

RESEARCH ARTICLE

A diagnostic accuracy study validating cardiovascular ICD-9-CM codes in healthcare administrative databases. The Umbria Data-Value Project

Francesco Cozzolino^{1,2}, Alessandro Montedori¹, Iosief Abraha^{1,3*}, Paolo Eusebi¹, Chiara Grisci⁴, Anna Julia Heymann⁵, Guido Lombardo⁶, Anna Mengoni², Massimiliano Orso^{1,2}, Giuseppe Ambrosio²

1 Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy, **2** Division of Cardiology, Santa Maria della Misericordia Hospital, University of Perugia School of Medicine, Perugia, Italy, **3** Centro Regionale Sangue, Servizio Immunotrasfusionale, Azienda Ospedaliera di Perugia, Perugia, Italy, **4** Department of Experimental Medicine, Section of Public Health, University of Perugia, Perugia, Italy, **5** Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche "Togo Rosati", Perugia, Italy, **6** Department of Surgical and Biomedical Sciences, University of Perugia, Perugia, Italy

* iosief_a@yahoo.it



OPEN ACCESS

Citation: Cozzolino F, Montedori A, Abraha I, Eusebi P, Grisci C, Heymann AJ, et al. (2019) A diagnostic accuracy study validating cardiovascular ICD-9-CM codes in healthcare administrative databases. The Umbria Data-Value Project. PLoS ONE 14(7): e0218919. <https://doi.org/10.1371/journal.pone.0218919>

Editor: Gianni Virgili, Università degli Studi di Firenze, ITALY

Received: January 2, 2019

Accepted: June 13, 2019

Published: July 8, 2019

Copyright: © 2019 Cozzolino et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was developed within the Data-Value Project ("Progetto Data-Value: valorizzazione del dato sanitario regionale per la Ricerca dei Servizi Sanitari (Health Services Research)" – D.G.R. No 1798 of 29/12/2014) supported by funding from the Regional Health Authority of Umbria. The funder had no role in

Abstract

Background

Administrative healthcare databases are useful and inexpensive tools that can provide a comprehensive assessment of the burden of diseases in terms of major outcomes, such as mortality, hospital readmissions, and use of healthcare resources. However, a crucial issue is the reliability of information gathered. The aim of this study was to validate ICD-9 codes for several major cardiovascular conditions, i.e., acute myocardial infarction (AMI), atrial fibrillation/flutter (AF), and heart failure (HF), in order to use them for epidemiological, outcome, and health services research.

Methods

Data from the centralised administrative database of the Umbria Region (890,000 residents, located in Central Italy) were considered. Patients with a first hospital discharge for AMI, AF/flutter, and HF, between 2012 and 2014, were identified using ICD-9-CM codes in primary position. A sample of cases and non-cases was randomly selected, and the corresponding medical charts reviewed by specifically trained investigators. For each disease, case ascertainment was based on all clinical, laboratory, and instrumental examinations available in medical charts. Sensitivity, specificity, and predictive values with 95% confidence intervals (CIs), were calculated.

Results

We reviewed 458 medical charts, 128 for AMI, 127 for AF/flutter, 127 for HF, and 76 of non-cases for each condition. Diagnostic accuracy measures of the original discharge diagnosis

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

were as follows. AMI: sensitivity 98% (95% CI, 94–100%), specificity 91% (95% CI, 83–97%), positive predictive value (PPV) 95% (95% CI, 89–98%), negative predictive value (NPV) 97% (95% CI, 91–100%). AF/flutter: sensitivity 95% (95% CI, 90–98%), specificity 95% (95% CI, 87–99%), PPV 97% (95% CI, 92–99%), NPV 92% (95% CI, 84–97%). HF: sensitivity 96% (95% CI, 91–99%), specificity 90% (95% CI, 81–96%), PPV 94% (95% CI, 88–97%), NPV 93% (95% CI, 85–98%).

Conclusion

The case ascertainment for AMI, AF and flutter, and HF, showed a high level of accuracy ($\geq 90\%$). The healthcare administrative database of the Umbria Region can be confidently used for epidemiological, outcome, and health services research.

Introduction

Administrative databases are considerable data repositories that are increasingly being used within healthcare systems[1, 2]. These databases are organized and maintained at different administrative levels including hospitals, local health units, and at regional level. The continuous collection of demographic data together with diagnosis, therapeutic interventions as well as prescription information makes these databases attractive for comprehensive assessment of the burden of diseases in terms of major outcomes, such as mortality, hospital readmissions, and use of healthcare resources[1, 2].

In addition to maintaining a rigorous anonymity of patient's demographic, the most relevant data that makes these healthcare databases interesting for research purposes is the diagnosis provided to the patient at hospital discharge. This diagnosis is coded according to the International Classification of Diseases (ICD) which is a standardized diagnostic tool planned to map health conditions. When individual patient data are linked with other data (prescription and laboratory data) it is possible to explore a wide range of clinical issues, research questions as well as quality performance evaluations. To reach this target, administrative databases need to be validated, which means the diagnoses that correspond to the ICD-9 code need to be ascertained according to a defined disease criteria by consulting a reference standard which is usually the medical chart[1, 3, 4].

Acute myocardial infarction (AMI), heart failure (HF), and atrial fibrillation (AF) are the most common cardiovascular diseases[5–7] in developed countries and they are the leading cause of morbidity and mortality, thereby representing a major social and economic problem [5, 8]. Administrative databases are significant tools that can provide the best assessment of the incidence, prevalence and general prognosis of cardiovascular diseases[3, 9–11]. In addition, they can contribute to identifying the risk factors in the development of cardiovascular diseases as well as the outcomes, including mortality, that they can determine.

According to our published protocol[3], the objective of the present study was to evaluate the accuracy of the ICD-9-CM codes related to AMI, AF and flutter, and HF in the administrative database of the Regional Health Authority of Umbria.

Materials and methods

Setting and data source

Administrative database and source population. Details regarding the regional healthcare administrative database of Umbria have been reported elsewhere[3]. Briefly, this database

constantly collects demographic and hospital data as well as prescription data regarding all residents. The Umbria administrative healthcare database has been used for research purposes [12–15] and it has been validated for several oncological diagnoses [16–21].

Any permanent resident in the Umbria Region aged 18 years or older that has been discharged from a hospital with the diagnosis of AMI, AF or atrial flutter, and HF was eligible for inclusion.

Residents that have been hospitalised outside the regional territory of Umbria were excluded from analysis [3].

Case selection and sampling method. From the entire discharge abstract database of Umbria we identified three cohorts of “cases”, that is patients having the ICD-9 codes located in primary position of acute myocardial infarction (ICD-9 codes 410.x), atrial fibrillation (code 427.31) and flutter (code 427.32), and heart failure (codes 428.x), between 2012 and 2014. According to Italian legislation, the primary diagnosis constitutes the main cause of the need for treatment and/or diagnostic tests, and is mainly responsible for the use of resources. We excluded estimated prevalent cases, i.e. patients discharged from hospitals with the same diagnosis in the five years before. Repeated hospital admissions were also excluded. From each cohort, we extracted a random sample of 130 cases (see Statistical Methods for details). At the same time, we identified a cohort of “non-cases”, that is patients who had been discharged in the same period of time, also with a diagnosis of cardiovascular disease (ICD-9 codes 390–459), but other than AMI, AF and flutter, and HF. From this cohort of non-cases, we extracted a random sample of 80 patients. This sample of non-cases was used as control for each of the three conditions.

Chart abstraction and case ascertainment

Medical charts of the samples of cases and non-cases were selected by simple randomization method using SAS 9.4 procedures and were obtained from hospitals for case ascertainment. The following data were abstracted from the medical charts: unique patient identification code, age and sex, hospital admission and discharge dates, diagnostic procedures as well as interventions that contributed to the diagnosis of the disease. In addition, clinical, laboratory and instrumental procedures data, including the date of performance, were abstracted.

Two physicians acting as chart reviewers received specific training on data abstraction. An initial chart review was then performed, with the same medical charts ($n = 20$) being independently examined by each reviewer. The inter-reviewer agreement was very high ($\kappa > 0.9$). To further ensure consistency among the reviewers, findings of this initial assessment were discussed in review; data extraction was then completed independently by each reviewer using a standardised data collection form. Reviewers were not blinded to the diagnosis due to limitation of resources. However, the validation process was not biased as the reviewers independently abstracted the data.

Uncertainties were discussed and resolved through third party involvement (GA).

Case ascertainment of disease within medical charts was based on symptoms, laboratory, and diagnostic tests, as described below.

Validation criteria

Criteria used for validation of the considered ICD-9-CM codes were those derived from international guidelines or systematic reviews published on these topics, as previously defined [3]. To validate ICD-9 codes 410.x for AMI, we considered the criteria defined by the European Society of Cardiology [22].

For validation of the ICD-9 code 427.31 for AF and 427.32 for atrial flutter, we required at least one ECG tracing documenting, the presence of AF or atrial flutter, according to the ESC Guidelines [23].

To validate the ICD-9 codes 428.x relating to HF, we considered the European Society of Cardiology Heart Failure Guidelines [24]. These guidelines consider the algorithm for the diagnosis of HF in the non-acute setting and in the acute setting. As the ICD-9 codes do not distinguish between acute and non-acute settings, we combined the clinical presentation of both settings. Accordingly, diagnosis of HF was adjudicated when, in addition to the presence of symptoms (such as dyspnoea, orthopnoea), or presence of signs at physical examination (rales, bilateral ankle oedema, increased jugular venous pressure, displaced apical beat), at least one of the following conditions were found in the medical chart: 1. any abnormalities in resting ECG; 2. brain natriuretic peptide (BNP) concentration ≥ 35 pg/mL, and/or N-terminal pro-brain natriuretic peptide (NT-proBNP) ≥ 125 pg/mL; 3. echocardiography abnormalities (ventricular and atrial volumes and function) attributable to heart failure.

Statistical analysis

As reported elsewhere [3], for each index condition (ICD-9 code) we anticipated a sample of 121 charts of cases to obtain an expected sensitivity of 80%, with a half-width of the 95% CI equals to 8% [25]. For specificity, we calculated that a sample of 73 charts of non-cases (i.e., cardiovascular patients without the target diseases of our interest) was necessary to obtain an expected specificity of 90%, with a half-width of the 95% CI equals to 8%, according to binomial exact calculation [25]. Expected accuracy figures were based on published literature [10, 11, 26–29]. We then chose to review a higher number of medical charts compared to the anticipated sample size to allow for potential missing medical charts from hospital archives.

Sensitivity and specificity, with their corresponding 95% CI, were calculated for each ICD-9-CM code by producing 2×2 tables. Positive and negative predicting values were also calculated, along with their 95% CI.

Reporting

To ensure quality and thoroughness of reporting, this paper followed the criteria of Standards for Reporting Diagnostic Accuracy (STARD) 2015 [30] (S1 Table).

Ethics statement

Ethics approval has been obtained from the Regional Ethics Committee of Umbria (CEAS), registry No 2695/15 of 16/12/2015.

Results

From each cohort, we extracted a random sample of 130 cases, while from the cohort of non-cases we extracted a random sample of 80 patients. Of these, we were able to retrieve and analyse 128 medical charts for AMI, 127 for AF and flutter, and 127 for HF. For non-cases, we retrieved and analysed 76 medical charts. **Tables 1–3** show the characteristics of the patients for each disease. A minimal anonymized dataset is available as supporting information file (S1 Dataset).

Diagnostic accuracy measures are reported in **Table 4**.

Table 1. Characteristics of patients with acute myocardial infarction who were identified in the Regional Administrative Database of Umbria.

Acute myocardial infarction	
Incident cases (N medical charts reviewed)	128
ICD-9 code, N (%)	
410.0 Anterolateral wall	3 (2)
410.1 Other anterior wall	31 (24)
410.2 Inferolateral wall	8 (6)
410.3 Inferoposterior wall	1 (1)
410.4 Other inferior wall	22 (17)
410.5 Other lateral wall	1 (1)
410.6 True posterior wall infarction	-
410.7 Subendocardial infarction	53 (41)
410.8 Other specified sites	1 (1)
410.9 Unspecified site	8 (6)
Sex, N (%)	
Male	85 (66)
Female	43 (34)
Age, N (%)	
< 60	29 (23)
60–79	61 (48)
≥ 80	38 (30)
Instrumental examinations, N (%)	
Electrocardiography (ECG)	125 (98)
Echocardiography	110 (86)
Coronary arteriography	91 (71)
Laboratory analyses, N (%)	
Troponin	124 (97)
CPK-MB	60 (47)
Creatine phosphokinase (CPK)	28 (22)
Lactate dehydrogenase (LDH)	31 (24)
Surgical procedures, N (%)	
Percutaneous transluminal coronary angioplasty (PTCA) or coronary atherectomy	79/128 (62)
PTCA in STEMI	54/79 (68)
PTCA in NSTEMI	25/79 (32)

<https://doi.org/10.1371/journal.pone.0218919.t001>

Acute myocardial infarction

We identified a cohort of 4,682 patients with AMI, from which we extracted a sample of 130 cases, 128 clinical charts were analysed (two clinical charts were not available). **Table 1** shows the basic characteristics of the patients with AMI. The most common ICD-9 subgroup was code 410.7 (subendocardial infarction) (41%), followed by code 410.1 (other anterior wall) (24%), and by code 410.4 (other inferior wall) (17%). Two thirds of patients were males. Most patients (48%) were in the age class 60–79 years. Instrumental examinations performed for the diagnosis included electrocardiography, echocardiography, and coronary angiography. Troponin levels were measured in almost all patients. Sixty-two percent of patients underwent percutaneous transluminal coronary angioplasty (PTCA).

Table 2 shows the diagnostic accuracy measures for AMI: sensitivity 98% (95% CI: 94% - 100%), specificity 91% (95% CI: 83% - 97%), positive predictive value (PPV) 95% (95% CI: 89% - 98%), and negative predictive value (NPV) 97% (95% CI: 91% - 100%).

Table 2. Characteristics of patients with atrial fibrillation and flutter who were identified in the Regional Administrative Database of Umbria.

Atrial fibrillation and flutter	
Incident cases (N medical charts reviewed)	127
ICD-9 code, N (%)	
427.31 Atrial fibrillation	107 (84)
427.32 Atrial flutter	20 (16)
Sex	
Male	67 (53)
Female	60 (47)
Age, N (%)	
< 60	14 (11)
60–79	66 (52)
≥ 80	47 (37)
Instrumental examinations, N (%)	
Electrocardiography (ECG)	126 (99)
Echocardiography	92 (72)
Procedures, N (%)	
Electrical cardioversion	37 (29)

<https://doi.org/10.1371/journal.pone.0218919.t002>

Misclassification of cases and non-cases is described in [Table 5](#). False positives were patients with troponin levels within the normal range, or not reported in the medical chart. The false negatives were non-cases actually found to have increased troponin concentration and presence of symptoms or instrumental confirmation of myocardial infarction.

Atrial fibrillation/flutter

We identified a cohort of 2,792 patients with AF/flutter, from which we extracted a sample of 130 cases, 127 clinical charts were analysed (three clinical charts were not available).

[Table 3](#) shows the basic characteristics of the patients with AF/flutter. In this sample, 84% percent of patients had atrial fibrillation, while 16% had atrial flutter according to ICD codes in primary position. Fifty-three percent of patients were males. Most patients (52%) were in the age class 60–79 years. During hospital stay, almost all patients underwent electrocardiography, and most of them (72%) echocardiography. Twenty-nine percent of the patients underwent electrical cardioversion ([Table 2](#)).

The diagnostic accuracy measures for AF/flutter were: sensitivity 95% (95% CI: 90% - 98%), specificity 95% (95% CI: 87% - 99%), PPV 97% (95% CI: 92% - 99%), and NPV 92% (95% CI: 84% - 97%) ([Table 2](#)).

All patients without an ECG in the medical chart, or with ECG negative for AF/flutter, were considered false positives. All false positives had a clinical history of AF or flutter, while, between the non-cases, false negatives were patients having other cardiovascular diseases even with an ECG confirmation of AF or flutter in the medical chart ([Table 5](#)).

Heart failure

We identified a cohort of 6,196 patients with AMI, from which we extracted a sample of 130 cases, 127 clinical charts were analysed (three clinical charts were not available).

[Table 4](#) shows the basic characteristics of patients with HF. The most common ICD-9 subgroup was code 428.0 (congestive heart failure, unspecified) (64%), followed by code 428.1 (left heart failure) (24%). Fifty-three percent of patients were females. Most patients (61%)

Table 3. Characteristics of patients with heart failure who were identified in the Regional Administrative Database of Umbria.

Heart failure	
Incident cases (N medical charts reviewed)	127
ICD-9 code, N (%)	
428.0 Congestive heart failure, unspecified	81 (64)
428.1 Left heart failure	30 (24)
428.2 Systolic heart failure	
428.20 Unspecified	-
428.21 Acute	-
428.22 Chronic	-
428.23 Acute on chronic	3 (2)
428.3 Diastolic heart failure	
428.30 Unspecified	2 (2)
428.31 Acute	-
428.32 Chronic	-
428.33 Acute on chronic	1 (1)
428.4 Combined systolic and diastolic heart failure	
428.40 Unspecified	-
428.41 Acute	3 (2)
428.42 Chronic	-
428.43 Acute on chronic	5 (4)
428.9 Heart failure, unspecified	2 (2)
Sex, N (%)	
Male	60 (47)
Female	67 (53)
Age, N (%)	
< 60	4 (3)
60–79	45 (35)
≥ 80	78 (61)
Instrumental examinations, N (%)	
Electrocardiography (ECG)	123 (97)
Echocardiography	81 (64)
Chest radiography	61 (48)
Coronary arteriography	10 (8)
Laboratory analysis, N (%)	
Brain natriuretic peptide (BNP)	59 (46)

<https://doi.org/10.1371/journal.pone.0218919.t003>

were 80 years or older. The most frequent diagnostic instrumental examinations were electrocardiography, echocardiography, and chest radiography. Brain natriuretic peptide (BNP) levels were measured in 46% of patients (Table 3).

The diagnostic accuracy measures for HF were: sensitivity 96% (95% CI: 91% - 99%), specificity 90% (95% CI: 81% - 96%), PPV 94% (95% CI: 88% - 97%), and NPV 93% (95% CI: 85% - 98%) (Table 2).

Patients with signs or symptoms but without instrumental or laboratory confirmation of HF in their medical chart were considered false positives. Most false positives (5/8) were patients ≥ 85 having polypathologies that during the hospital admission underwent only treatment procedures, no instrumental examinations were performed. Among the non-cases, false negatives were patients that, in addition to other cardiovascular diseases, also had signs, symptoms and instrumental examinations confirming heart failure (Table 5).

Table 4. Cross tabulation of the index test (ICD-9-CM code) for the results of the reference standard (medical chart).

	True Positive	False Positive	True Negative	False Negative
Myocardial infarction	121	7	74	2
Atrial fibrillation and flutter	123	4	70	6
Heart failure	119	8	71	5

<https://doi.org/10.1371/journal.pone.0218919.t004>

Discussion

To our knowledge this is the first study that has validated ICD-9 related to the most frequent cardiovascular diseases, namely myocardial infarction, heart failure, and atrial fibrillation, using a large administrative database in the Italian population. Findings from the present study indicate that the assessed ICD-9 codes are highly predictive of myocardial infarction, atrial fibrillation and heart failure. Consequently, the Umbria administrative database can be used to identify the three cardiovascular diseases either as an outcome or as a covariate.

Myocardial infarction

A systematic review published in 2014 [11] identified 30 studies that across 1984–2010 evaluated the validity of ICD-9 or ICD-10 codes related to MI and found that in more than half of the studies, sensitivity and specificity exceeded 81%, and PPV exceeded 92%; at the same time, a large number of studies reported unsatisfactory performance. The review underlined that accuracy might be influenced by the reference standard used. For example, 40% of studies used creatine kinase, lactate dehydrogenase, and aspartate transaminase enzymes as part of MONICA criteria to diagnose MI [31], whereas other studies used cardiac troponin in addition to creatine kinase according to American Heart Association (AHA) criteria, while two studies [31, 32] used the Joint European Society of Cardiology/American College of Cardiology (ESC/ACC) criteria, which is primarily based on troponin levels. The review author concluded that the introduction of cardiac troponin testing in clinical practice may have positively influenced the increase in the PPV of MI codes over time. Cardiac troponins are a more sensitive and specific indicators of myocardial injury with respect to the other enzymes, and its use is supported by many clinical and epidemiological evidence [33–35]. In our assessment, the primary requirement criterion was the increased troponin level in addition to at least one other clinical or instrumental (ECG, echocardiography) resulting in a PPV of 95%. Of the seven false positive cases, five (two case with missing troponin data and three with normal troponin value (single measurement of troponin) had evidence of occlusion or severe stenosis of the coronary vessels that required treatment with PTCA or eluting stents (DES). Two other

Table 5. Reasons for incorrect identification of cases and controls.

	Myocardial infarction	Atrial fibrillation and flutter	Heart failure
FALSE POSITIVES	- Troponin levels not reported with the presence of signs or symptoms, or instrumental confirmation of myocardial infarction (n. 3); - Troponin levels not reported without the presence of signs or symptoms, or instrumental confirmation of myocardial infarction (n. 1); - Troponin levels within the normal range with the presence of signs or symptoms, or instrumental confirmation of myocardial infarction (n. 3).	- Atrial fibrillation in the medical history without the presence in the medical chart of an ECG confirmation (n. 4).	- Signs and symptoms of heart failure without instrumental or laboratory confirmation in the medical chart (n. 7); - Patient with instrumental confirmation of heart failure without signs or symptoms, admitted for implantation of cardiac resynchronization defibrillator (n. 1).
FALSE NEGATIVES	- Patients with altered troponin levels and the presence of symptoms with instrumental confirmation of myocardial infarction (n. 2).	- Patients with an ECG confirmation of atrial fibrillation or flutter (n. 6).	- Patients with signs and symptoms with instrumental confirmation of heart failure (n. 5).

<https://doi.org/10.1371/journal.pone.0218919.t005>

cases had missing troponin data but reported symptoms (retrosternal chest pain radiating to the left arm and the neck ($n = 1$), epigastralgia associated with dyspnoea and sweating ($n = 1$)) and signs in the ECG (3 mm ST segment elevation in V1 and V2 with reciprocal changes in D2 and D3 ($n = 1$); complete bundle branch block with subsequent regression, followed by antero-lateral reciprocal changes and subsequent death due to cardiac arrest ($n = 1$). Myocardial infarction can well be confirmed in these cases and if we had used broader criteria we could have obtained a PPV of 99%.

Heart failure

Regarding heart failure, three reviews systematically summarized the results of the primary studies that validated ICD codes related to this disease [10, 28, 29]. The reviews differed in terms of inclusion criteria and obtained different results. Quach et al [28] identified 25 studies published between 1992 and 2008 that used 70 different ICD codes for defining HF. Accuracy measures were highly variable among the 21 studies validating hospital discharge abstract data, with sensitivity widely varying, ranging from 29% to 89% and PPV ranging from 12% to 100%. Conversely, Saczynski et al [29] identified 35 studies in most of which the PPV was higher than 90%. These higher accuracy measures for HF could be explained by a higher baseline risk of the populations that resulted to be older, especially in the North American databases. The most recent review by McCormick et al [10] identified 19 studies published from 1999–2009, and confirmed a high specificity ($\geq 95\%$) in all studies and a PPV $\geq 87\%$, but the authors concluded that administrative databases are less-than-optimal to identify HF cases, since values of sensitivities were low ($\geq 69\%$ in $\geq 50\%$ of studies). In general, the studies used different types of reference standards including patient self-report, chart reviews by either clinicians or non-clinicians, two distinct disease registries, a single measurement of left ventricular ejection fraction, and the application of several sets of standard diagnostic criteria including the Framingham, Carlson, and European Society of Cardiology. In addition, the studies provided different algorithms to validate their criteria to adjudicate the diagnosis of HF [10].

These include the use of diagnosis in secondary position, outpatient database, the use of prescription claims of drugs such as angiotensin converting enzyme (ACE) inhibitor, angiotensin-II receptor antagonist, loop diuretic, or digoxin, and laboratory data such as BNP. Despite it is not specific to HF, Alqaisi et al [36] suggest that elevated values of BNP (≥ 200 pg/ml) in addition to any hospital discharge diagnosis + at least 2 HF outpatient encounters, improves the ability of ICD-9 codes to identify HF subjects. Our algorithm was based on a combination of signs and symptoms with at least one other conditions that includes abnormality in the ECG, BNP concentration ≥ 35 pg/mL and/or NT-proBNP ≥ 125 pg/mL, or echocardiography abnormalities. These components were derived from the algorithm proposed by the European Society of Cardiology to diagnose HF in the non-acute setting and in the acute setting [24]. As the ICD-9 codes do not distinguish between acute and non-acute settings, we combined the clinical presentation of both settings. The algorithm we used was not proposed by any study included in the systematic reviews but in our setting it provided a high sensitivity and PPV accuracy measures.

Atrial fibrillation/flutter

A systematic review included 16 studies that validated algorithms to identify AF from administrative databases [26]. Methods used to identify AF varied across the studies and included the use of inpatient and/or outpatient data, with or without the use of ECG as an important element to validate the data, and the position of the diagnosis. Most studies used code 427.31 (atrial fibrillation), whereas four studies explicitly included atrial flutter (ICD-9 code 427.32). In terms of accuracy measures, while sensitivity ranged from 57% to 95%, the PPVs from 14 of

the most relevant and comparable studies ranged from 70% to 96%. The PPV resulted lower for incident than for prevalent AF and was affected by characteristics of the algorithm, the database, and the validation criteria.

Strengths and limitations

This study evaluated the validity of the ICD-9-CM codes for cardiovascular diseases using the administrative database of the Umbria Region. For this purpose, we reviewed original source documentation available in medical charts to adjudicate cases of cardiovascular disease.

Our study was based on a pre-published protocol, and no deviation from protocol occurred during study development.

We followed recommended guidelines based on the STARD 2015 [30] for the accurate reporting of investigations of diagnostic studies. Hence, we used detailed and explicit validation criteria, as well as duplicate and independent processes for medical chart review, data abstraction and analysis.

One limitation of our study is that we did not evaluate the accuracy of ICD-9 codes in secondary position. Another limitation of the present study concerns generalizability. In general, results that originate from a healthcare database are immediately applicable only to the setting in which such database has been validated, due to obvious differences that may exist with respect to demographics, ethnicity, disease prevalence, and standards of care, among different populations. In Italy, healthcare provision and management of healthcare databases are regulated at a national level, and therefore we speculate that results from our database, nonetheless, can be generalized to other regions in this country, also because we do not expect major differences in demographics and disease prevalence. However, replication of our validation study in other national settings may allow to explore and refine possible differences in accuracy results, and explore potential factors that might explain heterogeneity.

Conclusion

In administrative databases, the diagnosis of a given disease is associated with a specific ICD code. Despite some limitations, the ICD code is an important tool designed to map health conditions to corresponding general disease categories, along with specific variations. These codes have the advantage of being widely available and require a much lower effort and cost than consulting medical charts [4]. Recently, the Regional Health Authority of Umbria has started a research activity regarding case definitions of several diseases [16, 17, 37, 38], and has validated ICD-9 codes for several oncological diseases [18–21]. Considering the importance of validation processes in public health, in the present study we evaluated for the first time the validity of diagnoses related to ICD-9 codes for important cardiovascular diseases namely AMI, AF/flutter, and HF, using the Regional Health database of Umbria. The case ascertainment for AMI, AF and flutter, and HF, was based on the criteria stated in the best available evidence and showed a high level of all diagnostic accuracy measures ($\geq 90\%$). The healthcare administrative database of Umbria, validated for these diseases, can be confidently used for epidemiological, outcome, and health services research.

Supporting information

S1 Table. STARD-2015-Checklist.
(DOCX)

S1 Dataset. Minimal anonymized dataset.
(PDF)

Author Contributions

Conceptualization: Francesco Cozzolino, Alessandro Montedori, Iosief Abraha, Massimiliano Orso, Giuseppe Ambrosio.

Data curation: Francesco Cozzolino, Iosief Abraha, Massimiliano Orso.

Formal analysis: Francesco Cozzolino, Iosief Abraha, Paolo Eusebi, Massimiliano Orso.

Funding acquisition: Alessandro Montedori, Giuseppe Ambrosio.

Investigation: Francesco Cozzolino, Iosief Abraha, Anna Mengoni, Massimiliano Orso.

Methodology: Francesco Cozzolino, Iosief Abraha, Massimiliano Orso.

Project administration: Alessandro Montedori, Iosief Abraha, Giuseppe Ambrosio.

Resources: Alessandro Montedori, Iosief Abraha, Chiara Grisci, Anna Julia Heymann.

Software: Alessandro Montedori, Giuseppe Ambrosio.

Supervision: Alessandro Montedori, Iosief Abraha, Giuseppe Ambrosio.

Validation: Francesco Cozzolino, Alessandro Montedori, Iosief Abraha, Massimiliano Orso, Giuseppe Ambrosio.

Visualization: Francesco Cozzolino, Iosief Abraha, Anna Julia Heymann, Massimiliano Orso.

Writing – original draft: Francesco Cozzolino, Iosief Abraha, Massimiliano Orso, Giuseppe Ambrosio.

Writing – review & editing: Francesco Cozzolino, Alessandro Montedori, Iosief Abraha, Paolo Eusebi, Chiara Grisci, Anna Julia Heymann, Guido Lombardo, Anna Mengoni, Massimiliano Orso, Giuseppe Ambrosio.

References

1. Abraha I, Orso M, Grilli P, Cozzolino F, Eusebi P, Casucci P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. *International Journal of Statistics in Medical Research*. 2014; 3:309–20.
2. Gavriellov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014; 68(3):283–7. Epub 2013/11/20. <https://doi.org/10.1136/jech-2013-202744> PMID: 24248997.
3. Cozzolino F, Abraha I, Orso M, Mengoni A, Cerasa MF, Eusebi P, et al. Protocol for validating cardiovascular and cerebrovascular ICD-9-CM codes in healthcare administrative databases: the Umbria Data Value Project. *BMJ Open*. 2017; 7(3):e013785. <https://doi.org/10.1136/bmjopen-2016-013785> PMID: 28360241; PubMed Central PMCID: PMC5372118.
4. West SL, Ritchey ME, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. Wiley-Blackwell, 2012:757–94.
5. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe 2014: epidemiological update. *Eur Heart J*. 2014; 35(42):2950–9. Epub 2014/08/21. <https://doi.org/10.1093/eurheartj/ehu299> PMID: 25139896.
6. Farre N, Vela E, Cleries M, Bustins M, Cainzos-Achirica M, Enjuanes C, et al. Real world heart failure epidemiology and outcome: A population-based analysis of 88,195 patients. *PloS one*. 2017; 12(2): e0172745. Epub 2017/02/25. <https://doi.org/10.1371/journal.pone.0172745> PMID: 28235067; PubMed Central PMCID: PMC5325273.
7. Lip GYH, Brechin CM, Lane DA. The global burden of atrial fibrillation and stroke: a systematic review of the epidemiology of atrial fibrillation in regions outside North America and Europe. *Chest*. 2012; 142(6):1489–98. Epub 2012/03/31. <https://doi.org/10.1378/chest.11-2888> PMID: 22459778.
8. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. *Journal of the American College of Cardiology*. 2014; 63(12):1123–33. Epub 2014/02/05. <https://doi.org/10.1016/j.jacc.2013.11.053> PMID: 24491689.

9. McCormick N, Bhole V, Lacaille D, Avina-Zubieta JA. Validity of Diagnostic Codes for Acute Stroke in Administrative Databases: A Systematic Review. *PLoS one*. 2015; 10(8):e0135834. Epub 2015/08/21. <https://doi.org/10.1371/journal.pone.0135834> PMID: 26292280; PubMed Central PMCID: PMC4546158.
10. 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. Epub 2014/08/16. <https://doi.org/10.1371/journal.pone.0104519> PMID: 25126761; PubMed Central PMCID: PMC4134216.
11. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS one*. 2014; 9(3):e92286. Epub 2014/04/01. <https://doi.org/10.1371/journal.pone.0092286> PMID: 24682186; PubMed Central PMCID: PMC3969323.
12. Abraha I, Montedori A, Stracci F, Rossi M, Romagnoli C. Statin compliance in the Umbrian population. *European journal of clinical pharmacology*. 2003; 59(8–9):659–61. <https://doi.org/10.1007/s00228-003-0675-2> PMID: 14508622.
13. Menniti-Ippolito F, Maggini M, Raschetti R, Da Cas R, Traversa G, Walker AM. Ketorolac use in outpatients and gastrointestinal hospitalization: a comparison with other non-steroidal anti-inflammatory drugs in Italy. *European journal of clinical pharmacology*. 1998; 54(5):393–7. PMID: 9754982.
14. Raschetti R, Maggini M, Da Cas R, Popoli P, Rossi A. Time trends in the coprescribing of cisapride and contraindicated drugs in Umbria, Italy. *Jama*. 2001; 285(14):1840–1. Epub 2001/04/20. PMID: 11308395.
15. Traversa G, Bianchi C, Da Cas R, Abraha I, Menniti-Ippolito F, Venegoni M. Cohort study of hepatotoxicity associated with nimesulide and other non-steroidal anti-inflammatory drugs. *Bmj*. 2003; 327(7405):18–22. <https://doi.org/10.1136/bmj.327.7405.18> PMID: 12842950; PubMed Central PMCID: PMC164233.
16. Abraha I, Giovannini G, Serraino D, Fusco M, Montedori A. Validity of breast, lung and colorectal cancer diagnoses in administrative databases: a systematic review protocol. *BMJ Open*. 2016; 6(3):e010409. <https://doi.org/10.1136/bmjopen-2015-010409> PMID: 26993624; PubMed Central PMCID: PMC4800131.
17. Abraha I, Serraino D, Giovannini G, Stracci F, Casucci P, Alessandrini G, et al. Validity of ICD-9-CM codes for breast, lung and colorectal cancers in three Italian administrative healthcare databases: a diagnostic accuracy study protocol. *BMJ Open*. 2016; 6(3):e010547. <https://doi.org/10.1136/bmjopen-2015-010547> PMID: 27016247; PubMed Central PMCID: PMC4809074.
18. Abraha I, Serraino D, Montedori A, Fusco M, Giovannini G, Casucci P, et al. Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study. *BMJ Open*. 2018; 8(7):e020627. <https://doi.org/10.1136/bmjopen-2017-020627> PMID: 30037866; PubMed Central PMCID: PMC6059298.
19. Cozzolino F, Bidoli E, Abraha I, Fusco M, Giovannini G, Casucci P, et al. Accuracy of colorectal cancer ICD-9-CM codes in Italian administrative healthcare databases: a cross-sectional diagnostic study. *BMJ Open*. 2018; 8(7):e020630. <https://doi.org/10.1136/bmjopen-2017-020630> PMID: 29980543.
20. Montedori A, Bidoli E, Serraino D, Fusco M, Giovannini G, Casucci P, et al. Accuracy of lung cancer ICD-9-CM codes in Umbria, Napoli 3 Sud and Friuli Venezia Giulia administrative healthcare databases: a diagnostic accuracy study. *BMJ Open*. 2018; 8(5):e020628. <https://doi.org/10.1136/bmjopen-2017-020628> PMID: 29773701; PubMed Central PMCID: PMC5961589.
21. Orso M, Serraino D, Abraha I, Fusco M, Giovannini G, Casucci P, et al. Validating malignant melanoma ICD-9-CM codes in Umbria, ASL Napoli 3 Sud and Friuli Venezia Giulia administrative healthcare databases: a diagnostic accuracy study. *BMJ Open*. 2018; 8(4):e020631. <https://doi.org/10.1136/bmjopen-2017-020631> PMID: 29678984; PubMed Central PMCID: PMC5914898.
22. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD, et al. Third universal definition of myocardial infarction. *Eur Heart J*. 2012; 33(20):2551–67. <https://doi.org/10.1093/eurheartj/ehs184> PMID: 22922414.
23. Camm AJ, Lip GY, De Caterina R, Savelieva I, Atar D, Hohnloser SH, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *Eur Heart J*. 2012; 33(21):2719–47. <https://doi.org/10.1093/eurheartj/ehs253> PMID: 22922413.
24. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Bohm M, Dickstein K, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2012; 33(14):1787–847. <https://doi.org/10.1093/eurheartj/ehs104> PMID: 22611136.

25. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*. 1927; 22(158):209–12. <https://doi.org/10.1080/01621459.1927.10502953>
26. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. A systematic review of validated methods for identifying atrial fibrillation using administrative data. *Pharmacoepidemiology and drug safety*. 2012; 21 Suppl 1:141–7. Epub 2012/01/25. <https://doi.org/10.1002/pds.2317> PMID: 22262600; PubMed Central PMCID: PMC3674852.
27. Metcalfe A, Neudam A, Forde S, Liu M, Drosler S, Quan H, et al. Case definitions for acute myocardial infarction in administrative databases and their impact on in-hospital mortality rates. *Health Serv Res*. 2013; 48(1):290–318. <https://doi.org/10.1111/j.1475-6773.2012.01440.x> PMID: 22742621; PubMed Central PMCID: PMC3589967.
28. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. *Can J Cardiol*. 2010; 26(8):306–12. [https://doi.org/10.1016/s0828-282x\(10\)70438-4](https://doi.org/10.1016/s0828-282x(10)70438-4) PMID: 20931099; PubMed Central PMCID: PMC2954539.
29. Saczynski JS, Andrade SE, Harrold LR, Tjia J, Cutrona SL, Dodd KS, et al. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiology and drug safety*. 2012; 21 Suppl 1:129–40. <https://doi.org/10.1002/pds.2313> PMID: 22262599; PubMed Central PMCID: PMC3808171.
30. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Bmj*. 2015; 351:h5527. <https://doi.org/10.1136/bmj.h5527> PMID: 26511519; PubMed Central PMCID: PMC4623764.
31. Investigators WMPP. The world health organization monica project (monitoring trends and determinants in cardiovascular disease): A major international collaboration. *Journal of clinical epidemiology*. 1988; 41(2):105–14. [https://doi.org/10.1016/0895-4356\(88\)90084-4](https://doi.org/10.1016/0895-4356(88)90084-4) PMID: 3335877
32. Varas-Lorenzo C, Castellsague J, Stang MR, Tomas L, Aguado J, Perez-Gutthann S. Positive predictive value of ICD-9 codes 410 and 411 in the identification of cases of acute coronary syndromes in the Saskatchewan Hospital automated database. *Pharmacoepidemiology and drug safety*. 2008; 17(8):842–52. <https://doi.org/10.1002/pds.1619> PMID: 18498081
33. Kontos MC, Fritz LM, Anderson FP, Tatum JL, Ornato JP, Jesse RL. Impact of the troponin standard on the prevalence of acute myocardial infarction. *American Heart Journal*. 2003; 146(3):446–52. [https://doi.org/10.1016/S0002-8703\(03\)00245-X](https://doi.org/10.1016/S0002-8703(03)00245-X) PMID: 12947361
34. Salomaa V, for the FSG, Koukkunen H, for the FSG, Ketonen M, for the FSG, et al. A new definition for myocardial infarction: what difference does it make? *European Heart Journal*. 2005; 26(17):1719–25. <https://doi.org/10.1093/eurheartj/ehi185> PMID: 15814567
35. Kavsak PA, MacRae AR, Lustig V, Bhargava R, Vandersluis R, Palomaki GE, et al. The impact of the ESC/ACC redefinition of myocardial infarction and new sensitive troponin assays on the frequency of acute myocardial infarction. *American Heart Journal*. 2006; 152(1):118–25. <https://doi.org/10.1016/j.ahj.2005.09.022> PMID: 16824840
36. Alqaisi F, Williams LK, Peterson EL, Lanfear DE. Comparing methods for identifying patients with heart failure using electronic data sources. *BMC health services research*. 2009; 9(1):237. <https://doi.org/10.1186/1472-6963-9-237> PMID: 20021648
37. Montedori A, Abraha I, Chiatti C, Cozzolino F, Orso M, Luchetta ML, et al. Validity of peptic ulcer disease and upper gastrointestinal bleeding diagnoses in administrative databases: a systematic review protocol. *BMJ Open*. 2016; 6(9):e011776. <https://doi.org/10.1136/bmjopen-2016-011776> PMID: 27633635; PubMed Central PMCID: PMC5030614.
38. Rimland JM, Abraha I, Luchetta ML, Cozzolino F, Orso M, Cherubini A, et al. Validation of chronic obstructive pulmonary disease (COPD) diagnoses in healthcare databases: a systematic review protocol. *BMJ Open*. 2016; 6(6):e011777. <https://doi.org/10.1136/bmjopen-2016-011777> PMID: 27251687; PubMed Central PMCID: PMC4893853.