recorded. These included demographics, symptoms (including New York Heart Association symptom severity), signs, medical history, medication use, and diagnostic studies such as electrocardiogram, chest x-ray, and standard blood test results. At the time of enrollment, a 5-mL sample of blood was collected into a tube containing EDTA for blinded measurement of NT-proBNP. This was performed with a commercially available immunoassay (Elecsys proBNP assay, Roche Diagnostics, Indianapolis, IN) on a Roche Elecsys 2010 analyzer, using established methodology.

Diagnostic score development

Potential predictors of a final diagnosis of acute HF were identified from the PRIDE study database.⁷ Univariable screening was performed on 38 baseline clinical and biochemical variables to identify potential predictors of acute HF. Odds ratios for the presence of acute HF were generated and expressed with 95% CI. Candidate variables identified through univariable analysis (those with *P* values <.1) were entered into a multivariable model, using forward stepping logistic regression, to identify independent predictors of a final diagnosis of acute HF. Verification of goodness of fit was confirmed with the Hosmer-Lemeshow test. All candidate variables were started out of the logistic model and each was entered in order

of largest to smallest test χ^2 statistic only, whereas the maximum likelihood estimate of the corresponding regression parameter was significantly different from zero at *P* < .05. An independent variable was removed from the model only when its corresponding regression parameter was not significantly different from zero at *P* > .1. First-order testing for interactions was performed to evaluate for interaction(s) between candidate independent variables. Although not statistically significant in the multivariable model, a prior diagnosis of HF was forced into the multivariable model because of its strong potential as a predictor of acute HF.

After the identification of significant independent predictors of a final diagnosis of acute HF, the β coefficients generated for each variable from logistic regression were rounded to the nearest whole integer to generate a relative weight for each variable. An integer score, composed of the sum of applicable weighted variables, was calculated for each individual.

Assessment of score performance

The observed prevalence of HF at each score level was first compared with the predicted prevalence of HF for each score parameter from the logistic model generated from the derivation cohort. Receiver operating characteristic curve analysis was used to determine the score value with the optimal sensitivity and specificity for the diagnosis of acute HF. The score was further examined among several important subgroups including patients with and without acute HF, in those patients in whom clinicians reported uncertainty of the presence of acute HF (certainty range >10%-<90%) at the time of first evaluation, as well as in those patients in whom NT-proBNP was either falsely elevated or low.

As an elevated NT-proBNP was the dominant variable in the HF score model, its role in the prediction of acute HF was examined both as a component of the score as well as an independent variable. The relationship between NT-proBNP values and the absolute score results was examined using Pearson correlation, whereas the median NT-proBNP value for each score category was calculated and compared with other score categories with nonparametric Kruskal-Wallis testing. The additive value of the score relative to NT-proBNP testing alone was examined with respect to reduction in false-negative and false-positive diagnoses of acute HF.

Lastly, to externally validate the PRIDE Acute HF Score, we calculated the HF scores of patients in a prior prospective study of NT-proBNP testing in patients with dyspnea from Christchurch, New Zealand.³ These patients were compared with the PRIDE cohort, using χ^2 tests for categorical variables and the Wilcoxon rank sum test for continuous variables. Factors from the PRIDE Acute HF Score were generated from the validation set, and the observed proportion of patients with acute HF was determined.

All *P* values are 2-sided, with values <.05 considered significant. All statistical analyses were conducted with SYSTAT software (SPSS, Inc, Chicago, IL).

Results

The PRIDE Acute HF score

As previously reported, of the original 38 candidate variables examined in multivariable logistic regression

Table I. Independent clinical and biochemical predictors of acute HF, expressed with their respective odds ratios, 95% CI, β coefficient, and the integeric score derived from each

Predictor	Odds ratio	95% CI	β Coefficient	Integeric score
Elevated NT-proBNP*	44	21.0-91.0	3.8	4
Interstitial edema on chest x-ray	11	4.5-26.0	2.4	2
Orthopnea	9.6	4.0-23.0	2.26	2
Lack of fever	6.0	2.0-18.0	1.80	2
Current loop diuretic use	3.4	1.8-6.4	1.23	1
Age >75 y	2.7	1.4-5.2	1.0	1
Rales on lung exam	2.4	1.2-4.7	0.86	1
Lack of cough	2.3	1.2-4.3	0.81	1

A final possible score of 0 to 14 points was possible.

*Elevated NT-proBNP was defined as >450 pg/mL if age <50 years and >900 pg/mL if age \geq 50 years as previously described.

analysis for the identification of a final diagnosis of acute HF, 8 remained statistically significant after multivariable analysis. These included an elevated NT-proBNP (defined as an NT-proBNP >450 pg/mL among patients <50 years and an NT-proBNP >900 pg/mL for those \geq 50 years), interstitial edema on chest x-ray, orthopnea, absence of fever, current loop diuretic use, age >75 years, rales on lung examination, and absence of cough. In the present study, forcing a prior history of HF into the model did not yield any further significant independently useful information.

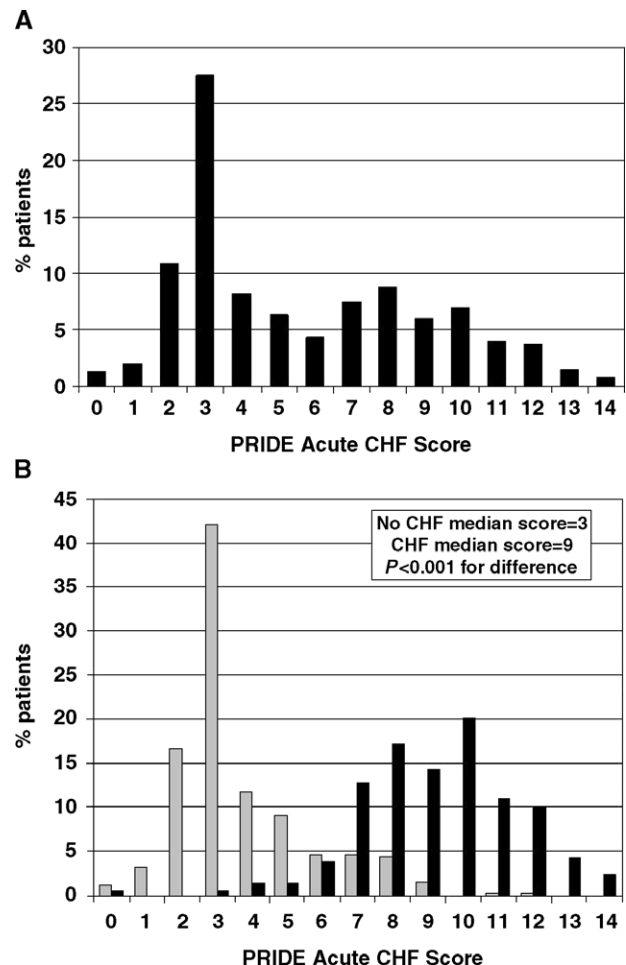
The diagnostic predictors used as the final components of the PRIDE Acute HF Score with their respective β coefficients and rounded integeric weights are listed in Table I. This analysis yielded a final score of 0 to 14 points.

Distribution of scores in the PRIDE study

The distribution of the PRIDE Acute HF Score among the patients in the derivation cohort, as well as a function of the presence or absence of acute HF is demonstrated in Figure 1. The overall median score for the cohort was 5 points. As a function of the presence or absence of acute HF, a strongly bimodal distribution was detected (Figure 1, B). The median score among patients with a final diagnosis of acute HF was statistically significantly higher than those without acute HF (9 vs 3 points, $P < .001$).

The PRIDE Acute HF Score: Internal validation

When applied to the overall PRIDE cohort (N = 599), for all score values, the predicted probability of HF matched very well with the observed incidence (Figure 2). For all score values, the predicted probability of acute HF was within the 95% CI of observed proportions.

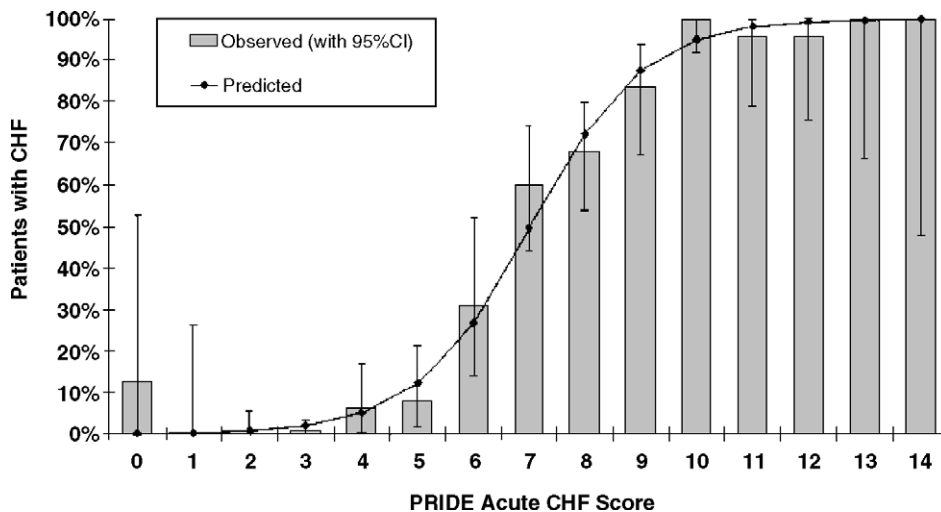
Figure 1

Distribution of PRIDE Acute HF Scores among (A) all patients in the entire PRIDE study and (B) expressed as a function of the presence (black bars) or absence (gray bars) of acute HF. The overall median score was 5 points. A score threshold of ≥ 6 points for the diagnosis of acute HF yielded a sensitivity and specificity of 96% and 84%, respectively.

The PRIDE Acute HF Score: Determining a diagnostic cut point

Receiver operating characteristic curve analysis was used to evaluate the analytical relationship between different scores and the diagnosis of acute HF. Intermediate values selected as potential diagnostic cut points are shown in Table II. In the derivation cohort, a PRIDE Acute HF Score of ≥ 6 points was found to maximize overall diagnostic accuracy with a sensitivity of 96%, a specificity of 84%, a positive predictive value (PPV) of 77%, and a negative predictive value (NPV) of 98% and was deemed ideal as a threshold for diagnosis of HF;

Figure 2



Observed (with 95% CI) versus predicted proportion of patients with acute HF, stratified by the PRIDE Acute HF Score. The observed proportion of acute HF parallels the predicted proportion.

Table II. Results of receiver operating characteristic curve analysis defining sensitivity, specificity, PPV, and NPV for intermediate score values selected as potential cut points for the diagnosis of acute HF

PRIDE Acute HF Score	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
5	97	75	68	98
6	96	84	77	98
7	92	89	82	96
8	79	89	87	89
9	70	93	92	82

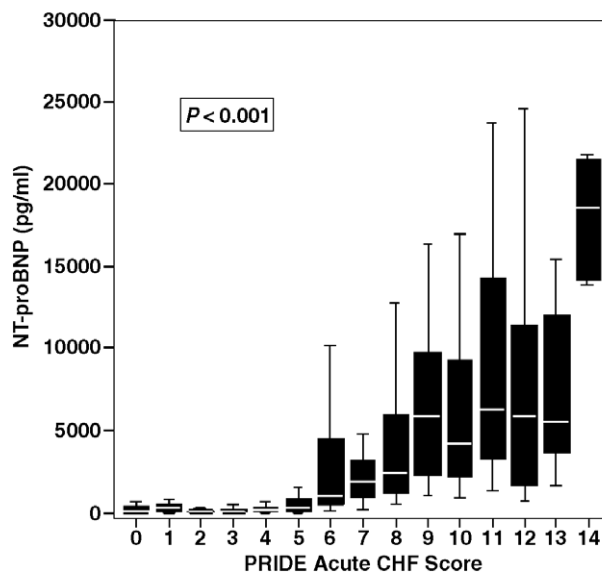
notably, as the score rose, the PPV rose in tandem, illustrating the value of the score not only as a dichotomous variable, but also as a continuous variable as well.

Relationship between the PRIDE Acute HF Score and the NT-proBNP results

As NT-proBNP represented the dominant variable in the PRIDE Acute HF Score, the relationship between NT-proBNP levels and a rising PRIDE Acute HF Score was established. In addition, the additive value of the PRIDE Acute HF Score to NT-proBNP testing alone was examined.

A strong direct relationship existed between NT-proBNP values and HF score (Pearson coefficient = 0.43, $P < .001$). In addition, the median values of NT-proBNP, expressed as a function of the overall score and are detailed in Figure 3, demonstrating that as the score

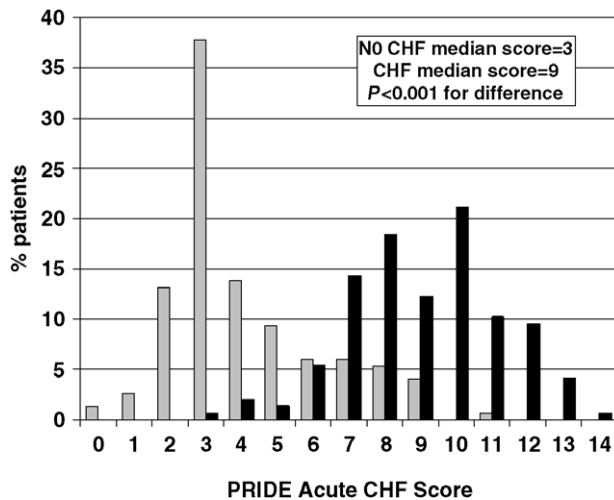
Figure 3



Median NT-proBNP results, expressed as a function of the PRIDE Acute HF Score. Boxes indicate the interquartile range, whereas whiskers indicate the 5th and 95th percentile values in each category.

value increased there was a parallel increase in the overall median NT-proBNP value ($P < .001$, test for trend), confirm the previously demonstrated correlation

Figure 4



Performance of the PRIDE Acute HF Score among patients for whom the diagnosis was uncertain at the time of clinical evaluation. A score ≥ 6 points for the diagnosis of acute HF yielded a sensitivity of 90% and a specificity of 87% in this subpopulation, consistent with the overall score performance in the PRIDE study.

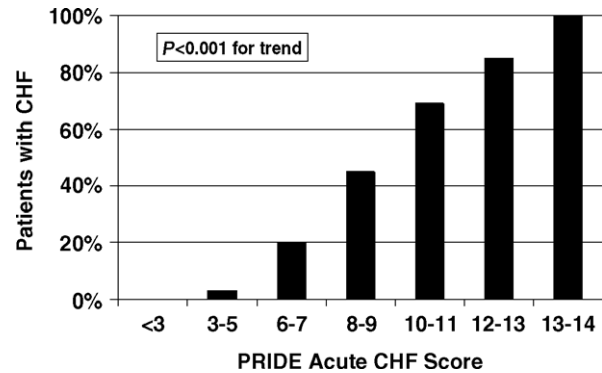
between rising NT-proBNP levels and more prevalent symptoms/signs of acute HF.⁷

The PRIDE Acute HF Score and patient subgroups

Among the 304 patients whose likelihood of HF was diagnosed by their managing physician to be between 10% and 90%, 147 (48%) had a final diagnosis of acute HF. Figure 4 shows the performance of the PRIDE Acute HF Score among these patients in whom diagnostic uncertainty was present. Similar to the group as a whole, the median HF score was 9 points among patients with acute HF, compared with 3 points among patients without acute HF ($P < .001$). The sensitivity of the scoring system for patients in the 10% to 90% range of clinical certainty was 93%, with a specificity of 87%.

Patients with an elevated NT-proBNP but without acute HF ($n = 59$) had lower median score values than the 188 patients with elevated NT-proBNP and acute HF (6 vs 9 points, P value for difference in median $< .001$). Notably, among those patients in whom a falsely elevated NT-proBNP was present, 91% of patients had scores < 7 points, whereas among those patients with a high NT-proBNP in the presence of HF, 88% had scores ≥ 7 points. Among these patients, using the PRIDE Acute HF Score, 18 more patients (31%) with elevated NT-proBNP levels but without acute HF would have been ruled out than if using NT-proBNP results alone for the diagnosis.

Figure 5



Rate of acute HF diagnosis in a previously described patient population from Christchurch, New Zealand,³ stratified by PRIDE Acute HF Score results, demonstrating a higher prevalence of acute HF diagnosis with higher scores ($P < .001$ for trend). In this data set, the median score of patients with acute HF was higher than those without acute HF (11 vs 5 points, $P < .001$).

Similarly, among the small number of patients diagnosed with acute HF whose NT-proBNP was not elevated ($n = 18$), the diagnostic yield of the PRIDE Acute HF Score was still superior to NT-proBNP alone, with 6 additional patients (33%) correctly identified as having acute HF by the score than were identified by the use of NT-proBNP results alone. Among these patients, the median score was 5 points, which was significantly higher than those patients without HF with low NT-proBNP values, whose median score was 2 points ($P < .001$).

The PRIDE Acute HF Score: External validation

Of 205 patients previously reported in a prior study from Christchurch, New Zealand,³ 195 had complete data and were analyzed for this study. When compared with the patients in the PRIDE study, patients from the Christchurch data set were more likely to be older (mean age 70.5 vs 63.3 years, $P < .001$); however, there were no significant differences between the PRIDE and Christchurch patients with respect to numerous other characteristics known to affect NT-proBNP levels, such as sex, body mass index, left ventricular ejection fraction (when known), or a history of prior HF or left ventricular dysfunction (all $P = NS$). In addition, renal function, as assessed by serum creatinine levels, was comparable between groups.

When applied to the cohort of patients from Christchurch, a rising PRIDE Acute HF Score was associated with a higher percentage of patients with a final diagnosis of acute HF (Figure 5). The median score of patients with acute HF in the validation set was

significantly higher than those without acute HF (11 vs 5 points, $P < .001$).

Receiver operating characteristic curve analyses in the Christchurch validation cohort confirm the overall sensitivity and specificity of the PRIDE Acute HF Score in a demographically distinct, older patient population, with an area under the curve of 0.92 ($P < .001$), which was superior to the area under the curve for NT-proBNP testing of 0.89 previously reported³ from this patient group (P value for difference = .005).

Discussion

The clinical diagnosis of acute HF may be challenging to identify or exclude in the ED and may lead to under- or overdiagnosis with attendant ramifications in terms of hospital costs and potential adverse effects on patient outcomes.¹¹ Brain-type natriuretic peptide and NT-proBNP, secreted in tandem from ventricular myocardium in response to increases in transmural pressure, are both useful in differentiating the dyspnea of HF from dyspnea attributable to other causes. As such, both peptides have emerged as diagnostic indicators of acute HF indicators and their use in patients suspected of having this syndrome has increased. Despite the value of natriuretic peptide testing for urgent patient evaluation, some critics assert that the biomarkers lack sufficient specificity and PPV value to be used in isolation, and some decry the rising use of these tests as a replacement for clinical judgment.¹²

In a previous report,⁷ it was shown that the combination of clinical evaluation and NT-proBNP testing is the superior strategy for the evaluation of patients with dyspnea in the ED. However, a simple and reproducible strategy to integrate routine clinical assessment with biomarker testing for the diagnosis of acute HF had not yet been developed. Accordingly, we sought to develop a simple, rapid, and accurate bedside tool for this purpose, one that emphasizes the strengths and minimizes the weaknesses of both clinical and biochemical diagnostic approaches.

The PRIDE Acute HF Score includes 8 readily available variables found to be statistically significant independent predictors of acute HF at presentation: elevated NT-proBNP results (considered elevated if >450 pg/mL in patients aged <50 years or >900 pg/mL in patients ≥ 50 years), interstitial edema on chest x-ray, orthopnea, absence of fever, current loop diuretic use, age >75 years, rales on lung examination, and absence of cough. The integrative score derived from these factors yielded a simple, highly accurate method for correctly identifying or excluding acute HF at the time of first patient evaluation with improved accuracy over isolated NT-proBNP testing. In addition, the PRIDE Acute HF Score was useful for patients in whom diagnostic uncertainty existed and for those patients in whom the

NT-proBNP was elevated but acute HF was not present. Lastly, we demonstrate validity of the scoring strategy among a distinct patient population.³ Although valuable as a dichotomous variable (with a “positive” score defined as ≥ 6 points), the rising prevalence of HF among patients with a rising score suggests that it could be also used as a continuous variable, with increasing graded likelihood for a diagnosis of HF. In aggregate, we propose the optimal method for categorizing patients would be to divide patients into those with low (score = 0-5 points), intermediate (score = 6-8 points), and high (score = 9-14 points) likelihood for HF. Using this strategy, one maintains the value of a dichotomous score of ≥ 6 points (a score below which delivers an NPV of 98%), while emphasizing the continuously increasing likelihood of HF with rising scores.

The PRIDE Acute HF Score offers several distinct applications for clinical use. First, it provides clinicians with a simple and accurate method of integrating NT-proBNP assay results with conventional clinical assessment. In doing so, it reduces the diagnostic inaccuracy inherent in both strategies used in isolation. Second, the score result may be used as a guide to the correct selection of initial therapeutic management, thereby potentially limiting mistakes in diagnostic evaluation, indicating appropriate venues for therapies with proven benefit for HF, while potentially reducing improper medication administration for those without HF. Lastly, the score may be useful for assisting in correct triage decisions.

Although our study is the first to examine the role of a clinical and biochemical risk score for the diagnosis and triage of patients with suspected acute HF, important limitations to our study exist. First, the experience in the PRIDE study reflected that of a single urban hospital, and, as such, the patient characteristics used to derive the PRIDE Acute HF Score may not reflect those of patients with acute HF in other centers. However, the derivation cohort in our study is large and diverse, similar to other cohorts in trials of acute HF with dyspnea.^{3,4} Furthermore, an external validation using a distinct patient cohort from a medical center in another continent³ revealed that a rising score was still strongly associated with an increased likelihood of acute HF. Other issues include the method in which NT-proBNP levels were handled in the score. To preserve simplicity, we examined NT-proBNP results in a dichotomous fashion rather than as a continuous variable. As such, we did not examine the role of marked elevations of NT-proBNP levels in the design of the score. However, a direct relationship existed between NT-proBNP and HF score values, reflecting the relationship between NT-proBNP values and prevalent HF symptoms/signs, and suggested that more marked elevation in NT-proBNP would be more likely to be accompanied by a parallel

rise in HF score, without making adjustment for the absolute NT-proBNP value.

In conclusion, we report the derivation of a simple, accurate, and portable diagnostic scoring system, which combines NT-proBNP results and clinical factors for the diagnosis of acute HF. The PRIDE Acute HF Score will be useful for the accurate diagnosis or exclusion of acute HF among patients presenting to the ED with dyspnea.

References

1. Stevenson LW, Perloff JK. The limited reliability of physical signs for estimating hemodynamics in chronic heart failure. *JAMA* 1989;261:884-8.
2. Remes J, Miettinen H, Reunanen A, et al. Validity of clinical diagnosis of heart failure in primary health care. *Eur Heart J* 1991;12:315-21.
3. Lainchbury JG, Campbell E, Frampton CM, et al. 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;42:728-35.
4. Maisel AS, Krishnaswamy P, Nowak RM, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med* 2002;347:161-7.
5. Mueller C, Scholer A, Laule-Kilian K, et al. Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea. *N Engl J Med* 2004;350:647-54.
6. Raymond I, Groenning BA, Hildebrandt PR, et al. 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;89:745-51.
7. Januzzi JL, Camargo CA, Anwaruddin S, et al. N-terminal ProBNP for urgent evaluation of shortness of breath: the ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) study. *Am J Cardiol* 2005;95:948-54.
8. Mehta RH, Suzuki T, Hagan PG, et al. Predicting death in patients with acute type a aortic dissection. *Circulation* 2002;105:200-6.
9. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: a method for prognostication and therapeutic decision making. *JAMA* 2000;284:835-42.
10. Mehran R, Aymong ED, Nikolsky E, et al. A simple risk score for prediction of contrast-induced nephropathy after percutaneous coronary intervention: development and initial validation. *J Am Coll Cardiol* 2004;44:1393-9.
11. Wuerz RC, Meador SA. Effects of prehospital medications on mortality and length of stay in congestive heart failure. *Ann Emerg Med* 1992;21:669-74.
12. Packer M. Should B-type natriuretic peptide be measured routinely to guide the diagnosis and management of chronic heart failure? *Circulation* 2003;108:2950-3.