ORIGINAL RESEARCH

Identification of Incident Atrial Fibrillation From Electronic Medical Records

Alanna M. Chamberlain [ID], PhD, MPH; Véronique L. Roger [ID], MD, MPH; Peter A. Noseworthy [ID], MD; Lin Y. Chen [ID], MD, MS; Susan A. Weston, MS; Ruoxiang Jiang, BS; Alvaro Alonso [ID], MD, PhD

BACKGROUND: Electronic medical records are increasingly used to identify disease cohorts; however, computable phenotypes using electronic medical record data are often unable to distinguish between prevalent and incident cases.

METHODS AND RESULTS: We identified all Olmsted County, Minnesota residents aged ≥18 with a first-ever *International Classification of Diseases, Ninth Revision (ICD-9)* diagnostic code for atrial fibrillation or atrial flutter from 2000 to 2014 (N=6177), and a random sample with an *International Classification of Diseases, Tenth Revision (ICD-10)* code from 2016 to 2018 (N=200). Trained nurse abstractors reviewed all medical records to validate the events and ascertain the date of onset (incidence date). Various algorithms based on number and types of codes (inpatient/outpatient), medications, and procedures were evaluated. Positive predictive value (PPV) and sensitivity of the algorithms were calculated. The lowest PPV was observed for 1 code (64.4%), and the highest PPV was observed for 2 codes (any type) >7 days apart but within 1 year (71.6%). Requiring either 1 inpatient or 2 outpatient codes separated by >7 days but within 1 year had the best balance between PPV (69.9%) and sensitivity (95.5%). PPVs were slightly higher using *ICD-10* codes. Requiring an anticoagulant or antiarrhythmic prescription or electrical cardioversion in addition to diagnostic code(s) modestly improved the PPVs at the expense of large reductions in sensitivity.

CONCLUSIONS: We developed simple, exportable, computable phenotypes for atrial fibrillation using structured electronic medical record data. However, use of diagnostic codes to identify incident atrial fibrillation is prone to some misclassification. Further study is warranted to determine whether more complex phenotypes, including unstructured data sources or using machine learning techniques, may improve the accuracy of identifying incident atrial fibrillation.

Key Words: atrial fibrillation ■ computable phenotype ■ electronic medical records

Atrial fibrillation (AF) is the most common cardiac arrhythmia affecting between 2.7 million and 6.1 million Americans,[1–3] and with the aging population, the prevalence of AF is projected to increase to ≈12 million by the year 2050.[2–4] The outcomes of AF are of major clinical consequence as AF causes substantial morbidity, including a 5-fold increased risk of stroke,[5] a tripling of risk for heart failure,[6–8] a doubling of the risk for dementia,[9] and a nearly 2-fold increased risk of mortality.[7,8,10]

Electronic medical records (EMR) are increasingly used to identify disease cohorts; study trends in incidence of a disease; characterize the demographics,

clinical characteristics, and treatment strategies for a disease cohort; identify risk factors for and outcomes related to a disease; conduct comparative effectiveness research studies; and identify patients for enrollment in pragmatic clinical trials. However, accurate identification of a disease cohort is paramount to ensure valid study results. Several computable phenotypes for AF have been developed using EMR data.[11–15] Some phenotypes for AF were created using cohorts of older populations, included only inpatient diagnoses, included only *International Classification of Diseases, Ninth Revision (ICD-9)* diagnostic codes, or were validated in a small number of cases, limiting the generalizability of previously

CLINICAL PERSPECTIVE

What Is New?

- We developed simple, exportable, computable phenotypes to identify incident atrial fibrillation using structured data from the electronic medical record, including number and type of diagnosis code.
- Addition of prescriptions for anticoagulant or antiarrhythmic drugs or procedure codes for electrical cardioversion procedures did not improve the performance of our computable phenotypes, because more than one-third of the patients with atrial fibrillation did not receive one of these therapies within the year after their first diagnosis of atrial fibrillation.

What Are the Clinical Implications?

- We have developed atrial fibrillation computable phenotypes that are easily exportable and can be implemented using electronic health records, claims data, and large research networks.

Nonstandard Abbreviations and Acronyms

EMR	electronic medical records
REP	Rochester Epidemiology Project

published computable phenotypes for AF. Furthermore, the majority of phenotypes were developed to identify AF cases and controls, and thus are unable to distinguish between prevalent and incident cases. Although identification of prevalent cases may be appropriate for genomic studies, accurate ascertainment of the onset of disease (incidence date) is necessary for many types of epidemiologic studies. Identification of incident AF is crucial for studies on temporal trends in incidence or outcomes of AF, case-control studies investigating risk factors for AF, or cohort studies developing a risk prediction tool for AF. Therefore, we sought to define and validate computable phenotypes for incident AF using structured EMR data from inpatient and outpatient visits over the past 2 decades in a well-characterized, geographically-defined community in Minnesota in which all diagnoses were confirmed by manual chart review.

METHODS

Study Population

Because of the sensitive nature of the data collected for this study, requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be sent to the Rochester Epidemiology Project (REP) at Mayo Clinic at info@rochesterproject.org.

This study used the REP, a medical records-linkage system allowing virtually complete capture of health care provided to residents of Olmsted County, Minnesota.[16–19] The retrieval of nearly all health care related events is possible through the linkage of information from the main health care providers in Olmsted County, which includes Mayo Clinic and its 2 affiliated hospitals and Olmsted Medical Center and its branch offices and affiliated hospital. Data captured include demographics, diagnostic codes, surgical procedure codes, outpatient drug prescriptions, and laboratory results from all participating institutions. In addition, the medical records from all providers in the REP are available to researchers for chart review to validate events, to identify the onset (incidence date) of events, and to abstract other relevant clinical information. Furthermore, demographic and ethnic characteristics of Olmsted County are representative of the state of Minnesota and the Midwest region of the United States, supporting the generalizability of REP data.[17] This study was approved by the Mayo Clinic and Olmsted Medical Center Institutional Review Boards. The study was considered minimal risk by both Institutional Review Boards; therefore, the requirement for informed consent was waived. However, records of any patient who had not provided authorization for their medical records to be used for research, as per Minnesota statute 144.335, were not reviewed.

Validation of Incident AF Diagnoses

An existing community cohort of people diagnosed with incident AF in Olmsted County, Minnesota has been established as part of ongoing surveillance.[20] Briefly, incident AF or atrial flutter from 2000 to 2014 among adults (aged ≥18) was ascertained using *ICD-9* codes 427.31 and 427.32 from all providers in the REP and electrocardiograms (ECGs) from Mayo Clinic. All medical records, including the clinical notes, tests, and procedures from inpatient and outpatient encounters, were reviewed in detail by trained nurse abstractors to validate the events. Evidence of AF or atrial flutter on 1 or more of the following was required to validate the diagnosis of AF: (1) on ECG or rhythm strip, (2) on Holter monitor, event monitor, or telemetry, (3) on monitor during an emergency department visit or hospitalization, (4) on the ECG during an echocardiogram, (5) on pacemaker interrogation, or (6) a physician diagnosis. The source(s) of documentation of AF were recorded and the date of onset (incidence date) was determined based on the documentation in the medical record(s). Consistent with other epidemiologic studies of AF, postoperative AF that occurred within 30 days of a

cardiac surgery (coronary artery bypass graft, surgery to repair or replace a heart valve, surgery to repair an atrial septal defect, or other open heart surgery) was excluded. However, if a postoperative patient with AF went on to experience a future episode of AF unrelated to a surgery (or occurring >30 days after a surgery), this episode of AF was considered incident AF. In addition to ascertainment of all incident AF occurring between 2000 and 2014, we also selected a random sample of 200 people with *International Classification of Diseases, Tenth Revision (ICD-10)* diagnostic code I48 between 2016 and 2018. The medical records of this random sample of 200 people were reviewed to validate the diagnosis as described previously.

AF Computable Phenotype Algorithm Development

Several computable phenotype algorithms were evaluated using combinations of the number (≥1, ≥2) and type (inpatient, outpatient) of diagnostic codes. The diagnostic codes used were *ICD-9* codes 427.31 and 427.32 for the 2000 to 2014 time frame and *ICD-10* code I48 for the 2016 to 2018 time frame. Codes that were preceded by cardiac surgery (such as coronary artery bypass graft, surgery to repair or replace a heart valve, surgery to repair an atrial septal defect, surgery on chordae tendineae, or other open heart surgery) within 30 days before the code were not used in the algorithms. When multiple diagnostic codes were required, different restrictions on timing between diagnostic codes was imposed. For example, some algorithms required ≥2 diagnostic codes that occurred at least 1 day apart whereas others required the ≥2 diagnostic codes to occur >30 days apart. In addition, we tested algorithms that required both diagnostic code(s) along with 1 of the following treatments: outpatient prescription for an anticoagulant drug, outpatient prescription for an antiarrhythmic drug, or electrical cardioversion. These algorithms were employed in a subset of patients starting in 2005 because of the availability of prescription data from the REP (prescription data are not routinely available before 2004). The anticoagulant drugs included warfarin, unfractionated heparin, low molecular weight heparin, and the nonvitamin K antagonist oral anticoagulants dabigatran, rivaroxaban, apixaban, and edoxaban. The antiarrhythmic drugs included amiodarone, disopyramide, dofetilide, dronedarone, flecainide, propafenone, and sotalol. Electrical cardioversions were identified using procedure codes (Current Procedural Terminology codes 92960 and 92961, *ICD-9* procedure codes 99.61 and 99.62, and *ICD-10* procedure code 5A2204Z).

Statistical Analysis

Analyses were performed using SAS, version 9.4 (SAS Institute Inc., Cary, NC). Patients with diagnostic
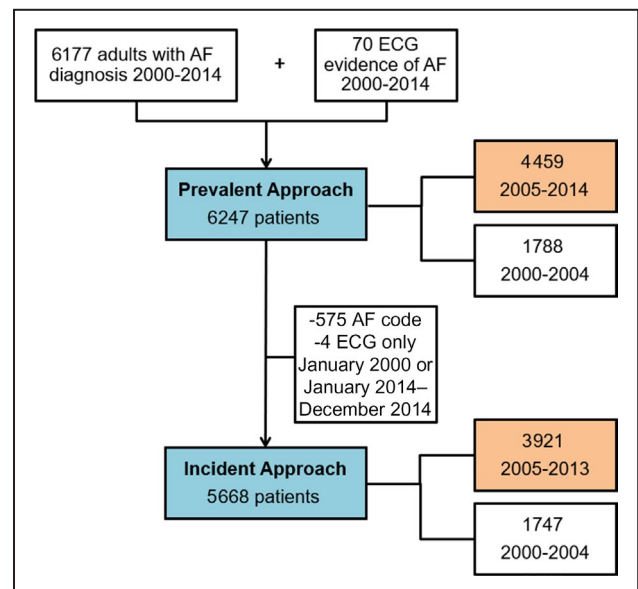

Figure 1. Flow diagram of patients included in the analysis for the prevalent approach and the incident approach.
Number of patients included in the code only phenotypes are shown in the blue boxes. For the computable phenotypes requiring medications and procedures, the analysis was restricted to 2005 forward (number of patients included are shown in orange boxes). AF indicates atrial fibrillation.

code(s) before January 1, 2000 or who died within 30 days of their first diagnostic code were excluded. Between January 1, 2000 and December 31, 2014 there were a total of 6177 patients with at least 1 diagnostic code for AF, and an additional 70 patients with an ECG indicating AF (but without a diagnostic code). These 6247 patients were included in our analyses. For all analyses, we intended to replicate the development of our incident AF cohort where postoperative AF was ignored and a future episode of AF unrelated to a surgery was considered incident AF.

The index dates of the algorithms were defined differently for each type of algorithm. For the algorithms that did not specify the type of diagnostic code, the index date was the first AF code date. For the algorithms that specified either 1 inpatient or 2 outpatient codes, the date of the first outpatient AF code served as the index date if the criterion of 2 outpatient codes was met, the date of the inpatient code served as the index date if the criterion of 1 inpatient code was met, and the earlier of the 2 dates served as the index date if both criteria were met. As described, the records of all patients with a diagnosis code for AF were reviewed to validate the diagnosis, and the date of incident AF was recorded based on review of the medical records.

Two approaches were used to analyze the data (Figure 1). In the first approach (prevalent approach), we included all patients who had a diagnostic code for AF from January 1, 2000 through December 31, 2014.

The positive predictive value (PPV, also referred to as precision) of each computable phenotype was calculated as the proportion of patients identified as having AF from the algorithm who have validated AF (true +/algorithm +). In addition, because electronic interpretations of ECGs were available in patients without a diagnostic code (N=70), we were able to estimate an upper limit of sensitivity (also referred to as recall) of the algorithms, defined as the proportion of validated AF that are identified as having AF from the algorithm (algorithm +/true +). Of note, we ignored any differences in dates between the algorithm date and the date of validated AF. Thus, any algorithm that was met between January 1, 2000 and December 31, 2014 was defined as algorithm +, and people who had validated incident AF during this same time frame were defined as true +.

In the second approach (incident approach), we aimed to determine whether the algorithms accurately identified the date of incident AF. For this analysis, we restricted to patients with a diagnostic code for AF, or an ECG indicating AF without any diagnostic codes, from February 1, 2000 through December 31, 2013 (n=5668). We imposed a window (−30 days to +1 year) around the index date of the algorithm to identify patients who had validated AF (gold standard). Specifically, patients with validated AF with an incidence date within 30 days before to 1 year after the index date determined by the algorithm were considered as true +, whereas those with validated AF that occurred outside this window were not included as true +. The PPV and upper limit of sensitivity were estimated using the formulas previously described. Analyses using the incident approach were repeated stratified by sex and age groups (<65, 65–74, 75–84, ≥85). Finally, analyses using the incident approach were completed for the random sample of 200 patients who had *ICD-10* codes between February 1, 2016 and December 31, 2018. For this analysis, only the PPV was estimated because of the sampling of patients and the lack of information on patients who had AF but did not receive a diagnosis code in this time period.

RESULTS

Using EMR data from the providers captured in the REP, a total of 6177 patients had a diagnostic code for AF and an additional 70 people had an ECG indicating AF without any diagnostic codes between January 1, 2000 and December 31, 2014; thus 6247 people contributed to the analyses (Figure 1). Of the 6247 patients, the median (interquartile range) age was 73 (62–82) years, and 54% were male. Of the 6177 patients with a diagnostic code for AF between January 1, 2000 and December 31, 2014, 42 had a diagnostic code preceded by a cardiac surgery within 30 days before the code with no subsequent diagnostic codes for AF.

Therefore, all algorithms classified them as "no AF." In addition to these 42, another 1759 patients (28%) did not have incident AF during the same time period (993 had no evidence of AF, 269 had isolated postoperative AF, and 497 had AF diagnosed before 2000, which was outside the time window of our study).

Using the prevalent approach, the PPVs ranged from 71.3% for the phenotype requiring 1 code of any type to 83.0% for the phenotype requiring 2 codes separated by more than 30 days but within 1 year (Figure 2; Table 1). However, the algorithm with the highest PPV (2 codes separated by >30 days but within 1 year) had the lowest sensitivity (78.7%). As expected, the sensitivity was highest for the model requiring a single code (98.4%). Requiring a medication or procedure in addition to diagnostic code(s) modestly improved the PPVs at the expense of large reductions in the sensitivity (Figure 2; Table 1).

In an attempt to distinguish between prevalent and incident cases and accurately identify not only the presence of AF but the diagnosis date of AF, we conducted analyses requiring the date of validated AF to occur within 30 days before to 1 year after the index date of the algorithm (incident approach). To allow for the time windows around the index date, the analysis was restricted to individuals with codes for AF (N=5602) or an ECG without any diagnostic codes (N=66) between February 1, 2000 and December 31, 2013 (Figure 1). Of the 5668 patients contributing to the incident analyses, 34 had a diagnostic code preceded by a cardiac surgery (within 30 days prior) with no subsequent codes for AF. Therefore, all algorithms classified these 34 as "no AF." An additional 1982 (35%) patients did not have incident AF within the −30 days to +1 year time window. Records for 100 of these 1982 patients were reviewed. Nearly one-quarter had another type of cardiac dysrhythmia (n=23) and did not have AF despite receiving an AF diagnostic code. Other common reasons for incorrect AF diagnoses included prevalent AF (n=24) and failure to identify presence or date of incident AF after cardiac postoperative AF (n=21).

The PPVs were lower when attempting to more accurately ascertain the date of diagnosis of AF (incident approach; Figure 2; Table 2) compared with the prevalent approach. The lowest PPV was observed for the phenotype requiring 1 code (64.4%), and the highest PPV was observed for the phenotype requiring 2 codes (any type) >7 days apart but within 1 year (71.6%). The phenotype requiring either 1 inpatient code or 2 outpatient codes separated by >7 days but within 1 year appeared to have the best balance between PPV (69.9%) and sensitivity (95.5%). The sensitivities were quite similar between the incident and prevalent approaches. In addition, the PPVs and sensitivities were similar for men and women and across all age categories except for the phenotypes requiring a medication
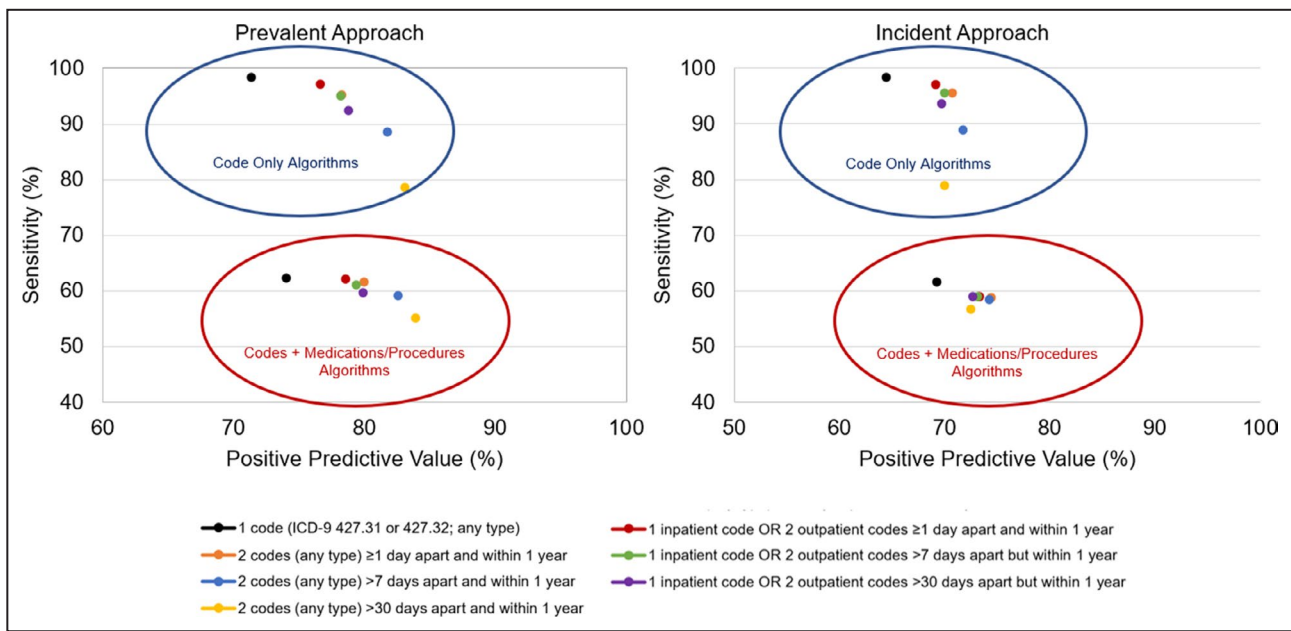
Figure 2. Positive predictive value and upper limit of sensitivity of computable phenotypes for atrial fibrillation using the prevalent approach (left panel) and incident approach (right panel).

The sensitivity reported is an upper limit of the true sensitivity of the algorithms and was estimated including patients with electrocardiographic evidence of atrial fibrillation without a diagnostic code. The study period for the computable phenotypes including medications and procedures was January 1, 2005 to December 31, 2014 for the prevalent approach and January 1, 2005 to December 31, 2013 for the incident approach. *ICD-9* indicates *International Classification of Diseases, Ninth Revision*.

or procedure in addition to the diagnostic code(s), for which the sensitivities were lower in women compared with men. Finally, the PPVs were slightly higher for the computable phenotypes using *ICD-10* codes (Table 3). For example, the PPV for the phenotype requiring 1 code was 67.5% (compared with 64.4% using *ICD-9* codes), and the PPV for the phenotype requiring 2 codes (any type) >7 days apart but within 1 year was 74.1% (compared with 71.6% using *ICD-9* codes).

To better understand how well the algorithms identified the incidence date of AF, we compared the algorithm date and the validated incidence date of AF for the computable phenotype of 1 inpatient code or 2 outpatient codes separated by >7 days but within 1 year. Out of the 4936 patients meeting the algorithm definition, 3860 had validated incident AF. Of these 3860 with validated AF, 3451 (89%) occurred within our required time window (the date of validated AF had to occur within 30 days before to 1 year after the index date of the algorithm) with 1942 (50.3%) having the same date for the algorithm and the validated AF incidence date. The median (Q1, Q3) time between the algorithm date and validated date was 0 (−3, 0) days.

DISCUSSION

In this large study spanning decades, we observed that use of diagnostic codes from the EMR to identify

AF is prone to some misclassification. We developed computable phenotypes with varying number and type of diagnoses as well as inclusion of AF treatment. The use of different time restrictions on the phenotypes requiring at least 2 diagnostic codes was employed to test varying degrees of certainty that the codes were assigned during separate encounters as an attempt to rule out false positive diagnoses. We observed that the best balance of PPV and sensitivity was achieved for the phenotype requiring 1 inpatient code or 2 outpatient codes separated by >7 days but within 1 year. Requiring a medication or procedure in addition to diagnostic code(s) modestly improved the PPVs at the expense of reductions in the sensitivity of the computable phenotypes, indicating that a large proportion of patients who have AF were missed when these additional data elements were required. As indicated by the low sensitivities, more than one-third of the patients with AF did not receive an anticoagulant or antiarrhythmic drug or an electrical cardioversion procedure within the year after their first diagnosis of AF. In addition, the sensitivities of the phenotypes incorporating AF treatment were lower in women compared with men, indicating that women were less likely to be treated with an anticoagulant drug, an antiarrhythmic drug, or an electrical cardioversion procedure. Nevertheless, the choice of the most appropriate computable phenotype may vary depending on the goal of

Table 1. Positive Predictive Value and Upper Limit of Sensitivity* of Various Computable Phenotypes for Atrial Fibrillation Using the Prevalent Approach, January 1, 2000 to December 31, 2014

Computable phenotype	Code only algorithms				Codes + medications/procedures†			
	PPV		Sensitivity		PPV		Sensitivity	
	True+ / algorithm+	%	Algorithm+ / true+	%	True+ / algorithm+	%	Algorithm+ / true+	%
1 code (*International Classification of Diseases, Ninth Revision* 427.31 or 427.32; any type)	4376/6135	71.3	4376/4446	98.4	1958/2645	74.0	1958/3143	62.3
2 codes (any type) ≥1 d apart and within 1 y	4237/5419	78.2	4237/4446	95.3	1936/2422	79.9	1936/3143	61.6
2 codes (any type) >7 d apart and within 1 y	3933/4814	81.7	3933/4446	88.5	1862/2258	82.5	1862/3143	59.2
2 codes (any type) >30 d apart and within 1 y	3499/4216	83.0	3499/4446	78.7	1737/2074	83.8	1737/3143	55.3
1 inpatient code OR 2 outpatient codes ≥1 d apart and within 1 y	4316/5638	76.6	4316/4446	97.1	1949/2484	78.5	1949/3143	62.0
1 inpatient code OR 2 outpatient codes >7 d apart but within 1 y	4223/5407	78.1	4223/4446	95.0	1920/2420	79.3	1920/3143	61.1
1 inpatient code OR 2 outpatient codes >30 d apart but within 1 y	4112/5226	78.7	4112/4446	92.5	1880/2355	79.8	1880/3143	59.8

PPV indicates positive predictive value.

*The sensitivity was estimated including patients with ECG evidence of atrial fibrillation without a diagnostic code. However, patients with true atrial fibrillation without an ECG or diagnostic code may have been missed, resulting in an overestimate of the sensitivity. Thus, our sensitivity estimates are an upper limit of the true sensitivity of the algorithms.

†The study period for this analysis was January 1, 2005 to December 31, 2014 (N=4459; 4430 with codes and 29 with ECG only). The algorithms incorporating both codes and medications/procedures require the patient to have the listed number and type of diagnostic codes plus a prescription for an anticoagulant drug or an antiarrhythmic drug or an electrical cardioversion procedure within 1 year of the first diagnostic code.

Table 2. Positive Predictive Value and Upper Limit of Sensitivity* of Various Computable Phenotypes for Atrial Fibrillation Using the Incident Approach, February 1, 2000 to December 31, 2013

Computable phenotype	Code only algorithms				Codes + medications/procedures†			
	PPV		Sensitivity		PPV		Sensitivity	
	True+ / algorithm+	%	Algorithm+ / true+	%	True+ / algorithm+	%	Algorithm+ / true+	%
1 code (*International Classification of Diseases, Ninth Revision* 427.31 or 427.32; any type)	3586/5568	64.4	3586/3652	98.2	1581/2288	69.1	1581/2572	61.5
Sex								
Male	1919/2976	64.5	1919/1954	98.2	893/1302	68.6	893/1398	63.9
Female	1667/2592	64.3	1667/1698	98.2	688/986	69.8	688/1174	58.6
Age group, y								
Age <65	985/1626	60.6	985/1011	97.4	431/661	65.2	431/728	59.2
Age 65–74	848/1314	64.5	848/861	98.5	395/573	68.9	395/615	64.2
Age 75–84	1052/1608	65.4	1052/1064	98.9	519/720	72.1	519/719	72.2
Age ≥85	701/1020	68.7	701/716	97.9	236/334	70.7	236/510	46.3
2 codes (any type) ≥1 d apart and within 1 y	3480/4929	70.6	3480/3645	95.5	1515/2036	74.4	1515/2578	58.8
Sex								
Male	1865/2627	71.0	1865/1956	95.3	860/1157	74.3	860/1407	61.1
Female	1615/2302	70.2	1615/1689	95.6	655/879	74.5	655/1171	55.9
Age group								
Age <65	937/1349	69.5	937/997	94.0	421/575	73.2	421/718	58.6
Age 65–74	834/1160	71.9	834/866	96.3	382/509	75.0	382/620	61.6
Age 75–84	1027/1470	69.9	1027/1062	96.7	491/651	75.4	491/724	67.8
Age ≥85	682/950	71.8	682/720	94.7	221/301	73.4	221/516	42.8
2 codes (any type) >7 d apart and within 1 y	3160/4411	71.6	3160/3554	88.9	1485/2004	74.1	1485/2536	58.6
Sex								
Male	1681/2358	71.3	1681/1908	88.1	837/1137	73.6	837/1385	60.4
Female	1479/2053	72.0	1479/1646	89.9	648/867	74.7	648/1151	56.3
Age group, y								
Age <65	802/1131	70.9	802/964	83.2	397/539	73.7	397/697	57.0
Age 65–74	764/1053	72.6	764/838	91.2	373/503	74.2	373/608	61.3
Age 75–84	962/1346	71.5	962/1038	92.7	491/659	74.5	491/715	68.7
Age ≥85	632/881	71.7	632/714	88.5	224/303	73.9	224/516	43.4
2 codes (any type) >30 d apart and within 1 y	2732/3909	69.9	2732/3461	78.9	1411/1949	72.4	1411/2490	56.7

(Continued)

Table 2. Continued

Computable phenotype	Code only algorithms				Codes + medications/procedures†			
	PPV		Sensitivity		PPV		Sensitivity	
	True+ / algorithm+	%	Algorithm+ / true+	%	True+ / algorithm+	%	Algorithm+ / true+	%
Sex								
Male	1459/2100	69.5	1459/1858	78.5	793/1099	72.2	793/1362	58.2
Female	1273/1809	70.4	1273/1603	79.4	618/850	72.7	618/1128	54.8
Age group								
Age <65	665/965	68.9	665/931	71.4	370/510	72.5	370/684	54.1
Age 65–74	680/947	71.8	680/818	83.1	360/500	72.0	360/600	60.0
Age 75–84	860/1219	70.6	860/1012	85.0	466/643	72.5	466/700	66.6
Age ≥85	527/778	67.7	527/700	75.3	215/296	72.6	215/506	42.5
1 inpatient code OR 2 outpatient codes ≥1 d apart and within 1 y	3544/5125	69.2	3544/3645	97.2	1520/2073	73.3	1520/2576	59.0
Sex								
Male	1901/2735	69.5	1901/1955	97.2	863/1182	73.0	863/1403	61.5
Female	1643/2390	68.7	1643/1690	97.2	657/891	73.7	657/1173	56.0
Age group								
Age <65	960/1416	67.8	960/998	96.2	425/590	72.0	425/719	59.1
Age 65–74	845/1205	70.1	845/866	97.6	385/520	74.0	385/620	62.1
Age 75–84	1042/1518	68.6	1042/1062	98.1	489/657	74.4	489/723	67.6
Age ≥85	697/986	70.7	697/719	96.9	221/306	72.2	221/514	43.0
1 inpatient code OR 2 outpatient codes >7 d apart but within 1 y	3451/4936	69.9	3451/3614	95.5	1509/2066	73.0	1509/2559	59.0
Sex								
Male	1847/2636	70.1	1847/1943	95.1	853/1173	72.7	853/1399	61.0
Female	1604/2300	69.7	1604/1671	96.0	656/893	73.5	656/1160	56.6
Age group								
Age <65	901/1314	68.6	901/981	91.8	414/576	71.9	414/708	58.5
Age 65–74	829/1168	71.0	829/859	96.5	383/523	73.2	383/618	62.0
Age 75–84	1026/1478	69.4	1026/1055	97.3	490/662	74.0	490/719	68.2
Age ≥85	695/976	71.2	695/719	96.7	222/305	72.8	222/514	43.2
1 inpatient code OR 2 outpatient codes >30 d apart but within 1 y	3348/4801	69.7	3348/3578	93.6	1497/2058	72.7	1497/2543	58.9
Sex								
Male	1784/2569	69.4	1784/1917	93.1	843/1169	72.1	843/1387	60.8

(Continued)

Table 2. Continued

Computable phenotype	Code only algorithms				Codes + medications/procedures†			
	PPV		Sensitivity		PPV		Sensitivity	
	True+ / algorithm+	%	Algorithm+ / true+	%	True+ / algorithm+	%	Algorithm+ / true+	%
Female	1564/2232	70.1	1564/1661	94.2	654/889	73.6	654/1156	56.6
Age group								
Age <65	846/1244	68.0	846/962	87.9	405/565	71.7	405/700	57.9
Age 65–74	810/1146	70.7	810/850	95.3	381/525	72.6	381/614	62.1
Age 75–84	1005/1445	69.6	1005/1048	95.9	488/664	73.5	488/717	68.1
Age ≥85	687/966	71.1	687/718	95.7	223/304	73.4	223/512	43.6

PPV indicates positive predictive value.

*The sensitivity was estimated including patients with ECG evidence of atrial fibrillation without a diagnostic code. However, patients with true atrial fibrillation without an ECG or diagnostic code may have been missed, resulting in an overestimate of the sensitivity. Thus, our sensitivity estimates are an upper limit of the true sensitivity of the algorithms.

†The study period for this analysis was January 1, 2005 to December 31, 2013 (N=3921; 3894 with codes and 27 ECG only). The algorithms incorporating both codes and medications/procedures require the patient to have the listed number and type of diagnostic codes plus a prescription for an anticoagulant drug or an antiarrhythmic drug or an electrical cardioversion procedure within 1 year of the first diagnostic code.

a research study. For example, surveillance studies or studies aiming to develop an early detection system may opt for a computable phenotype with the highest sensitivity, whereas a computable phenotype with higher PPV may be more appropriate for identification of patients for a survey study or a clinical trial.

A distinct advantage of our study was the medical record review to confirm diagnosis of AF as well as to identify the date of onset (incidence date). We observed that the computable phenotypes had lower PPVs when attempting to more accurately ascertain the date of diagnosis of AF (incident approach) compared with the prevalent approach. These findings have important implications when interpreting the performance of published computable phenotypes, which are often developed to identify prevalent cases. For example, the Electronic MEdical Records and G nomics Network has developed more than 50 computable phenotypes to identify cases and controls for genome-wide association studies, and these computable phenotypes have been published on the Phenotype Knowledgebase allowing other investigators to reuse the published algorithms.[13,21] Accurate identification of cases and controls is appropriate for genomic studies. However, the ability to distinguish between prevalent and incident cases and to accurately identify the diagnosis date of a disease is often essential for epidemiologic studies. For example, identification of incident AF is needed to study temporal trends in incidence or outcomes of AF, for case-control studies investigating risk factors for AF, and for cohort studies developing a risk prediction tool for AF. Nevertheless, the majority of published computable phenotypes for AF have been developed to identify prevalent AF[11–13,15] or were developed to identify incident AF but lacked record review to validate the diagnoses and assess algorithm performance.[14]

The simplicity of our computable phenotypes for AF, which include only number and type of diagnosis code, is another advantage of our study. Our AF phenotypes are easily exportable and can be implemented using electronic health records, claims data, and large research networks such as the National Patient-Centered Clinical Research Network, allowing large-scale national surveillance.[22–24] Although inclusion of more than 1 type of structured data element may generally increase the ability to correctly identify a disease of interest,[14,25,26] we found that for AF, inclusion of prescriptions for anticoagulant or antiarrhythmic drugs or procedure codes for electrical cardioversion procedures did not improve the performance of our computable phenotypes. Nevertheless, more complex phenotypes that include natural language processing of the unstructured text of the medical record, or machine learning approaches applied to the EMR, could be developed to improve the accuracy of identifying

Table 3. Positive Predictive Value of Various Computable Phenotypes for Atrial Fibrillation Using *ICD-10* codes, February 1, 2016 to December 31, 2018

Computable phenotype	Code only algorithms		Codes + medications/procedures	
	PPV		PPV	
	True+ / algorithm+	%	True+ / algorithm+	%
1 code (*ICD-10* I48; any type)	135/200	67.5	99/143	69.2
2 codes (any type) ≥1 d apart and within 1 y	130/180	72.2	97/133	72.9
2 codes (any type) >7 d apart and within 1 y	120/162	74.1	94/125	75.2
2 codes (any type) >30 d apart and within 1 y	112/148	75.7	89/117	76.1
1 inpatient code OR 2 outpatient codes ≥1 d apart and within 1 y	133/187	71.1	98/138	71.0
1 inpatient code OR 2 outpatient codes >7 d apart but within 1 y	131/184	71.2	97/136	71.3
1 inpatient code OR 2 outpatient codes >30 d apart but within 1 y	125/175	71.4	93/130	71.5

ICD-10 indicates *International Classification of Diseases, Tenth Revision*; PPV, positive predictive value.

incident AF.[26–28] However, these complex phenotypes may be more difficult to implement across large data networks or in nonacademic settings, or may not improve the ability to identify incident AF over simpler approaches.[29]

Limitations and Strengths

We acknowledge the following limitations. AF is often asymptomatic and may go undiagnosed for some time before coming to clinical attention. Our algorithms were developed to identify the date of diagnosis of AF but may not accurately identify the true onset of disease. The performance of the algorithms may differ across medical systems due to differences in coding practices, patient populations, or completeness of data capture. We used a medical records-linkage system (the REP) that captures the majority of medical care delivered to local residents. Thus, when the computable phenotypes are implemented in other medical systems or data sources such as claims, the PPVs and sensitivities may be lower than reported herein. We did not have an external data set with validated AF events to estimate how the performance of the algorithms may differ across data sources. Finally, due to the limited geographic and racial diversity of our local population, our results may not be generalizable to all patient populations.

Our study also has a number of strengths, including the large sample size and inclusion of all people with a diagnosis code for AF over a 15-year period. In addition, we included a random sample of patients with *ICD-10* diagnosis codes to allow for comparison of the PPVs over time. The records of all patients were reviewed to manually validate all AF events and to ascertain the date of first diagnosis of AF (incidence date). As a result, we created computable phenotypes not only to identify prevalent AF but also to identify incident AF.

We also had the ability to identify patients with ECG evidence of AF without a diagnosis code in order to estimate an upper limit of sensitivity of the computable phenotypes. Finally, all computable phenotypes included structured data and are easy to implement in other research settings.

CONCLUSIONS

We developed simple, exportable, computable phenotypes to identify incident AF using structured data from the EMR. However, use of diagnostic codes to identify incident diagnoses of AF remains prone to some misclassification. Creation and validation of more complex algorithms including unstructured data sources, such as natural language processing of the medical record notes, or machine learning approaches applied to the EMR, may improve the accuracy of identifying AF and warrant further study.

REFERENCES

1. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA*. 2001;285:2370–2375. doi: 10.1001/jama.285.18.2370

2. Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, Seward JB, Tsang TSM. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation*. 2006;114:119–125. doi: 10.1161/CIRCULATIONAHA.105.595140

3. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. *Am J Cardiol*. 2013;112:1142–1147. doi: 10.1016/j.amjcard.2013.05.063

4. Naccarelli GV, Varker H, Lin J, Schulman KL. Increasing prevalence of atrial fibrillation and flutter in the United States. *Am J Cardiol*. 2009;104:1534–1539. doi: 10.1016/j.amjcard.2009.07.022

5. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke*. 1991;22:983–988. doi: 10.1161/01.STR.22.8.983

6. Wang TJ, Larson MG, Levy D, Vasan RS, Leip EP, Wolf PA, D'Agostino RB, Murabito JM, Kannel WB, Benjamin EJ. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. *Circulation*. 2003;107:2920–2925. doi: 10.1161/01.CIR.0000072767.89944.6E

7. Stewart S, Hart CL, Hole DJ, McMurray JJ. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. *Am J Med*. 2002;113:359–364. doi: 10.1016/S0002-9343(02)01236-6

8. Krahn AD, Manfreda J, Tate RB, Mathewson FA, Cuddy TE. The natural history of atrial fibrillation: incidence, risk factors, and prognosis in the Manitoba follow-up study. *Am J Med*. 1995;98:476–484. doi: 10.1016/S0002-9343(99)80348-9

9. Ott A, Breteler MM, de Bruyne MC, van Harskamp F, Grobbee DE, Hofman A. Atrial fibrillation and dementia in a population-based study. The Rotterdam Study. *Stroke*. 1997;28:316–321. doi: 10.1161/01.STR.28.2.316

10. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation*. 1998;98:946–952. doi: 10.1161/01.CIR.98.10.946

11. 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. *Pharmacoepidemiol Drug Saf*. 2012;21:141–147.

12. Khurshid S, Keaney J, Ellinor PT, Lubitz SA. A simple and portable algorithm for identifying atrial fibrillation in the electronic medical record. *Am J Cardiol*. 2016;117:221–225. doi: 10.1016/j.amjcard.2015.10.031

13. Kirby JC, Speltz P, Rasmussen LV, Basford M, Gottesman O, Peissig PL, Pacheco JA, Tromp G, Pathak J, Carrell DS, et al. PheKB: a catalog and workflow for creating electronic phenotype algorithms for

14. Morley KI, Wallace J, Denaxas SC, Hunter RJ, Patel RS, Perel P, Shah AD, Timmis AD, Schilling RJ, Hemingway H. Defining disease phenotypes using national linked electronic health records: a case study of atrial fibrillation. *PLoS One*. 2014;9:e110900. doi: 10.1371/journal.pone.0110900

15. Wang SV, Rogers JR, Jin Y, Bates DW, Fischer MA. Use of electronic healthcare records to identify complex patients with atrial fibrillation for targeted intervention. *J Am Med Inform Assoc*. 2017;24:339–344. doi: 10.1093/jamia/ocw082

16. Rocca WA, Yawn BP, St Sauver JL, Grossardt BR, Melton LJ III. History of the Rochester Epidemiology Project: half a century of medical records linkage in a US population. *Mayo Clin Proc*. 2012;87:1202–1213. doi: 10.1016/j.mayocp.2012.08.012

17. St. Sauver JL, Grossardt BR, Leibson CL, Yawn BP, Melton LJ III, Rocca WA. Generalizability of epidemiological findings and public health decisions: an illustration from the Rochester Epidemiology Project. *Mayo Clin Proc*. 2012;87:151–160. doi: 10.1016/j.mayocp.2011.11.009

18. St. Sauver JL, Grossardt BR, Yawn BP, Melton LJ, Rocca WA. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester Epidemiology Project. *Am J Epidemiol*. 2011;173:1059–1068. doi: 10.1093/aje/kwq482

19. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ, Pankratz JJ, Brue SM, Rocca WA. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. *Int J Epidemiol*. 2012;41:1614–1624. doi: 10.1093/ije/dys195

20. Chamberlain AM, Gersh BJ, Alonso A, Chen LY, Berardi C, Manemann SM, Killian JM, Weston SA, Roger VL. Decade-long trends in atrial fibrillation incidence and survival: a community study. *Am J Med*. 2015;128:260–267. doi: 10.1016/j.amjmed.2014.10.030

21. Crawford DC, Crosslin DR, Tromp G, Kullo IJ, Kuivaniemi H, Hayes MG, Denny JC, Bush WS, Haines JL, Roden DM, et al. eMERGEing progress in genomics—the first seven years. *Front Genet*. 2014;5.184. doi: 10.3389/fgene.2014.00184

22. Forrest CB, McTigue KM, Hernandez AF, Cohen LW, Cruz H, Haynes K, Kaushal R, Kho AN, Marsolo KA, Nair VP, et al. PCORnet® 2020: current state, accomplishments, and future directions. *J Clin Epidemiol*. 2021;129:60–67. doi: 10.1016/j.jclinepi.2020.09.036

23. Pletcher MJ, Fontil V, Carton T, Shaw KM, Smith M, Choi S, Todd J, Chamberlain AM, O'Brien EC, Faulkner M, et al. The PCORnet Blood Pressure Control Laboratory: a platform for surveillance and efficient trials. *Circ Cardiovasc Qual Outcomes*. 2020;13:e006115. doi: 10.1161/CIRCOUTCOMES.119.006115

24. Roger VL, Sidney S, Fairchild AL, Howard VJ, Labarthe DR, Shay CM, Tiner AC, Whitsel LP, Rosamond WD, American Heart Association Advocacy Coordinating Committee. Recommendations for cardiovascular health and disease surveillance for 2030 and beyond: a policy statement from the American Heart Association. *Circulation*. 2020;141:e104–e119. doi: 10.1161/CIR.0000000000000756

25. Denaxas S, Gonzalez-Izquierdo A, Direk K, Fitzpatrick NK, Fatemifar G, Banerjee A, Dobson RJB, Howe LJ, Kuan V, Lumbers RT, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. *J Am Med Inform Assoc*. 2019;26:1545–1559. doi: 10.1093/jamia/ocz105

26. Wei WQ, Denny JC. Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome Med*. 2015;7:41. doi: 10.1186/s13073-015-0166-y

27. Banda JM, Seneviratne M, Hernandez-Boussard T, Shah NH. Advances in electronic phenotyping: from rule-based definitions to machine learning models. *Annu Rev Biomed Data Sci*. 2018;1:53–68. doi: 10.1146/annurev-biodatasci-080917-013315

28. Richesson RL, Sun J, Pathak J, Kho AN, Denny JC. Clinical phenotyping in selected national networks: demonstrating the need for high-throughput, portable, and computational methods. *Artif Intell Med*. 2016;71:57–61. doi: 10.1016/j.artmed.2016.05.005

29. Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a machine learning model applied to harmonized electronic health record data for the prediction of incident atrial fibrillation. *JAMA Netw Open*. 2020;3:e1919396. doi: 10.1001/jamanetworkopen.2019.19396

transportability. *J Am Med Inform Assoc*. 2016;23:1046–1052. doi: 10.1093/jamia/ocv202