Congestive Heart Failure in the United States

Is There More Than Meets the I(CD Code)? The Corpus Christi Heart Project

David C. Goff, Jr, MD, PhD; Dilip K. Pandey, MD, PhD; Frances A. Chan, MPH; Carmen Ortiz; Milton Z. Nichaman, MD, ScD

Background: Congestive heart failure (CHF) is increasing as a public health problem in the United States. The ability to quantify this problem has been limited by a lack of data regarding the validity of CHF identification.

Objective: To assess the validity of the use of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD) codes to identify hospitalizations with clinical evidence of an episode of acute CHF in data of The Corpus Christi Heart Project, a population-based surveillance program for hospitalized coronary heart disease.

Methods: The validation standard was a composite variable including the presence of physician diagnosed acute CHF or radiographic evidence of pulmonary edema. Data were abstracted from the medical records of 5083 patients identified as hospitalized for possible acute myocardial infarction, aortocoronary bypass surgery, percutaneous transluminal coronary angioplasty, and related revascularization procedures in the Corpus Christi Heart Project. Discharge diagnoses, a secondary source of data, were used to apply 3 computer algorithms to assess the assignment of ICD codes.

Results: The prevalence of clinically documented CHF was 27.1% (1376/5083). The ICD code 428 (CHF), assigned as the primary or a secondary discharge diagnosis, was associated with 62.8% sensitivity, 95.4% specificity, 83.5% positive predictive value, 87.4% negative predictive value, and a 24.8% underenumeration of CHF-related hospitalizations. An algorithm based on a series of ICD codes was associated with 67.1% sensitivity, 92.6% specificity, 77.1% positive predictive value, 88.3% negative predictive value, and a 13.0% underenumeration of CHF-related hospitalizations.

Conclusions: Reliance on ICD codes results in the exclusion of one third of the patients with clinical evidence of acute CHF. This underenumeration is compounded by the typical reliance on the first listed diagnosis. Congestive heart failure may be a greater public health problem than currently recognized. The allocation of resources for relevant surveillance, research, medical care, and preventive efforts should be reevaluated.

Arch Intern Med. 2000;160:197-202

Congestive heart failure (CHF) has been recognized as a major public health problem in the United States, is associated with substantial morbidity and mortality, and increases exponentially in occurrence with aging. Both the human suffering and the financial burden associated with CHF are substantial. Yet, we know little about the incidence of CHF. Most reports that have documented the occurrence or other aspects of CHF have, by necessity, relied on administrative data, such as the National Hospital Discharge Survey, Medicare databases, and mortality tapes. When using these data sources, CHF is defined on the basis of diagnostic coding according to the International Classification of Diseases (ICD) system. To the extent that the application and use of the ICD coding system is inaccurate, the current impression of the magnitude of the public health problem posed by CHF may be erroneous. The potential policy implications of reliance on erroneous data are worrisome and include inappropriate allocation of resources for research and provision of preventive and therapeutic services. Therefore, we assessed the validity of the use of ICD codes to identify hospitalizations with clinical evidence of an episode of acute CHF in data of The Corpus Christi Heart Project, a population-based surveillance program for hospitalized coronary heart disease (CHD).
PARTICIPANTS, METHODS, AND MATERIALS

STUDY GOAL
The overall aim of The Corpus Christi Heart Project was to elucidate the natural history of CHD as it occurs in 2 distinct populations, Mexican Americans and non-Hispanic whites, residing in a single community. Thus, surveillance was conducted with the primary aim of detecting hospitalizations for myocardial infarction, aortocoronary bypass surgery, percutaneous transluminal coronary angioplasty, and related revascularization procedures. The design of The Corpus Christi Heart Project has been described in detail previously13,16 and is summarized below.

STUDY POPULATION
The target population of this surveillance program included all residents of Nueces County, Texas, aged 25 through 74 years. Data for this article were collected from May 1, 1988, through April 30, 1994. Nueces County, located on the Texas Gulf Coast, had an estimated population of approximately 291,145 in 1990, with approximately 95% of the county’s population residing within the central city of Corpus Christi.17 Medical care in Nueces County was self-contained, a situation that is essential to the accurate population-based ascertainment of cardiac events. Surveillance occurred at all times in all 7 hospitals.

CASE ASCERTAINMENT AND DATA COLLECTION
Potentially eligible cases were ascertained on a concurrent basis from May 1, 1988, through April 30, 1994, by monitoring admissions to special care units at the 7 hospitals. This process is similar to the “hot pursuit” procedure used in the World Health Organization’s Monitoring of Trends and Determinants of Cardiovascular Disease (MONICA) Project.18 Special care units were monitored regularly to identify patients admitted with diagnoses possibly indicative of CHD and those who underwent bypass surgery or revascularization. To ensure complete ascertainment of cases of hospitalized myocardial infarction, bypass surgery, and revascularization, lists of discharges with selected CHD-related diagnoses (Table 1) were obtained from each hospital. Those hospitalizations, which had not already been identified through the concurrent procedures described earlier, were reviewed. Using capture-mark-recapture techniques, completeness of ascertainment of myocardial infarction was estimated to be 97.9% and did not differ by sex or ethnicity. For all potential cases, hospital records were abstracted when first available for review, usually 4 to 6 months after discharge. Data were collected on the course and care of the patient during the hospital stay, sociodemographic factors, medical history, electrocardiograms, and hospital discharge diagnostic codes.

CLASSIFICATION OF ETHNICITY, MYOCARDIAL INFARCTION, AND CHF
During the first 2 years of ascertainment, the main method of classification of ethnicity was by self-identification during an in-hospital interview as to “Spanish origin” and for the subcategory of “Mexican origin,” as adopted in the 1980 US Bureau of the Census,19 supplemented by medical record abstraction. Ethnicity of study subjects who were not interviewed was obtained from the hospital records. Among 683 individuals interviewed during the first 2 years of the project, the medical record and the interview data concerning

RESULTS
Between May 1, 1988, and April 30, 1994, a total of 5083 hospitalizations of interest were identified, including, in nonexclusive categories, 1154 hospitalizations (22.7%) for definite myocardial infarction, 2916 (57.4%) for possible myocardial infarction, 1696 (33.4%) for aortocoronary bypass surgery, and 1402 (27.6%) for percutaneous transluminal coronary angioplasty. The mean age (SD) of the population was 60.4 (10.1) years, with a median age of 63 years. Thirty-seven percent (1883/5083) of admissions were women, including 903 Mexican Americans, 798 non-Hispanic whites, and 182 of other ethnicity; 63.0% (3200/5083) of admissions were men, including 903 Mexican Americans, 798 non-Hispanic whites, and 182 of other ethnicity.

The prevalence of a medical record documented episode of acute CHF was 27.1% (1376/5083) when defined either by a medical record notation in a physician’s progress note or the discharge summary or by a notation of the presence of pulmonary edema by chest radiograph. Evidence of an episode of acute CHF was found in the progress notes or discharge summaries for 24.0% (1222/5083) of hospitalizations and in a report from a chest radiograph for 20.5% (1040/5083) of hospitalizations. The 2 sources of information agreed for 64.4% (886/1376) of the patients classified as having clinical evidence of an episode of acute CHF, 24.4% (336/1376) were classified on the basis of hospital diagnoses (Table 1), and 11.2% (154/1376) were classified on the basis of data from the chest radiograph alone. Overall agreement of classification was 90.4% (4593/5083, κ = 0.72, P < .001). The prevalence of clinical evidence of an episode of acute CHF varied by ascertainment of ICD code, being 33.4%, 16.0%, 37.5%, 14.5%, 23.8%, 40.2%, 43.2%, 34.8%, and 8.6%, among discharges with a 410, 411, 412, 413, 414, 427, 429, 440, or 786.5 ICD discharge diagnostic code, respectively.

The prevalence of CHF-related ICD discharge codes is given in Table 2. One or more CHF-related ICD discharge diagnostic codes were assigned to 1197 (23.5%) hospitalizations; 367 were assigned as the primary diagnosis. The other 830 hospitalizations with a CHF-related code assigned, representing 17.6% of all remaining hospitalizations, had another diagnosis listed first. Only ICD codes 428 (20.4%) and 402 (2.6%) were assigned to more than 1% of these CHF-related hospitalizations. Given the almost exclusive use of these 2 codes, subsequent analyses focused on the performance of 3 ICD discharge code-based classification algorithms for
CHF (Table 3): (1) ICD code 428 alone, (2) either ICD code 428 or 402, and (3) any CHF-related ICD code.

The sensitivity of these ICD code–based algorithms ranged from 62.8% to 67.1%; hence, approximately one third of the patients with clinical evidence of an episode of acute CHF documented in the medical record were missed by these 3 algorithms. Although specificity exceeded 92% for all 3 algorithms, the magnitude of this misclassification was sufficient to result in positive predictive values that ranged from a high of 83.5% to a low of 77.1%. The negative predictive value was approximately 88% for all 3 algorithms. The net effect of this misclassification was to underenumerate hospitalizations with clinical evidence of an episode of acute CHF. Whereas 1376 patients had clinical evidence of an episode of acute CHF, or a notation of pulmonary edema in a discharge summary that the patient experienced an episode of acute CHF, the 3 algorithms were examined: (1) the presence of ICD code 428, (2) the presence of either ICD code 428 or 402, and (3) the presence of any of the ICD codes listed in Table 2. The sensitivity, specificity, predictive values (positive and negative), and extent of underenumeration or overenumeration were examined for the ICD discharge diagnosis code–based classification algorithms in comparison with the clinical documentation classification. Sex, ethnic, and age differences in the performance of the 3 ICD code–based algorithms were examined using cross tabulations with χ² testing.

In this article, 3 different ICD code–based classification algorithms for CHF were shown to underenumerate hos-
Congestive heart failure (CHF) is a common condition that increases substantially by CHF. The costs related to CHF are also great. Congestive heart failure is the single most frequent cause of hospitalization in the Medicare-recipient population; the estimated direct and indirect costs attributable to CHF exceeded $18.8 billion in 1997 in the United States. As the size of the elderly population increases, the substantial morbidity and mortality attributable to CHF will continue to increase.

Congestive heart failure is an important public health problem, in part, because survival following diagnosis is poor. Only 80% survived 3 months and 66% survived 1 year in one population-based series. Survival was 65.3% and 31.0% at 1 and 5 years, respectively, in a nationally representative series; 5-year survival was 38% for women and 25% for men in the Framingham study. Although survival can be improved with the use of effective therapy such as angiotensin-converting enzyme inhibitors and β-adrenergic receptor blockers, mortality remains substantial. Readmission is common among survivors, occurring in almost half of the patients with CHF within 6 months of hospital discharge. Quality of life is impaired substantially by CHF. The costs related to CHF are also great. Congestive heart failure is the single most frequent cause of hospitalization in the Medicare-recipient population; the estimated direct and indirect costs attributable to CHF exceeded $18.8 billion in 1997 in the United States. The findings of this study support the contention that all of these figures are substantial underestimates of the true burden of morbidity, mortality, and cost of CHF.

These results were generated from data collected in a population-based surveillance program for hospitalized CHD. Consequently, hospitalizations that were not assigned a CHD-related discharge diagnosis were not examined. The likely effect of this limitation to patients with a CHD diagnosis is an underestimate of the number of hospitalizations with CHF that are missed by reliance on the identification of CHF-related discharge diagnostic ICD codes. The CHF-related codes were not part of the case ascertainment procedures for CHD; thus, we were limited to an examination of the agreement between ICD coding and medical record documentation in a subset of patients with CHF, those with a CHD-related diagnosis. The performance of these CHF classification algorithms might differ systematically between patients with and without CHD. Furthermore, since subgroup differences in per-

### Table 1. Coronary Heart Disease–Related Discharge International Classification of Diseases (ICD), Ninth Revision, Clinical Modification Codes Used for Case Ascertainment and Prevalence Among Study Population: The Corpus Christi Heart Project

<table>
<thead>
<tr>
<th>ICD Code</th>
<th>Diagnosis</th>
<th>Prevalence, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>410</td>
<td>Acute myocardial infarction</td>
<td>34.8</td>
</tr>
<tr>
<td>411</td>
<td>Other acute and subacute forms of ischemic heart disease</td>
<td>24.8</td>
</tr>
<tr>
<td>412</td>
<td>Old myocardial infarction</td>
<td>0.2</td>
</tr>
<tr>
<td>413</td>
<td>Angina pectoris</td>
<td>12.5</td>
</tr>
<tr>
<td>414</td>
<td>Other forms of chronic ischemic heart disease</td>
<td>56.1</td>
</tr>
<tr>
<td>427</td>
<td>Cardiac dysrhythmias</td>
<td>27.3</td>
</tr>
<tr>
<td>429</td>
<td>Ill-defined descriptions and complications of heart disease</td>
<td>5.7</td>
</tr>
<tr>
<td>440</td>
<td>Atherosclerosis</td>
<td>3.1</td>
</tr>
<tr>
<td>786.5</td>
<td>Chest pain</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Note: CHF indicators congestive heart failure, any of the codes listed in Table 2.

### Table 2. Prevalence of Assignment of Congestive Heart Failure–Related Discharge International Classification of Diseases (ICD), Ninth Revision, Clinical Modification Codes Among 5083 Hospitalizations for Coronary Heart Disease: The Corpus Christi Heart Project

<table>
<thead>
<tr>
<th>ICD Code</th>
<th>Diagnosis</th>
<th>Prevalence, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>398.91</td>
<td>Rheumatic heart failure</td>
<td>0.0</td>
</tr>
<tr>
<td>402.x1</td>
<td>Hypertensive heart disease with congestive heart failure</td>
<td>2.6</td>
</tr>
<tr>
<td>404.x</td>
<td>Hypertensive heart and renal disease</td>
<td>0.3</td>
</tr>
<tr>
<td>415.0</td>
<td>Acute cor pulmonale</td>
<td>0.0</td>
</tr>
<tr>
<td>416.9</td>
<td>Chronic pulmonary heart disease, unspecified</td>
<td>0.0</td>
</tr>
<tr>
<td>425.4</td>
<td>Other primary cardiomyopathies</td>
<td>0.0</td>
</tr>
<tr>
<td>428.x</td>
<td>Heart failure</td>
<td>20.4</td>
</tr>
<tr>
<td>429.4</td>
<td>Functional disturbances following cardiac surgery</td>
<td>0.0</td>
</tr>
<tr>
<td>514</td>
<td>Pulmonary congestion and hypostasis</td>
<td>0.1</td>
</tr>
<tr>
<td>518.4</td>
<td>Acute edema of lung, unspecified</td>
<td>0.0</td>
</tr>
<tr>
<td>786.0</td>
<td>Dyspnea and respiratory abnormalities</td>
<td>0.6</td>
</tr>
<tr>
<td>Any of the above</td>
<td></td>
<td>23.5</td>
</tr>
</tbody>
</table>

Note: CHF indicates congestive heart failure, any of the codes listed in Table 2.

### Table 3. Performance of Three International Classification of Diseases (ICD), Ninth Revision, Clinical Modification Discharge Code–Based Classification Algorithms for Congestive Heart Failure Among 5083 Hospitalizations for Coronary Heart Disease: The Corpus Christi Heart Project

<table>
<thead>
<tr>
<th>ICD Discharge Codes</th>
<th>Sensitivity (n = 1376)</th>
<th>Specificity (n = 3707)</th>
<th>Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>428.x</td>
<td>864 (62.8)</td>
<td>3536 (95.4)</td>
<td>1035 (83.5)</td>
</tr>
<tr>
<td>428.x or 402.x</td>
<td>911 (66.2)</td>
<td>3458 (93.3)</td>
<td>1160 (78.5)</td>
</tr>
<tr>
<td>Any CHF† code</td>
<td>923 (67.1)</td>
<td>3433 (92.6)</td>
<td>1197 (77.1)</td>
</tr>
</tbody>
</table>

Note: *Data collected among Nueces County, Texas, residents, aged 25 through 74 years, from May 1, 1988, through April 30, 1994. All values are given as number (percentage).

†CHF indicates congestive heart failure, any of the codes listed in Table 2.
formance of the algorithms were observed in this study, especially by age, these algorithms might be expected to perform differently in populations that differ in sociodemographic characteristics or across geographic regions. Furthermore, performance may change over time. These considerations influence the generalizability of these findings. Fisher et al\textsuperscript{31} examined the accuracy of \textit{ICD} coding as performed in 1985 for the identification of CHF among Medicare recipients. The referent standard were \textit{ICD} codes applied by accredited medical records technicians who were masked to the diagnoses assigned by the hospitals. Thus, this study represents an assessment of intercoder agreement. Given this caveat, sensitivity for any of the following \textit{ICD} codes, ie, 402.01, 402.11, 402.91, and 428-428.9, was reported to be 85\% when the principal diagnosis was examined for the ability to detect the primary reason for hospitalization and 89\% when any of the diagnoses were examined for the ability to detect a medically important condition identified and treated during the hospitalization; specificity was 99\% and 95\%, respectively. The methods used by Fisher et al\textsuperscript{31} would be expected to result in a more optimistic estimate of sensitivity in comparison with the methods used in our investigation.

In support of the generalizability of these results are several strengths of this surveillance effort. First, as described earlier, the ascertainment of CHD hospitalizations is estimated to be almost complete. Second, the population under surveillance is served by community hospitals, rather than academic medical centers; therefore, the quality of care, medical record documentation, and \textit{ICD} coding may be more typical of that available at the vast majority of hospitals in the United States than would be the results from the experience at a major academic institution. Finally, both women and men and all ethnic groups represented in the community were examined. The quality of the medical record evidence used to document CHF is another potential limitation of this study. It would be preferable to collect prospective, standardized data to identify episodes of acute CHF in population-based surveillance programs. In the absence of those data, the evidence examined in this article seems sufficiently valid to call attention to the potential inadequacies present in the \textit{ICD} code–based approach to surveillance of CHF.

If these results can be generalized to all CHD-related hospitalizations, then current estimates of the number of hospitalizations for CHF are substantially misleading. The failure to capture evidence of CHF through the assignment of appropriate \textit{ICD} codes may be compounded by decisions regarding the ordering of the \textit{ICD} codes. According to data from the National Hospital Discharge Survey for 1995,\textsuperscript{3,6} there were approximately 2.7 million discharges with the \textit{ICD} code 428 among the diagnoses, but only 874,000 with \textit{ICD} code 428 as the first listed diagnosis. This latter number is cited to describe the burden of CHF. For all CHDs, there were approximately 4.0 million discharges for which a CHD diagnosis was listed first\textsuperscript{7} and approximately 8.5 million discharges with at least 1 CHD diagnosis (L. J. Kozack, National Center for Health Statistics, oral communication, February 1999). Thus, there were approximately 1.4 million discharges with a CHD diagnosis listed first and approximately 3.8 million discharges with any CHD diagnosis but no CHF diagnosis. The results of this study support the contention that as much as 8.4\% of the 1.4 million to 5.8 million, or approximately 114,000 to 485,000 additional hospitalizations, had clinical evidence of an episode of acute CHF, for a total of approximately 2.8 million to 3.2 million CHF-related hospitalizations. Presumably, an additional number of hospitalizations with clinical evidence of an episode of acute CHF would come from other diagnosis groups (non–CHD-related hospitalizations). The combined effects of underutilization of CHF-related \textit{ICD} codes and reliance on the first listed diagnosis provide an indicator of CHF that may capture fewer than 27.5\% (874,000/3,183,000) to 31.1\% (874,000/2,812,000) of hospitalizations with clinical evidence of acute CHF. Using the number of hospitalizations with \textit{ICD} code 428 in any diagnostic position, the “all-listed” approach would still result in an undercount of hospitalizations with clini-
CONCLUSIONS

Our findings support the contention that CHF is a much greater public health problem than currently recognized. If the previously reported increasing trends are an accurate reflection of the underlying changes in the natural history of CHF at the population level, then the future morbidity and mortality attributable to CHF will be much greater than envisioned. The potential policy implications of this underenumeration are worrisome. Given falsely low estimates of the current and future occurrence of CHF, insufficient resources may be allocated for relevant research, medical care, and preventive efforts. The results of our study should stimulate action to develop and implement more accurate surveillance systems for CHF. The collection of accurate data regarding the occurrence, care, outcomes, and costs of CHF should become a national priority. Furthermore, greater emphasis should be devoted toward the development and implementation of effective preventive and therapeutic interventions for this increasingly important, but poorly delineated, public health challenge.

Accepted for publication March 23, 1999.

This study was supported by grant HL38429 from the National Heart, Lung, and Blood Institute, Bethesda, Md (Dr Nichaman).

We thank the Corpus Christi Advisory Committee, the administrators and staffs of the 7 participating hospitals, and the Corpus Christi—Nueces County Department of Public Health for their support of this project.

Reprints: David C. Goff, Jr, MD, PhD, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157-1063 (e-mail: dgoff@wfuhs .wfubmc.edu).

REFERENCES

21. CCSP Coordinating Center Program. Final Report to the NHLBI. Division of Clinical Investigation, Dept of Epidemiology and Preventive Medicine, University of Maryland, Baltimore. Community Cardiovascular Surveillance Program, June 1, 1984.