The Diagnosis of Heart Failure in the Community

Comparative Validation of Four Sets of Criteria in Unselected Older Adults: The ICARe Dicomano Study

Mauro Di Bari, MD, PhD,*† Claudia Pozzi, MD,* Maria Chiara Cavallini, MD, PhD,*† Francesca Innocenti, MD,*† Giorgio Baldareschi, MD,*† Walter De Alfieri, MD,*† Enrico Antonini, MS,* Riccardo Pini, MD, FACC,*† Giulio Masotti, MD,*† Niccolò Marchionni, MD*†

Florence and Grosseto, Italy

OBJECTIVES

We sought to compare construct and predictive validity of four sets of heart failure (HF) diagnostic criteria in an epidemiologic setting.

BACKGROUND

The prevalence estimates of HF vary broadly depending on the diagnostic criteria.

METHODS

Data were collected in a survey of community dwellers who were ≥65 years of age living in Dicomano, Italy. At baseline, HF was diagnosed with the criteria of the Framingham, Boston, and Gothenburg studies and of the European Society of Cardiology (ESC). Left ventricular mass index and ejection fraction, left atrium systolic dimension, lower extremity mobility disability, summary physical performance score, and 6-min walk test were compared between HF and non-HF participants to test for construct validity of each set of criteria. Predictive validity was evaluated with follow-up assessment of cardiovascular mortality, incident disability, and HF-related hospitalizations. Comparisons were adjusted for demographics, comorbidity, and psychoaffective status.

RESULTS

Of 553 participants, 11.9%, 10.7%, 20.8%, and 9.0% had HF, according to Framingham, Boston, Gothenburg, and ESC criteria, respectively. In terms of construct validity, Framingham and Boston criteria discriminated HF from non-HF participants better than Gothenburg and ESC criteria across the measures of cardiac function and global performance. The Boston criteria showed a superior predictive validity because they indicated a significantly greater adjusted risk of cardiovascular death (hazard ratio 3.9, 95% confidence interval 1.2 to 13.2), incident disability, and hospitalizations in participants with HF.

CONCLUSIONS

The Boston criteria are preferable to Framingham, Gothenburg, and ESC criteria for the diagnosis of HF in older community dwellers because they have good construct validity and more accurately predict cardiovascular death, incident disability, and hospitalizations. (J Am Coll Cardiol 2004;44:1601–8) © 2004 by the American College of Cardiology Foundation

Heart failure (HF) poses a major burden on public health systems of industrialized countries. Its prevalence and incidence sharply increase in older individuals (1,2), in whom this condition is a leading cause of death, morbidity, disability, and hospital admissions (3–5). However, randomized clinical trials of angiotensin-converting enzyme inhibitors (6), angiotensin receptor antagonists (7), beta-blockers (8), and spironolactone (9) have clearly shown that this condition is a leading cause of death, morbidity, disability, and hospital admissions (3–5). However, randomized clinical trials of angiotensin-converting enzyme inhibitors (6), angiotensin receptor antagonists (7), beta-blockers (8), and spironolactone (9) have clearly shown that survival and functional outcomes can be effectively improved when HF is recognized in a timely manner.

Unfortunately, the clinical recognition of HF can be difficult, especially at a primary care level. Even experienced physicians frequently disagree on the diagnosis of mildest cases (10), whereas ≤50% of the diagnoses made by primary care physicians are confirmed after further cardiac assessment (11). Because of comorbidity and atypical clinical manifestations, the diagnosis is particularly challenging in older persons: a recent study reported that primary care physicians often feel uncomfortable in diagnosing HF and in differentiating it from other diseases that are common in late life (12). Thus, the syndrome remains unrecognized and poorly managed in many older patients (12).

Besides clinical practice, difficulties in the diagnosis of HF also impact on its epidemiologic assessment in the population. Prevalence figures reported in the medical literature vary widely (13–17), mainly because different sets of diagnostic criteria have been used. Most of these criteria rarely have been comparatively evaluated and, therefore, it is unknown which set provides the most accurate estimates. The present study was conducted to compare construct and predictive validity of four established sets of criteria for the diagnosis of HF in unselected older community dwellers.

METHODS

In the validation process of a diagnostic instrument, the term construct validity indicates the ability of that instrument to represent characteristics (constructs) of the condi-
tion under examination, whereas predictive validity is the ability of the instrument to predict clinically relevant outcomes (18). In this study, HF constructs were considered abnormalities in cardiac structure and function that are typical of the syndrome and limitations in physical performance that are its ultimate consequences. Predictive validity was evaluated by comparing cardiovascular mortality, hospital admissions, and incident disability in subjects with and without the diagnosis of HF.

**Study population and protocol.** Data were obtained in the “Insufficienza Cardiaca negli Anziani Residenti a Dicomano” (ICARe Dicomano) Study, a longitudinal epidemiologic survey on HF in the elderly that was conducted in Dicomano, a small rural town near Florence, Italy. The methods of the study have been previously detailed (19,20).

Briefly, the ICARe Dicomano Study, which is consistent with the principles of the Declaration of Helsinki on clinical research involving human subjects, enrolled in 1995 the entire unselected, community-dwelling elderly (65 years) population recorded in the City Registry Office. The only exclusion criterion was living in a nursing home.

In a follow-up study, the City Registry Office was consulted to define vital status. Causes of death were ascertained from the International Classification of Diseases-9th edition-coded death certificates and structured questionnaires, which were answered by primary care physicians. Deaths due to International Classification of Diseases codes 410 (myocardial infarction), 430 to 438 (stroke), and 428 (HF) were considered cardiovascular deaths. Clinical data were obtained by interviewing survivors and their primary care physicians.

**Data collection.** After informed consent, multidimensional, geriatric assessment data, including complete clinical examination, physical performance tests, 12-lead electrocardiogram, and echocardiography, were collected at baseline.

**Diagnosis of HF.** The diagnosis of HF was based on four previously published instruments: Framingham (16), Boston (21), and Gothenburg criteria (22) and the European Society of Cardiology (ESC) principles (23), as operationalized by Davies et al. (14) and Fischer et al. (24).

Prelimarily, the presence of dyspnea as the main symptom limiting physical activity was investigated using Goldman’s Specific Activity Scale (25), which explores tolerance to exertion in standardized daily activities, grouped into four classes according to the level of energy expenditure in metabolic equivalents (METS), from >7 (class I) to <2 METS or at rest (class IV). The occurrence of orthopnea, paroxysmal nocturnal dyspnea, and pulmonary edema also was recorded.

With the Framingham instrument (16), HF was diagnosed in the presence of two major or one major and two minor criteria, representing signs and symptoms typical of HF. Circulation time ≥25 s, reduced vital capacity, and weight loss after diuretics, which were enlisted in the original Framingham diagnostic tool, were not considered in this study.

The Boston instrument (21) explores the presence of symptoms and signs typical of HF in the three categories of history, physical findings, and chest X-ray. A score of 0 to 4 is assigned to each category, providing a summary score from 0 to 12 (Appendix). The original distinction between a definite (score >8) and a probable (score 5 to 7) diagnosis of HF was not maintained in this study, where a score >5 was considered as diagnostic. Chest X-ray was not performed when the sum of scores assigned to history and physical findings was either 0 or >5 because in these two extremes chest X-ray alone would have allowed neither to establish nor to exclude the diagnosis. When indicated, chest X-ray was scored independently by a cardiologist and a pulmonologist; disagreement (<10% of cases) was resolved by consensus.

Gothenburg criteria (22) take into account history and physical findings to calculate a cardiac score contributing to define, together with specific drug treatment, an HF stage, ranging from 0 (HF absent) to 4 (death due to HF). In this study, a diagnosis of HF based on this instrument was assigned only to participants resulting to have “overt HF” (stages 2 to 3). Stage 4, which can be assigned only in follow-up studies, was not applicable at baseline in the present investigation.

The ESC principles require a combination of symptoms, objective evidence of cardiac dysfunction, and response to treatment in doubtful cases (23). As suggested by the ESC, echocardiography was used to assess the presence of cardiac dysfunction (see the next section). In agreement with Davies et al. (14), reduced left ventricular ejection fraction (LVEF, %), valve dysfunction, and atrial fibrillation were considered cardiac abnormalities supporting the diagnosis. Isolated diastolic abnormality also was taken into account (24) to acknowledge the possible occurrence of HF with preserved systolic function (24,26). Combining Davies’ categories of definite and probable HF (14), the ESC diagnosis was assigned to participants with dyspnea (Specific Activity Scale classification >2) and echocardiographic evidence of cardiac dysfunction.

**Echocardiographic examination.** Echocardiography was performed with a mechanical sector scanner (Challenger,
3.5 to 2.5 MHz dynamically focused transducer, ESAOTE Biomedica, Genoa, Italy) according to standard methods (27,28). Images were videotaped and stop frames were digitized (TomTecP® System, TomTec Imaging Systems, Munich, Germany) for quantitative analysis by examiners who were blinded to clinical data. At least three measures were averaged in participants in sinus, and five in those in non-sinus rhythm.

Technical acceptability of M-mode images was judged by standardized criteria (29). When M-mode orientation was suboptimal, linear measures were taken from two-dimensional images. Left ventricular (LV) mass was calculated from LV wall thickness and internal dimensions (30) and indexed by body surface area (left ventricular mass index [LVMI], g/m²).

The LV volumes were calculated with the area-length formula, from manually traced endocardial end-diastolic and end-systolic borders from apical four-chamber view images. When this view was unavailable, volumes were calculated with Teichholz’s formula from M-mode measures (31). An LVEF <50% was considered as indicative of systolic dysfunction.

Valvular and LV diastolic function was assessed with color, pulsed-wave, and continuous-wave Doppler signals. Aortic stenosis with peak gradient >20 mm Hg, mitral stenosis with area <1.5 cm², or regurgitant lesions >2/4+ were considered significant valve abnormalities. Peak early and late velocities were measured from the pulsed-wave Doppler mitral inflow pattern recorded at the mitral leaflet tips. Isovolumetric relaxation time was measured as the time interval between the end of aortic outflow and the beginning of mitral inflow signal. Isolated diastolic abnormality was suggested in the presence of an LVEF >45% associated with an early/late ratio less than or an isovolumetric relaxation time greater than the age-appropriate limits of 0.50 and 105 ms, respectively (32).

Functional assessment. Because HF impairs blood supply to performing skeletal muscles, functional assessment was considered a relevant component in the evaluation of construct validity of the criteria for the diagnosis of HF.

Baseline functional status was assessed subjectively in terms of self-reported lower extremity mobility disability (LEMD) and objectively by using performance tests. A modified version of the World Health Organization disability questionnaire (33) was used to identify LEMD as difficulty in one or more of six activities involving lower extremity function (transferring in and out of bed, using the toilet, moving in the house, bathing or showering, walking a quarter of mile, climbing stairs) that are pivotal to maintain autonomy in late life (34,35). The number of items in which the participant reported difficulty was taken as a subjective measure of functional limitation.

Objective measures of functional status were the 6-min walk test (6MWT) (36) and a modified version of Guralnik’s lower extremity physical performance battery (37). Studies have shown that the distance (m) walked in the 6MWT is a strong prognostic predictor in HF and other clinical conditions (36). The lower extremity physical performance battery is a tool well established in comprehensive geriatric assessment, which includes three tests assessing balance, a walk of short distance, and lower extremity strength (37). Balance is evaluated as the time, up to a maximum of 10 s, the participant is able to maintain standing equilibrium in five tasks of increasing difficulty. Walking ability is assessed as speed in a 4-m path, taking the best result in two trials. Lower extremity strength is indirectly evaluated as time required to stand up five times from a chair. The performance is evaluated by assigning to each test a score from 0 (worst performance) to 4 (best performance) based on the quartile distribution of the test results in a reference older population. A summary performance score (SPS, range 0 to 12) is then calculated as the sum of individual test scores. The SPS predicts important outcomes such as incident physical disability, nursing home admission, and death (37).

In survivors who were not disabled at baseline, incident disability in basic activities of daily living (BADL) was evaluated in the follow-up as onset of need for help in at least one of the following: walking in the house, washing and dressing self, toileting, transferring from bed to chair, and eating (33).

Assessment of comorbidity. The previously described measures of functional status are nonspecific for HF and can be influenced by noncardiac diseases. Thus, they can be used to validate a diagnosis of HF only when other causes of functional limitation are accounted for in multivaried analyses. To this purpose, comorbidities were thoroughly ascertained and quantified.

Diagnostic algorithms, based on questionnaires, physical examination, laboratory tests, and instrumental data (19), were used to identify 14 coexisting diseases. The individual disease severity was weighted using Greenfield’s index, which ranges from 0 (disease absent) to 4 (life threatening or uncontrolled disease). The burden of comorbidity was then summarized as Index of Coexisting Diseases (ICED), which is the sum of the individual severity scores (38).

Cognitive impairment and depressive symptoms, which are frequent non-somatic comorbidities of late life not included in the ICED, were evaluated with the Mini Mental State Examination (MMSE) (39) and the Geriatric Depression Scale (GDS) (40).

Analytic procedures. Statistical analysis was performed with SPSS for Windows 10.1 (SPSS Inc., Chicago, Illinois). Continuous variables are reported as mean ± SEM. Relative frequencies were compared with the chi-squared test. Kappa statistics was used to analyze the agreement between pairs of sets of diagnostic criteria.

To evaluate construct validity, measures of cardiac structure and function (LVMI, LVEF, and left atrial systolic dimension [LASD]) and global functional status (LEMD, 6MWT, and SPS) were compared between HF and non-HF participants in age-adjusted and gender-adjusted
Prevalence of Heart Failure According to Different Diagnostic Criteria by Age and by Gender

<table>
<thead>
<tr>
<th></th>
<th>Total n (%)</th>
<th>65–74 n (%)</th>
<th>75+ n (%)</th>
<th>Men n (%)</th>
<th>Women n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Framingham</td>
<td>66 (11.9)</td>
<td>34 (9.0)</td>
<td>32 (18.4)</td>
<td>31 (13.4)</td>
<td>35 (10.9)</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td>0.002</td>
<td></td>
<td>0.379</td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td></td>
<td>2.3 (1.4–3.8)</td>
<td></td>
<td>0.8 (0.5–1.3)</td>
<td></td>
</tr>
<tr>
<td>Boston</td>
<td>59 (10.7)</td>
<td>32 (8.4)</td>
<td>27 (15.5)</td>
<td>23 (9.9)</td>
<td>36 (11.2)</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td>0.012</td>
<td></td>
<td>0.625</td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td></td>
<td>2.0 (1.2–3.4)</td>
<td></td>
<td>1.1 (0.7–2.0)</td>
<td></td>
</tr>
<tr>
<td>Gothenburg</td>
<td>115 (20.8)</td>
<td>67 (17.7)</td>
<td>48 (27.6)</td>
<td>49 (21.1)</td>
<td>66 (20.6)</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td>0.008</td>
<td></td>
<td>0.873</td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td></td>
<td>1.8 (1.2–2.7)</td>
<td></td>
<td>1.0 (0.6–1.5)</td>
<td></td>
</tr>
<tr>
<td>ESC</td>
<td>50 (9.0)</td>
<td>15 (4.0)</td>
<td>35 (20.1)</td>
<td>19 (8.2)</td>
<td>31 (9.7)</td>
</tr>
<tr>
<td>p value</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td>0.553</td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>6.1 (3.2–11.5)</td>
<td></td>
<td>1.2 (0.7–2.2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI = confidence interval; ESC = European Society of Cardiology; OR = odds ratio.

Results

Of 864 subjects eligible as of April 25, 1995, 614 underwent cardiologic examination. Reasons for nonparticipation were death or nursing home admission before data collection in 21 cases and refusal in 229. After exclusion of another 61 participants with incomplete data, the final study sample included 553 participants (64.0% of the original cohort), of whom 232 (41.9%) were men. Mean age was 73.0 ± 0.4 years; p < 0.001 but had a similar proportion of men (45%; p = 0.383). In most eligible nonparticipating (286 of 311) and participating subjects (538 of 553), clinical data could be obtained from primary care physicians, who reported a diagnosis of HF less frequently in those who did participate (56 of 538; 10.4%), as compared with those who did not (59 of 286; 20.6%; p < 0.001).

Prevalence of HF and diagnostic concordance. HF was diagnosed in 11.9%, 10.7%, 20.8%, and 9.0% of participants using Framingham, Boston, Gothenburg, and ESC criteria, respectively. Prevalence figures increased with age but were similar in men and women (Table 1). As shown by kappa values between 0.05 and 0.59, the agreement between diagnostic criteria ranged from poor to only moderate even when comparing criteria that yielded similar prevalence (Framingham, Boston, and ESC criteria) (Fig. 1).

Predictive validity was evaluated by comparing cardiovascular mortality, incident BADL disability, and HF-related hospitalizations between HF and non-HF participants. Cox's proportional hazards survival model was used to analyze cardiovascular mortality. The proportionality of hazards was assessed with visual inspection of the survival curves. Logistic regression was used to analyze the risk of incident BADL disability and hospital admission. In all analyses, age, gender, and ICED, MMSE, and GDS scores were entered as covariates. A two-tailed p value < 0.05 was considered statistically significant.

Predictive validity. In a three-year follow-up, 18 cardiac and 29 noncardiac deaths were recorded (cumulative mor-
Cardiovascular mortality was 7.6%, 10.2%, 3.5%, and 6.0% in participants labeled as HF using Framingham, Boston, Gothenburg, and ESC criteria, respectively. In Cox’s regression models adjusted for age, gender, MMSE, GDS, and ICED, only Boston criteria predicted cardiovascular death (Fig. 4).

Of the survivors, 446 who were not disabled at baseline were interviewed in the follow-up to detect changes in their functional status. New BADL disability was identified in 35 cases (7.8%). Cumulative incidence of BADL disability was 21.3%, 26.2%, 10.1%, and 26.5% in participants diagnosed with HF using Framingham, Boston, Gothenburg, and ESC criteria, respectively; the corresponding figures for non-HF participants were 6.3%, 5.9%, 7.3%, and 6.3%. In logistic regression models, adjusted for age, gender, MMSE, GDS, and ICED, only Boston criteria predicted incident BADL disability (Table 2).

The absolute risk of HF-related hospital admission in the follow-up was 6.1%, 6.8%, 3.5%, and 0% in participants with and without HF according to different criteria. Data are least square means, adjusted for age and gender. Δ is the difference between non-HF and HF participants. Open bars = non-HF; closed bars = HF. ESC = European Society of Cardiology; SEM = standard error of the mean.

**DISCUSSION**

In this community-based sample of older persons, the prevalence of HF obtained with four sets of established diagnostic criteria varied between 9.0% and 20.8%. Even when instruments that yielded similar prevalence were compared, the diagnostic agreement was only poor to...
pressure commonly is increased in patients with HF, yet after diuretic or vasodilator treatment, it can be normal at rest and increase only on exertion. Left ventricular ejection fraction is low when HF is due to systolic dysfunction, but by definition it is normal in the presence of diastolic dysfunction. Finally, even cardiologists may disagree on the presence of HF, at least in mildest cases and, therefore, the judgment of a single clinician can be hardly assumed as a gold standard.

Following a different approach, we aimed at identifying the instrument that would provide the best discrimination between participants labeled as HF versus non–HF across a spectrum of measures pertinent to the two domains of construct and predictive validity. In the domain of construct validity, LVEF, LVMI, and LASD were selected as indicators of the underlying structural and functional abnormalities of the heart. On average, LVEF is indeed expected to be lower, and LVMI and LASD greater, in subjects with HF than in those without HF. Measures of physical functioning also were considered because the ultimate effect of HF is to reduce blood supply to metabolically active tissues, such as skeletal muscles, therefore limiting physical performance. The three measures considered in the domain of predictive validity represent expected consequences of HF, which reduces life expectancy, compromises functional independence, and causes frequent hospitalizations. Because these outcomes are not exclusive of HF and can be due to a variety of conditions, the comparisons between HF and non–HF participants were adjusted for demographics, comorbidity, and cognitive and affective status. Thus, our validation procedure encompassed thoroughly the clinical complexity of HF while controlling for a variety of possible confounders. We believe that this analytic approach and the population-based nature of our sample represent major strengths of this study. It should be further emphasized that, to our knowledge, no previous study examined predictive validity of diagnostic criteria for HF.

The superiority of Boston criteria likely depends on their well-balanced scoring of history, physical findings, and chest X-ray. However, it also should be pointed out that this instrument, although valid and preferable over the others tested in this study, relies heavily on the presence of dyspnea as the leading complaint. This symptom is neither specific nor sensitive for HF in older patients (43). On one hand, shortness of breath is reported in the presence of many noncardiac diseases (44), including a condition common but usually overlooked in the elderly, such as thoracic kyphosis (45). On the other hand, in late life HF may present atypically, with a sudden decline of cognitive or physical functioning and negligible respiratory symptoms (46). We would argue that diagnostic criteria for HF, specifically targeting older persons, should rely less on dyspnea and rather include items on atypical presentations. Therefore, the Boston criteria might be refined along this line, to further improve their diagnostic accuracy.

Table 2. Multivariate Prediction of Incident Disability in Basic Activities of Daily Living in Participants With a Baseline Diagnosis of Heart Failure According to Different Criteria

<table>
<thead>
<tr>
<th></th>
<th>OR*</th>
<th>95% CI*</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framingham</td>
<td>2.4</td>
<td>0.8–6.8</td>
<td>0.102</td>
</tr>
<tr>
<td>Boston</td>
<td>4.2</td>
<td>1.4–12.6</td>
<td>0.011</td>
</tr>
<tr>
<td>Gothenburg</td>
<td>1.1</td>
<td>0.4–3.2</td>
<td>0.826</td>
</tr>
<tr>
<td>ESC</td>
<td>2.6</td>
<td>0.9–7.5</td>
<td>0.077</td>
</tr>
</tbody>
</table>

*From separate logistic regression models, each adjusted for age, gender, Mini Mental State Examination, Geriatric Depression Scale, and Index of Coexisting Diseases. Abbreviations as in Table 1.

Table 3. Multivariate Prediction of Hospital Admissions for Heart Failure in Participants With a Baseline Diagnosis of Heart Failure According to Different Criteria

<table>
<thead>
<tr>
<th></th>
<th>OR*</th>
<th>95% CI*</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framingham</td>
<td>6.9</td>
<td>1.3–36.1</td>
<td>0.022</td>
</tr>
<tr>
<td>Boston</td>
<td>8.7</td>
<td>1.5–51.5</td>
<td>0.017</td>
</tr>
<tr>
<td>Gothenburg</td>
<td>5.1</td>
<td>0.9–27.7</td>
<td>0.058</td>
</tr>
<tr>
<td>ESC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*From separate logistic regression models, each adjusted for age, gender, Mini Mental State Examination, Geriatric Depression Scale, and Index of Coexisting Diseases. OR could not be calculated for ESC criteria because no participant with an ESC-based diagnosis had hospital admissions due to heart failure in the follow-up.

NA = not applicable; other abbreviations as in Table 1.
Study limitations. Some study limitations must be acknowledged. The study sample was relatively small, which might have limited the assessment of the validity for some of the analyzed diagnostic criteria. On the other hand, it is likely that a larger sample size would have increased the differences in performance of the diagnostic instruments, therefore leaving the overall results of the study substantially unchanged. Eligible residents in Di Comano who did not participate may pose a selection bias to data interpretation because they were older than those who participated and more often were reported by their physicians as having HF. However, the accuracy of the diagnosis reported by primary care physicians should be questioned because it did not rely on uniform assessments and criteria. The mortality of participants labeled as HF was lower than expected in such an old population, suggesting that milder forms of HF prevailed in participants who were assessed. Yet, this possibly reinforces the validity of our findings because mildest cases are more difficult to recognize. Even with these limitations, we believe that our findings have remarkable implications in the clinical and public health fields. The demonstration that Boston criteria are more accurate than others may provide clinicians with an instrument of proven validity, therefore possibly overcoming the difficulties they admittedly encounter in the diagnosis of HF in older persons. In the public health arena, these results suggest that the Boston criteria should be preferred in future epidemiologic surveys of HF. To some extent, they might also challenge the available information on HF epidemiology obtained with less accurate instruments.

Reprint requests and correspondence: Dr. Mauro Di Bari, Department of Critical Care Medicine and Surgery, Unit of Gerontology and Geriatrics, University of Florence, via delle Oblate, 4, 50141 Florence, Italy. E-mail: dibari@unif.it.

REFERENCES


Comorbidity and outcomes after hip replacement. Med Care 1993;31:141–54.

APPENDIX
For the Boston criteria for the diagnosis of heart failure, please see the October 20, 2004, issue of JACC at www.onlinejacc.org.