HHS Public Access

Author manuscript

Int J Med Inform. Author manuscript; available in PMC 2019 December 01.

Published in final edited form as:

Int J Med Inform. 2018 December ; 120: 1–7. doi:10.1016/j.ijmedinf.2018.09.016.

Identifying Heart Failure using EMR-based algorithms

Geoffrey H. Tison, MD, MPH[*], **Alanna M. Chamberlain, PhD, MPH**[†], **Mark J. Pletcher, MD, MPH**[‡], **Shannon M. Dunlay, MD, MS**[†,§], **Susan A. Weston, MS**[†], **Jill M. Killian, BS**[†], **Jeffrey E. Olgin, MD**[*], and **Véronique L. Roger, MD, MPH**[†,§,‖]

[*]Division of Cardiology, University of California, San Francisco

[†]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA

[‡]Department of Epidemiology and Biostatistics, University of California, San Francisco, USA

[§]Department of Cardiovascular Diseases, Mayo Clinic, Rochester, MN, USA

[‖]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA.

Abstract

BACKGROUND—Heart failure (HF) is a major clinical and public health problem, the management of which will benefit from large-scale pragmatic research that leverages electronic medical records (EMR). Requisite to using EMRs for HF research is the development of reliable algorithms to identify HF patients. We aimed to develop and validate computable phenotype algorithms to identify patients with HF using standardized data elements defined by the Patient Centered Outcomes Research Network (PCORnet) Common Data Model (CDM).

METHODS—We built HF computable phenotypes utilizing the data domains of HF diagnosis codes, prescribed HF-related medications and N-terminal B-type natriuretic peptide (NT-proBNP). Algorithms were validated in a cohort (n=76,254) drawn from Olmsted County, MN between 2010–2012 a sample of whose records were manually reviewed to confirm HF according to Framingham criteria.

Correspondence: Geoffrey H. Tison, MD MPH; Universty of California, San Francisco; 555 Mission Bay Blvd South, Box 3120, San Francisco, CA 94158; Telephone 415-502-0992; Fax (415) 502-7949; geoff.tison@ucsf.edu.

Author Statement:

All authors have contributed significantly to constitute authorship on the current work. Tison, Chamberlain, Pletcher, Dunlay, Weston, Olgin and Roger were involved in the conception and design of the project. Tison, Chamberlain, Pletcher, Dunlay, Weston, Killian, Olgin and Roger were involved in the analysis and critical interpretation of the data. Tison, Roger, Chamberlain and Weston were involved in the drafting of the manuscript, and all authors contributed to the critical revision for intellectual content, give approval and agree to be held accountable for all aspects of the manuscript. G Tison is the corresponding author, to whom all editorial correspondence can be directed.

DISCLOSURES

Relationships with Industry:

Jeff Olgin—Research Grant, Zoll Medical Corporation

Conflict of Interest:

Relevant industry disclosures are as follows: Jeff Olgin—Research Grant, Zoll Medical Corporation.

Conflicts of Interest: None, for all authors.

RESULTS—The different algorithms we tested provided different tradeoffs between sensitivity and positive predictive value (PPV). The highest sensitivity (78.7%) algorithm utilized one HF diagnosis code and had the lowest PPV (68.5%). The addition of more algorithm components, such as additional HF diagnosis codes, HF medications or elevated NT-proBNP, improved the PPV while reducing sensitivity. When added to a diagnostic code, the addition of NT-proBNP (>450 pg/mL) had a similar impact compared to additional HF medication criteria, increasing PPV by ~3–4% and decreasing sensitivity by ~7–10%.

CONCLUSIONS—Algorithms derived from PCORnet CDM elements can be used to identify patients with HF without manual adjudication with reasonable sensitivity and PPV. Algorithm choice should be driven by the goal of the research.

Keywords

heart failure; outcomes assessment; electronic health records; cohort studies; learning health system

INTRODUCTION

The rapid adoption of electronic medical records (EMR) in the United States is prompting a reengineering of clinical research systems, where aggregated data from clinical care can contribute to large-scale research. The Patient Centered Outcomes Research Institute created a nation-wide infrastructure platform for trials and observational studies, known as the Patient Centered Outcomes Research Network (PCORnet)[1]. This “network of networks” of nearly 100 million people from all 50 states in the United States enables large-scale patient recruitment into clinical trials[1] and longitudinal follow-up using a set of data standards, known as the PCORnet Common Data Model (CDM). For this infrastructure to serve its purpose, validated disease-specific algorithms, known as computable phenotypes, are critical to accurately identify candidates for participation in research studies.

Studying the performance of EMR-based CDM data for this purpose is an essential prerequisite to the conduct of research that relies on the PCORnet CDM[2]. To examine this matter, we elected to study heart failure (HF), which affects 6.4 million US adults, is projected to increase in prevalence by 46% by 2030[3], and is the most common cause for hospital admissions in the Medicare population[4]. To identify HF patients using the EMR, billing codes are often used but vary widely in sensitivity, specificity and positive predictive value when compared to validated HF definitions[5–7]. Algorithms with more criteria or that are designed within specific institutions or databases[5,6,8], while informative, must be adapted for use in other institutions since EMR systems may differ and contain non-standardized data elements[9,10]. Relying on a CDM-based HF algorithm standardizes data elements and is attractive in being EMR-agnostic and deployable across networks like PCORnet, providing access to millions of patients across numerous institutions.

Our goal was to develop and validate computable phenotype algorithms to identify patients with prevalent HF using the PCORnet CDM, while leveraging an established community-based epidemiologic cohort of patients with validated HF.

METHODS

Study Setting and Design

Multiple algorithms were developed to identify heart failure using data elements from the PCORnet CDM. Algorithm validation was conducted amongst a population from Olmsted County, Minnesota (2010 population: 144, 248), which has similar age- and sex-specific mortality rates when compared to the entire United States [11]. The provider-linked medical records from each institution are indexed through the Rochester Epidemiology Project, resulting in the linkage of clinical and demographic information from nearly all sources of care for local residents[11]. Mayo Clinic is part of the PCORnet Learning Health Systems Clinical Data Research Network, which has been described elsewhere[12]. The HF computable phenotype algorithms were validated in a cohort of patients determined to have HF from manual medical record review. Participants granted Minnesota research authorization for use of their medical records to conduct research. This study was approved by the Mayo Clinic and Olmsted Medical Center Institutional Review Boards.

Case Identification and Validation

The validation cohort was drawn from a larger cohort built as part of ongoing HF community surveillance work[13,14]. In the community surveillance cohort, Olmsted County residents were identified who had a diagnosis code for HF defined by the ICD-9 code 428.XX between January 1, 1979 and December 31, 2014. During this period, only ICD-9 codes were used. A 50% random sample of patient records with a code 428 identified in 1979–2006 and 100% of patient records with a code 428 identified from 2007–2014 were selected for review. Trained nurse abstractors then manually reviewed sampled records to confirm HF according to Framingham criteria, which require the presence of either two major criteria to confirm HF, or one major and two minor criteria[15]. This approach has been applied previously, showing minimal missing data and excellent inter-observer agreement[13].

The validation cohort used in this study was drawn from the community surveillance cohort, and consisted of residents 30 years of age or older who had an ICD-9 code of any type (inpatient or outpatient) from January 1, 2010 through December 31, 2012. This ensured that residents were alive and accessing healthcare. The index date was defined as the date of the first ICD-9 code of any type during the study period. HF cases were defined as those patients who had validated HF by manual review of the medical record prior to or within 2 years after the index date; those without validated HF were defined as non-cases. Data on additional algorithm criteria, including International Classification of Diseases – 9th Revision, Clinical Modification (ICD-9) 428.XX codes, prescribed HF medications, and NT-proBNP occurring within two years after the index date were obtained electronically.

In addition, a random sample of 200 residents who did not have an ICD-9 code 428.XX during the study period were identified and underwent manual review of their medical records to ascertain whether any met Framingham criteria for HF. None of these ICD-9 code 428.XX negative residents had either a clinical diagnosis of HF or HF confirmed by

Framingham criteria. Consequently, all residents during the study period who did not have an ICD-9 code 428.XX are classified as not having validated HF.

PCORnet CDM Characteristics

PCORnet is comprised of 13 Clinical Data Research Networks, each of which is a collaboration of various institutions including academic health systems, community hospitals, individual providers, health centers and regional health information exchanges. The PCORnet CDM (https://pcornet.org/pcornet-common-data-model/) is based on the FDA Sentinel Initiative Common Data Model (www.sentinelsystem.org) and is continually updated. The broad data domain categories that are included within the PCORnet CDM include: demographics, healthcare encounters, diagnoses and procedure codes, vital signs, select lab results, prior or current health conditions or diseases, patient reported outcomes (when available), medication dispensing and prescribing, death and cause of death and whether the patient is enrolled in a PCORnet clinical trial (Supplementary Figure 1). The two CDM data domains identified a priori as being relevant to identify individuals as having HF were HF diagnosis codes and prescribed HF-related medications. N-terminal B-type natriuretic peptide (NT-proBNP) is not currently included as a common measure that is uniformly populated in the lab results domain of the CDM, though some individual sites may contribute these data based on their readiness and preferences. We therefore built and validated algorithms including and excluding NT-proBNP. Echocardiogram data are not currently available in the CDM, and thus were not incorporated into the algorithms.

Heart Failure Computable Phenotype Algorithm Development

Several computable phenotype algorithms were evaluated using combinations of HF diagnosis codes, medications, and NT-proBNP. Because each data element has variable sensitivity and specificity to identify HF, we developed algorithms employing various permutations of data elements and report their performance. Within the diagnosis codes data domain, ICD-9 code 428.XX is used most commonly for HF diagnosis[16,17], though there are reports of variability in performance when used alone to identify HF[6,8]. The presence of more than one ICD-9 code 428.XX across discrete episodes of care improves the specificity and positive predictive value, at the cost of lower sensitivity[18]. In addition, some argue that inclusion of outpatient diagnosis codes may increase the sensitivity of algorithms, since a significant portion of HF care is provided in outpatient settings[5,18]. Therefore, we tested algorithms employing ICD-9 code 428.XX identified during ≥1 or ≥2 episodes of care regardless of inpatient/outpatient status, as well as a combination of inpatient and outpatient codes. For algorithms that required two or more HF diagnosis codes, the two codes were required to be separated by more than 30 days, in order to capture diagnosis codes from two separate encounters.

Medication categories were identified that suggest a HF diagnosis, including: aldosterone antagonists (eplerenone, spironolactone), HF specific beta blockers (bisoprolol, carvedilol, metoprolol succinate), loop diuretics (bumetanide, ethacrynic acid, furosemide, torsemide), digoxin, angiotensin converting enzyme inhibitors and angiotensin receptor blockers. Additional medications, including sacubatril/valsartan and Ivabradine were not approved during the study period, but should be considered for future algorithms utilizing more recent

data. We examined the impact of including potassium supplementation as a category of HF medication. Since the performance of all algorithms was nearly identical with and without potassium supplementation medication, we excluded this medication category for algorithm simplicity.

Finally, we designed and validated algorithms which included thresholds of NT-proBNP in light of existing literature suggesting its utility in heart failure diagnosis, particularly in database cohort identification[9,19,20]. We employed NT-proBNP in this study as opposed to B-natriuretic peptide (BNP) due to availability in the validation cohort.

Statistical Analysis

The HF computable phenotype algorithms were evaluated with sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). Sensitivity is the proportion of HF cases that are identified as having HF by the algorithm; specificity is the proportion of non-cases that are identified as not having HF by the algorithm. PPV is defined as the proportion of subjects identified as having HF from the algorithm who are HF cases; NPV is the proportion of subjects who are identified as not having HF from the algorithm who are non-cases. Since sampling was employed to identify validated HF cases prior to 2007, analyses were weighted accordingly. Patients in the 50% random sample were assigned a sample weight of 2 (and non-sampled persons were dropped from the analysis) to account for this sampling scheme.

Validation results for computable phenotype algorithms incorporating NT-proBNP are presented both in the entire validation population, as well as only amongst the sub-population who had NT-proBNP measured. Elevated NT-proBNP algorithm criteria were met if participants had NT-proBNP measured and the value was above the specified threshold; those who did not have NT-proBNP measured or had levels below the threshold did not meet criteria. Elevated levels of NT-proBNP were defined using the following thresholds, which have been described previously: >300, >450, >900 and >1800 pg/mL[19]. To allow for direct comparisons between algorithms with and without NT-proBNP, the performance of the algorithms that include only diagnostic codes and medications was also evaluated among the subset of people with NT-proBNP measured. A sensitivity analysis was also conducted that restricted analyses to 30-day survivors. Analyses were performed using SAS statistical software, version 9.4 (SAS InstituteInc., Cary, North Carolina).

RESULTS

During the validation cohort study period (January 1, 2010 and December 31, 2012), 76,254 Olmsted County residents received a diagnosis code of any kind, of which 4,956 (6.5%) had a HF diagnosis code. Taking into account the sampling strategy used to validate HF, 2,201 (44.4%) of those with a HF diagnosis code had validated HF by manual review of the medical record; 71,298 without a HF diagnosis code were classified as not having HF based on manual review of a sampling of these records, and 2,755 with a HF diagnosis code did not meet Framingham criteria (Figure 1). The mean (SD) age in the cases (validated HF) and non-cases were 76.6 (13.3) and 51.3 (14.8), respectively; 53.3% and 52.9% of the cases and non-cases were female, respectively.

The performance of the algorithms based on HF diagnosis codes, HF medications and NT-proBNP are shown in Table 1, and PPV is plotted against sensitivity for these algorithms in Figure 2. The simplest algorithm required at least one HF diagnosis code (either inpatient or outpatient) and demonstrated the highest sensitivity (78.7%) but the lowest PPV (68.5%). The sensitivity for this algorithm is not 100% since some subjects in the cohort did not have an ICD-9 428.XX code during the study period but had validated HF from manual adjudication of records prior to the study period. Requiring more HF codes improved PPV (ie 79.3% for ≥2 HF codes) but reduced sensitivity (61.7%). Similar changes were observed for the addition of medication criteria as for NT-proBNP criteria. Specificity was also increased for algorithms with more HF criteria. In sensitivity analysis, subjects who died within 30 days after their index date were excluded and results were not materially different.

Because only a fraction of the validation population had NT-proBNP measured, we additionally present results of validation restricted to the sub-population of patients who had NT-proBNP measured (n=3,230; 1335 or 60.7% of the cases, and 1895 or 2.6% of the non-cases). The addition of elevated NT-proBNP to the algorithms incorporating diagnosis codes and medications further increased the PPV and specificity while also decreasing the sensitivity and NPV (Table 2). Sensitivity analyses were performed using cutpoints (ng/mL) of >300, >900, and >1800 to define elevated NT-proBNP (Supplemental Table S1). As expected, PPV and specificity increased with increasing cutpoints for elevated NT-proBNP, while sensitivity and NPV decreased.

As a demonstration of a large-scale pilot implementation of our HF algorithms within PCORnet, we report the results of deploying several algorithms within the Learning Health Systems Clinical Data Research Network (Table 3), which is one of the 13 total Clinical Data Research Networks within PCORnet. The Learning Health Systems Clinical Data Research Network is comprised of 6 health systems (Mayo Clinic, Allina Health System, Essentia Health, Intermountain Health Care, University of Michigan and Ohio State University); 1 health plan (Medica Research Institute); 1 data partner based in a university (Arizona State University); and 1 local public health department (Olmsted County Public Health Services)[21]. Two health systems did not have medication data available and were not included in queries requiring medication data. The total population within which the algorithms were deployed was 7,755,117. For the simplest algorithm comprised of ≥1 HF code (ie Algorithm 1 from Table 1), 254,552 individuals met the HF algorithm criteria (3.28%, Table 3), whereas 134,015 (1.73%) met the criteria for ≥2 HF codes (ie Algorithm 4 from Table 1). Of those sites able to report medication data, 1.12% of the population met criteria for ≥2 HF codes + any HF medication (Table 3).

DISCUSSION

Herein we report on the development and validation of several computable phenotype algorithms based on the PCORnet CDM in a large community-based cohort, and we demonstrate varying performance as measured by levels of sensitivity, specificity, PPV and NPV. Due to their adherence to PCORnet CDM data elements, the algorithms we present can be deployed throughout PCORnet, enabling assembly of large cohorts of individuals with HF. We demonstrated a pilot implementation of these algorithms within a single

PCORnet Clinical Data Research Network. Because these algorithms are simple in design, their strength lies in the broad availability of their component data elements and their ability to be deployed in any institution containing PCORnet CDM data elements, without the need to adapt the algorithm for deployment in each new institution.

The identification of an "ideal" computable phenotype algorithm to employ for a given application will be driven by the goal of the research. For example, to identify a population of patients with HF from which to recruit into a clinical trial, a higher PPV would be preferable despite lower sensitivity, as false positives may be resource intensive to exclude (Figure 2). Algorithm 5 (Table 1; green triangle in Figure 2) which requires ≥2 HF diagnostic codes and any HF medication would perform well in this setting, providing a PPV of 80.7% and sensitivity of 56.1%. Conversely, for an epidemiologic study where complete ascertainment of cases might be desired, algorithm 1 (≥1 HF code; sensitivity=78.7%) or 7 (≥1 inpatient or ≥2 outpatient HF codes; sensitivity=73.3%) might be used. Our pilot implementation included several algorithms across a range of algorithm sensitivities (Table 3), demonstrating the "real-world" trade-off of a decreasing population size with decreasing algorithm sensitivity.

Our results extend prior work to build algorithms to identify patients with HF which often rely heavily on diagnostic codes [5,6,8]. The simplest algorithms use diagnostic codes as the only criteria, presumably due to their availability[5,6]. The performance of such algorithms is highly variable, depending on the population from which the validation dataset was drawn and on the criteria used as the gold-standard[6,8]. Generally, however, the PPV improves when more than one diagnostic code is required, for HF as well as for other conditions[10,18,22,23]. On the other extreme, more complex algorithms have been developed—for example in the eMERGE network[24]—which rely on techniques such as natural language processing or machine learning to make use of unstructured free text data[24,25]. While these complex algorithms have been shown to out-perform simpler rule-based algorithms[24,25], such as we present, due to their complexity they are less suitable for large-scale deployment across large data networks or across multiple institutions. Specifically, they are not deployable in PCORnet since unstructured data is not available in the CDM. Our algorithms strike a balance between maximal performance and simplicity, favoring algorithm scalability and ease of deployment since they were constrained at the outset to the data elements available in the PCORnet CDM.

Among validation methodologies, our study employs a large population (n=76,254) and relies on the most rigorous validation method of manual review for Framingham HF criteria[5,6]. Less rigorous approaches (physician diagnosis or non-validated symptom lists) may misclassify HF cases. Our methodology also enables reporting of algorithm sensitivity, which is less commonly reported and critical to estimate the proportion of patients not captured, which affects an algorithm's usefulness[6,9]. An algorithm with high PPV and lower sensitivity may be desirable for clinical trials to optimize identification of true cases but would not be suitable for surveillance studies where broad capture of candidate cases would be important. The incremental contribution of medication prescription data to algorithm performance is also less well studied [23,26]. As a whole in our study, adding medication to diagnostic codes increases specificity and PPV of algorithms, while

decreasing sensitivity. Rector et al.[23] reported a similar increase in specificity with a major decrease in sensitivity, but did not report PPV. Similarly, though BNP and NT-proBNP can be used to aid in the identification of decompensated HF [20,27], its use in clinical practice is heterogeneous. Nearly 40% of validated HF cases in our cohort never had NT-proBNP measured. In our study, adding NT-proBNP as an algorithm criterion at the threshold of >450 pg/mL exerted similar influence when added to HF diagnostic codes as compared to additional HF medication criteria by increasing PPV and decreasing sensitivity. This is analogous to the change in sensitivity and PPV reported by Rosenman et al. when adding BNP at a single threshold to HF diagnostic codes[9], though they did not evaluate other BNP thresholds and did not report results for HF medication in addition to BNP. Alqaisi et al. also examined BNP criteria at several thresholds, but constructed algorithms containing either diagnostic codes or BNP alone, rather than in combination, nor did they utilize algorithm criteria containing HF medications[10]. The higher sensitivity of Table 2 algorithms is likely due to the higher likelihood of HF in this sub-population where a clinician decided to order an NT-proBNP test.

Our study has several limitations to mention. First, we performed our algorithm validation retrospectively within an existing community surveillance cohort which may introduce attendant biases or may be subject to secular changes in heart failure over time. Similarly, this validation cohort is limited to data derived from a single county, although the cohort is fairly large and encompasses multiple institutions. Our use of existing data from this community surveillance cohort facilitated our rigorous adjudication process, but for both of these reasons our algorithm generalizability would be strengthened by prospective validation of these results in an external validation cohort. We relied on an approach similar to prior research efforts [9,26,28] in assuming that all true HF cases will have a HF diagnostic code due to the cost constraints of the manual chart review process, though this remains an important limitation. We addressed this limitation by reviewing 200 charts without HF diagnostic codes and found no HF cases, thus suggesting that false negatives are less than 0.5% (1/200); our reported sensitivities may, however, still be impacted by this limitation. An additional limitation is that ICD-10 codes were not used during the defined study period. The relatively straightforward mapping of HF codes from ICD-9 to ICD-10 suggests that use of ICD-10 codes should yield similar performance, although future studies are warranted to verify this. Similarly, in our cohort, we did not have sufficient data on BNP, and so used NT-proBNP instead; also the combination of Hydralazine and Isosorbide dinitrate or sacubitril/valsartan was not frequent enough in the validation dataset to be tested in this study. Finally, our algorithms are not designed to achieve identification of acute decompensated HF, active HF hospitalization or differentiation between HF with preserved ejection fraction and HF with reduced ejection fraction, each of which would likely require different sets of criteria.

By developing and validating a HF computable phenotype built upon the PCORnet CDM, we begin to operationalize the vision of harnessing EMR data for a learning healthcare system[29] by facilitating the identification of participants from clinical care for research and intervention. Phenotypes such as these, combined with networks like PCORnet, offer the potential to identify large patient populations in the tens of millions, thereby transforming how we approach clinical research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding and Acknowledgments

This work was made possible by support from the National Institutes of Health (R01 HL 120859, R01 AG034676 and K23 HL135274), the Patient Centered Outcomes Research Institute Learning Health System CDRN (1501–26638), the Health eHeart Alliance Patient Powered Research Network (1306– 04709) and the PCORnet Cardiovascular Health Collaborative Research Group. The funding sources played no role in the design, conduct, or reporting of this study. Olgin and Pletcher –5U2CEB021881 The Health ePeople Resource for Mobilized Research. The content is solely the responsibility of the authors.

We acknowledge the help of Ellen Koepsell, RN, Sandra Severson, RN, Dawn Schubert, RN, and Deborah Strain.

ABBREVIATIONS:

EMR	Electronic Medical Record
HF	Heart Failure
PCORnet	Patient Centered Outcomes Research Network
CDM	Common Data Model
NT-proBNP	N-terminal B-type natriuretic peptide
BNP	B-natriuretic peptide
PPV	Positive predictive value
NPV	Negative predictive value
ICD-9	International Classification of Diseases – 9th Revision, Clinical Modification (ICD-9)

References

1. Fleurence RL, Curtis LH, Califf RM, et al. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc 2014;21:578–82. 10.1136/amiajnl-2014-002747 [PubMed: 24821743]
2. Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. J Am Med Inform Assoc 2013;20:117–21. 10.1136/amiajnl-2012-001145 [PubMed: 22955496]
3. Benjamin EJ, Mozaffarian D, Go AS, et al. Heart Disease and Stroke Statistics--2017 Update: A Report From the American Heart Association 2017 10.1161/CIR.0000000000000152
4. Jencks SF, Williams MV., Coleman EA. Rehospitalizations among Patients in the Medicare Fee-for-Service Program. N Engl J Med 2009;360:1418–28. 10.1056/NEJMsa0803563 [PubMed: 19339721]
5. Saczynski JS, Andrade SA, Harrold JT, et al. A systematic review of validated methods for identifying heart failure using administrative data. Parmacoepidemiology Drug Saf 2012;21:129–40. 10.1002/pds
6. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. Can J Cardiol 2010;26:306–12. 10.1016/S0828-282X(10)70438-4 [PubMed: 20931099]

7. Cox ZL, Lewis CM, Lai P, et al. Validation of an automated electronic algorithm and 'dashboard' to identify and characterize decompensated heart failure admissions across a medical center. Am Heart J 2017;183:40–8. 10.1016/j.ahj.2016.10.001 [PubMed: 27979040]
8. McCormick N, Lacaille D, Bhole V, et al. Validity of heart failure diagnoses in administrative databases: A systematic review and meta-analysis. PLoS One 2014;9 10.1371/journal.pone.0104519
9. Rosenman M, He J, Martin J, et al. Database queries for hospitalizations for acute congestive heart failure: flexible methods and validation based on set theory. J Am Med Inform Assoc 2014;21:345–52. 10.1136/amiajnl-2013-001942 [PubMed: 24113802]
10. Alqaisi F, Williams LK, Peterson EL, et al. Comparing methods for identifying patients with heart failure using electronic data sources. BMC Health Serv Res 2009;9:237 10.1186/1472-6963-9-237 [PubMed: 20021648]
11. St Sauver JL, Grossardt BR, Leibson CL, et al. Generalizability of epidemiological findings and public health decisions: An illustration from the Rochester Epidemiology Project. Mayo Clin Proc 2012;87:151–60. 10.1016/j.mayocp.2011.11.009 [PubMed: 22305027]
12. Finney RL, Alexander A, Haug P, et al. Patient-Centered Network of Learning Health Systems: Developing a resource for clinical translational research. J Clin Transl Sci 2017;1:40–4. 10.1017/cts.2016.11 [PubMed: 28515960]
13. Roger VL, Weston SA, Redfield MM, et al. Trends in Heart Failure Incidence and Survival in a Community-Based Population. J Am Med Assoc 2004;292:344–50. 10.1001/jama.292.3.344
14. Gerber Y, Weston SA, Redfield MM, et al. A Contemporary Appraisal of the Heart Failure Epidemic in Olmsted County, Minnesota, 2000 to 2010. JAMA Intern Med 2015;175:996–1004. 10.1001/jamainternmed.2015.0924 [PubMed: 25895156]
15. Ho KKL, Pinsky JL, Kannel WB, et al. The epidemiology of heart failure: The Framingham Study. J Am Coll Cardiol 1993;22:A6–13. 10.1016/0735-1097(93)90455-A
16. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. Med Care 2005;43:182–8. 10.1097/00005650-200502000-00012 [PubMed: 15655432]
17. Vermeulen MJ, Tu JV., Schull MJ. ICD-10 adaptations of the Ontario acute myocardial infarction mortality prediction rules performed as well as the original versions. J Clin Epidemiol 2007;60:971–4. 10.1016/j.jclinepi.2006.12.009 [PubMed: 17689814]
18. Schultz SE, Rothwell DM, Chen Z, et al. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. Chronic Dis Inj Can 2013;33:160–6. [PubMed: 23735455]
19. Januzzi JL, Van Kimmenade R, Lainchbury J, et al. NT-proBNP testing for diagnosis and short-term prognosis in acute destabilized heart failure: An international pooled analysis of 1256 patients: The international collaborative of NT-proBNP study. Eur Heart J 2006;27:330–7. 10.1093/eurheartj/ehi631 [PubMed: 16293638]
20. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: Executive summary: A report of the American college of cardiology foundation/American Heart Association task force on practice guidelines. Circulation 2013;128:1810–52. 10.1161/CIR.0b013e31829e8807 [PubMed: 23741057]
21. Rutten LJF, Alexander A, Embi PJ, et al. Patient-Centered Network of Learning Health Systems: Developing a resource for clinical translational research. J Clin Transl Sci 2017;1:40–4. 10.1017/cts.2016.11 [PubMed: 28515960]
22. Tu K, Campbell NR, Chen Z-L, et al. Accuracy of administrative databases in identifying patients with hypertension. Open Med 2007;1:e18– 26.http://www.ncbi.nlm.nih.gov/pubmed/20101286%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2801913 [PubMed: 20101286]
23. Rector TS, Wickstrom SL, Shah M, et al. Specificity and sensitivity of claims-based algorithms for identifying members of Medicare+Choice health plans that have chronic medical conditions. Health Serv Res 2004;39:1839–57. 10.1111/j.1475-6773.2004.00321.x [PubMed: 15533190]
24. Bielinski SJ, Pathak J, Carrell D, et al. A Robust e-Epidemiology Tool in Phenotyping Heart Failure with Differentiation for Preserved and Reduced Ejection Fraction: the Electronic Medical

Records and Genomics (eMERGE) Network. J Cardiovasc Transl Res 2015;8:475–83. 10.1007/s12265-015-9644-2 [PubMed: 26195183]

25. Blecker S, Katz SD, Horwitz LI, et al. Comparison of Approaches for Heart Failure Case Identification From Electronic Health Record Data. JAMA Cardiol 2016;1:1014–20. 10.1001/jamacardio.2016.3236 [PubMed: 27706470]
26. Li Q, Glynn R, Dreyer NA, et al. Validity of claims-based definitions of left ventricular systolic dysfunction in Medicare patients. Pharmacoepidemiol Drug Saf 2011;20:700–8. 10.1002/pds [PubMed: 21608070]
27. Maisel A B-type natriuretic peptide levels: Diagnostic and prognostic in congestive heart failure: What's next? Circulation 2002;105:2328–31. 10.1161/01.CIR.0000019121.91548.C2 [PubMed: 12021215]
28. Goff DC, Pandey DK, Chan FA, et al. Congestive heart failure in the United States: Is there more than meads the I(CD Code)? The Corpus Christi Heart Project 2000;160:197–202.
29. Smith M, Saunders R, Stuckhardt L, et al. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America The Institute of Medicine. The National Academies Press 2012 10.17226/13444

Summary Table:

What was already known on the topic:

- Prior algorithms to identify heart failure from medical records have been developed, and vary in composition and complexity—many contain non-standardized data elements and must be adapted for use at each new site.
- Heart failure algorithms most commonly use billing codes alone, but vary in performance compared to validated heart failure definitions.
- A validated heart failure algorithm using PCORnet common data model elements does not currently exist.

What this study added to our knowledge:

- We develop and validate multiple heart failure algorithms whose strength lie in the broad availability of their component data elements and can be deployed in any institution containing these elements.
- Adding additional algorithm components can improve algorithm performance in some aspects, at the expense of others.
- Choice of heart failure algorithm should be guided by the goals of the research application.

Highlights:

- We developed multiple algorithms to identify heart failure from medical record data
- Various algorithms have tradeoffs between sensitivity and positive predictive value
- Simpler algorithms have high sensitivity but lower positive predictive value
- Additional components, like medication or BNP, impact the algorithm similarly
- Algorithm choice should be guided by the goals of the research application

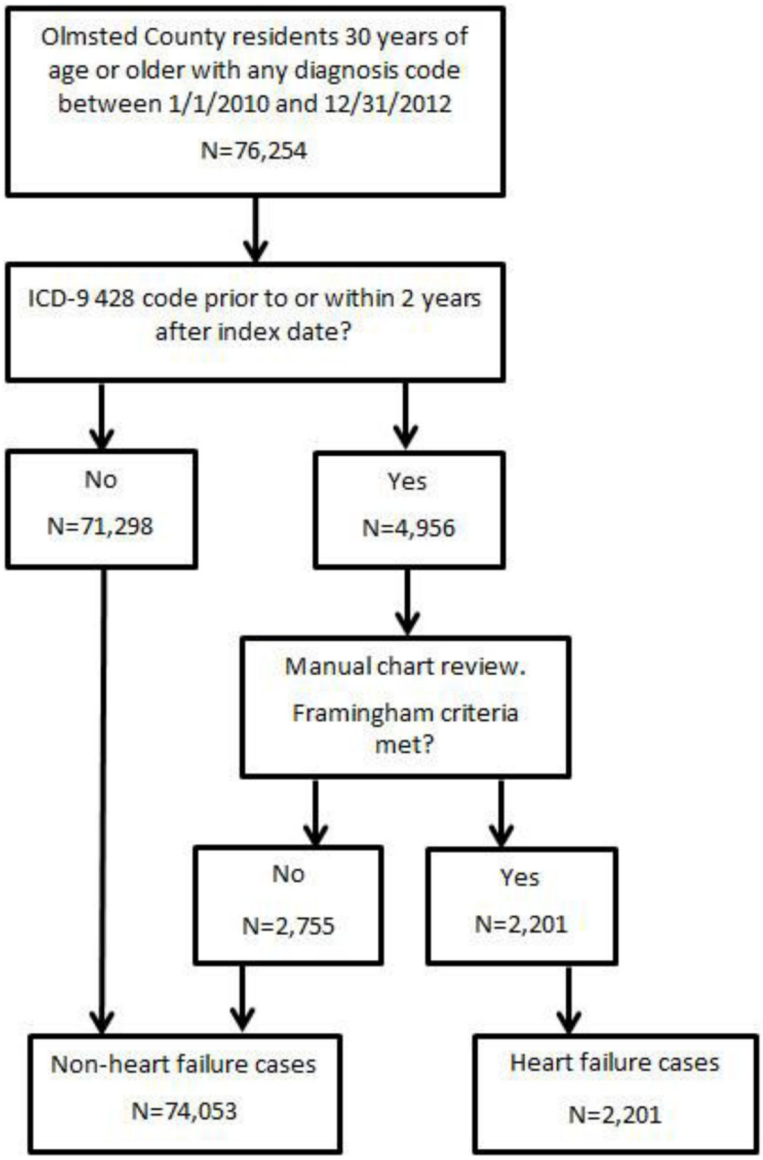

Figure 1: Validation cohort

Classification of the cohort used for algorithm validation.

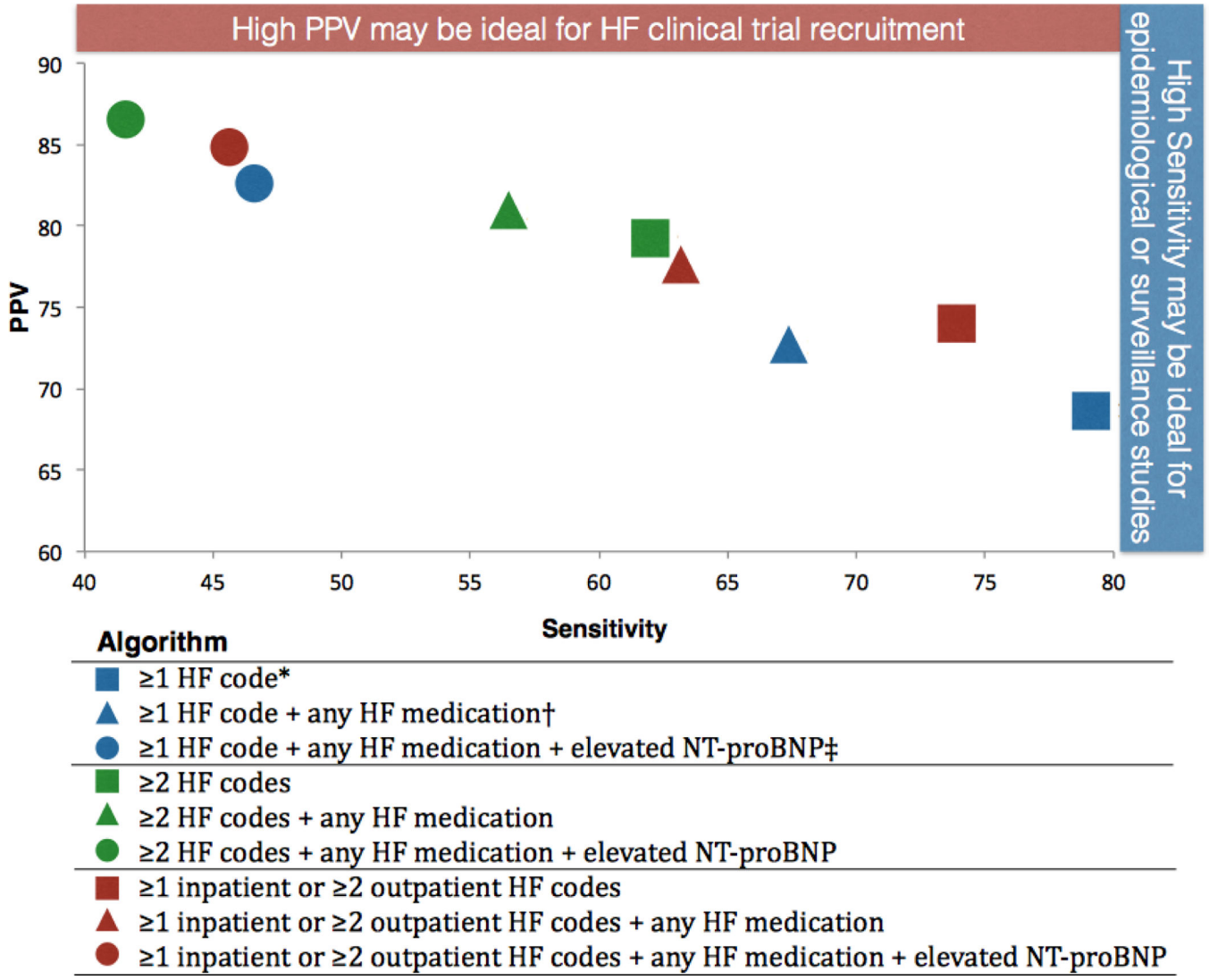

Figure 2: Positive predictive value (PPV) and Sensitivity for HF computable phenotype algorithms

PPV vs. Sensitivity for HFphenotype algorithms incorporating HF diagnosis codes, HF medications and elevated NT-proBNP

Table 1:

Performance of heart failure computable phenotype algorithms compared to validated heart failure in Olmsted County, Minnesota, 2010–2012 (N=76,254)

		Sensitivity		Specificity		PPV		NPV	
	Algorithm	**Rate**	**%**	**Rate**	**%**	**Rate**	**%**	**Rate**	**%**
1	≥1 HF code *	1733/2201	78.7	73257/74053	98.9	1733/2529	68.5	73257/73725	99.4
2	≥1 HF code + any HF medication †	1479/2201	67.2	73485/74053	99.2	1479/2047	72.3	73485/74207	99.0
3	≥1 HF code + any HF medication + elevated NT-proBNP ‡	1018/2201	46.3	735837/74053	99.7	1018/1234	82.5	73837/75020	98.4
4	≥2 HF codes	1357/2201	61.7	73698/74053	99.5	1357/1712	79.3	73698/74542	98.9
5	≥2 HF codes + any HF medication	1234/2201	56.1	73757/74053	99.6	1234/1530	80.7	73757/74724	98.9
6	≥2 HF codes + any HF medication + elevated NT-proBNP	916/2201	41.6	73910/74053	99.8	916/1059	86.5	73910/74077	98.3
7	≥1 inpatient or ≥2 outpatient HF codes	1614/2201	73.3	73490/74053	99.2	1614/2177	74.1	73490/74077	99.2
8	≥1 inpatient or ≥2 outpatient HF codes + any HF medication	1386/2201	63.0	73638/74053	99.4	1386/1801	77.0	73638/74453	98.9
9	≥1 inpatient or ≥2 outpatient HF codes + any HF medication + elevated NT-proBNP	1000/2201	45.4	73870/74053	99.8	1000/1183	84.5	73870/75071	98.4

*Heart failure code is International Classification of Diseases – 9th Revision, Clinical Modification code 428.

†Heart failure medications include aldosterone antagonists (eplerenone, spironolactone), HF specific beta blockers (bisoprolol, carvedilol, metoprolol succinate), loop diuretics (bumetanide, ethacrynic acid, furosemide, torsemide), digoxin, angiotensin converting enzyme inhibitors and angiotensin receptor blockers.

‡Elevated NT-proBNP criteria defined as >450 pg/mL; those who did not have NT-proBNP measured or had NT-proBNP <450 pg/mL did not meet this criteria.

HF, heart failure; NPV, negative predictive value; PPV, positive predictive value.

Table 2:

Performance of heart failure computable phenotype algorithms in the sub-population of patients who had NT-proBNP measured compared to validated heart failure in Olmsted County, Minnesota, 2010– 2012 (N=3,230)

	Sensitivity		Specificity		PPY		NPY	
Algorithm	**Rate**	**%**	**Rate**	**%**	**Rate**	**%**	**Rate**	**%**
Elevated NT-proBNP *	1178/1335	88.2	1178/1895	62.2	1178/1895	62.2	1178/1335	88.2
≥1 HF codet †	1267/1335	94.9	1502/1895	79.3	1267/1660	76.3	1502/1570	95.7
≥1 HF code + any HF medication ‡	1141/1335	85.5	1586/1895	83.7	1141/1450	78.7	1586/1780	89.1
≥1 HF code + elevated NT-proBNP	1136/1335	85.1	1618/1895	85.4	1136/1413	80.4	1618/1817	89.0
≥1 HF code + any HF medication + elevated NT-proBNP	1018/1335	76.3	1679/1895	88.6	1018/1234	82.5	1679/1996	84.1
≥2 HF codes	1073/1335	80.4	1665/1895	87.9	1073/1303	82.3	165/1927	86.4
≥2 HF codes + any HF medication	1000/1335	74.9	1696/1895	89.5	1000/1199	83.4	1696/2031	83.5
≥2 HF codes + elevated NT-proBNP	984/1335	73.7	1726/1895	91.1	984/1153	85.3	1726/2077	83.1
≥2 HF codes + any HF medication + elevated NT-proBNP	916/1335	68.6	1752/1895	92.5	916/1059	86.5	1752/2171	80.7
≥1 inpatient or >2 outpatient HF codes	1220/1335	91.4	1574/1895	83.1	1220/1541	79.2	1574/1689	93.2
≥1 mpatient or >2 outpatient HF codes + any HF medication	1098/1335	82.2	1642/1895	86.6	1098/1351	81.3	1642/1879	87.4
≥1 inpatient or >2 outpatient codes + elevated NT-proBNP	1116/1335	83.6	1656/1895	87.4	1116/1355	82.4	1656/1875	88.3
≥1 inpatient or >2 outpatient codes + any HF medication + elevated NT-proBNP	1000/1335	74.9	1712/1895	90.3	1000/1183	84.5	1712/2047	83.6

* Elevated NT-proBNP criteria defined as >450 pg/mL; those who did not have NT-proBNP measured or had NT-proBNP <450 pg/mL did not meet this criteria.

† Heart failure code is International Classification of Diseases – 9th Revision, Clinical Modification code 428.

‡ Heart failure medications include aldosterone antagonists (eplerenone, spironolactone), HF specific beta blockers (bisoprolol, carvedilol, metoprolol succinate), loop diuretics (bumetanide, ethacrynic acid, furosemide, torsemide), digoxin, angiotensin converting enzyme inhibitors and angiotensin receptor blockers.

HF, heart failure; NPV, negative predictive value; NT-proBNP, N-terminal B-type natriuretic peptide; PPV, positive predictive value.

Table 3:

Heart failure computable phenotype algorithms deployed within the Learning Health Systems[‡] PCORnet Clinical Data Research Network, Query period 2013–2016.

Algorithm	Age Group	Heart failure population identified		Total Population
		N	%	
≥1 HF code[*] (Table 1, Algorithm 1)	30–49	26,903	0.79%	3,423,297
	50–64	61,532	2.4%	2,493,470
	65+	166,117	9.0%	1,838,350
	All	**254,552**	**3.28%**	**7,755,117**
≥2 HF codes (Table 1, Algorithm 4)	30–49	13,686	0.40%	3,423,297
	50–64	31,471	1.26%	2,493,470
	65+	88,858	4.83%	1,838,350
	All	**134,015**	**1.73%**	**7,755,117**
≥2 HF codes + any HF medicationf[†] (Table 1. Algorithm 5)	30–49	6,797	0.23%	2,954,319
	50–64	18,951	0.88%	2,143,829
	65+	49,028	3.08%	1,592,416
	All	**74,776**	**1.12%**	**6,690,564**

[*] Heart failure code is International Classification of Diseases – 9th Revision, Clinical Modification code 428.

[†] Heart failure medications include aldosterone antagonists (eplerenone, spironolactone), HF specific beta blockers (bisoprolol, carvedilol, metoprolol succinate), loop diuretics (bumetanide, ethacrynic acid, furosemide, torsemide), digoxin, angiotensin converting enzyme inhibitors and angiotensin receptor blockers.

[‡] These algorithms were deployed within the Learning Health Systems (LHSnet) Clinical Data Research Network of PCORnet. LHSnet is comprised of 6 health systems (Mayo Clinic, Allina Health System, Essentia Health, Intermountain Health Care, University of Michigan and Ohio State University); 1 health plan (Medica Research Institute); 1 data partner based in a university (Arizona State University); and 1 local public health department (Olmsted County Public Health Services). Two health systems did not have medication data available and were not included in queries requiring medication data.