Original Investigation | Cardiology

A Natural Language Processing–Based Approach for Identifying Hospitalizations for Worsening Heart Failure Within an Integrated Health Care Delivery System

Andrew P. Ambrosy, MD; Rishi V. Parikh, MPH; Sue Hee Sung, MPH; Anand Narayanan, MD; Rajeev Masson, MD; Phuong-Quang Lam, MD; Kevin Kheder, MD; Alan Iwahashi, MD; Alexander B. Hardwick, MD; Jesse K. Fitzpatrick, MD; Harshith R. Avula, MD, MPH; Van N. Selby, MD; Xian Shen, PhD; Navneet Sanghera, MPharm; Joaquim Cristino, MSc; Alan S. Go, MD

Abstract

IMPORTANCE The current understanding of epidemiological mechanisms and temporal trends in hospitalizations for worsening heart failure (WHF) is based on claims and national reporting databases. However, these data sources are inherently limited by the accuracy and completeness of diagnostic coding and/or voluntary reporting.

OBJECTIVE To assess the overall burden of and temporal trends in the rate of hospitalizations for WHF.

DESIGN, SETTING, AND PARTICIPANTS This cohort study, performed from January 1, 2010, to December 31, 2019, used electronic health record (EHR) data from a large integrated health care delivery system.

EXPOSURES Calendar year trends.

MAIN OUTCOMES AND MEASURES Hospitalizations for WHF (ie, excluding observation stays) were defined as 1 symptom or more, 2 objective findings or more including 1 sign or more, and 2 doses or more of intravenous loop diuretics and/or new hemodialysis or continuous kidney replacement therapy. Symptoms and signs were identified using natural language processing (NLP) algorithms applied to EHR data.

RESULTS The study population was composed of 118 002 eligible patients experiencing 287 992 unique hospitalizations (mean [SD] age, 75.6 [13.1] years; 147 203 [51.1%] male; 1655 [0.6%] American Indian or Alaska Native, 28 451 [9.9%] Asian or Pacific Islander, 34 903 [12.1%] Black, 23 452 [8.1%] multiracial, 175 840 [61.1%] White, and 23 691 [8.2%] unknown), including 65 357 with a principal discharge diagnosis and 222 635 with a secondary discharge diagnosis of HF. The study population included 59 868 patients (20.8%) with HF with a reduced ejection fraction (HFrEF) (<40%), 33 361 (11.6%) with HF with a midrange EF (HFmrEF) (40%-49%), 142 347 (49.4%) with HF with a preserved EF (HFpEF) (≥50%), and 52 416 (18.2%) with unknown EF. A total of 58 042 admissions (88.8%) with a primary discharge diagnosis of HF and 62 764 admissions (28.2%) with a secondary discharge diagnosis of HF met the prespecified diagnostic criteria for WHF. Overall, hospitalizations for WHF identified on NLP-based algorithms increased from 5.2 to 7.6 per 100 hospitalizations per year during the study period. Subgroup analyses found an increase in hospitalizations for WHF based on NLP from 1.5 to 1.9 per 100 hospitalizations for HFrEF, from 0.6 to 1.0 per 100 hospitalizations for HFmrEF, and from 2.6 to 3.9 per 100 hospitalizations for HFpEF.

CONCLUSIONS AND RELEVANCE The findings of this cohort study suggest that the burden of hospitalizations for WHF may be more than double that previously estimated using only principal

(continued)

Key Points

Question What are the overall burden and temporal trends in the rate of hospitalizations for worsening heart failure (WHF)?

Findings This cohort study of 118 002 patients found that applying rigorous and prespecified diagnostic criteria to electronic health record data was associated with a more than 2-fold increase in the number of hospitalizations for WHF identified compared with estimates using a principal discharge diagnosis alone. There has been a gradual increase in the rate of hospitalizations for WHF over time, with a more noticeable increase observed among patients with heart failure with a preserved ejection fraction.

Meaning These findings suggest that population temporal trends based on a principal hospital discharge diagnosis of heart failure may underreport the increasing burden of hospitalizations for WHF, particularly among those with heart failure with a preserved ejection fraction, compared with a comprehensive approach using structured and unstructured electronic health record data.

+ **Invited Commentary**

+ **Supplemental content**

Author affiliations and article information are listed at the end of this article.

Abstract (continued)

discharge diagnosis. There has been a gradual increase in the rate of hospitalizations for WHF with a more noticeable increase observed for HFpEF.

Introduction

Heart failure (HF) is a global public health problem that affects an estimated 64 million people worldwide.[1-3] In the US alone, more than 6 million patients have been diagnosed with HF and in excess of 1 million annual hospital admissions have had HF as a primary discharge diagnosis.[4] Despite the increasing burden of hospitalizations for worsening HF (WHF), our understanding of epidemiological factors, patient clinical characteristics, care management, and outcomes in patients admitted for WHF is based almost entirely on observational studies relying on administrative or claims data and/or national reporting databases or quality improvement registries.[4] However, these data sources are intrinsically limited by the accuracy and completeness of diagnostic coding and/or voluntary reporting. Furthermore, signs and symptoms of HF inherently lack sensitivity and specificity, and overlap occurs with other common cardiovascular and noncardiovascular conditions, raising the possibility of misclassification.[5]

In contrast, randomized clinical trials that tested novel drugs and/or devices have applied a standardized multidimensional definition (ie, requiring symptoms, signs, and a change in HF-related therapy) to enroll a well-defined target population (ie, HF) as well as to validate outcomes of interest (ie, WHF).[6,7] However, this approach has traditionally been time-consuming and resource intensive, requiring the assessment of patient eligibility at the point of care and/or manual medical record review and adjudication of potential end points. The widespread availability of electronic health records (EHRs) and advances in data science now make it possible to automate this rigorous approach for observational research on a health care system–wide scale that leverages the full range of structured and unstructured data.

Thus, the main objective of this analysis of the UTILIZE-WHF (Understanding Inpatient and Outpatient Care and Outcomes Within a Learning and Integrated Health Care Delivery Organization Due to Worsening Heart Failure) study was to use rule-based natural language processing (NLP) algorithms applied to state-of-the-art EHR data to describe epidemiological mechanisms and temporal trends in the rate of hospitalizations for WHF (defined as changes in symptoms, signs, and/or objective criteria and HF-related therapy) overall and by the degree of systolic dysfunction (ie, reduced vs midrange vs preserved).

Methods

Setting and Source Population

Kaiser Permanente Northern California (KPNC) is a large integrated health care delivery system with 21 hospitals and more than 260 freestanding clinics where 4.5 million members receive comprehensive care (ie, inpatient, emergency department, and ambulatory encounters). The KPNC membership is highly representative of the local and statewide population with respect to age, sex, race, ethnicity, and socioeconomic status.[8-10] This cohort study was approved by the KPNC Institutional Review Board, and a waiver of informed consent was obtained because this was a retrospective, data-only study. All data were deidentified. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Study Overview and Cohort Assembly

The study sample included all members 18 years or older with a diagnosis of HF hospitalized at KPNC between January 1, 2010, and December 31, 2019. For the purposes of this study, qualifying clinical events included all hospitalizations that lasted more than 24 hours, excluding emergency department visits and observation stays. A diagnosis of HF is based on having been hospitalized with a primary discharge diagnosis of HF and/or having 3 ambulatory visits or more coded for HF based on *International Classification of Diseases, Ninth Edition (ICD-9)* codes (398.91, 402.x1, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, and 428.x) and *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)* codes (I09.81, I11.0, I11.9, I13.0, I13.1, I13.10, I13.11, I13.2, I50, I50.1, I50.2, I50.20, I50.21, I50.22, I50.23, I50.3, I50.30, I50.31, I50.32, I50.33, I50.4, I50.40, I50.41, I50.42, I50.43, I50.9, I97.13). These codes have been validated in multiple health care delivery systems and have a positive predictive value of 95% or higher.[11,12] Patients with end-stage kidney disease (defined as receipt of long-term dialysis or kidney transplant) before hospitalization were excluded.

Data Sources and Covariates

The KPNC Epic-based EHR system was the primary source for hospitalization data, patient progress notes, and cardiac imaging reports (ie, echocardiography and chest radiography). In addition, the KPNC Virtual Data Warehouse was used to ascertain concurrent comorbidities, outpatient medications, and outpatient laboratory values, as previously described and validated.[13,14] Demographic information including age, self-reported gender, self-reported race, and self-reported ethnicity was obtained from the EHR. Comorbid conditions were ascertained within 5 years of the hospital admission date. Laboratory results (except for B-type natriuretic peptide [BNP]) were defined as the most recent outpatient, nonemergency value within 7 to 365 days before the admission date. Baseline medication use is based on outpatient dispensed prescriptions within 120 days before the admission date. Data on left ventricular ejection fraction (LVEF), if available, were ascertained using the most recent value within 2 years of admission from structured results of echocardiography, radionuclide scintigraphy, other nuclear imaging modalities, and left ventriculography or extracted from echocardiogram reports using rule-based NLP algorithms. We categorized patients as having HF with reduced EF (HFrEF; defined as an EF <40%), HF with midrange EF (HFmrEF; defined as an EF 40%-49%), or HF with preserved EF (HFpEF; defined as an EF ≥50%) using criteria congruent with current US and European guidelines for the management of HF.[15,16] Data on inpatient medications, BNP, vital signs, and procedures were obtained from EHR databases. We characterized the principal discharge diagnostic category of hospitalizations with a secondary code for HF using standardized diagnosis-related group codes mapped to major diagnostic categories.

Definition of WHF

Hospitalizations for WHF, including in patients with preexisting and new-onset HF, were identified using EHR data and defined as including 1 symptom or more (dyspnea, orthopnea, paroxysmal nocturnal dyspnea, fatigue, weight gain, and/or tachypnea), 2 objective findings or more (tachycardia [heart rate >100 beats/min], elevated BNP [≥100 pg/mL; to convert to nanograms per liter, multiply by 1], and/or chest radiographic findings [pulmonary edema, pleural effusion, and/or cardiomegaly]) including 1 sign or more (lower extremity edema, pulmonary rales or wheezing, jugular venous distension, third heart sound [S_3 gallop], hepatomegaly, and/or abdominal swelling), and new administration of intravenous (IV) loop diuretics (≥2 doses) and/or new hemodialysis or continuous kidney replacement therapy (eTable 1 in the Supplement). These diagnostic criteria are based on a standardized definition of inpatient and outpatient WHF previously developed and validated by a consensus panel of trialists with expertise in clinical end point classification in collaboration with the US Food and Drug Administration.[6,7] This multidimensional definition is specific for WHF and was

accurate and reproducible in multiple pivotal trials of investigational drugs and devices for which regulatory approval was being sought.[17-21]

Applying NLP Algorithms

Natural language processing was used to parse relevant unstructured notes in the EHR (practitioner notes, discharge summaries, and imaging reports) taken within 72 hours of admission for all hospitalizations with a primary or secondary discharge diagnosis of HF. First, all notes were separated using a set of regular expressions (ie, regex) to detect line breaks and prespecified section headings. Only the Chief Complaint, History of Present Illness, Review of Systems, Physical Exam, Assessment and Plan, Hospital Course, and Subjective sections were searched in queries; Prior Medical History, Problem List, Family History, and others were excluded. Second, notes were indexed within I2E software, version 6.2.0 (Linguamatics), an ontology-based interactive information extraction system.[22,23] Indexing notes using I2E labels organizes text using syntactic (ie, nouns, verbs, sentences, and so on) and semantic (ie, diseases and relationships) classes for faster querying. We developed rule-based queries for each diagnostic criterion (except heart rate, BNP, and receipt of therapies, which were assessed using structured data) using custom ontologies, which included clinical negations and accounted for common spelling errors and word variations (see eTable 2 in the Supplement for basic ontologies). All queries are available on request and are viewable in YAML format (yaml.org) using any text editor.

The NLP algorithms for WHF were derived and validated against a criterion standard that consisted of manual record review and validation by 2 physicians (among A.N., R.M., P.Q.L., K.K., A.I., A.B.H., and J.K.F.), with final adjudication by a board-certified cardiologist (A.P.A.) when discrepancies occurred. Medical record review was recorded through an electronic survey tool with reviewers noting the presence or absence of each diagnostic criterion as well as providing an overall assessment of WHF based on our operational definition. We initially identified a random sample of 50 hospitalizations with a primary discharge diagnosis of HF to derive the queries. We then identified a random validation set of 200 hospitalizations equally distributed between hospitalizations with a primary vs secondary discharge diagnosis of HF to validate the NLP queries, and we observed a sensitivity of 98%, a specificity of 95%, a positive predictive value of 95%, a negative predictive value of 98%, and an accuracy of 97% for overall WHF (performance of each specific criterion is given in eTable 3 in the Supplement). In the validation set, cases for which NLP results differed from reviewer consensus were re-reviewed because reviewers sometimes missed positive mentions of criteria in prolonged hospitalizations with extensive written documentation.

Statistical Analysis

Descriptive characteristics are presented as means (SDs), medians (IQRs), and numbers (percentages). We calculated rates of HF and WHF hospitalizations per 100 KPNC hospitalizations per year (with normal 95% CIs), overall and stratified by LVEF category. We calculated the risks of readmission (with 95% CIs) because of all-cause hospitalization, hospitalization with a principal diagnosis code of HF, and hospitalization for WHF within 30 days of discharge, stratified by the HF diagnosis position and WHF classification of the initial hospitalization. We also conducted a sensitivity analysis examining hospitalization rates for WHF using alternative LVEF cutoffs of 40% to 60% and greater than 60%. All analyses were performed using SAS software, version 9.4 (SAS Institute Inc) at KPNC's Division of Research.

Results

Cohort Assembly and Clinical Characteristics

Between January 1, 2010, and December 31, 2019, we identified 118 002 eligible patients experiencing 287 992 unique hospitalizations (mean [SD] age, 75.6 [13.1] years; 147 203 [51.1%] male; 1655 [0.6%] American Indian or Alaska Native, 28 451 [9.9%] Asian or Pacific Islander, 34 903 [12.1%]

Black, 23 452 [8.1%] multiracial, 175 840 [61.1%] White, and 23 691 [8.2%] of unknown race or ethnicity), including 65 357 with a principal discharge diagnosis and 222 635 with a secondary discharge diagnosis of HF (**Table**). The study population included 59 868 patients (20.8%) with HFrEF, 33 361 (11.6%) with HFmrEF, 142 347 (49.4%) with HFpEF, and 52 416 (18.2%) with unknown EF. The burden of cardiac and noncardiac comorbidities was high, and 222 163 patients (77.1%) were receiving a diuretic as an outpatient.

In general, patients with a primary discharge diagnosis of HF were more likely to have a reduced LVEF, exhibit signs and symptoms of volume overload, have objective findings (eg, elevated BNP and evidence of pulmonary congestion on chest radiography) suggestive of WHF, and receive an IV diuretic compared with patients with a secondary discharge diagnosis of HF (Table). In contrast, patients with a primary vs secondary discharge diagnosis of HF were otherwise similar in terms of demographic characteristics, medical history, vital signs, laboratory values, and outpatient medications (Table).

Identification of Hospitalizations for WHF

After applying NLP-based algorithms for WHF to EHR data, 58 042 admissions (88.8%) with a primary discharge diagnosis of HF and 62 764 admissions (28.2%) with a secondary discharge diagnosis of HF were identified as meeting the prespecified diagnostic criteria for WHF (**Figure 1**). These findings were consistent across LVEF categories including HFrEF, HFmrEF, HFpEF, and unknown LVEF (**Figure 2**).

Patients hospitalized for WHF were more likely to have a reduced LVEF and less likely to have an unknown LVEF compared with patients hospitalized for reasons other than WHF (Table). In addition, patients hospitalized for WHF vs without WHF were more likely to manifest signs and symptoms of hypervolemia and have objective evidence of HF and were significantly more likely to have received 2 doses or more of an IV diuretic (119 833 [99.2%] vs 6974 [4.2%]). In contrast, patients hospitalized for WHF vs without WHF were otherwise similar in terms of demographic characteristics, medical history, vital signs, laboratory values, and outpatient medications.

The baseline clinical characteristics of patients stratified by principal discharge diagnosis and WHF status are given in eTable 4 in the Supplement. Few systematic differences were found between patients hospitalized for WHF who had a primary vs secondary discharge diagnosis of HF, with the notable exception that patients with a secondary discharge diagnosis of HF were more likely to have preserved systolic function (ie, defined as LVEF ≥50%).

The breakdown of the major categories for the principal discharge diagnosis for admissions with HF as a secondary discharge diagnosis code that met the definition for WHF is shown in eFigure 1 in the Supplement. Among these hospitalizations, the principal discharge diagnosis is most commonly classified as circulatory, respiratory, or infectious, with other major diagnostic categories each constituting less than 5.0% of cases. A complete list of the 300 most frequently listed principal discharge *ICD-9* or *ICD-10* diagnosis codes for hospitalizations with a secondary discharge diagnosis of HF that met the definition for WHF is given in eTable 5 in the Supplement.

Temporal Trends in Hospitalizations for WHF

Overall, a gradual increase was observed between January 1, 2010, and December 31, 2019, in total hospitalizations for WHF based on NLP from 5.2 to 7.6 per 100 hospitalizations per year with balanced increases in hospitalizations for WHF among primary vs secondary discharge diagnoses of HF (**Figure 3**). A subgroup analysis stratified by LVEF found a gradual increase in the rate of hospitalizations for WHF from 1.5 to 1.9 per 100 hospitalizations per year for HFrEF and from 0.6 to 1.0 per 100 hospitalizations per year for HFmrEF. In contrast, a more substantial increase occurred in total hospitalizations for WHF from 2.6 to 3.9 per 100 hospitalizations per year for HFpEF (**Figure 4**). Point estimates with 95% CIs for all rates are given in eTable 6 in the Supplement.

As a sensitivity analysis, we also examined temporal trends in total hospitalizations for WHF based on NLP among patients with HF with an LVEF of 40% to 60% and an LVEF of 60% or higher

Table. Baseline Characteristics of the Cohort Based on Principal or Secondary Discharge Diagnosis of Heart Failure and Confirmed Worsening Heart Failure Status by NLP Algorithm[a]

Characteristic	Overall hospitalizations (N = 287 992)	Discharge diagnosis		Worsening heart failure by NLP	
		Principal (n = 65 357)	Secondary (n = 222 635)	Yes (n = 120 806)	No (n = 167 186)
Age, mean (SD), y	75.6 (13.1)	75.0 (13.8)	75.8 (12.9)	75.1 (13.4)	76.0 (12.8)
Self-reported sex					
Male	147 203 (51.1)	34 408 (52.6)	112 795 (50.7)	62 362 (51.6)	84 841 (50.7)
Female	140 765 (48.9)	30 946 (47.3)	109 819 (49.3)	58 434 (48.4)	82 331 (49.2)
Other	4 (0.0)	1 (0.0)	3 (0.0)	2 (0.0)	2 (0.0)
Unknown	20 (0.0)	2 (0.0)	18 (0.0)	8 (0.0)	12 (0.0)
Self-reported race					
American Indian or Alaska Native	1655 (0.6)	331 (0.5)	1324 (0.6)	675 (0.6)	980 (0.6)
Asian or Pacific Islander	28 451 (9.9)	7059 (10.8)	21 392 (9.6)	13 103 (10.8)	15 348 (9.2)
Black	34 903 (12.1)	10 087 (15.4)	24 816 (11.1)	15 473 (12.8)	19 430 (11.6)
Multiracial	23 452 (8.1)	5204 (8.0)	18 248 (8.2)	9538 (7.9)	13 914 (8.3)
White	175 840 (61.1)	36 854 (56.4)	138 986 (62.4)	71 451 (59.1)	104 389 (62.4)
Unknown	23 691 (8.2)	5822 (8.9)	17 869 (8.0)	10 566 (8.7)	13 125 (7.9)
Self-reported Hispanic ethnicity	14 945 (5.2)	3519 (5.4)	11 426 (5.1)	7047 (5.8)	7898 (4.7)
Left ventricular ejection fraction category					
Reduced (<40%)	59 868 (20.8)	21 303 (32.6)	38 565 (17.3)	32 172 (26.6)	27 696 (16.6)
Midrange (40%-49%)	33 361 (11.6)	8235 (12.6)	25 126 (11.3)	14 990 (12.4)	18 371 (11.0)
Preserved (≥50%)	142 347 (49.4)	30 730 (47.0)	111 617 (50.1)	62 728 (51.9)	79 619 (47.6)
Unknown	52 416 (18.2)	5089 (7.8)	47 327 (21.3)	10 916 (9.0)	41 500 (24.8)
Left ventricular ejection fraction					
Mean (SD), %	49.8 (15.3)	46.2 (16.9)	51.1 (14.6)	48.5 (16.1)	51.0 (14.6)
Missing	126 878 (44.1)	22 657 (34.7)	104 221 (46.8)	42 585 (35.3)	84 293 (50.4)
Symptoms					
Shortness of breath	225 143 (78.2)	63 080 (96.5)	162 063 (72.8)	117 186 (97.0)	107 957 (64.6)
Orthopnea	60 141 (20.9)	30 287 (46.3)	29 854 (13.4)	44 773 (37.1)	15 368 (9.2)
Paroxysmal nocturnal dyspnea	28 424 (9.9)	14 696 (22.5)	13 728 (6.2)	21 056 (17.4)	7368 (4.4)
Tachypnea	81 100 (28.2)	24 681 (37.8)	56 419 (25.3)	50 476 (41.8)	30 624 (18.3)
Weight gain	25 531 (8.9)	15 206 (23.3)	10 325 (4.6)	20 659 (17.1)	4872 (2.9)
Fatigue	62 493 (21.7)	15 745 (24.1)	46 748 (21.0)	30 335 (25.1)	32 158 (19.2)
Signs					
Hypoxia (SpO$_2$ <94%)	137 619 (47.8)	32 461 (49.7)	105 158 (47.2)	70 676 (58.5)	66 943 (40.0)
Rales	170 968 (59.4)	55 350 (84.7)	115 618 (51.9)	102 073 (84.5)	68 895 (41.2)
S3 gallop	3933 (1.4)	2032 (3.1)	1901 (0.9)	2897 (2.4)	1036 (0.6)
Lower-extremity edema	189 303 (65.7)	57 037 (87.3)	132 266 (59.4)	103 065 (85.3)	86 238 (51.6)
Jugular vein distention	52 430 (18.2)	25 118 (38.4)	27 312 (12.3)	39 138 (32.4)	13 292 (8.0)
Hepatomegaly	2407 (0.8)	733 (1.1)	1674 (0.8)	1315 (1.1)	1092 (0.7)
Abdominal swelling	39 727 (13.8)	11 334 (17.3)	28 393 (12.8)	20 950 (17.3)	18 777 (11.2)
Objective findings					
Tachycardia (>100 beats/min)	145 510 (50.5)	30 922 (47.3)	114 588 (51.5)	66 870 (55.4)	78 640 (47.0)
B-type natriuretic peptide ≥100 pg/mL	171 540 (59.6)	60 317 (92.3)	111 223 (50.0)	105 242 (87.1)	66 298 (39.7)
On chest radiography					
Pulmonary edema	147 348 (51.2)	43 235 (66.2)	104 113 (46.8)	83 240 (68.9)	64 108 (38.3)
Cardiomegaly	133 973 (46.5)	42 835 (65.5)	91 138 (40.9)	76 578 (63.4)	57 395 (34.3)
Pleural effusion	194 771 (67.6)	52 740 (80.7)	142 031 (63.8)	99 017 (82.0)	95 754 (57.3)

(continued)

Table. Baseline Characteristics of the Cohort Based on Principal or Secondary Discharge Diagnosis of Heart Failure and Confirmed Worsening Heart Failure Status by NLP Algorithm[a] (continued)

Characteristic	Overall hospitalizations (N = 287 992)	Discharge diagnosis		Worsening heart failure by NLP	
		Principal (n = 65 357)	Secondary (n = 222 635)	Yes (n = 120 806)	No (n = 167 186)
In-hospital therapies					
≥2 Administrations of IV diuretic	126 807 (44.0)	59 081 (90.4)	67 726 (30.4)	119 833 (99.2)	6974 (4.2)
Dialysis or CKRT	2339 (0.8)	399 (0.6)	1940 (0.9)	2025 (1.7)	314 (0.2)
≥1 Symptom	244 170 (84.8)	63 896 (97.8)	180 274 (81.0)	120 806 (100.0)	123 364 (73.8)
≥1 Sign	250 062 (86.8)	64 054 (98.0)	186 008 (83.5)	120 806 (100.0)	129 376 (77.4)
≥1 Objective	257 326 (89.4)	64 347 (98.5)	192 979 (86.7)	119 855 (99.2)	137 471 (82.2)
≥1 Therapy	128 040 (44.5)	59 191 (90.6)	68 849 (30.9)	120 806 (100.0)	7234 (4.3)
Smoker					
Current	21 882 (7.6)	5293 (8.1)	16 589 (7.5)	9919 (8.2)	11 963 (7.2)
Former	148 390 (51.5)	33 221 (50.8)	115 169 (51.7)	62 282 (51.6)	86 108 (51.5)
Never	117 720 (40.9)	26 843 (41.1)	90 877 (40.8)	48 605 (40.2)	69 115 (41.3)
Medical history					
Atrial fibrillation or flutter	143 163 (49.7)	34 663 (53.0)	108 500 (48.7)	61 625 (51.0)	81 538 (48.8)
Ventricular fibrillation or tachycardia	16 463 (5.7)	4160 (6.4)	12 303 (5.5)	6618 (5.5)	9845 (5.9)
Ischemic stroke or transient ischemic attack	32 655 (11.3)	6037 (9.2)	26 618 (12.0)	10 943 (9.1)	21 712 (13.0)
Acute myocardial infarction	37 043 (12.9)	7213 (11.0)	29 830 (13.4)	16 054 (13.3)	20 989 (12.6)
Prior documented heart failure	200 457 (69.6)	52 925 (81.0)	147 532 (66.3)	84 843 (70.2)	115 614 (69.2)
Mitral or aortic valvular disease	73 259 (25.4)	19 453 (29.8)	53 806 (24.2)	33 915 (28.1)	39 344 (23.5)
Venous thromboembolism	37 946 (13.2)	7956 (12.2)	29 990 (13.5)	15 280 (12.6)	22 666 (13.6)
Hospitalization for bleeding	32 374 (11.2)	5695 (8.7)	26 679 (12.0)	11 270 (9.3)	21 104 (12.6)
Diabetes	138 864 (48.2)	33 375 (51.1)	105 489 (47.4)	60 777 (50.3)	78 087 (46.7)
Hypertension	237 820 (82.6)	54 924 (84.0)	182 896 (82.2)	99 884 (82.7)	137 936 (82.5)
Dyslipidemia	245 190 (85.1)	56 414 (86.3)	188 776 (84.8)	102 497 (84.8)	142 693 (85.3)
Hyperthyroidism	17 786 (6.2)	3665 (5.6)	14 121 (6.3)	6727 (5.6)	11 059 (6.6)
Hypothyroidism	70 001 (24.3)	15 804 (24.2)	54 197 (24.3)	28 770 (23.8)	41 231 (24.7)
Chronic kidney disease	159 954 (55.5)	41 574 (63.6)	118 380 (53.2)	71 457 (59.2)	88 497 (52.9)
Chronic liver disease	23 059 (8.0)	5071 (7.8)	17 988 (8.1)	9413 (7.8)	13 646 (8.2)
Chronic lung disease	151 180 (52.5)	34 121 (52.2)	117 059 (52.6)	64 649 (53.5)	86 531 (51.8)
Depression	65 319 (22.7)	13 423 (20.5)	51 896 (23.3)	24 836 (20.6)	40 483 (24.2)
Dementia	32 292 (11.2)	5980 (9.1)	26 312 (11.8)	10 746 (8.9)	21 546 (12.9)
BMI					
Mean (SD)	28.9 (8.3)	29.4 (8.7)	28.7 (8.2)	29.6 (8.7)	28.4 (8.0)
Missing	24 399 (8.5)	5564 (8.5)	18 835 (8.5)	11 017 (9.1)	13 382 (8.0)
Systolic blood pressure					
Mean (SD), mm Hg	124.3 (22.0)	123.1 (22.9)	124.6 (21.8)	124.0 (22.3)	124.5 (21.8)
Missing	19 218 (6.7)	4620 (7.1)	14 598 (6.6)	9033 (7.5)	10 185 (6.1)
Hemoglobin					
Mean (SD), g/dL	11.7 (2.1)	11.6 (2.1)	11.8 (2.1)	11.6 (2.1)	11.8 (2.1)
Missing	41 801 (14.5)	9543 (14.6)	32 258 (14.5)	18 864 (15.6)	22 937 (13.7)
Hemoglobin A_{1c}					
Mean (SD), %	6.7 (1.5)	6.8 (1.5)	6.7 (1.5)	6.8 (1.5)	6.7 (1.5)
Missing	99 878 (34.7)	21 689 (33.2)	78 189 (35.1)	41 051 (34.0)	58 827 (35.2)

(continued)

Table. Baseline Characteristics of the Cohort Based on Principal or Secondary Discharge Diagnosis of Heart Failure and Confirmed Worsening Heart Failure Status by NLP Algorithm[a] (continued)

Characteristic	Overall hospitalizations (N = 287 992)	Discharge diagnosis		Worsening heart failure by NLP	
		Principal (n = 65 357)	Secondary (n = 222 635)	Yes (n = 120 806)	No (n = 167 186)
Serum creatinine					
Mean (SD), mg/dL	1.5 (1.0)	1.6 (1.1)	1.4 (0.9)	1.6 (1.1)	1.4 (0.9)
Missing	28 736 (10.0)	6509 (10.0)	22 227 (10.0)	13 153 (10.9)	15 583 (9.3)
Estimated glomerular filtration rate					
Mean (SD), mL/min/1.73 m^2	53.3 (24.7)	47.7 (23.9)	54.9 (24.7)	50.2 (24.6)	55.5 (24.5)
Missing	28 736 (10.0)	6509 (10.0)	22 227 (10.0)	13 153 (10.9)	15 583 (9.3)
B-type natriuretic peptide					
Median (IQR), pg/mL	488 (225-1033)	746 (370-1471)	428 (197-896)	646 (317-1302)	387 (178-821)
Missing	26 312 (9.1)	2987 (4.6)	23 325 (10.5)	5818 (4.8)	20 494 (12.3)
Medications					
Angiotensin-converting enzyme inhibitor	135 560 (47.1)	32 119 (49.1)	103 441 (46.5)	56 352 (46.6)	79 208 (47.4)
Angiotensin II receptor blocker	76 315 (26.5)	18 637 (28.5)	57 678 (25.9)	32 587 (27.0)	43 728 (26.2)
Angiotensin receptor-neprilysin inhibitor	1177 (0.4)	461 (0.7)	716 (0.3)	603 (0.5)	574 (0.3)
Aldosterone receptor antagonist	41 043 (14.3)	12 287 (18.8)	28 756 (12.9)	18 804 (15.6)	22 239 (13.3)
Diuretic	222 163 (77.1)	54 881 (84.0)	167 282 (75.1)	97 258 (80.5)	124 905 (74.7)
β-Blocker	217 251 (75.4)	51 635 (79.0)	165 616 (74.4)	91 409 (75.7)	125 842 (75.3)
Calcium channel blocker	116 195 (40.3)	27 982 (42.8)	88 213 (39.6)	51 458 (42.6)	64 737 (38.7)
Antiarrhythmic drug	41 785 (14.5)	10 627 (16.3)	31 158 (14.0)	18 040 (14.9)	23 745 (14.2)
Oral anticoagulant	114 486 (39.8)	27 704 (42.4)	86 782 (39.0)	48 670 (40.3)	65 816 (39.4)
Antiplatelet drug	53 159 (18.5)	11 763 (18.0)	41 396 (18.6)	21 334 (17.7)	31 825 (19.0)
Any antihypertensive drugs	252 501 (87.7)	58 120 (88.9)	194 381 (87.3)	105 437 (87.3)	147 064 (88.0)
Statins	208 341 (72.3)	48 905 (74.8)	159 436 (71.6)	87 787 (72.7)	120 554 (72.1)
Other lipid-lowering drugs	17 272 (6.0)	3815 (5.8)	13 457 (6.0)	6728 (5.6)	10 544 (6.3)
Nitrates	93 255 (32.4)	25 875 (39.6)	67 380 (30.3)	42 415 (35.1)	50 840 (30.4)
Vasodilators	112 775 (39.2)	30 696 (47.0)	82 079 (36.9)	51 012 (42.2)	61 763 (36.9)
Any diabetic therapy	106 562 (37.0)	25 929 (39.7)	80 633 (36.2)	47 204 (39.1)	59 358 (35.5)
Sodium-glucose cotransporter 2 inhibitors	220 (0.1)	66 (0.1)	154 (0.1)	98 (0.1)	122 (0.1)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CKRT, continuous kidney replacement therapy; IV, intravenous; NLP, natural language processing.

SI conversion factors: To convert hemoglobin to grams per liter, multiply by 10; hemoglobin A$_{1c}$ to proportion of total hemoglobin, multiply by 0.01; B-type natriuretic peptide to nanograms per liter, multiply by 1; and serum creatinine to micromoles per liter, multiply by 88.4.

[a] Data are presented as number (percentage) unless otherwise indicated.

(eFigure 2 in the Supplement). We observed the greatest increase among the subgroup of patients with HF with an LVEF of 40% to 60%, which suggests that the increasing burden of hospitalizations for WHF was among patients with HF with an LVEF of 50% to 60%.

All-Cause 30-Day Readmissions After Hospitalizations for WHF

The rate of all-cause 30-day readmissions was 20.1% overall and was only marginally higher for hospitalizations with a principal discharge diagnosis of HF and those that met the definition for WHF (eFigure 3 in the Supplement). In contrast, after an index hospitalization with a principal discharge diagnosis of HF and/or adjudicated by NLP as being for WHF, the risk of HF-specific 30-day readmissions was substantially higher based on discharge diagnostic coding and NLP algorithms for WHF. Notably, the risk of HF-related 30-day readmission after an index hospitalization with a

principal discharge diagnosis of HF and those that met the definition of WHF differed by less than 1% to 2%.

Discussion

To our knowledge, this cohort study is the first comprehensive analysis to use a systematic approach that leveraged state-of-the-art EHR data to better capture and more accurately characterize hospitalizations for WHF within an integrated and learning health care delivery system. The NLP-based algorithms, complemented by structured data elements, revealed that approximately 90% of hospitalizations with a primary discharge diagnosis of HF and approximately 30% of hospitalizations with a secondary discharge diagnosis of HF met criteria for WHF. Of importance, the inclusion of admissions with a primary and secondary discharge diagnosis of HF that satisfied systematic, prespecified diagnostic criteria resulted in a more than 2-fold increase in the perceived population burden of hospitalizations for WHF. Finally, there has been a gradual increase in the rate of hospitalizations for WHF, with a more prominent increase observed among patients with HFpEF.

It is well established that primary discharge diagnostic codes for HF are highly specific for the condition but have suboptimal sensitivity, particularly for HFpEF, when compared with a criterion standard of manual medical record review and validation based on clinical criteria.[24] We hypothesized that NLP-based algorithms, supplemented by structured data sources and using a multidimensional definition for WHF, would boost sensitivity, without sacrificing specificity, and improve overall diagnostic accuracy for identifying hospitalized episodes of WHF. This assertion is strongly supported by the excellent test performance characteristics seen in the derivation and validation cohorts. Furthermore, evidence of the validity of this approach includes the observations that (1) this analysis confirmed a hospitalization for WHF in approximately 90% of admissions with a primary discharge diagnosis of HF (ie, maintaining specificity) and (2) the clinical characteristics and 30-day readmission rates for patients hospitalized for WHF were virtually identical to those with a principal discharge diagnosis of HF (ie, enhancing sensitivity). Although the diagnostic criteria used in this study are based on a standardized definition for WHF, the decision was made a priori to purposefully exclude long-term dialysis and/or other IV vasoactive medications (ie, inotropes, vasopressors, and/or vasodilators) given that these interventions inherently lack specificity for WHF.

Figure 1. Mosaic Diagram of Hospitalizations That Met Prespecified Diagnostic Criteria for Worsening Heart Failure (HF) by HF Diagnosis Type
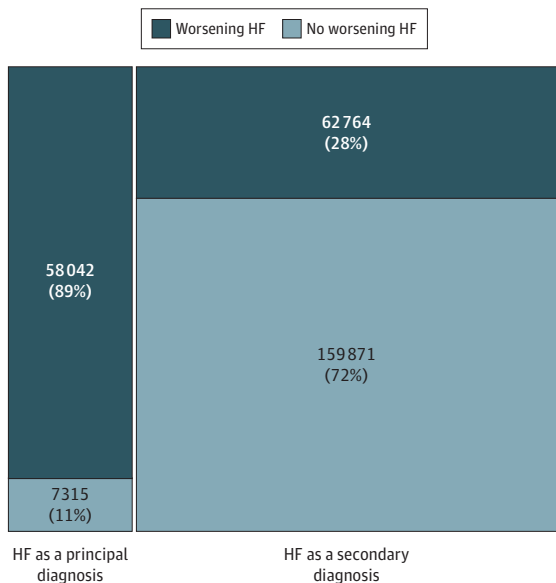
The total number of hospitalizations was 287 992.

In contrast, administration of multiple doses of IV diuretics within a short time frame is a sine qua non with an episode of WHF, and this highly specific finding was documented in fewer than 5% of patients hospitalized without WHF.

Several novel epidemiological observations stem from this analysis. First, this study found that the burden of hospitalizations for WHF is more than 2-fold higher compared with estimates using only a principal discharge diagnosis. Second, the rate of hospitalizations for WHF based on NLP has gradually increased over time, which is congruent with the increasing prevalence and incidence of HF.[25] Third, the increase in the rate of hospitalizations for WHF is predominantly associated with a diagnosis of HFpEF, particularly in patients with systolic function below normal (ie, LVEF of 50%-60%). Historically, HFpEF has not been well characterized, and therapeutic options have been limited for this heterogeneous population. More recently, it has been proposed that a prior hospitalization for WHF may denote a high-risk subset of patients who are more likely to derive a robust benefit from several emerging classes of medications regardless of LVEF.[26-28] Thus, the use of NLP algorithms and the full range of EHR data to identify hospitalizations for WHF, a well-established enrichment factor, is an incremental advancement over sole reliance on claims data and voluntary participation in disease-based registries. This innovative approach may serve as the basis

Figure 2. Mosaic Diagram of Hospitalizations That Met Prespecified Diagnostic Criteria for Worsening Heart Failure (HF) by HF Diagnosis Type Across Different Ejection Fraction Categories

for a prospective, EHR-based, practitioner-level intervention aimed at improving the implementation of guideline-directed medical therapy after an index hospitalization for WHF.

The use of validated, rule-based NLP algorithms with full access to structured and unstructured EHR data also has a range of potential research applications, beyond quality improvement and implementation science, including traditional and pragmatic clinical trials. The US Food and Drug

Figure 3. Temporal Trends in Hospitalizations for Heart Failure (HF) Based on Principal Discharge Diagnosis and Diagnostic Criteria for Worsening HF
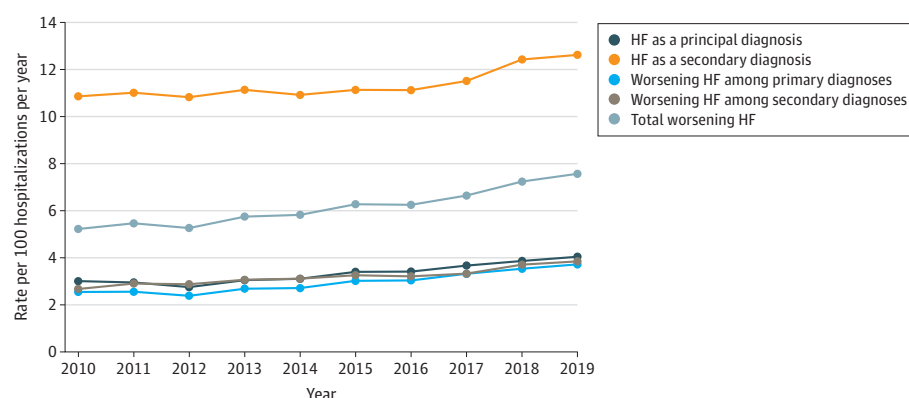
Figure 4. Temporal Trends in Hospitalizations for Heart Failure (HF) Based on Principal Discharge Diagnosis and Diagnostic Criteria for Worsening HF Across Different Ejection Fraction Categories
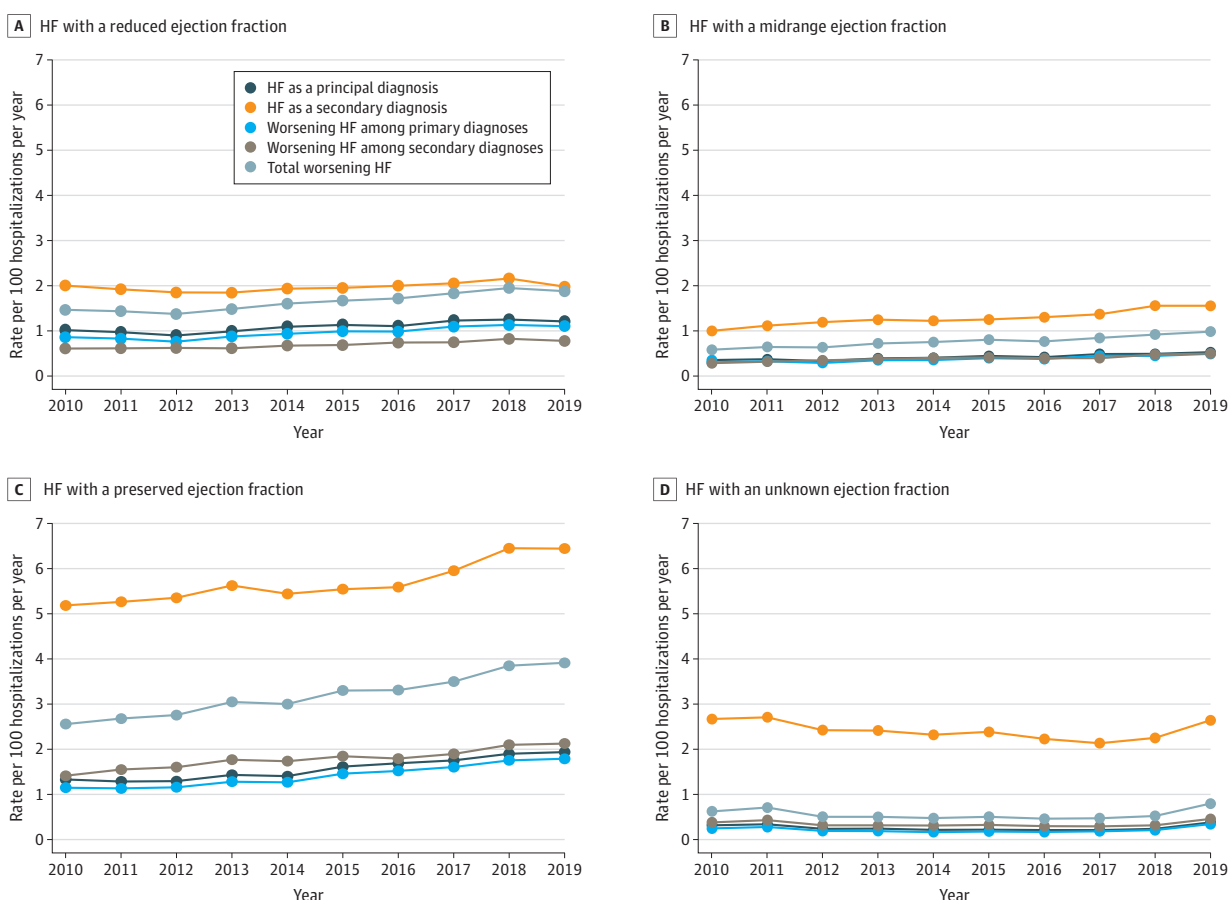
Administration typically requires randomized clinical trials of investigational drugs and devices for which regulatory approval is sought that involve an independent and blinded clinical end point classification to adjudicate all outcomes according to prespecified end point definitions. This adjudication requires surveillance for event triggers, obtaining source documentation, and manual verification by trained physician reviewers, making clinical end point classification adjudication a time-consuming and resource-intensive process. An automated approach that leverages the full range of EHR data has the competitive advantages of being efficient, cost-effective, and highly accurate and reproducible. Furthermore, a software-based solution is completely objective and eliminates the subjectivity of physician review, in which discrepancy is relatively common, possibly contributing to additional downstream time requirements and costs. Developing validated diagnostic criteria for common medical conditions (eg, acute myocardial infarction, stroke, and HF) based on EHR data has the potential to revolutionize the design and conduct of pragmatic clinical trials from subject identification to outcome ascertainment.[29]

Limitations

Our study has several limitations. Although this study incorporated a rigorous consensus definition for WHF based on expert opinion and historical precedent,[6,7] it was not technically feasible to incorporate several potential therapeutic interventions for WHF, including augmentation of oral diuretics, other IV vasoactive medications, temporary mechanical circulatory support, and/or long-term kidney replacement therapy, because these treatments were inconsistently documented and/or lacked diagnostic specificity for WHF. However, the provision of multiple doses of IV diuretics within a short time frame is the preferred treatment modality in more than 90% of patients hospitalized for WHF.[30] This study also excluded data beyond the first 72 hours of admission to distinguish between hospitalizations for WHF and prolonged hospitalizations complicated by WHF. Another possibility is that evolving trends in the use of diagnostic tests (ie, BNP) alone could create the appearance of temporal changes in the incidence of hospitalizations for WHF. However, the definition of WHF is by design multidimensional and requires symptoms, signs, and/or objective findings and a change in HF-related therapy to diminish the potential risk of ascertainment bias. Other potential concerns are that a subset of admissions may have originated at outside hospitals and EHR data on the initial care provided before transfer to a KPNC facility may not be available for analysis. However, at KPNC there is an exclusive relationship among the insurer, members, and practitioners, and prior studies[12,31,32] have found that 95% of events are captured in the KPNC EHR. Additionally, there is increasing interest in the field in disentangling WHF from location of care and moving away from using hospitalization as a surrogate for actual WHF. Thus, future studies should apply validated NLP-based algorithms to ambulatory encounters for HF to better understand the patient journey and full impact of this emerging clinical entity across the entire spectrum of care settings.

Conclusions

This study found that applying validated NLP-based algorithms to structured and unstructured EHR data was not only technically feasible but also highly accurate, leading to a more than 2-fold increase in the burden of hospitalizations for WHF compared with estimates based entirely on the principal discharge diagnosis. In addition, a gradual increase in the rate of hospitalizations for WHF occurred, with a more notable increase seen in patients with HFpEF. Future research should disassociate WHF from clinical settings to better capture the magnitude of WHF across the care continuum, from ambulatory encounters to hospitalizations.

Corresponding Author: Andrew P. Ambrosy, MD, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612 (andrew.p.ambrosy@kp.org).

Author Affiliations: Department of Cardiology, Kaiser Permanente San Francisco Medical Center, San Francisco, California (Ambrosy, Kheder, Hardwick, Fitzpatrick, Selby); Division of Research, Kaiser Permanente Northern California, Oakland (Ambrosy, Parikh, Sung, Go); Department of Health Systems Science, Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, California (Ambrosy, Go); Department of Medicine, Kaiser Permanente San Francisco Medical Center, San Francisco, California (Narayanan, Masson, Lam, Iwahashi); Department of Cardiology, Kaiser Permanente Walnut Creek Medical Center, Walnut Creek, California (Avula, Cristino); Novartis Pharmaceuticals Corporation, East Hanover, New Jersey (Shen, Sanghera); Departments of Epidemiology, Biostatistics and Medicine, University of California, San Francisco (Go); Department of Medicine, Stanford University, Palo Alto, California (Go).

Author Contributions: Drs Ambrosy and Go had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Ambrosy, Parikh, Masson, Shen, Sanghera, Cristino, Go.

Acquisition, analysis, or interpretation of data: Ambrosy, Parikh, Sung, Narayanan, Masson, Lam, Kheder, Iwahashi, Hardwick, Fitzpatrick, Avula, Selby, Shen, Cristino, Go.

Drafting of the manuscript: Ambrosy, Masson, Kheder, Sanghera.

Critical revision of the manuscript for important intellectual content: Ambrosy, Parikh, Sung, Narayanan, Masson, Lam, Iwahashi, Hardwick, Fitzpatrick, Avula, Selby, Shen, Cristino, Go.

Statistical analysis: Parikh, Lam, Kheder, Cristino.

Obtained funding: Ambrosy, Sung, Shen, Sanghera.

Administrative, technical, or material support: Ambrosy, Sung, Narayanan, Hardwick, Avula, Sanghera, Cristino, Go.

Supervision: Ambrosy, Avula, Sanghera, Cristino, Go.

Conflict of Interest Disclosures: Dr Ambrosy reported receiving grants from Novartis AG; Abbott; and the National Heart, Lung, and Blood Institute during the conduct of the study and grants from Amarin outside the submitted work. Dr Masson reported receiving funding from Novartis Pharmaceuticals Corporation during the conduct of the study. Dr Shen reported receiving funding from Novartis Pharmaceuticals Corporation outside the submitted work. Dr Cristino reported being an employee of Novartis Pharmaceuticals Corporation during the conduct of the study. Dr Go reported receiving grants from Novartis AG during the conduct of the study and grants from the National Institute of Diabetes, Digestive and Kidney Diseases and the National Heart, Lung and Blood Institute outside the submitted work. No other disclosures were reported.

Funding/Support: The UTILIZE-WHF study is funded by Novartis AG (East Hanover, New Jersey) and is an ongoing research project of the Division of Research at Kaiser Permanente Northern California (sponsor) in Oakland, California, and Novartis AG (funder) in East Hanover, New Jersey.

Role of the Funder/Sponsor: The funder approved the study in advance but had no formal role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. Braunwald E. The war against heart failure: the Lancet lecture. *Lancet*. 2015;385(9970):812-824. doi:10.1016/S0140-6736(14)61889-4

2. Ambrosy AP, Gheorghiade M, Chioncel O, Mentz RJ, Butler J. Global perspectives in hospitalized heart failure: regional and ethnic variation in patient characteristics, management, and outcomes. *Curr Heart Fail Rep*. 2014;11 (4):416-427. doi:10.1007/s11897-014-0221-9

3. Disease GBD, Injury I, Prevalence C; GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789-1858. doi:10.1016/S0140-6736(18)32279-7

4. Virani SS, Alonso A, Benjamin EJ, et al; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2020 update: a report from the American Heart Association. *Circulation*. 2020;141(9):e139-e596. doi:10.1161/CIR.0000000000000757

5. Gheorghiade M, Follath F, Ponikowski P, et al; European Society of Cardiology; European Society of Intensive Care Medicine. Assessing and grading congestion in acute heart failure: a scientific statement from the Acute Heart Failure Committee of the Heart Failure Association of the European Society of Cardiology and endorsed by the European Society of Intensive Care Medicine. *Eur J Heart Fail*. 2010;12(5):423-433. doi:10.1093/eurjhf/hfq045

6. Hicks KA, Mahaffey KW, Mehran R, et al; Standardized Data Collection for Cardiovascular Trials Initiative (SCTI). 2017 Cardiovascular and stroke endpoint definitions for clinical trials. *J Am Coll Cardiol*. 2018;71(9):1021-1034. doi:10.1016/j.jacc.2017.12.048

7. Hicks KA, Mahaffey KW, Mehran R, et al; Standardized Data Collection for Cardiovascular Trials Initiative (SCTI). 2017 Cardiovascular and stroke endpoint definitions for clinical trials. *Circulation*. 2018;137(9):961-972. doi:10.1161/CIRCULATIONAHA.117.033502

8. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health*. 1992;82(5):703-710. doi:10.2105/AJPH.82.5.703

9. Gordon NP. *Characteristics of Adult Health Plan Members in the Northern California Region Membership, as Estimated from the 2011 Member Health Survey.* Division of Research, Kaiser Permanente Medical Care Program; 2013. Accessed October 1, 2021. https://divisionofresearch.kaiserpermanente.org/projects/memberhealthsurvey/SiteCollectionDocuments/mhs11reg.pdf

10. Koebnick C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *Perm J*. 2012;16(3):37-41. doi:10.7812/TPP/12-031

11. Smith DH, Thorp ML, Gurwitz JH, et al. Chronic kidney disease and outcomes in heart failure with preserved versus reduced ejection fraction: the Cardiovascular Research Network PRESERVE Study. *Circ Cardiovasc Qual Outcomes*. 2013;6(3):333-342. doi:10.1161/CIRCOUTCOMES.113.000221

12. Go AS, Lee WY, Yang J, Lo JC, Gurwitz JH. Statin therapy and risks for death and hospitalization in chronic heart failure. *JAMA*. 2006;296(17):2105-2111. doi:10.1001/jama.296.17.2105

13. Go AS, Magid DJ, Wells B, et al. The Cardiovascular Research Network: a new paradigm for cardiovascular quality and outcomes research. *Circ Cardiovasc Qual Outcomes*. 2008;1(2):138-147. doi:10.1161/CIRCOUTCOMES.108.801654

14. Magid DJ, Gurwitz JH, Rumsfeld JS, Go AS. Creating a research data network for cardiovascular disease: the CVRN. *Expert Rev Cardiovasc Ther*. 2008;6(8):1043-1045. doi:10.1586/14779072.6.8.1043

15. Yancy CW, Jessup M, Bozkurt B, et al; American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2013;62(16):e147-e239. doi:10.1016/j.jacc.2013.05.019

16. Ponikowski P, Voors AA, Anker SD, et al; ESC Scientific Document Group. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2016;37(27):2129-2200. doi:10.1093/eurheartj/ehw128

17. Gheorghiade M, Konstam MA, Burnett JC Jr, et al; Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan (EVEREST) Investigators. Short-term clinical effects of tolvaptan, an oral vasopressin antagonist, in patients hospitalized for heart failure: the EVEREST Clinical Status Trials. *JAMA*. 2007;297(12):1332-1343. doi:10.1001/jama.297.12.1332

18. Konstam MA, Gheorghiade M, Burnett JC Jr, et al; Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan (EVEREST) Investigators. Effects of oral tolvaptan in patients hospitalized for worsening heart failure: the EVEREST Outcome Trial. *JAMA*. 2007;297(12):1319-1331. doi:10.1001/jama.297.12.1319

19. O'Connor CM, Starling RC, Hernandez AF, et al. Effect of nesiritide in patients with acute decompensated heart failure. *N Engl J Med*. 2011;365(1):32-43. doi:10.1056/NEJMoa1100171

20. Teerlink JR, Cotter G, Davison BA, et al; RELAXin in Acute Heart Failure (RELAX-AHF) Investigators. Serelaxin, recombinant human relaxin-2, for treatment of acute heart failure (RELAX-AHF): a randomised, placebo-controlled trial. *Lancet*. 2013;381(9860):29-39. doi:10.1016/S0140-6736(12)61855-8

21. Massie BM, O'Connor CM, Metra M, et al; PROTECT Investigators and Committees. Rolofylline, an adenosine A1-receptor antagonist, in acute heart failure. *N Engl J Med*. 2010;363(15):1419-1428. doi:10.1056/NEJMoa0912613

22. Milward D, Bjäreland M, Hayes W, et al. Ontology-based interactive information extraction from scientific abstracts. *Comp Funct Genomics*. 2005;6(1-2):67-71. doi:10.1002/cfg.456

23. Cormack J, Nath C, Milward D, Raja K, Jonnalagadda SR. Agile text mining for the 2014 i2b2/UTHealth Cardiac risk factors challenge. *J Biomed Inform*. 2015;58(suppl):S120-S127. doi:10.1016/j.jbi.2015.06.030

24. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS One*. 2014;9(8):e104519. doi:10.1371/journal.pone.0104519

25. Heidenreich PA, Albert NM, Allen LA, et al; American Heart Association Advocacy Coordinating Committee; Council on Arteriosclerosis, Thrombosis and Vascular Biology; Council on Cardiovascular Radiology and Intervention; Council on Clinical Cardiology; Council on Epidemiology and Prevention; Stroke Council. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail*. 2013;6(3):606-619. doi:10.1161/HHF.0b013e318291329a

26. Pitt B, Pfeffer MA, Assmann SF, et al; TOPCAT Investigators. Spironolactone for heart failure with preserved ejection fraction. *N Engl J Med*. 2014;370(15):1383-1392. doi:10.1056/NEJMoa1313731

27. Solomon SD, McMurray JJV, Anand IS, et al; PARAGON-HF Investigators and Committees. Angiotensin-neprilysin inhibition in heart failure with preserved ejection fraction. *N Engl J Med*. 2019;381(17):1609-1620. doi:10.1056/NEJMoa1908655

28. Bhatt DL, Szarek M, Steg PG, et al; SOLOIST-WHF Trial Investigators. Sotagliflozin in patients with diabetes and recent worsening heart failure. *N Engl J Med*. 2021;384(2):117-128. doi:10.1056/NEJMoa2030183

29. Zhao Y, Sivaswamy A, Lee MK, et al. A feasibility study for CODE-MI: high-sensitivity cardiac troponin—optimizing the diagnosis of acute myocardial infarction/injury in women. *Am Heart J*. 2021;234:60-70. doi:10.1016/j.ahj.2021.01.008

30. Adams KF Jr, Fonarow GC, Emerman CL, et al; ADHERE Scientific Advisory Committee and Investigators. Characteristics and outcomes of patients hospitalized for heart failure in the United States: rationale, design, and preliminary observations from the first 100,000 cases in the Acute Decompensated Heart Failure National Registry (ADHERE). *Am Heart J*. 2005;149(2):209-216. doi:10.1016/j.ahj.2004.08.005

31. Goldberg RJ, Gurwitz JH, Saczynski JS, et al; CVRN PRESERVE HF Investigators. Comparison of medication practices in patients with heart failure and preserved versus those with reduced ejection fraction (from the Cardiovascular Research Network [CVRN]). *Am J Cardiol*. 2013;111(9):1324-1329. doi:10.1016/j.amjcard.2013.01.276

32. Go AS, Yang J, Ackerson LM, et al. Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure: the Anemia in Chronic Heart Failure: Outcomes and Resource Utilization (ANCHOR) Study. *Circulation*. 2006;113(23):2713-2723. doi:10.1161/CIRCULATIONAHA.105.577577

SUPPLEMENT.

eTable 1. Specific Criteria Used to Assess for the Presence of Worsening Heart Failure

eTable 2. Operational Ontologies for Specific Criteria Included the Definition of Worsening Heart Failure Using Natural Language Processing (NLP)

eTable 3. Performance of Natural Language Processing (NLP) Queries for Specific Criteria Included the Definition of Worsening Heart Failure as Compared to Physician Manual Adjudication of Medical Records in the Validation Sample of 200 Hospitalizations

eTable 4. Baseline Characteristics of the Cohort Based on Principal or Secondary Discharge Diagnosis of Heart Failure and Confirmed Worsening Heart Failure (WHF) Status by Natural Language Processing (NLP) Algorithm (ie, Mutually Exclusive Categories)

eTable 5. List of Top 300 Principal Diagnoses for Hospitalizations for Worsening Heart Failure With a Secondary Discharge Diagnosis Code for Heart Failure

eTable 6. Point Estimates and Associated 95% Confidence Intervals for Temporal Trends in Hospitalizations for Worsening Heart Failure, Overall and Stratified by Left Ventricular Ejection Fraction Category (Corresponding to Manuscript Figures 3 And 4)

eFigure 1. Major Diagnostic Categories for the Principal Discharge Diagnosis for Admissions With Heart Failure as a Secondary Discharge Diagnosis Identified as Meeting the Definition for Worsening Heart Failure

eFigure 2. Temporal Trends in Hospitalizations for Heart Failure Based on Principal Discharge Diagnosis and Diagnostic Criteria for Worsening Heart Failure for (A) Heart Failure With an Ejection Fraction Between 40-60% and (B) Heart Failure With and Ejection Fraction ≥60%

eFigure 3. All-Cause and Cause-Specific Readmission Rates Based on Principal Discharge Diagnosis and Diagnostic Criteria for Worsening Heart Failure