Methodology

Accuracy of Self-Reported Heart Failure. The Atherosclerosis Risk in Communities (ARIC) Study

RICKY CAMPLAIN, PhD,[1,2] ANNA KUCHARSKA-NEWTON, PhD,[2] LAURA LOEHR, PhD,[2] THOMAS C. KEYSERLING, MD, MPH,[3] J. BRADLEY LAYTON, PhD,[2] LISA WRUCK, PhD,[4] AARON R. FOLSOM, PhD,[5] ALAIN G. BERTONI, MD,[6] AND GERARDO HEISS, PhD[2]

Flagstaff, Arizona; Chapel Hill, Durham and Winston-Salem, North Carolina; and Minneapolis, Minnesota

ABSTRACT

Objective: The aim of this work was to estimate agreement of self-reported heart failure (HF) with physician-diagnosed HF and compare the prevalence of HF according to method of ascertainment.

Methods and Results: ARIC cohort members (60–83 years of age) were asked annually whether a physician indicated that they have HF. For those self-reporting HF, physicians were asked to confirm their patients' HF status. Physician-diagnosed HF included surveillance of hospitalized HF and hospitalized and outpatient HF identified in administrative claims databases. We estimated sensitivity, specificity, positive predicted value, kappa, prevalence and bias–adjusted kappa (PABAK), and prevalence. Compared with physician-diagnosed HF, sensitivity of self-report was low (28%–38%) and specificity was high (96%–97%). Agreement was poor (kappa 0.32–0.39) and increased when adjusted for prevalence and bias (PABAK 0.73–0.83). Prevalence of HF measured by self-report (9.0%), ARIC-classified hospitalizations (11.2%), and administrative hospitalization claims (12.7%) were similar. When outpatient HF claims were included, prevalence of HF increased to 18.6%.

Conclusions: For accurate estimates HF burden, self-reports of HF are best confirmed by means of appropriate diagnostic tests or medical records. Our results highlight the need for improved awareness and understanding of HF by patients, because accurate patient awareness of the diagnosis may enhance management of this common condition. (*J Cardiac Fail 2017;23:802–808*)

Key Words: Heart failure, administrative claims, medical records, self-report.

From the [1]Center for Health Equity, Northern Arizona University, Flagstaff, Arizona; [2]Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, United States of America; [3]Department of Medicine, The University of North Carolina, Chapel Hill, North Carolina, United States of America; [4]Center for Predictive Medicine, Duke Clinical Research Institute, Durham, North Carolina, United States of America; [5]Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, Minnesota, United States of America and [6]Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America.

Manuscript received February 23, 2017; revised manuscript received July 19, 2017; revised manuscript accepted September 1, 2017.

Reprint requests: Ricky Camplain, PhD, Center for Health Equity, Northern Arizona University, PO Box 4065, Flagstaff, AZ 86011-4065. Tel: 928 523 5165; Fax: 928-523-4648. E-mail: ricky.camplain@nau.edu.

Funding: The Atherosclerosis Risk in Communities Study is a collaborative study supported by National Heart, Lung, and Blood Institute contracts HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C. RC was supported by the National Heart, Lung, and Blood Institute Graduate Research Supplement Diversity Supplement (HHSN268201100007C).

1071-9164/$ - see front matter

Published by Elsevier Inc.

https://doi.org/10.1016/j.cardfail.2017.09.002

Patient report of having physician-diagnosed heart failure (HF) has been used clinically and to quantify the burden of HF in the community. In the National Health and Nutrition Examination Survey (NHANES), self-report of a physician diagnosis of HF is used to estimate the prevalence of HF in the United States (US). HF is difficult to diagnose and identify in population research. Although this difficulty is not limited to estimating the accuracy of self-reported HF, estimates comparing self-reported HF and clinically diagnosed HF,[1] medical records,[2–8] and health administrative data[4,9] may be affected. Given the complexity of diagnosing and classifying HF, it may be difficult for health professionals to accurately inform patients of the diagnosis, which may limit the accuracy of self-report of physician-diagnosed HF and therefore the practical advantage in using self-report to estimate prevalence of HF.

Compared with self-reports of HF, self-reports of coronary heart disease and myocardial infarction have greater validity. However, most prior studies compared self-reported

HF with only a single benchmark. Because no consensus exists on a single HF classification scheme, examination of the agreement and validity of self-reported HF against different benchmark definitions of HF is desirable. Therefore, in the Atherosclerosis Risk in Communities (ARIC) study cohort, we addressed an individual's ability to convey prior diagnoses of HF by estimating the agreement of self-reported HF with confirmation of HF by the participant's health care provider, hospital medical record extraction, and the presence of HF International Classification of Disease, 9th Edition, Clinical Modification (ICD-9-CM) codes in administrative claims. Estimates of the prevalence of HF based on these methods of ascertainment were compared.

Methods

Study Population

The ARIC study is an ongoing prospective cohort of 15,792 men and women aged 45–64 years at baseline (1987–1989) recruited from the following 4 US communities: Forsyth County, North Carolina; Jackson, Mississippi; Washington County, Maryland; and suburban Minneapolis, Minnesota.[10] Standardized physical examinations and interviewer-administered questionnaires were conducted at baseline and at 4 follow-up visits through 2012. Participants were additionally followed annually (from 1987) and semiannually (beginning in 2012) through telephone interviews and review of hospitalization and vital status records. Institutional Review Boards at each participating institution approved the study, and every participant provided written informed consent at each examination.

Self-Reported Heart Failure

Starting in 2005, participants were asked questions regarding their HF status during annual follow-up telephone interviews. Participants who reported a diagnosis of HF or who reported that their heart was weak, that their heart did not pump as strongly as it should, or that they had fluid in their lungs before 2005 were classified as having prevalent HF. Participants free of self-reported HF before 2005 were asked at the initial (2005) and subsequent annual telephone interviews whether a doctor said that they had HF, or if their heart was weak or did not pump as strongly as it should, since the last time they were contacted. The approximate date of diagnosis and whether or not the participant reported an HF-related hospitalization was also collected. Participants were classified as having new self-reported HF if they answered "yes" to either of the above questions.

Physician-Confirmed Heart Failure

In parallel with the ARIC participants' self-reported HF status, confirmation of HF from the participants' physicians was sought. If a participant reported being diagnosed with HF, or told by a physician that their heart was pumping weakly, they were asked to authorize a release of medical information from their physician. Once a signed authorization was obtained, the provider of medical care was sent a survey to confirm their patient's HF status, HF characteristics, and treatment status.

ARIC-Classified Heart Failure

Before 2005, ARIC recorded ICD-9-CM codes but did not abstract HF records; we excluded participants who had an HF-related ICD-9-CM discharge code of 428.x in any position. Starting in 2005 ARIC conducted continuous surveillance of hospitalized HF events, including acute decompensated HF (ADHF) and chronic stable HF among the cohort participants. The medical records of all cohort hospitalizations are abstracted by trained study staff adhering to a common protocol.[11] Each record is reviewed for any evidence of relevant HF symptoms or mention by the physician of HF in the hospital record. If the hospital record contains such confirmation, a detailed abstraction is completed. Abstracted data include the elements required by 4 diagnostic criteria commonly used (Framingham, modified Boston, NHANES, Gothenburg) and ICD-9-CM codes. Each hospitalization eligible for full abstraction is independently reviewed by 2 physicians who are provided portions of the medical record and a report of the abstracted data. Reviewers then classify hospitalizations as definite ADHF, possible ADHF, chronic stable HF, HF unlikely, or unclassifiable. Hospitalizations classified as definite or possible ADHF and chronic stable HF were considered to be confirmation of HF for our study.

Heart Failure Identified From Administrative Claims

ARIC cohort participants' identifiers were linked with Centers for Medicare and Medicaid Services (CMS) Medicare claims for the years 1991–2013 by means of a finder file that included participants' Social Security Numbers, sex, and date of birth. From the total number of study participants with available Social Security Numbers (n = 15,744), 238 died before 1991 and 607 died after 1991 but before reaching the Medicare eligibility age of 65 years, leaving 14,899 eligible ARIC participants. A crosswalk file was used to identify ARIC cohort participants eligible for Medicare coverage. The crosswalk file between the ARIC study finder file and the Medicare Beneficiary Summary file yielded 14,702 ARIC cohort IDs for analysis (98.7% match).

Information regarding ARIC participant enrollment in fee-for-service (FFS) Medicare was obtained from monthly indicators of enrollment in Part A, Part B, and Medicaid buy-in available from annual Medicare Beneficiary Summary files. Continuous enrollment periods were created to indicate uninterrupted FFS Medicare coverage, defined as enrollment in Medicare Part A and Part B as well as lack of enrollment in a Medicare Advantage (HMO) plan. All inpatient and outpatient claims were linked.

Hospitalized HF was identified from Medicare Provider Annual Review records by ICD-9-CM code 428.x in any position. Outpatient HF was identified from claims with

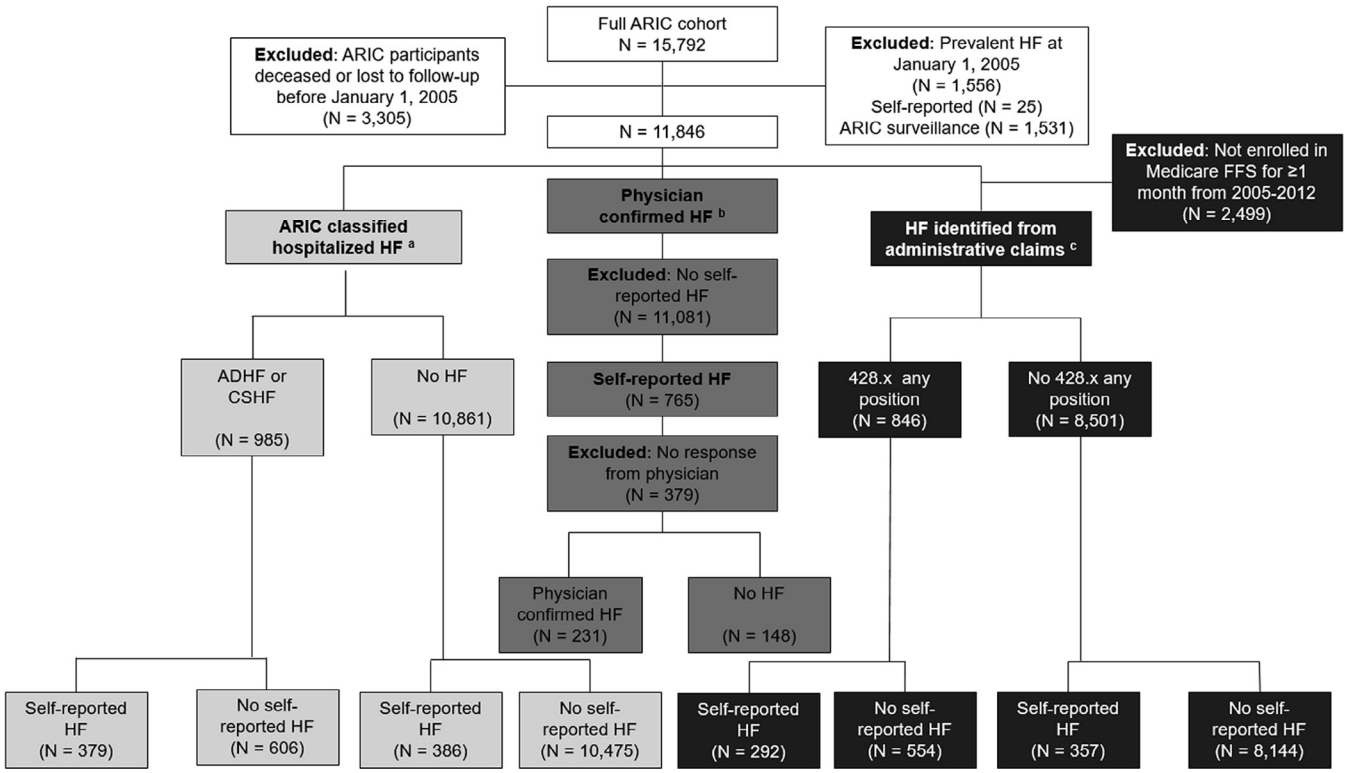

Fig. 1. Study design, Atherosclerosis Risk in Communities (ARIC) study (2005–2012). Self-reported heart failure (HF) ascertained from ARIC annual follow-up telephone interview. FFS, fee-for-service. [a]Definite and probable acute decompensated heart failure (ADHF) and chronic stable heart failure (CSHF) ascertained from reviewed abstracted medical records. [b]Ascertained from HF surveys sent to participants' physicians for confirmation of HF diagnosis. [c]Centers for Medicare and Medicaid Services Medicare claims, International Classification of Disease, 9th Edition, Clinical Modification (ICD-9-CM) code 428.x in any position present in the hospital record, ICD-9-CM 428.x in any position present in the hospital record, or ICD-9-CM 428.x present in the outpatient record.

Evaluation and Management service codes for new and established outpatient visits, consultations, and established preventive medicine visits matched with date of service found in the Carrier (Part B) claims. Similarly to hospitalized HF, outpatient HF events were identified by HF-specific ICD-9-CM codes 428.0–428.9.

Eligibility

To assess measurement properties of self-reported HF, participants were excluded if deceased or lost to follow-up before January 1, 2005 (Fig. 1). Participants were also excluded if they were hospitalized with an HF code (defined by an ICD-9-CM discharge code of 428.x in any position) or self-reported having HF before January 1, 2005.

To estimate and compare prevalence of self-reported HF, ARIC-classified HF hospitalizations, and HF identified from Medicare administrative claims diagnoses of participants alive from 2005 to 2012, we constructed a cohort of participants who had information available for all methods of HF ascertainment. Participants were included if they responded to any telephone interview questions regarding HF and were continuously enrolled in FFS Medicare for ≥1 calendar month from 2005 to 2012. No exclusions were made based on prevalent HF before 2005.

Statistical Analysis

To verify self-reported HF, we calculated the proportion of self-reported HF confirmed by physicians (verification). Because only ARIC participants who self-reported having HF (yes) were followed up for physician-confirmed HF (yes/no), other measures of validity were not possible with the use of physician confirmation (Table 1). To directly compare self-reported HF with other methods of HF ascertainment, we also calculated verification for ARIC-classified HF and HF ascertained from administrative claims.

To compare self-reported HF (yes/no) with ARIC-classified HF (yes/no) or administrative claims (yes/no), participants were classified as true positive, true negative, false positive, or false negative for each comparison. We evaluated agreement (the number of true positives and negatives divided by all participants), positive predictive value (PPV; probability that a participant has recorded HF if they self-reported HF), sensitivity (probability of positively self-reporting HF among participants with recorded HF), and specificity (probability of negatively self-reporting HF among participants without recorded HF; Table 1). Finally, we computed prevalence and bias–adjusted kappa (PABAK).[12] Because both prevalence and bias play a part in determining the magnitude of the kappa coefficient, and the prevalence of HF is relatively low (~10%),

Table 1. Agreement Estimates of Self-Reported Heart Failure Versus Objective Measures of Heart Failure—Atherosclerosis Risk in Communities (ARIC) study

Estimate	Physician-Confirmed Heart Failure	ARIC-Classified Heart Failure*	Heart Failure Identified From Administrative Claims†
Verification‡	✓	✓	✓
Kappa		✓	✓
PABAK		✓	✓
Sensitivity		✓	✓
Specificity		✓	✓
PPV		✓	✓

PABAK, prevalence and bias–adjusted kappa; PPV, positive predictive value.
✓, estimate calculated for objective measure of heart failure.
*Definite and probable acute decompensated heart failure and chronic stable heart failure ascertained from reviewed abstracted medical records.
†Centers for Medicare and Medicaid Services Medicare claims.
‡Confirmed heart failure among self-reporters of heart failure.

PABAK is considered to be an appropriate complement to the traditional kappa statistic. PABAK is calculated as $2I_o - 1$ where I_o is the observed agreement.

Prevalence and 95% confidence intervals (CIs) for self-reported HF, ARIC-classified HF hospitalizations, and HF identified from Medicare administrative claims diagnoses were estimated from 2005 through 2012. We made no attempt to align the years of HF report with HF hospitalization or administrative claim, but rather took affirmative responses at any time as having agreed. All analyses were conducted with the use of SAS V9.4 (Cary, North Carolina).

Results

Of the 15,792 members of the ARIC cohort, 3,305 were deceased (n = 2,884) or lost to follow-up (n = 421) before January 1, 2005, and 1,556 participants were classified as having prevalent HF, leaving 11,846 participants for the present analyses (Fig. 1). Among eligible ARIC participants, 59% were female, 25% were African American (Table 2), and ages ranged from 60 to 83 years in 2005. More than 40% of participants had hypertension, and 12% had diabetes at baseline (2005). ARIC participants with FFS Medicare coverage were similar in age, sex, race, and comorbidity to all eligible ARIC participants. Participants who self-reported HF were older and more likely to be male, and had higher prevalence of diabetes, hypertension, myocardial infarction, stroke, and coronary heart disease at baseline (Table 2).

Prevalence

Of the eligible participants enrolled in FFS Medicare, 649 self-reported HF (6.9%, 95% CI 6.4%–7.5%) during ≥1 telephone interview (n = 9,347), 824 had hospitalizations with a diagnosis of HF confirmed by ARIC (8.8%, 95% CI 8.2%–9.4%), 846 were hospitalized with a HF-related ICD-9-CM discharge code of 428.x identified from administrative claims (9.1%, 95% CI 8.5%–9.6%), and 1,391 were hospitalized or seen in the outpatient setting for HF identified from administrative claims (14.9%, 95% CI 14.2%–15.6%; Supplemental Table S1).

Table 2. Descriptive Characteristics by Heart Failure Ascertainment, n (%)—Atherosclerosis Risk in Communities (ARIC) study, 2005–2012

Characteristic	ARIC Participants Who May Be Classified as Having Heart Failure		
	Physician-Confirmed Heart Failure (n = 379)	ARIC-Classified Heart Failure* (n = 11,846)	Heart Failure Identified From Administrative Claims† (n = 9,347)
Age, y, mean ± SD‡	72 ± 5.6	71 ± 5.7	71 ± 5.5
Female	181 (48)	6,960 (59)	5,584 (60)
African-American	100 (26)	2,912 (25)	2,492 (27)
Center			
Forsyth County	58 (15)	3,104 (26)	2,049 (22)
Jackson	93 (24)	2,532 (21)	2,219 (24)
Minneapolis	104 (27)	3,236 (27)	2,266 (24)
Washington County	124 (33)	2,974 (25)	2,813 (30)
Comorbidities			
Diabetes§	82 (22)	1,386 (12)	1,161 (12)
Hypertension¶	236 (62)	5,270 (44)	4,340 (46)
Myocardial infarction**	67 (18)	551 (5)	452 (5)
Stroke††	24 (6)	331 (3)	288 (3)
Coronary heart disease‡‡	90 (24)	654 (6)	536 (6)

*Definite and probable acute decompensated heart failure and chronic stable heart failure ascertained from reviewed abstracted medical records.
†Centers for Medicare and Medicaid Services Medicare (CMS) fee-for-service inpatient and outpatient claims.
‡Age at start of follow-up (January 1, 2005).
§Fasting glucose ≥126 mg/dl, nonfasting glucose ≥200 mg/dl, using medication for diabetes, or self-reported physician diagnosis of diabetes before start of follow-up (January 1, 2005).
¶Systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or on medication for high blood pressure before start of follow-up (January 1, 2005).
**Prevalent myocardial infarction ascertained from ARIC surveillance.
††Prevalent stroke ascertained from ARIC surveillance.
‡‡Prevalent coronary heart disease ascertained from ARIC surveillance.

Table 3. Agreement Between Self-Report of Heart Failure and Objective Measures of Heart Failure, % (95% Confidence Interval)—Atherosclerosis Risk in Communities (ARIC) study, 2005–2012

Measure	Physician-Confirmed Heart Failure (n = 379)	ARIC-Classified Heart Failure* (n = 11,846)	Administrative Claims†	
			Heart Failure Hospitalizations‡ (n = 9,347)	Heart Failure Hospitalizations and Outpatient Visits§ (n = 9,347)
Verification¶	60.9 (56.0–65.9)	49.5 (46.0–53.1)	45.0 (41.1–48.8)	60.9 (57.1–64.6)
Sensitivity	–	38.5 (35.4–41.6)	34.5 (31.3–37.8)	28.4 (26.1–30.9)
Specificity	–	96.4 (96.1–96.8)	95.8 (95.3–96.2)	96.8 (96.4–97.2)
Positive Predicted Value	–	49.5 (45.9–53.1)	45.0 (41.3–48.9)	60.9 (57.0–64.6)
Kappa	–	0.39 (0.35–0.42)	0.34 (0.31–0.37)	0.32 (0.29–0.35)
PABAK	–	0.83 (0.82–0.85)	0.81 (0.79–0.82)	0.73 (0.71–0.75)

PABAK, prevalence and bias–adjusted kappa. Confirmed heart failure: physician-confirmed heart failure: n = 231; ARIC-classified heart failure: n = 379; hospitalized heart failure identified from administrative claims: n = 292; hospitalized and outpatient heart failure identified from administrative claims: n = 395. Prevalence: self-reported heart failure: 6.9%; ARIC-classified heart failure: 8.8%; hospitalized heart failure identified from administrative claims: 9.1%; hospitalized and outpatient heart failure identified from administrative claims: 14.9%.

*Definite and probable acute decompensated heart failure and chronic stable heart failure ascertained from reviewed abstracted medical records.

†Centers for Medicare and Medicaid Services Medicare claims.

‡International Classification of Disease, 9th Edition, Clinical Modification (ICD-9-CM) code 428.x in any position present in the hospital record.

§ICD-9-CM 428.x in any position present in the hospital record or ICD-9-CM 428.x present in the outpatient record.

¶Confirmed heart failure among self-reporters of heart failure.

Verification of Heart Failure

From January 1, 2005, to December 31, 2012, 379 of the 765 participants who self-reported HF (49.5%; Supplemental Table S2) had information regarding confirmed HF status from their physicians. For 231 of those, the participants' physicians confirmed that their patients had HF (60.9% agreement, 95% CI 56.0%–65.9%; Table 3).

Comparison of Self-Reported and ARIC-Classified Heart Failure

Of the 765 ARIC participants who self-reported HF, 379 (50%) were classified as having HF according to ARIC study criteria (Supplemental Table S2). The probability that a participant was classified by ARIC as having HF if they self-reported HF (PPV) was 49.5% (95% CI 45.9%–53.1%; Table 3). The sensitivity of self-reported HF versus ARIC-classified HF was low (38.5%, 95% CI 35.4%–41.6%) and specificity was high (96.4%, 95% CI 96.1%–96.8%). The agreement of self-reported HF with HF classified by ARIC was quantified as kappa 0.39 (CI 0.35–0.42) and PABAK 0.83 (95% CI 0.82–0.85).

Comparison of Self-Reported Heart Failure and Heart Failure Identified From Administrative Claims

Of the 11,846 eligible ARIC participants, 9,347 were enrolled in FFS Medicare for ≥1 calendar month from January 1, 2005, to December 31, 2012. Among them, 45% (95% CI 41.2%–48.8%) of those who self-reported HF had been previously hospitalized with an HF code according to administrative claims, and 60.9% (95% CI 57.1%–64.6%) of participants who self-reported HF had been hospitalized or were seen in the outpatient setting with an HF code (Table 3). Overall agreement between self-reported HF and hospitalized HF identified from administrative claims, as measured by kappa, was 0.34 (95% CI 0.31–0.37), whereas the agreement between self-reported HF and with HF identified from administrative hospitalization and outpatient claims decreased to 0.32 (95% CI 0.29–0.34). Similarly, the sensitivity, PPV, and PABAK decreased when outpatient claims were included in the comparison definition, partly owing to the shift of participants from true negative to false negative (Supplemental Table S2).

Discussion

Because health-status questionnaires continue to be important tools in clinical settings and in public health research, we assessed the accuracy of self-reported HF compared with physician-diagnosed HF and evaluated the agreement between self-reported HF and a diagnosis of HF by the individual's health care provider, prior indications of HF in an individual's hospital records, and HF diagnostic codes in administrative claims. We observed low agreement (kappa 0.32–0.39) between HF self-reported by participants and physician-diagnosed HF, and self-reports of HF were characterized by frequent false positives and false negatives. Adjustment of kappa statistics for prevalence and bias improved agreement to 0.73–0.83. Sensitivity was low (28%–38%) and specificity was high (96%–97%) for self-reported HF compared with all measures of physician-diagnosed HF. The prevalences of ARIC-classified HF and of hospitalized HF ascertained from administrative claims were similar. However, the prevalence of self-reported HF was lower, and the prevalence of pooled hospitalized and outpatient HF ascertained from administrative claims was higher than the prevalence of hospitalized HF.

Multiple Benchmarks

Previous studies have compared self-reported HF to a single validation benchmark, making comparisons across studies difficult. Therefore, our study was designed to directly compare the agreement between self-reported HF and

physician-diagnosed HF, including in-depth medical record review, administrative claims, and confirmation from a physician. Among previous reports estimating the agreement of self-reports of physician-diagnosed HF and medical records, agreement (kappa) ranged from 0.30 to 0.48 with low to fair sensitivity (31%–69%) and high specificity (91%–97%).[2–8] Our study yielded similar results with low sensitivity (39%) and agreement (0.32–0.39) and high specificity (96%).

Baumeister et al compared self-reported HF with physical examinations and laboratory data[1] and reported higher agreement (kappa 0.74) and sensitivity (89%) than other studies.[2–8] Although our data does not allow a direct comparison with those from Baumeister et al, our validation results were highest for self-reported HF compared with physician confirmation of HF (61%). Although we chose confirmation of HF by the interviewee's physician as our benchmark, we sought physician confirmation only from study participants who self-reported HF for the 1st time during an annual telephone interview. Approximately 50% of physicians ultimately returned the HF survey, and as a result 379 participant self-reports of HF were available for verification by a physician. Although the profile of study participant characteristics did not differ according to their physicians' responses to the survey, the low response constrains the generalizability of the observed rate of physician confirmation of self-reported HF.

Similarly to medical record review, studies comparing self-report with HF identified from hospital administrative claims returned low agreement (0.19–0.33) and sensitivity (26%), as well as high specificity (99%).[4,9] Among ARIC participants with FFS Medicare, sensitivity (28%–35%), specificity (96%–97%), and agreement (0.32–0.34) were similar. Although our results can be compared with other studies, administrative data have high variation in validity for recording HF,[13] and significant differences exist in the manner in which hospitalizations are recorded in administrative claims compared with medical records.[14,15] Despite such differences, there is a high degree of agreement between hospital records and Medicare administrative claims in the identification of individuals discharged from hospital with a HF diagnostic codes.

Patients' Awareness and Understanding of Heart Failure

Previous studies indicated that self-reported conditions characterized by complex and nonspecific symptoms—such as HF—have poor agreement compared with objective measures of the condition, in contrast to conditions that are better characterized and more easily diagnosed, such as myocardial infarction, stroke, and diabetes.[2–6,9,16] The low agreement of self-reported HF with physician-diagnosed HF may therefore reflect the complexity of HF as a syndrome and its varied presentations. The current lack of consensus on the combination of signs and symptoms to classify HF, resulting in several classification schema for use in clinical and research settings,[17] may lead clinicians and physicians to be cautious when conveying a diagnosis of HF to their patients. The increasing use of functional tests and biomarkers to diagnose and manage HF in primary care should assist clinicians in making more accurate diagnoses of HF, conveying this information to patients, and engaging them in an evidence-informed management plan. Similarly to the difficulties that practitioners face in conveying an accurate diagnosis of HF to a patient, patients face challenges discerning a diagnosis of HF from other medical conditions with similar symptoms and characteristics. As a case in point, the chronic morbidity that characterized many of the study participants who self-reported HF may have contributed to the lack of agreement with the practitioners' diagnoses: among the 148 individuals whose physicians indicated that their patients did not have the HF that they had self-reported, 69.6% had a diagnosis of atrial fibrillation (n = 103), angina pectoris (n = 37), previous myocardial infarction (n = 20), or another form of ischemic heart disease (n = 60). Conditions such as these may have led to self-reports of HF.

Prevalence and Bias–Adjusted Kappa

Because agreement measures may be influenced by the prevalence of HF (<10%), kappa coefficients may be decomposed into factors that reflect observed agreement, bias, and prevalence. Particularly for comparisons across studies, it is informative to report the kappa coefficient values and the effects of bias and prevalence on agreement. In contrast to the observed low kappa estimates, PABAK estimates were considerably higher (0.73–0.83), although it should be considered that although PABAK adjusts for prevalence it may overestimate agreement. Thus, the effects of bias and prevalence on the magnitude of kappa are of interest and, although it has been argued that they should not be adjusted for,[18] we provided such measures of agreement[19] alongside the value of kappa.[20]

Prevalence of HF

We estimated the prevalence of HF by means of ascertainment method in the study population and observed that 6.9% of ARIC participants, whose average age was 71 years, self-reported HF during annual follow-up telephone interviews. This is somewhat lower than the 8%–10% reported by other population-based surveys[21–23] and higher than the 4.6% self-reported by similar-age populations in the Health and Retirement Study.[23] Although the prevalence of ARIC-classified HF hospitalizations (8.8%) and hospitalized HF identified from administrative claims (9.1%) were similar, only 562 of the 1,108 participants with either type of HF hospitalizations were identified in both sources (Supplemental Table S3). In turn, only 244 participants (17.6%) who self-reported HF were identified by means of ARIC adjudication of medical records and had an HF hospitalization identified from administrative claims.

As hypothesized, including outpatient claims substantially increased the prevalence of HF ascertained from administrative claims. Although approximately one-half of HF patients are managed in the outpatient setting (ie, not as-

sociated with a hospitalization) and HF patients are increasingly diagnosed and treated in outpatient clinics,[24,25] reported population estimates of HF rarely include outpatient HF. These temporal trends in the medical care of HF, coupled with the variability in prevalence estimates mentioned above, underscore the importance of specifying the definition and source of HF events when reporting frequency estimates.

Conclusion

Our results suggest that agreement of self-reported HF with physician-confirmed HF, prior indications of HF in the patient's health record, and HF identified from administrative claims is low to fair, and that self-reported HF is insensitive. Although prevalence estimates of self-reported HF are similar to those from hospitalizations with HF discharge diagnoses, the agreement between these sources in the identification of "cases" of HF is low to poor. For accurate population estimates of HF, self-reported HF data should be coupled with other sources, such as diagnostic tests or medical records. The observed low accuracy of self-reported HF suggests that complexities in the diagnosis of HF make it challenging for health professionals to consistently and accurately convey this diagnosis to patients. These results highlight the need for improved awareness and understanding of HF by patients to enable their participation in the management of HF toward improved clinical outcomes.

Acknowledgments

The authors thank the staff and participants of the ARIC study for their important contributions.

Supplementary Data

Supplementary data related to this article can be found at doi:10.1016/j.cardfail.2017.09.002.

References

1. Baumeister H, Kriston L, Bengel J, Härter M. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental-physical comorbidity. J Clin Epidemiol 2010;63:558–65.
2. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004;57:1096–103.
3. Merkin SS, Cavanaugh K, Longenecker JC, Fink NE, Levey AS, Powe NR. Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. J Clin Epidemiol 2007;60:634–42.
4. Teh R, Doughty R, Connolly M, Broad J, Pillai A, Wilkinson T, et al. Agreement between self-reports and medical records of cardiovascular disease in octogenarians. J Clin Epidemiol 2013;66:1135–43.
5. Italian Longitudinal Study on Aging Working Group. Prevalence of chronic diseases in older Italians: comparing self-reported and clinical diagnoses. Int J Epidemiol 1997;26:995–1002.
6. Simpson CF, Boyd CM, Carlson MC, Griswold ME, Guralnik JM, Fried LP. Agreement between self-report of disease diagnoses and medical record validation in disabled older women: factors that modify agreement. J Am Geriatr Soc 2004;52:123–7.
7. Tisnado DM, Adams JL, Liu H, Damberg CL, Chen WP, Hu FA, et al. What is the concordance between the medical record and patient self-report as data sources for ambulatory care? Med Care 2006;44:132–40.
8. Heckbert SR, Kooperberg C, Safford MM, Psaty BM, Hsia J, McTiernan A, et al. Comparison of self-report, hospital discharge codes, and adjudication of cardiovascular events in the Women's Health Initiative. Am J Epidemiol 2004;160:1152–8.
9. Muggah E, Graves E, Bennett C, Manuel DG. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. BMC Public Health 2013;13:16.
10. ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. Am J Epidemiol 1989;129:687–702.
11. Rosamond WD, Chang PP, Baggett C, Johnson A, Bertoni AG, Shahar E, et al. Classification of heart failure in the Atherosclerosis Risk in Communities (ARIC) study: a comparison of diagnostic criteria. Circ Heart Fail 2012;5:152–9.
12. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46:423–9.
13. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. Can J Cardiol 2010;26:306–12.
14. Kucharska-Newton AM, Heiss G, Ni H, Stearns SC, Puccinelli-Ortega N, Wruck LM, et al. Identification of heart failure events in medicare claims: the Atherosclerosis Risk in Communities (ARIC) study. J Card Fail 2016;22:48–55.
15. Saczynski JS, Andrade SE, Harrold LR, Tjia J, Cutrona SL, Dodd KS, et al. A systematic review of validated methods for identifying heart failure using administrative data. Pharmacoepidemiol Drug Saf 2012;21(Suppl 1):129–40.
16. Colditz GA, Martin P, Stampfer MJ, Willett WC, Sampson L, Rosner B, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am J Epidemiol 1986;123:894–900.
17. Loehr LR, Agarwal SK, Baggett C, Wruck LM, Chang PP, Solomon SD, et al. Classification of acute decompensated heart failure: an automated algorithm compared with a physician reviewer panel: the Atherosclerosis Risk in Communities study. Circ Heart Fail 2013;6:719–26.
18. Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. J Clin Epidemiol 2000;53:499–503.
19. Chen G, Faris P, Hemmelgarn B, Walker RL, Quan H. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. BMC Med Res Methodol 2009;9:5.
20. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257–68.
21. Kitzman DW, Gardin JM, Gottdiener JS, Arnold A, Boineau R, Aurigemma G, et al, Cardiovascular Health Study Research Group. Importance of heart failure with preserved systolic function in patients ≥65 years of age. Am J Cardiol 2001;87:413–9.
22. Gillum RF. Epidemiology of heart failure in the United States. Am Heart J 1993;126:1042–7.
23. Gure TR, Kabeto MU, Blaum CS, Langa KM. Degree of disability and patterns of caregiving among older Americans with congestive heart failure. J Gen Intern Med 2008;23:70–6.
24. Ezekowitz JA, Kaul P, Bakal JA, Quan H, McAlister FA. Trends in heart failure care: has the incident diagnosis of heart failure shifted from the hospital to the emergency department and outpatient clinics? Eur J Heart Fail 2011;13:142–7.
25. Yeung DF, Boom NK, Guo H, Lee DS, Schultz SE, Tu JV. Trends in the incidence and outcomes of heart failure in Ontario, Canada: 1997 to 2007. CMAJ 2012;184:E765–73.