Published in final edited form as:
Congest Heart Fail. 2012 September ; 18(5): 262–271. doi:10.1111/j.1751-7133.2012.00288.x.

A Comparison of Criterion Standard Methods to Diagnose Acute Heart Failure

Sean P. Collins, MD, MSc[*],
Vanderbilt University, Department of Emergency Medicine

Christopher J. Lindsell, PhD,
University of Cincinnati, Department of Emergency Medicine

Donald M. Yealy, MD,
University of Pittsburgh, Department of Emergency Medicine

David J. Maron, MD,
Vanderbilt University, Departments of Medicine and Emergency Medicine

Allen J. Naftilan, MD,
Vanderbilt University, Division of Cardiology

John A. McPherson, MD, and
Vanderbilt University, Division of Cardiology

Alan B. Storrow, MD
Vanderbilt University, Department of Emergency Medicine

Abstract

Background—We sought to compare and contrast the clinical criterion standards currently used in a cohort of ED patients to diagnose acute heart failure syndromes (AHFS).

Methods—In a prospective observational study of patients with signs and symptoms of AHFS we examined three criterion standards: 1) the treating ED physician's diagnosis; 2) the hospital discharge diagnosis; and 3) a diagnosis based on a medical record review by a panel of cardiologists. Using Cohen's kappa (κ), we assessed agreement and then compared the different standards by repeatedly setting one as the criterion standard and the other two as index tests.

Results—483 patients were enrolled. Across all criterion standards those with AHFS were more likely to have a history of AHFS, congestion on physical exam and chest radiography, and elevated natriuretic peptide levels than those without AHFS. The standards agreed well (cardiology review vs. hospital discharge diagnosis, κ=0.74; cardiology review vs. ED diagnosis, κ=0.66; ED diagnosis vs. hospital discharge diagnosis κ =0.59). Each method had similar sensitivity, but differing specificities.

Conclusion—Different criterion standards identify different patients from among those being evaluated for AHFS. Researchers should consider this when choosing between the various criterion standard approaches when evaluating new index tests.

Introduction

Making an accurate diagnosis of acute heart failure syndromes (AHFS) in the emergency department (ED) is challenging. The misdiagnosis rate may be as high as 20%.[1, 2] Over the

[*]Corresponding author: sean.collins@vanderbilt.edu.

last decade, several new diagnostic methods (index tests) have been studied with an aim of improving early diagnosis.[2–9] The definition of acute heart failure varies in diagnostic studies (criterion standard) used to evaluate these new index tests, resulting in inconsistent results and difficulty comparing studies.

Traditionally, pulmonary capillary wedge pressure (PCWP) is the strongest criterion standard. However, placing a pulmonary artery catheter is time-consuming, is associated with risk, and it is frequently not practical.[10, 11] Further, most patients are treated correctly for a presumptive AHFS diagnosis based on much more readily available data, suggesting a pulmonary artery catheter may not be appropriate for routine clinical use.[1, 2, 5] These limits spurred searches for a noninvasive test which would perform the same or better than elevated PCWP and become a more practical criterion standard.

Previous studies suggest the *ED physician diagnosis* agrees with other AHFS assessments in about 70–75% of cases.[2] Emergency physicians deploy readily available tools to diagnose AHFS: history, physical examination, and simple tests such as chest radiography (CXR), ancillary lab studies, and electrocardiogram. The addition of commercially available natriuretic peptide assays alters the diagnostic performance of the ED physician's diagnosis by increasing diagnostic agreement another 5–10% in the extremes while adding some positive effect in the intermediate ranges.[12] *Hospital discharge diagnosis* may be obtained from record review or from billing records.

Previous studies suggest high agreement between a discharge diagnosis code, such as an ICD-9 code, and a physician review panel.[1314, 15] Finally, a review or "overread" of the medical record by a *panel of cardiologists* has been employed as a means of determining the presence of AHFS.[23, 5,16–22]

All of the aforementioned criterion standards have advantages and disadvantages. Understanding the strengths, limitations and potential biases of these criterion standards is fundamental to interpreting and comparing studies of diagnostic accuracy. As clinicians seek to evaluate the accuracy of a new index test, the strengths and weaknesses of each of the possible criterion standards are important to consider.

To better understand these concepts, we undertook a direct comparison of criterion standards that would be considered for diagnostic studies (ED physician diagnosis, hospital discharge diagnosis and cardiology chart review) in a cohort of subjects enrolled in a prospective, observational AHFS study. The purpose of this analysis was not to determine the ideal criterion standard, but to better understand how the different criterion standards compare to each other.

Methods

The authors of this manuscript have certified that they comply with the Principles of Ethical Publishing in the International Journal of Cardiology: Shewan LG and Coats AJ. Ethics in the authorship and publishing of scientific articles. Int J Cardiol 2010;144:1–2.

Setting and Patient Population

This was a secondary analysis of a prospective observational study of AHFS being conducted at two hospitals in Nashville, Tennessee (Vanderbilt University Hospital and Nashville VA Medical Center) and three in Cincinnati, Ohio (The University Hospital, The Christ Hospital and The Jewish Hospital). ED volumes per hospital range from 15,000 to 80,000 annual visits. The participating centers include both community and academic sites, with both emergency medicine house staff and board-certified attending physicians

responsible for patient care. The patient populations are heterogeneous, representing both black and white patients, men and women, and Medicare, Medicaid, self-pay and self-insured patients.

Patients were eligible for enrollment in the original study if they met the following criteria: 1) fulfilled the modified Framingham Criteria (Table 1); 2) were identified within 3 hours of initial ED evaluation; 3)≥ 18 years of age; and 4) provided written, informed consent. We modified the original Framingham Criteria to increase their sensitivity by adding a history of heart failure as one of the major criteria. In addition, we removed the 4 parameters not routinely available in the ED (circulation time, vital capacity, weight loss in response to treatment, and autopsy findings). To focus on initial presentations, we enrolled only patients providing informed consent within three hours of initial ED physician evaluation. The Institutional Review Boards at all participating centers approved the study.

Data Collection

Using trained study assistants, we approached patients with signs or symptoms of AHFS from the chief complaint (shortness of breath, leg swelling, weight gain, fatigue) and then determined whether they fulfilled the 4 study criteria as previously defined. Briefly, the study assistants collected ED data contemporaneously, followed the patient throughout their entire ED course, and followed up with them in the hospital until discharge. Treating physicians filled out structured data forms to document their history and physical examination findings. Data was recorded on a structured paper-based data collection instrument, which was then transferred to the electronic database. Chest radiograph results, as recorded by the radiology attending dictated medical record, were recorded in the study database. In addition to baseline ED variables, we tracked medications administered during the ED stay using the nurses' medication administration record. In those circumstances where an evaluation or measure was not performed on a patient, the study assistant requested missing measures be obtained.

Determination of Criterion Standards

We evaluated three criterion standards: ED diagnosis, hospital discharge diagnosis, and diagnosis based on cardiology review. Due to the limited number of patients who received pulmonary artery catheters and subsequent measurements of PCWP, we could not consider PCWP as a criterion standard in this analysis. Further, because of the use of the modified Framingham Criteria to identify and enroll patients in the ongoing study, we chose not to evaluate it as a criterion standard.

1) ED Diagnosis

Based on available data at the time of ED disposition the treating physician, blinded to the purpose of this study, completed a data form indicating whether AHFS was present. Available data included, for example, physical examination, CXR, BNP, troponin, medications administered, and medical records. The treating physician indicated one of three diagnoses: 1) AHFS was the primary diagnosis responsible for the patient's signs and symptoms; 2) AHFS was present and contributed to their presenting signs and symptoms (a secondary AHFS diagnosis); or 3) AHFS was not present at all. For the purposes of this investigation, we defined AHFS when the treating physician indicated AHFS was present, either as a primary or secondary diagnosis.

2) Hospital Discharge Diagnosis

Trained abstractors reviewed the patient's medical records after the index stay to determine the hospital discharge diagnosis. We defined the criterion standard for AHFS in this

approach when the hospital discharge diagnosis included AHFS as either a primary discharge diagnosis or a secondary diagnosis where AHFS was considered to be present by the inpatient team. When a patient was discharged directly from the ED, we used the ED diagnosis as the discharge diagnosis.

3) Cardiology Overread

We presented data from the index stay to 3 cardiologists who reviewed the data independently. The information included ED notes and diagnoses, inpatient progress notes, discharge summaries and results from diagnostic testing. We also provided all 30-day follow-up data which indicated whether the subject had an adverse event, including a hospital admission for AHFS or death. Each cardiologist then indicated whether AHFS was the primary diagnosis at presentation to the ED, a secondary diagnosis, or not present. For this criterion standard, we defined AHFS when at least two out of the three cardiologists agreed it was present, either as a primary or secondary diagnosis. In the rare situation that all three cardiologists had a different diagnosis, a fourth cardiologist adjudicated.

Statistical Analysis

Initial analysis characterized the cohort stratified by the presence and absence of AHFS for each criterion standard. We describe data using medians, ranges, frequencies and percentages as appropriate. To explore the variation in clinical and demographic characteristics among patients with and without AHFS, we tabulated summary statistics for groups of patients defined by their concordance or discordance on the different criterion standards. We did not perform statistical comparisons since our intent was not inferential but exploratory.

To assess agreement between the criterion standards, we used Cohen's kappa. Then, we computed diagnostic test statistics for each standard assuming the others as the "truth" and calculated 95% confidence intervals for the test statistics using the score method with continuity correction. REDCap (Research Electronic Data Capture) tools hosted at Vanderbilt University Data stored and managed all of the study-related information, including the diagnoses provided by the ED physician and the cardiology chart review.[23] REDCap is a secure, web-based application designed to support data capture for research studies, providing: 1) an intuitive interface for data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. We analyzed data using SPSS version 17.0 (SPSS Inc., Chicago, Il) and Microsoft Excel (Microsoft Corporation, Redmond, Wa).

Results

Patient Characteristics

Of the 1637 patients screened for enrollment, 483 patients qualified and were consented and enrolled over 22 months, of which 395 (82%) were admitted to the hospital.(Figure 1) This cohort of 483 patients had a median age of 63 (IQR 51–75), 55% were male and 66% were Caucasian. Multiple comorbidities existed including hypertension (85.9%), chronic obstructive pulmonary disease (COPD) (42.2%), previous myocardial infarction (34.0%), and chronic renal insufficiency (22.4%). Patients presented with a median systolic blood pressure and heart rate of 139 mmHg (IQR 121–158 mmHg) and 88 beats per minute (IQR 74–102), respectively. Signs and symptoms of congestion upon ED presentation included jugular venous distension (24.4%), extremity edema (60.9%), and rales or crackles on auscultation (52.7%). The median creatinine and B-type natriuretic peptide (BNP) were 1.30 mg/dl (IQR 0.96–1.96) and 460 pg/ml (IQR 122–1144), respectively. Chest radiograph

findings demonstrated cardiomegaly in 217 (45.7%), interstitial edema in 45 (9.5%) and pulmonary edema in 36 (7.6%). Loop diuretics were administered in 165 (34.2%) of patients while 147 (30.4%) patients received topical or sublingual nitroglycerin.

All three criterion standards had similar patterns of clinical characteristics associated with the presence of AHFS: a prior history of heart failure, congestion on physical exam and chest radiograph, and elevated BNP. Those classified as non-AHFS more often had a history of COPD. The occurrence of other comorbidities, such as hypertension and diabetes, was common in both the AHFS and non-AHFS patients.

Agreement

There were 301 (62%) cases with AHFS based on ED diagnosis, 292 (60.5%) based on hospital discharge diagnosis and 276 (57.1%) based on cardiology review. For the cardiology review, majority consensus of the three reviewers occurred in 478/483 cases (99.0%). There were 232 subjects classified as AHFS on all three measures, and 134 were classified as not having AHFS on all three. Agreement between the different standards was fair to good (cardiology review vs. hospital discharge diagnosis, κ=0.74, 95% CI 0.67–0.80; cardiology review vs. ED diagnosis, κ=0.66, 95% CI 0.59–0.73; ED diagnosis vs. hospital discharge diagnosis κ =0.59, 95%CI 0.52–0.67; Table 2).

Clinical Characteristics of Discordant Patients

Those subjects with discordant diagnoses between the ED physician and the cardiology panel (Table 3) had clinically significant differences with respect to: age; a history of renal disease and heart failure; the presence of a pacemaker or defibrillator; physical exam findings of congestion such as the presence of leg edema or pulmonary crackles; and laboratory findings such as BNP values and renal function.

Those subjects who had a discordant diagnosis between the ED physician and the inpatient discharge diagnosis had a similar pattern of differences in clinical characteristics.(Table 4) Those patients with a discordant diagnosis tended to have clinically significant differences with respect to: age; a history of heart failure, coronary artery bypass surgery or COPD; pulmonary crackles and diagnostic test findings such as pleural effusion on chest radiograph, BNP values and renal function (BUN and creatinine).

Those subjects who had a discordant diagnosis between the cardiology panel and the inpatient discharge diagnosis (Table 5) had a number of discrepancies related to past medical history: a history of heart failure, renal disease, COPD, coronary artery bypass surgery and a prior pacemaker implant; physical exam findings of extremity edema or pulmonary crackles; chest radiograph findings such as infiltrates, pleural effusion, and differences in laboratory such as BNP values and renal function.

Diagnostic Test Statistics

The diagnostic test statistics for each standard assuming the others as the truth are shown in Table 6. The overall accuracy of the discharge diagnosis and cardiology review for a prediction of an ED diagnosis of AHFS was very good (80.7% and 83.6%, respectively). Both the ED physician's diagnosis and the cardiology review had very good accuracy for prediction of a discharge diagnosis of AHFS (80.7% and 87.2%, respectively). The ED diagnosis had an accuracy of 83.6% for prediction of a cardiology review diagnosis of AHFS. The hospital discharge diagnosis was more accurate than the ED diagnosis for prediction of a cardiology review diagnosis of AHFS (87.2% vs. 83.6%). While both the hospital discharge diagnosis and ED diagnosis had similar sensitivity for prediction of a

cardiology review diagnosis of AHFS (91.7% vs. 90.2%), the improved accuracy of the discharge diagnosis was a result of greater specificity (81.2% vs. 74.9%).

Discussion

A variety of criterion standards make interpretation of the accuracy of new index tests difficult. In this head-to-head comparison, we considered three common clinical criterion standards in a group of well characterized patients being evaluated for possible AHFS. Our results suggest agreement is moderate amongst the ED diagnosis, hospital discharge diagnosis and cardiology review. Further, it appears that all three criterion standards heavily weight signs of congestion, natriuretic peptides and past medical history such as COPD and heart failure when considering a diagnosis of AHFS. Our purpose was not to declare one criterion standard as the “best”, but to compare and contrast them.

When considering the cardiology review as the truth and looking at the diagnostic test statistics of the ED and discharge diagnoses there are some specific points worthy of discussion (Table 6, column 3). The accuracy of the ED diagnosis (83.6%) is driven by its high sensitivity. The moderate specificity suggests that patients in the ED are over-diagnosed with AHFS. As more information becomes available during the hospital stay, such as echocardiography and angiography as well as response to therapy, the number of false positives decreases, as demonstrated by the increasing specificity and slightly higher accuracy of the hospital diagnosis. This has potential implications for future development of clinical evaluation and research in AHFS. As the amount of diagnostic information increases (i.e. from the limited information in the ED diagnosis to the comprehensive information used for cardiology review), or becomes more readily available (the use of limited echocardiography by ED physicians) the proportion of patients diagnosed with AHFS decreases. This suggests that the clinical diagnosis of AHFS is perhaps one of exclusion. However, the positives of a cardiology chart review must be balanced by its limitations, including the inability to account for the clinical gestalt of the physician at the bedside. When using a simple score (our subjects were included on the basis of modified Framingham Criteria) or the ED diagnosis as the criterion standard, investigators must recognize the inherent inclusion of false positives among those initially diagnosed as AHFS. Amending these criterion standards to emphasize more specific findings such as signs of congestion on chest radiography or significantly elevated natriuretic peptides may improve the accuracy of the criterion standard. Further, as more specific measures of congestion are available at the bedside, improved accuracy and a result of increased specificity would be expected.

Previous studies of the Framingham Criteria suggest a moderate sensitivity of 63% and good specificity of 94% for prediction of heart failure in the outpatient setting.[24] Four of the Framingham Criteria are problematic for use in the ED, however: 1) circulation time, 2) vital capacity, 3) weight loss in response to treatment, and 4) autopsy findings. While we had extensive data about the Framingham Criteria available to us, we did not conduct a comprehensive evaluation of these criteria because a modified version was used to include patients in the ongoing study. Baggish and colleagues looked at data from an AHFS study to derive and validate an AHFS prediction instrument for ED patients with dyspnea.[25] They found an area under the receiver operating characteristic (ROC) curve of 0.92 when they externally tested their 8-item criteria in a cohort of 195 patients. In a similar study, which used age, NT-proBNP levels and the treating ED physician’s pretest probability, external testing of this model resulted in an area under the ROC curve of 0.91. This predictive instrument was more informative when the pre-test probability was indeterminate.[26]

Emergency physicians use history, physical examination and diagnostic testing that is available in the first 3–6 hours of an ED stay to evaluate patients with possible AHFS. This generally includes chest radiography, ancillary lab studies, and the electrocardiogram. Despite its clear utility in AHFS, access to formal echocardiography performed in the ED outside of weekday daytime hours is rare. In general, the unstructured clinical assessment of the emergency physician creates three distinct groups initially: 1) those clearly with AHFS, 2) those clearly without AHFS, and 3) those where AHFS is possible but uncertain. Our current findings support those from previous studies, suggesting that the initial ED clinician's impression predicts a cardiology overread with 80–85% accuracy.[22, 27] The overall clinical impression remains the strongest predictor of AHFS presence.[28]

We used the hospital discharge diagnosis obtained from record review of the discharge summary. Billing and administrative data are often used instead of discharge diagnosis given the wide availability, but these cannot discriminate between the major cause of symptoms at admission versus the entire stay, and also may confuse the highest reward diagnosis with the most prominent. Additionally, heart failure is covered by many codes, and the code may assume one of many ordered positions, without any standardization regarding the impact of heart failure during that stay. One European study noted over 90% agreement between the discharge diagnosis code of heart failure with that of a panel review.[13] Research in the United States shows similar agreement between an assigned heart failure ICD9 diagnosis and panel review.[14, 15] Chart review to abstract the clinical diagnosis offers the potential ability to better classify the diagnosis, but requires more resources and clear adjudication rules. It should be noted that while the discharge diagnosis of heart failure strongly agrees with comparators, any lack of sensitivity is unknown since those without this diagnosis are not assessed to determine if they were true negatives.

Since no system of diagnostic criteria is agreed upon as the criterion standard for AHFS in the acute care setting, a cardiology review or overread of the medical record is often employed. The first major trial to use this method evaluated the diagnostic test characteristics of BNP in ED patients with acute dyspnea,[2] with several subsequent reports using similar methodology.[3, 5, 16–22] The diagnosis of AHFS may be determined independently or by consensus from two or more cardiologists blinded to the results of the index test, and sometimes the ED physician diagnosis. Unlike acute coronary syndromes, where a biomarker such as a cardiac troponin establishes the diagnosis, AHFS lacks a readily available and universally accepted criterion standard for the acute care setting. Blinded cardiology overread is attractive as a criterion standard given its reliance on expert review and preponderance of data. It is limited by work-up bias, the records of the primary treating physicians, and any biases introduced by the reviewers themselves. Its absolute accuracy when applied to the undifferentiated populations seen in the ED and other acute care settings remains accepted, but has not been extensively tested.

Limitations

While our data clearly demonstrate the differences and similarities of the various criterion standards used for studying diagnostic tests in AHFS, there are several limitations to consider when interpreting our results. While the cardiology review has increasingly become the criterion standard of choice over the last decade, chart review cannot always account for the physician experience gained by direct patient interaction and evaluation of response to therapy. This gestalt of patient care is not always accounted for in the documentation available to the chart reviewer. Furthermore, discussions amongst multiple practitioners and the nuances of rounding at the bedside, where several practitioners can provide input, are not always captured during chart review. Possible inaccuracies in the cardiology review are evidenced when complete agreement between three reviewers does not occur. The

cardiologists were not blinded to the ED or hospital discharge diagnosis when reviewing charts. This could have introduced bias. Further, while we have analyzed the accuracy of each method when considering the others to be the truth, the actual accuracy of each method is difficult to define without a study comparing them to a gold standard known to be correct. Diagnostic impressions can be right or wrong, but how wrong remains unknown. Short of measuring PCWP on a large cohort of undifferentiated dyspneic patients being evaluated for AHFS, the question of true accuracy may remain unanswered. For now, given that a clinical diagnosis of AHFS appears to be one of exclusion, it might be assumed that the criterion standard based on the most information is the most accurate available.

In this study, treating physicians did not follow a specific diagnostic or therapeutic protocol for trial purposes. The data originated from a purely observational trial. As a result, the study could have suffered from work-up bias where patients at lower likelihood of AHFS received less work-up. The lack of mandatory diagnostic testing, such as routine echocardiography, may have impacted the overall agreement and accuracy of the criterion standards. Further, those admitted with AHFS compared to those admitted for AHFS may have different work-ups performed and therapies administered.

Patients were enrolled based on the Framingham Criteria. While this may impact the generalizability of our findings, we feel this is an appropriate tradeoff. While they have been traditionally used to establish a longitudinal diagnosis of heart failure in the outpatient setting, they tend to be highly sensitive, facilitating enrollment of a broad cohort of patients with signs and symptoms of AHFS, as well as lend objectivity to our enrollment criteria.

The treating physician interacted with the study assistants to determine whether the patients qualified for the study. As a result of knowing the subject was being included in the study, the treating physician may have been biased to indicate more or less often that the subject had a diagnosis of AHFS.

Another possible limitation of our analysis is that when a patient was discharged to home from the ED, their ED diagnosis also served as their hospital discharge diagnosis. This will have increased the agreement between the hospital discharge diagnosis and ED diagnosis. There were 395 patients admitted (88 discharged) and in these patients only, the agreement between ED and hospital diagnosis was 0.49 (95%CI 0.40–0.58). When considering the hospital discharge diagnosis as the criterion standard among admitted patients only, the accuracy of the ED diagnosis was 76.7% (95% CI 72.2–80.8). This compares well with the previously calculated accuracy of the ED diagnosis for predicting a hospital discharge diagnosis of 80.7% (95% CI 76.9–84.1).

Conclusion

In conclusion, different criterion standards identify different patients from among those being evaluated for possible AHFS. The ED physician diagnosis is sensitive for AHFS, but the criterion standards become more selective, with fewer subjects being identified as having heart failure, as the quantity of information used in decision making increases. Each criterion standard carries a different cost and confers a different benefit. Researchers should consider this when choosing between the various criterion standard approaches when evaluating new index tests.

Acknowledgments

This work was supported by grants from the National Heart Lung and Blood Institute (K23HL085387, R01HL088459). Data collection was also supported in part by an investigator-initiated research grant from Abbott

POC. Database design and implementation was provided by the Vanderbilt Institute for Clinical and Translational Research with grant support from UL1 RR024975 from NCRR/NIH

References

1. Collins SP, Lindsell CJ, Storrow AB, Abraham WT. Prevalence of negative chest radiography results in the emergency department patient with decompensated heart failure. Ann Emerg Med. Jan; 2006 47(1):13–18. [PubMed: 16387212]
2. 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. Jul 18; 2002 347(3):161–167. [PubMed: 12124404]
3. Januzzi JL Jr, Camargo CA, Anwaruddin S, et al. The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. Am J Cardiol. Apr 15; 2005 95(8):948–954. [PubMed: 15820160]
4. Studler U, Kretzschmar M, Christ M, et al. Accuracy of chest radiographs in the emergency diagnosis of heart failure. Eur Radiol. Aug; 2008 18(8):1644–1652. [PubMed: 18369631]
5. Collins SP, Peacock WF, Lindsell CJ, et al. S3 detection as a diagnostic and prognostic aid in emergency department patients with acute dyspnea. Ann Emerg Med. Jun; 2009 53(6):748–757. [PubMed: 19232777]
6. Sharma GV, Woods PA, Lambrew CT, et al. Evaluation of a noninvasive system for determining left ventricular filling pressure. Arch Intern Med. Oct 14; 2002 162(18):2084–2088. [PubMed: 12374516]
7. Januzzi JL Jr, Peacock WF, Maisel AS, et al. Measurement of the interleukin family member ST2 in patients with acute dyspnea: results from the PRIDE (Pro-Brain Natriuretic Peptide Investigation of Dyspnea in the Emergency Department) study. J Am Coll Cardiol. Aug 14; 2007 50(7):607–613. [PubMed: 17692745]
8. Ely EW, Haponik EF. Using the chest radiograph to determine intravascular volume status: the role of vascular pedicle width. Chest. Mar; 2002 121(3):942–950. [PubMed: 11888980]
9. Ely EW, Smith AC, Chiles C, et al. Radiologic determination of intravascular volume status using portable, digital chest radiography: a prospective investigation in 100 patients. Crit Care Med. Aug; 2001 29(8):1502–1512. [PubMed: 11505116]
10. Sprung CL, Elser B, Schein RM, Marcial EH, Schrager BR. Risk of right bundle-branch block and complete heart block during pulmonary artery catheterization. Crit Care Med. Jan; 1989 17(1):1–3. [PubMed: 2909309]
11. Kearney TJ, Shabot MM. Pulmonary artery rupture associated with the Swan-Ganz catheter. Chest. Nov; 1995 108(5):1349–1352. [PubMed: 7587440]
12. Morrison LK, Harrison A, Krishnaswamy P, Kazanegra R, Clopton P, Maisel A. Utility of a rapid B-natriuretic peptide assay in differentiating congestive heart failure from lung disease in patients presenting with dyspnea. J Am Coll Cardiol. Jan 16; 2002 39(2):202–209. [PubMed: 11788208]
13. Ingelsson E, Arnlov J, Sundstrom J, Lind L. The validity of a diagnosis of heart failure in a hospital discharge register. Eur J Heart Fail. Aug; 2005 7(5):787–791. [PubMed: 15916919]
14. Philbin EF, Rocco TA, Lindenmuth NW, Ulrich K, McCall M, Jenkins PL. The results of a randomized trial of a quality improvement intervention in the care of patients with heart failure. The MISCHF Study Investigators. Am J Med. Oct 15; 2000 109(6):443–449. [PubMed: 11042232]
15. Philbin EF, DiSalvo TG. Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. J Am Coll Cardiol. May; 1999 33(6):1560–1566. [PubMed: 10334424]
16. Maisel A. B-type natriuretic peptide in the diagnosis and management of congestive heart failure. Cardiol Clin. Nov; 2001 19(4):557–571. [PubMed: 11715177]
17. Maisel A. B-type natriuretic peptide measurements in diagnosing congestive heart failure in the dyspneic emergency department patient. Rev Cardiovasc Med. 2002; 3(Suppl 4):S10–17. [PubMed: 12439426]

18. Maisel AS, Clopton P, Krishnaswamy P, et al. Impact of age, race, and sex on the ability of B-type natriuretic peptide to aid in the emergency diagnosis of heart failure: results from the Breathing Not Properly (BNP) multinational study. Am Heart J. Jun; 2004 147(6):1078–1084. [PubMed: 15199359]

19. Maisel AS, McCord J, Nowak RM, et al. Bedside B-Type natriuretic peptide in the emergency diagnosis of heart failure with reduced or preserved ejection fraction. Results from the Breathing Not Properly Multinational Study. J Am Coll Cardiol. Jun 4; 2003 41(11):2010–2017. [PubMed: 12798574]

20. McCullough PA, Duc P, Omland T, et al. 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. Mar; 2003 41(3):571–579. [PubMed: 12612980]

21. McCullough PA, Nowak RM, Foreback C, et al. Emergency evaluation of chest pain in patients with advanced kidney disease. Arch Intern Med. Nov 25; 2002 162(21):2464–2468. [PubMed: 12437406]

22. McCullough PA, Nowak RM, McCord J, et al. B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from Breathing Not Properly (BNP) Multinational Study. Circulation. Jul 23; 2002 106(4):416–422. [PubMed: 12135939]

23. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. Apr; 2009 42(2):377–381. [PubMed: 18929686]

24. Fonseca C, Oliveira AG, Mota T, et al. Evaluation of the performance and concordance of clinical questionnaires for the diagnosis of heart failure in primary care. Eur J Heart Fail. Oct; 2004 6(6): 813–820. 821–812. [PubMed: 15542422]

25. Baggish AL, Siebert U, Lainchbury JG, et al. 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. Am Heart J. Jan; 2006 151(1):48–54. [PubMed: 16368291]

26. Steinhart B, Thorpe KE, Bayoumi AM, Moe G, Januzzi JL Jr, Mazer CD. Improving the diagnosis of acute heart failure using a validated prediction model. J Am Coll Cardiol. Oct 13; 2009 54(16): 1515–1521. [PubMed: 19815122]

27. Doust JA, Glasziou PP, Pietrzak E, Dobson AJ. A systematic review of the diagnostic accuracy of natriuretic peptides for heart failure. Arch Intern Med. Oct 11; 2004 164(18):1978–1984. [PubMed: 15477431]

28. Wang CS, FitzGerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? JAMA. Oct 19; 2005 294(15):1944–1956. [PubMed: 16234501]

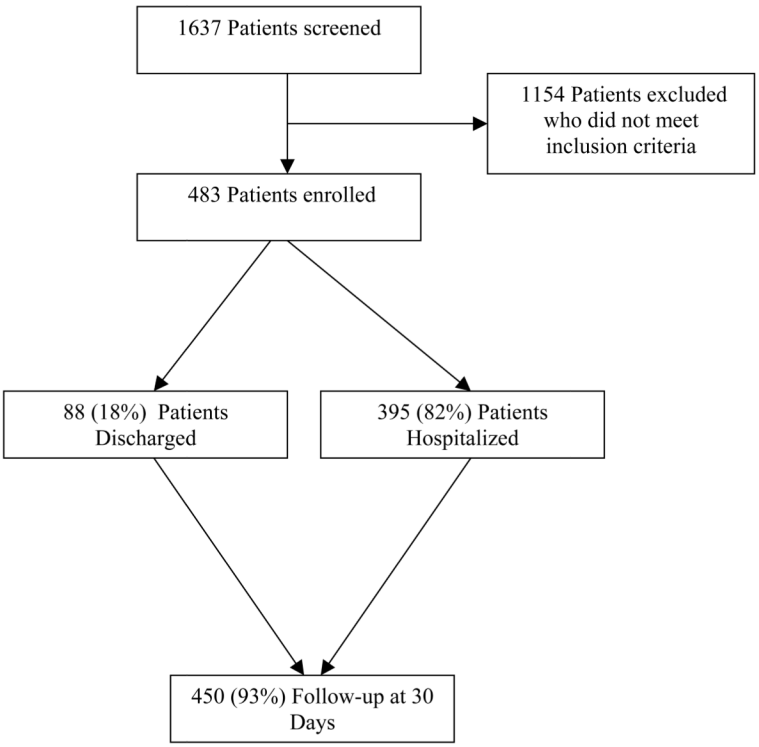

Figure 1.
Patient enrollment and flow through the study.

Table 1

Modified Framingham Criteria

Major	Minor
• History of Heart Failure	• Extremity edema
• Paroxysmal nocturnal dyspnea	• Night cough
• Pulmonary or interstitial edema (on CXR)	• Dyspnea on exertion
• Rales	• Hepatomegaly
• Cardiomegaly	• Pleural effusion
• S_3 gallop	• Tachycardia (≥130 beats/min)
• Jugular venous distention	
• Positive hepatojugular reflux	

* 2 Major or 1 Major and 2 Minor Criteria are required to establish a preliminary diagnosis of heart failure by the Framingham criteria.

Table 2

Levels of agreement between the criterion standard diagnoses. Cohen's kappa and associated 95% confidence intervals are presented.

Criterion Standard	Cardiology review		Discharge Diagnosis	
	κ	95% CI	κ	95% CI
ED Diagnosis	0.66	0.59–0.73	0.59	0.52–0.67
Cardiology review			0.74	0.67–0.80

Table 3

Population stratified by ED diagnosis and then subcategorized by a discrepant or congruent diagnosis based on cardiology review. Data are presented as median and **interquartile** ranges or frequencies and percentages as appropriate.

	ED Diagnosis							
	No AHFS				**AHFS**			
	What was the cardiology diagnosis?				**What was the cardiology diagnosis?**			
	No AHFS		**AHFS**		**No AHFS**		**AHFS**	
Demographics								
Age (years)	59	**50–72**	68	**49–79**	69	**55–79**	64	51–75
Female	**70**	**45.2**	**13**	**48.1**	22	**42.3**	**112**	45.0
African American	**53**	**34.2**	**9**	**33.3**	16	**30.8**	**86**	34.5
Caucasian	**102**	**65.8**	**18**	**66.7**	36	**69.2**	**162**	65.1
Vital Signs								
Heart Rate (beats/min)	89	**76–101**	89	**72–97**	82	**74–101**	87	74–104
Temperature (degrees F)	97.9	**97.4–98.6**	97.8	**97.3–98.3**	97.9	**97.5–98.2**	97.8	97.2–98.2
Systolic BP (mmHg)	140	**121–158**	146	**132–169**	137	**120–155**	139	121–158
Diastolic BP (mmHg)	78	**66–91**	80	**71–100**	74	**63–87**	81	69–93
Respiratory Rate	18	**18–22**	20	**18–24**	20	**17–22**	20	18–24
Oxygen Saturation	96	**94–98**	96	**94–98**	96	**94–98**	96	93–98
Lab Values								
BNP (pg/ml)	105	**38–293**	212	**87–941**	214	**79–518**	825	423–1656
Sodium (mEq/L)	138	**136–140**	137	**134–140**	139	**135–141**	138	136–140
Glucose (mg/dl)	117	**98–145**	102	**88–155**	117	**100–149**	120	103–151
BUN (mg/dl)	15	**11–27**	26	**14–40**	21	**13–33**	25	15–43
Creatinine (mg/dl)	1.2	**0.9–1.6**	1.2	**1.0–2.3**	1.4	**1.0–1.7**	1.4	1.0–2.2
Hemoglobin (g/dl)	12.4	**11.1–13.9**	12.5	**11.4–15.0**	12.5	**11.0–13.4**	12.3	10.8–13.9
Chest X-ray Findings								
Pulmonary Edema	1	0.7	0	0.0	6	11.5	29	11.8
Infiltrates	39	25.8	6	23.1	12	23.1	57	23.2
Pleural Effusion	28	18.5	4	15.4	10	19.2	81	32.9

	ED Diagnosis							
	No AHFS				AHFS			
	What was the cardiology diagnosis?				What was the cardiology diagnosis?			
	No AHFS		AHFS		No AHFS		AHFS	
Interstitial Edema	5	3.3	2	7.7	4	7.7	34	13.8
Exam Findings								
Rales or Crackles	69	44.8	5	20.0	31	59.6	147	59.5
S3 Gallop	7	5.1	1	4.2	3	6.3	33	15.9
JVD	21	16.4	4	18.2	10	22.7	61	30.5
Extremity Edema	94	61.4	14	51.9	30	60.0	152	61.8
Medical History								
Renal Disease	25	16.1	9	33.3	9	17.3	65	26.1
Heart Failure	68	43.9	24	88.9	29	55.8	215	86.3
COPD	76	49.0	12	44.4	24	46.2	92	36.9
Hypertension	129	83.2	22	81.5	49	94.2	215	86.3
Myocardial Infarction	35	22.6	10	37.0	14	26.9	105	42.2
CABG	26	16.8	6	22.2	9	17.3	77	30.9
Pacemaker	23	14.8	6	22.2	9	17.3	65	26.1
ICD	17	11.0	4	14.8	8	15.4	55	22.1

BP=blood pressure; BNP=B-type natriuretic peptide; BUN=blood urea nitrogen; JVP=jugular venous pressure; CRI=chronic renal insufficiency; COPD=chronic obstructive pulmonary disease; CABG=coronary artery bypass grafting; ICD=implantable cardioverter defibrillator

Congest Heart Fail. Author manuscript; available in PMC 2013 September 01.

Table 4

Population stratified by ED diagnosis and then subcategorized by a discrepant or congruent diagnosis based on discharge diagnosis. Data are presented as median and interquartile ranges or frequencies and percentages as appropriate.

	ED Diagnosis							
	No AHFS				AHFS			
	What was the discharge diagnosis?				What was the discharge diagnosis?			
	No AHFS		AHFS		No AHFS		AHFS	
Demographics								
Age (years)	60	**49–72**	63	**54–79**	68	**51–75**	64	**52–76**
Female	64	**45.7**	19	45.2	19	37.3	115	46.0
African American	48	**34.3**	14	33.3	13	25.5	89	35.6
Caucasian	92	**65.7**	28	66.7	38	74.5	160	64.0
Vital Signs								
Heart Rate (beats/min)	90	**78–99**	84	**71–102**	82	**71–98**	87	**74–104**
Temperature (degrees F)	98.00	**97.4–98.6**	97.75	**97.3–98.2**	97.90	**97.3–98.3**	97.80	**97.2–98.2**
Systolic BP (mmHg)	141	**126–159**	140	**113–159**	137	**121–154**	139	**121–158**
Diastolic BP (mmHg)	78	**67–92**	75	**65–95**	76	**63–88**	81	**69–93**
Respiratory Rate	20	**18–22**	19	**18–22**	20	**18–22**	20	**18–24**
Oxygen Saturation	96	**94–98**	96	**94–98**	95	**91–98**	96	**93–98**
Lab Values								
BNP (pg/ml)	109.00	**38–297**	168.00	**78–748**	474.00	**220–1029**	764.00	**357–1517**
Sodium (mEq/L)	138.00	**136–140**	137.00	**133–139**	139.00	**136–141**	138.00	**136–140**
Glucose (mg/dl)	115.00	**98–145**	116.50	**93–152**	120.00	**100–155**	120.00	**103–148**
BUN (mg/dl)	15.00	**41239.00**	25.50	**15–40**	26.00	**16–43**	23.00	**14–39**
Creatinine (mg/dl)	1.14	**0.9–1.6**	1.30	**1.0–1.8**	1.41	**1.1–2.1**	1.33	**1.0–2.1**
Hemoglobin (g/dl)	12.40	**11.1–13.9**	12.90	**11.4–14.0**	12.40	**10.4–13.3**	12.30	**11.0–13.9**
Chest X-ray Findings								
Pulmonary Edema	1	0.7	0	0.0	6	11.8	29	11.7
Infiltrates	36	26.5	9	22.0	15	29.4	54	21.9
Pleural Effusion	27	19.9	5	12.2	13	25.5	78	31.6

	ED Diagnosis							
	No AHFS				AHFS			
	What was the discharge diagnosis?				What was the discharge diagnosis?			
	No AHFS		AHFS		No AHFS		AHFS	
Interstitial Edema	6	4.4	1	2.4	10	19.6	28	11.3
Exam Findings								
Rales or Crackles	60	43.2	14	35.0	34	66.7	144	58.1
S3 Gallop	4	3.2	4	10.5	3	6.8	33	15.6
JVD	19	16.4	6	17.6	9	23.1	62	30.2
Extremity Edema	81	58.3	27	65.9	30	58.8	152	62.0
Medical History								
Renal Disease	24	17.1	10	23.8	13	25.5	61	24.4
Heart Failure	56	40.0	36	85.7	26	51.0	218	87.2
COPD	69	49.3	19	45.2	29	56.9	87	34.8
Hypertension	113	80.7	38	90.5	48	94.1	216	86.4
Myocardial Infarction	29	20.7	16	38.1	19	37.3	100	40.0
CABG	22	15.7	10	23.8	13	25.5	73	29.2
Pacemaker	17	12.1	12	28.6	6	11.8	68	27.2
ICD	13	9.3	8	19.0	6	11.8	57	22.8

BP=blood pressure; BNP=B-type natriuretic peptide; BUN=blood urea nitrogen; JVP=jugular venous pressure; CRI=chronic renal insufficiency; COPD=chronic obstructive pulmonary disease; CABG=coronary artery bypass grafting; ICD=implantable cardioverter defibrillator

Congest Heart Fail. Author manuscript; available in PMC 2013 September 01.

Table 5

Population stratified by discharge diagnosis and then subcategorized by a discrepant or congruent diagnosis based on cardiology diagnosis. Data are presented as median and **interquartile** ranges or frequencies and percentages as appropriate.

	Discharge Diagnosis							
	No AHFS				**AHFS**			
	What was the cardiology diagnosis?				**What was the cardiology diagnosis?**			
	No AHFS		**AHFS**		**No AHFS**		**AHFS**	
Demographics								
Age (years)	60	**49–73**	65	**51–73**	66	**56–82**	64	**51–76**
Female	74	44.0	9	39.1	18	46.2	116	45.8
African American	55	32.7	6	26.1	14	35.9	89	35.2
Caucasian	113	67.3	17	73.9	25	64.1	163	64.4
Vital Signs								
Heart Rate (beats/min)	88	**76–100**	90	**75–98**	83	**73–102**	87	**74–103**
Temperature (degrees F)	98.00	**97.4–98.6**	97.80	**97.1–98.1**	97.80	**97.3–98.1**	97.80	**97.2–98.2**
Systolic BP (mmHg)	140	**123–159**	136	**117–157**	129	**109–150**	140	**123–160**
Diastolic BP (mmHg)	78	**67–91**	72	**64–90**	70	**62–82**	82	**69–94**
Respiratory Rate	20	**18–22**	20	**18–26**	18	**17–20**	20	**18–24**
Oxygen Saturation	96	**94–98**	95	**90–98**	97	**95–98**	96	**93–98**
Lab Values								
BNP (pg/ml)	118.50	**38–318**	1001.50	**485–3109**	114.00	**63–459**	782.00	**366–1517**
Sodium (mEq/L)	138.00	**136–140**	138.00	**135–142**	137.00	**133–140**	138.00	**136–140**
Glucose (mg/dl)	116.00	**99–144**	127.00	**99–168**	123.00	**99–163**	118.50	**102–148**
BUN (mg/dl)	15.00	**11–27**	33.50	**22–62**	21.00	**15–38**	23.00	**14–40**
Creatinine (mg/dl)	1.16	**0.9–1.6**	1.99	**1.3–6.7**	1.31	**1.0–1.8**	1.32	**1.0–2.1**
Hemoglobin (g/dl)	12.50	**11.1–13.9**	11.80	**10.0–13.2**	12.60	**11.4–13.6**	12.40	**11.0–13.9**
Chest X-ray Findings								
Pulmonary Edema	6	3.7	1	4.3	1	2.6	28	11.2
Infiltrates	45	27.4	6	26.1	6	15.4	57	22.9
Pleural Effusion	33	20.1	7	30.4	5	12.8	78	31.3

	Discharge Diagnosis							
	No AHFS				AHFS			
	What was the cardiology diagnosis?				What was the cardiology diagnosis?			
	No AHFS		AHFS		No AHFS		AHFS	
Interstitial Edema	8	4.9	8	34.8	1	2.6	28	11.2
Exam Findings								
Rales or Crackles	81	48.5	13	56.5	19	48.7	139	55.8
S3 Gallop	6	4.0	1	5.3	4	10.8	33	15.6
JVD	24	17.3	4	25.0	7	21.2	61	29.6
Extremity Edema	98	58.7	13	56.5	26	72.2	153	61.2
Medical History								
Renal Disease	26	15.5	11	47.8	8	20.5	63	24.9
Heart Failure	63	37.5	19	82.6	34	87.2	220	87.0
COPD	84	50.0	14	60.9	16	41.0	90	35.6
Hypertension	141	83.9	20	87.0	37	94.9	217	85.8
Myocardial Infarction	36	21.4	12	52.2	13	33.3	103	40.7
CABG	28	16.7	7	30.4	7	17.9	76	30.0
Pacemaker	19	11.3	4	17.4	13	33.3	67	26.5
ICD	16	9.5	3	13.0	9	23.1	56	22.1

BP=blood pressure; BNP=B-type natriuretic peptide; BUN=blood urea nitrogen; JVP=jugular venous pressure; CRI=chronic renal insufficiency; COPD=chronic obstructive pulmonary disease; CABG=coronary artery bypass grafting; ICD=implantable cardioverter defibrillator

Congest Heart Fail. Author manuscript; available in PMC 2013 September 01.

Table 6

Diagnostic test characteristics for each of the Criterion Standards for Predicting an AHFS diagnosis when using of the Criterion Standards as the Index Test and the others as the Criterion Standards.

Diagnostic test statistics	Discharge diagnosis		Cardiology Review		ED diagnosis	
Cardiology review						
Accuracy	87.2	83.8 – 89.9			83.6	80.0 – 86.8
Sensitivity	86.6	82.1 – 90.2			82.7	77.9 – 86.7
Specificity	88.0	82.3 – 92.1			85.2	79.0 – 89.8
Negative LR	0.23	0.17 – 0.31			0.34	0.26 – 0.43
Positive LR	11	7.4 – 16.3			9.2	6.4 – 13.2
ED diagnosis						
Accuracy	80.7	76.9 – 84.1	83.6	80.0 – 86.8		
Sensitivity	85.6	80.9 – 89.3	90.2	85.0 – 93.3		
Specificity	73.3	66.3 – 79.3	74.9	68.0 – 80.5		
Negative LR	0.20	0.15 – 0.30	0.17	0.1 – 0.24		
Positive LR	4.9	3.8 – 6.3	4.8	3.0 – 6.2		
Discharge diagnosis						
Accuracy			87.2	83.0 – 89.9	80.7	76.9 – 84.1
Sensitivity			91.7	87.0 – 94.5	83.1	78.2 – 87.0
Specificity			81.2	75.0 – 86.1	76.9	70.0 – 82.7
Negative LR			0.14	0.0 – 0.20	0.36	0.29 – 0.46
Positive LR			6.5	4.0 – 8.7	5.95	4.5 – 7.9

LR=Likelihood ratio

Congest Heart Fail. Author manuscript; available in PMC 2013 September 01.