



Published in final edited form as:

Pharmacoepidemiol Drug Saf. 2012 January ; 21(0 1): 141–147. doi:10.1002/pds.2317.

Identifying atrial fibrillation from electronic medical data: a systematic review

Paul N. Jensen, BA^{1,2}, Karin Johnson, PhD³, James Floyd, MD^{1,2}, Susan R. Heckbert, MD, PhD^{1,2,3}, Ryan Carnahan, PharmD, MS⁴, and Sascha Dublin, MD, PhD^{2,3}

¹Cardiovascular Health Research Unit, University of Washington, Seattle, Washington

²Department of Epidemiology, University of Washington, Seattle, Washington

³Group Health Research Institute, Seattle, Washington

⁴University of Iowa College of Public Health, Department of Epidemiology, Iowa City, Iowa

Abstract

Purpose—To characterize the validity of algorithms to identify AF from electronic health data through a systematic review of the literature, and to identify gaps needing further research.

Methods—Two reviewers examined publications during 1997–2008 that identified patients with AF from electronic health data and provided validation information. We abstracted information including algorithm sensitivity, specificity, and positive predictive value (PPV).

Results—We reviewed 544 abstracts and 281 full-text articles, of which 18 provided validation information from 16 unique studies. Most used data from before 2000, and 10 of 16 used only inpatient data. Three studies incorporated electronic ECG data for case identification or validation. A large proportion of prevalent AF cases identified by ICD-9 code 427.31 were valid (PPV 70–96%, median 89%). Seven studies reported algorithm sensitivity (range, 57–95%; median 79%). One study validated an algorithm for incident AF and reported a PPV of 77%.

Conclusions—The ICD-9 code 427.31 performed relatively well, but conclusions about algorithm validity are hindered by few recent data, use of nonrepresentative populations, and a disproportionate focus on inpatient data. An optimal contemporary algorithm would likely draw on inpatient and outpatient codes and electronic ECG data. Additional research is needed in representative, contemporary populations regarding algorithms that identify incident AF and incorporate electronic ECG data.

Keywords

atrial fibrillation; cardiac arrhythmia; epidemiology; validation; positive predictive value; sensitivity

INTRODUCTION

Atrial fibrillation (AF) is an increasingly common arrhythmia that increases the risk of stroke and death.^{1,2} Electronic health data can be used to study the epidemiology of AF, assess quality of care, and monitor for AF as an adverse event related to newly approved

Correspondence: Paul Jensen, Cardiovascular Health Research Unit, 1730 Minor Avenue, Suite 1360, Seattle, WA 98101. Fax 206-287-2662; phone 206-287-2777; pjensen@uw.edu.

Conflict of interests: Dr. Dublin received a Merck/American Geriatrics Society New Investigator Award. No other authors reported conflicts of interest. This manuscript was derived from a longer report which is published on the Mini-Sentinel website.

medications. The latter is of interest to the Food and Drug Administration (FDA), which through its Sentinel Initiative seeks to monitor prospectively the safety of medical products using electronic health data for over 100 million people.³

These efforts depend on the accuracy of electronic health data to identify AF. To date, there has been no systematic review of the validity of algorithms used to identify AF. The primary objective of this project was to review all studies that have validated algorithms to identify AF from electronic health data, summarize their results, and identify gaps in knowledge for future research.

METHODS

Detailed methods of this systematic review can be found in the accompanying manuscript. (reference Carnahan methods paper) Briefly, as part of the FDA's Mini-Sentinel pilot, systematic reviews including this one were conducted for 20 health outcomes. The search strategy was developed by investigators at University of Iowa in collaboration with the FDA, based on prior work by the Observational Medical Outcomes Partnership.⁴ Details of the search strategy are available online at the Mini-Sentinel website. PubMed was searched on May 14, 2010 and July 6, 2010 and the Iowa Drug Information Service (IDIS) database on June 11, 2010. Mini-Sentinel collaborators were asked to identify relevant unpublished studies or other validation studies of which they were aware.

Two authors independently reviewed each abstract. Articles were selected for full review if AF was studied using electronic health data from the United States or Canada. If the reviewers disagreed or there was insufficient information, the article was selected for full-text review.

Two authors independently reviewed articles to identify validation studies described in the article or its references. Articles were excluded if they did not meet abstract inclusion criteria (above), if the algorithm for identifying AF was inadequately described, or if validity statistics were not provided or could not be calculated. If there was disagreement, reviewers attempted to reach consensus, and if they could not, a third author was consulted. One author (PJ) extracted information from articles for the evidence tables, and a second author (SD) checked it for accuracy.

Inter-rater agreement about inclusion of abstracts and articles was calculated using Cohen's kappa.⁵

RESULTS

Search Strategy

PubMed searches identified 527 citations and IDIS searches 49, for a total of 544 unique citations. Of these, 249 were selected for full-text review (Figure). Cohen's kappa for agreement between reviewers regarding article selection for full-text review was 0.62 (95% confidence interval [CI] 0.55–0.69). A total of 281 full text articles were reviewed: 249 from the original search and 32 more identified mainly from these articles' references. From the 281 articles, we identified 18 reporting 16 unique validation studies. Cohen's kappa for agreement on the presence of a validation study meeting inclusion criteria before consensus discussions was 0.83 (95% CI 0.68–0.98).

Summary of algorithms

The methods and data sources used to identify AF varied across studies. Of the 16 unique studies, ten used only inpatient data,^{6–15} two used only outpatient data,^{16,17} and four used

both.^{18–21} Two studies described algorithms incorporating electronic ECG data in addition to administrative data but did not actually validate these algorithms.^{17,21} Studies differed according to which ICD-9 codes were used, which coding positions were searched, and the number of codes required; algorithm details are provided in Tables 1 and 2. Most studies used the code 427.31 (atrial fibrillation). Four studies explicitly included atrial flutter (ICD-9 code 427.32).^{11,12,19,20} Three others^{15,16,18} probably included some cases of atrial flutter because they used a four-digit ICD-9 code, 427.3.

One study specifically validated an algorithm for incident AF.^{6,20} To ensure that people were free of AF prior to the event of interest, this study excluded people with any prior ICD-9 code for AF or atrial flutter during their HMO enrollment. Another study⁶ defined incident AF as the first study ECG showing AF or a hospital ICD-9 code for AF during follow-up and excluded people whose baseline study ECG showed AF, but the study lacked other information about AF history before baseline. Because information about prevalent AF was very limited, we did not classify this study as a relevant validation study of incident AF.

Two studies used a population or gold standard that made their findings less relevant to the aims of this review. Go et al.²² reviewed medical records to validate the presence of AF in a population not likely to have AF. Yuan et al.¹⁶ used as a gold standard a hospital database of diagnosis codes, rather than medical record review. Their results shed light on how the number of diagnosis codes in a database impacts sensitivity (not surprisingly, databases retaining a limited number of codes per hospitalization have lower sensitivity than those with a larger number) but are not informative about the actual validity of AF diagnosis codes. These studies are described in Tables 1 and 2, but their results are omitted from the summary statistics for PPV and sensitivity.

14 of 16 studies validated the diagnosis of AF by medical record review, including four studies^{6–8,11} in which a physician reviewed ECGs from the medical record. The definition used for the gold standard varied markedly, ranging from lax (any mention of AF in the chart) to strict (requiring that the chart contain an ECG or rhythm strip showing AF.) One study¹⁵ used as a gold standard a teaching hospital database with up to 27 diagnosis codes for each hospitalization (versus 5 in the primary database). The remaining study¹⁸ provided information from which we could calculate the sensitivity of their algorithm compared to their institution's electronic ECG database.

Summary of validity statistics and factors influencing algorithm validity

Validity statistics for individual studies are shown in Tables 1 and 2. We included in these summary measures all PPVs and sensitivities reported or that could be calculated from each study that we classified as a relevant validation study. In summary, the positive predictive value (PPV) of algorithms ranged from 70 to 96% (median 87%). Sensitivity was reported in 7 studies^{6,10,11,15,16,18,19} and ranged from 57 to 95% (median 79%).

Algorithm validity varied by characteristics of the algorithm, the data source, and the gold standard chosen for comparison. Borzecki¹⁶ found that increasing the required number of AF codes from one to two decreased the sensitivity from 80 to 67%, while specificity was unchanged. Extending the time period searched for these codes from one to two years increased the sensitivity from 80 to 86%, while specificity decreased from 99 to 97%.¹⁶

Two studies examined the impact of different validation criteria on algorithm validity. Brass⁸ reported a PPV of 97% using any documented history of AF within the medical record as the validation criterion, compared to a PPV of 90% when validation required dates and specific treatments for AF. Similarly, Flaker et al.⁹ reported that using a validation

criterion of any physician acknowledgement of AF yielded a PPV of 90.2%, while requiring an ECG or rhythm strip showing AF yielded a lower PPV of 70%.

One study reported that sensitivity was slightly lower for African Americans than for whites (80 vs. 85%), while specificity was similar.⁶ Most studies included predominantly white populations.

DISCUSSION

We identified 16 studies that validated algorithms to identify AF from electronic health data. Validation statistics varied by algorithm and data source, with the PPV from the 14 most relevant and comparable studies ranging from 70–96% and sensitivity from 57–95%. The PPV was lower for incident than for prevalent AF and was affected by characteristics of the algorithm, the database, and the validation criteria.

Reasons for variability between studies

The PPV of a test is highly influenced by the prevalence of the disease in the source population. Many studies focused on older individuals^{8,9,11,13–15} or other high-risk populations. Results from these studies may overestimate the true PPV in the general population. Variability in the racial make-up of the population between studies may have also played a role; studies performed on populations with a larger percentage of non-whites may result in a lower PPV, given that overall, the prevalence and incidence of AF are lower in non-white populations.⁶

Validation criteria varied substantially across studies. In clinical practice, the gold standard for diagnosing AF is a 12-lead ECG. Few studies required an ECG showing AF as part of their validation criteria. Instead, most considered AF confirmed based on any mention in the medical record. The use of this looser criterion will identify people with paroxysmal AF or a remote history of AF who would be incorrectly considered free of AF if validation criteria required a positive ECG. On the other hand, relying on any mention of AF in the chart could lead to misclassification, as some patients reported to have AF may not truly have such a history. The use of stricter validation criteria would be expected to result in lower apparent PPV for the algorithm being considered.

Limitations on the number of diagnosis codes per hospitalization in some databases may also limit the ability to identify AF, lowering the sensitivity of algorithms. This characteristic of the source database was not explicitly discussed in the majority of studies and so we could not assess its contribution to cross-study variation.

Additional considerations

No study specifically examined what was added by including 427.32 (atrial flutter). AF and atrial flutter share many features. They often occur within the same individual, and their potential complications and clinical management are similar.²² Physician documentation and encounter coding may not distinguish atrial flutter from AF when both arrhythmias have been present. For many purposes, an algorithm that includes codes for both AF and atrial flutter will be appropriate.

Some clinical trials monitor “serious” or symptomatic AF and do not investigate or monitor asymptomatic AF. The articles we reviewed did not distinguish between symptomatic and asymptomatic AF, so separate analyses of these AF subtypes were not possible. We believe that for the purposes of post-marketing surveillance, this distinction is not particularly helpful. Both symptomatic and asymptomatic AF are of interest, because both confer an increased risk of stroke and death and thus are clinically relevant. In addition, the decision to

treat AF (e.g. with antithrombotic agents) is not guided by symptoms alone but by overall stroke risk, and so for future studies seeking to identify cohorts of AF patients, the distinction between “serious” and “nonserious” AF will not be relevant.

Only one study specifically examined the PPV of an algorithm for incident, rather than prevalent, AF. Incident AF is of particular interest because incident outcomes are more useful than prevalent conditions when conducting surveillance for adverse effects of new medications. Also, comparative effectiveness studies may seek to include only newly diagnosed, treatment-naïve individuals to decrease some types of bias that arise in observational studies of treatment effectiveness. The few data available suggest that algorithms have considerably lower PPV for incident than for prevalent AF.

Limitations of this study

Our search strategy may have failed to identify some validation studies. Most studies provided limited information about the characteristics of the source population, which may influence risk of AF and thus algorithm PPV. Also, the use of different gold standards across studies inhibits our ability to compare directly the validation statistics for different algorithms. No included studies examined the validity of algorithms using ICD-10 codes, although we expect that algorithm validity may be similar in that setting because the general approach to categorizing these arrhythmias has not changed from ICD-9 to ICD-10.

Most of the studies we identified used data from 10 to 15 years ago. The PPV of these algorithms may be higher in the current era because the prevalence of AF has increased. In addition, the number of codes per encounter retained in electronic databases has generally increased over time. Thus, the sensitivity of algorithms may be higher in contemporary data than our results suggest.

No studies compared the validity of algorithms using only outpatient or only inpatient diagnosis codes to algorithms using both. Algorithms using data from only one setting are likely to be less sensitive than algorithms using data from both settings. Many of the studies we identified included only inpatient data, and their findings may not be generalizable to databases including both inpatient and outpatient codes.

Overall, the included studies had considerable heterogeneity in terms of their aims and methods. Most provided little information about their validation methods, which prevented us from ranking studies according to quality. Only one study set out specifically to validate an algorithm to identify AF from electronic medical data.¹⁶ Additionally, the populations studied may not be representative of the general US population today, in terms of many characteristics including comorbidity and race/ethnicity. Given the increasing use of electronic medical records, as well as the development of electronic ECG databases, the lack of recent data inhibits the generalizability of these results to contemporary studies. Further study is urgently needed given recent changes in the available data.

Conclusions and recommendations

Across a broad range of electronic health databases, an inpatient or outpatient ICD-9 diagnosis code of 427.31 correctly identifies prevalent AF in a substantial proportion of patients. Including code 427.32 (atrial flutter) may be desirable in many settings. Combining inpatient and outpatient AF diagnosis codes with AF diagnoses from an electronic ECG database appears especially promising for identifying prevalent AF. However, no study has validated this approach using medical record review as the gold standard.

Although available data are limited, we propose that the best algorithm to identify incident AF would use inpatient and outpatient diagnosis codes and also electronic ECG data, with

the availability of data for a multi-year lead-in period (for example, 2–5 years) to demonstrate that individuals were free of AF previously. Whether both an ICD-9 code and an ECG should be required needs further study. The validity of algorithms for incident AF warrants further study, because this outcome is of particular interest for purposes including prospective drug safety surveillance and comparative effectiveness research.

Acknowledgments

Funding: Food and Drug Administration contract HHSF223200910006I; NHLBI HL068986 and T32HL007902; NIA K23AG028954; Group Health Research Institute internal funds.

Mini-Sentinel is funded by the Food and Drug Administration (FDA) through Department of Health and Human Services (HHS) Contract Number HHSF223200910006I.

References

1. Kannel WB, Wolf PA, Benjamin EJ, et al. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *Am J Cardiol.* 1998; 82:2N–9N.
2. Naccarelli GV, Varker H, Lin J, et al. Increasing prevalence of atrial fibrillation and flutter in the United States. *Am J Cardiol.* 2009; 104:1534–1539. S0002-9149(09)01388-5 [pii]. 10.1016/j.amjcard.2009.07.022 [PubMed: 19932788]
3. Food and Drug Administration. Listing of DMEs is adapted from “Safety Reporting Requirements for Human Drug and Biological Products; Proposed Rule”. Department of Health and Human Services; Washington, DC: Mar 14, 2003 <http://www.fda.gov/OHRMS/DOCKETS/98fr/03-5204.pdf> [accessed December 3, 2010]
4. The Observational Medical Outcomes Partnership (OMOP). Health Outcomes of Interest. Foundation for the National Institutes of Health; Bethesda: <http://omop.fnih.org/HOI> [accessed June 7, 2010]
5. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960; 20:37–46.
6. Alonso A, Agarwal SK, Soliman EZ, et al. Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. *Am Heart J.* 2009; 158:111–117. S0002-8703(09)00362-7 [pii]. 10.1016/j.ahj.2009.05.010 [PubMed: 19540400]
7. Antani MR, Beyth RJ, Covinsky KE, et al. Failure to prescribe warfarin to patients with nonrheumatic atrial fibrillation. *J Gen Intern Med.* 1996; 11:713–720. [PubMed: 9016417]
8. Brass LM, Krumholz HM, Scinto JM, et al. Warfarin use among patients with atrial fibrillation. *Stroke.* 1997; 28:2382–2389. [PubMed: 9412618]
9. Flaker GC, McGowan DJ, Boechler M, et al. Underutilization of antithrombotic therapy in elderly rural patients with atrial fibrillation. *Am Heart J.* 1999; 137:307–312. [PubMed: 9924165]
10. Hravnak M, Hoffman LA, Saul MI, et al. Atrial fibrillation: prevalence after minimally invasive direct and standard coronary artery bypass. *Ann Thorac Surg.* 2001; 71:1491–1495. [PubMed: 11383788]
11. Psaty BM, Manolio TA, Kuller LH, et al. Incidence of and risk factors for atrial fibrillation in older adults. *Circulation.* 1997; 96:2455–2461. [PubMed: 9337224]
12. Shen AY, Yao JF, Brar SS, et al. Racial/Ethnic differences in ischemic stroke rates and the efficacy of warfarin among patients with atrial fibrillation. *Stroke.* 2008; 39:2736–2743. [PubMed: 18635860]
13. Shireman TI, Howard PA, Kresowik TF, et al. Combined anticoagulant-antiplatelet use and major bleeding events in elderly atrial fibrillation patients. *Stroke.* 2004; 35:2362–2367. [PubMed: 15331796]
14. Whittle J, Wickenheiser L, Venditti LN. Is warfarin underused in the treatment of elderly persons with atrial fibrillation? *Archives of Internal Medicine.* 1997; 157:441–445. [PubMed: 9046896]
15. Yuan Z, Bowlin S, Einstadter D, et al. Atrial fibrillation as a risk factor for stroke: a retrospective cohort study of hospitalized Medicare beneficiaries. *Am J Public Health.* 1998; 88:395–400. [PubMed: 9518970]

16. Borzecki AM, Wong AT, Hickey EC, et al. Identifying hypertension-related comorbidities from administrative data: what's the optimal approach? *Am J Med Qual.* 2004; 19:201–206. [PubMed: 15532912]
17. Go AS, Hylek EM, Phillips KA, et al. Implications of stroke risk criteria on the anticoagulation decision in nonvalvular atrial fibrillation: the Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. *Circulation.* 2000; 102:11–13. [PubMed: 10880408]
18. Brophy MT, Snyder KE, Gaehde S, et al. Anticoagulant use for atrial fibrillation in the elderly. *J Am Geriatr Soc.* 2004; 52:1151–1156. [PubMed: 15209654]
19. Dublin S, French B, Glazer NL, et al. Risk of new-onset atrial fibrillation in relation to body mass index. *Arch Intern Med.* 2006; 166:2322–2328. [PubMed: 17130384]
20. Glazer NL, Dublin S, Smith NL, et al. Newly detected atrial fibrillation and compliance with antithrombotic guidelines. *Arch Intern Med.* 2007; 167:246–252. [PubMed: 17296879]
21. Go AS, Hylek EM, Phillips KA, et al. 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. [PubMed: 11343485]
22. Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation): developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Circulation.* 2006; 114:e257–354. 114/7/e257 [pii]. 10.1161/CIRCULATIONAHA.106.177292 [PubMed: 16908781]
23. Gage BF, Boechler M, Doggette AL, et al. Adverse outcomes and predictors of underuse of antithrombotic therapy in medicare beneficiaries with chronic atrial fibrillation. *Stroke.* 2000; 31:822–827. [PubMed: 10753982]
24. Go AS, Hylek EM, Borowsky LH, et al. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *Ann Intern Med.* 1999; 131:927–934. [PubMed: 10610643]

Take-home messages

- No prior systematic review has examined the validity of algorithms to identify atrial fibrillation from electronic health data.
- A large proportion of prevalent cases identified by ICD-9 code 427.31 were valid (positive predictive value 70–96%, median 89%).
- Algorithm sensitivity ranged from 57–95% (median 79%).
- Few studies examined incident, rather than prevalent, atrial fibrillation, but PPV was lower for incident than prevalent atrial fibrillation.
- Further research is needed using contemporary data, particularly regarding incident atrial fibrillation and approaches that incorporate outpatient diagnosis data and electronic ECG data.

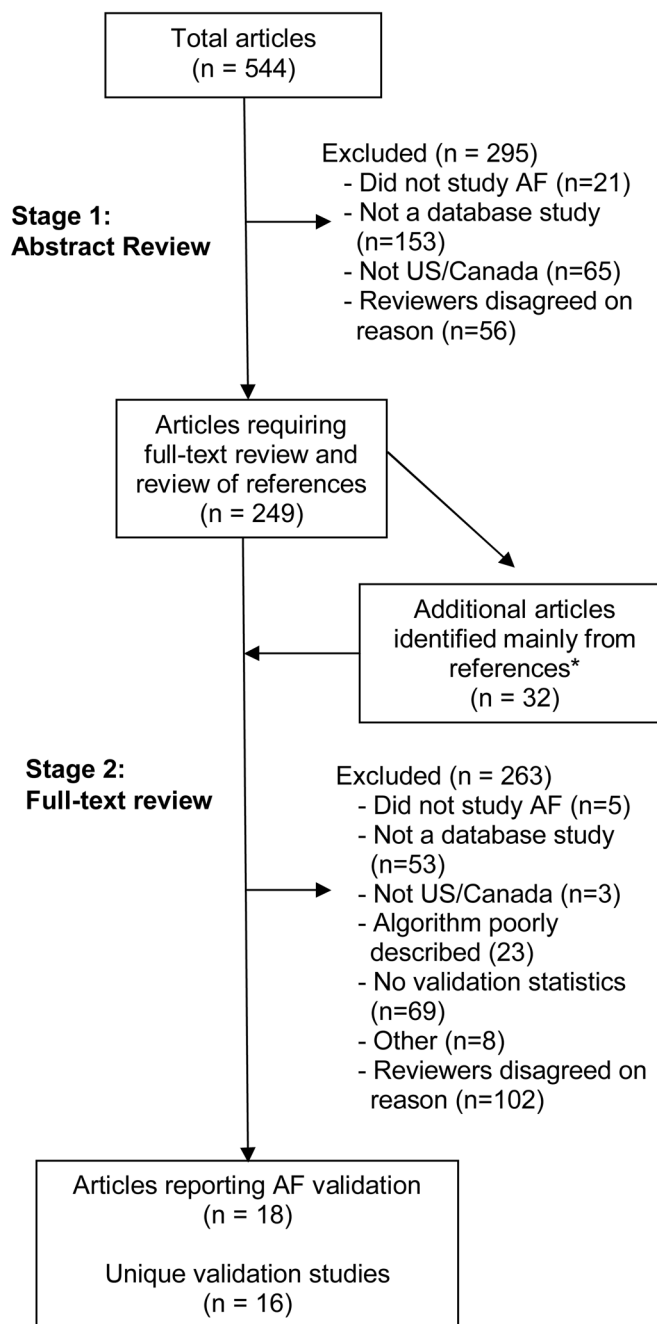


Figure. Selection of publications for inclusion. *One was identified by an external reviewer. A second was identified during article retrieval because its title and author substantially overlapped with an article selected for full review.

Table 1

Articles that identified AF from only inpatient data

Citation	Population and Time Period	Outcome(s) Studied	Algorithm * ICD-9 code(s)	Validation/Adjudication Procedure and Definition	Validation Statistics
Alonso (2009) ⁶	Atherosclerosis Risk in Communities cohort study Age 45–65 in 1987–9; excluded if baseline ECG showed AF Substudy 1: participants with a first inpatient ICD-9 code for AF Substudy 2: Suspected stroke cases 1987–2004	Prevalent and incident AF [†]	427.31	1. Physician review of hospital discharge summaries and inpatient ECGs 2. Medical record review	1. PPV 89% (111/125) for any AF PPV 62% (78/125) for incident AF ^{**} 2. SN 84% (135/161) SP 98% (1351/1385) for prevalent AF
Antani (1996) ⁷	Inpatients, 2 teaching hospitals 1992	Prevalent AF	427.31	Review of medical record including ECGs	PPV 90.8% (178/196)
Brass (1997) ⁸	Medicare patients age 65 from acute care hospitals; 50% with ischemic stroke; none with principal diagnosis of acute MI or non-stroke embolic event 1994	Prevalent AF	427.31 in primary or secondary position	Medical record review seeking ECG(s) interpreted as showing AF. If none found, cardiologist reviewed medical record “Detailed” history; notes included dates or specific treatments for AF	PPV 97% (686/707) for history of AF PPV 89.9% (635/707) for “detailed” history
Flaker (1998) ^{9b}	Hospitalized Medicare patients aged 65, without rheumatic heart disease or a recent cardiothoracic procedure who did not die and were not transferred during the admission 1993–4	Prevalent AF	427.31 in primary or secondary position	Medical record review Looser criterion: physician acknowledgement of AF in notes Stricter criterion: AF on ECG or rhythm strip	PPV 90.2% (1035/1147) for looser criterion PPV 70% (800/1147) for stricter criterion
Hravnak (2001) ¹⁰	CABG patients age >18 Exclusions: prior AF, other or prior cardiac surgery, perioperative or postoperative MI, death within 12 hrs of surgery 1996–8	New-onset AF	427.31	Word search run on the clinical database and from pharmacy data (receipt of procainamide) Medical records reviewed to verify AF and determine if new onset	SN 56.9% (148/260)
Psaty (1997) ¹¹	Cardiovascular Health Study (age 65, 4 geographic areas). Excluded if had a pacemaker or AF at baseline (from ECG, Holter or self-report) Validation study: participants with potential cerebrovascular or cardiovascular events ⁸ 1992 and 1996	Prevalent AF or atrial flutter	427.3, 427.31, 427.32	Physician review of all hospital ECGs	SN 70.7% (29/41)
Shen (2008) ¹²	HMO inpatients (Kaiser Permanente Southern California) 1995–2000	Prevalent AF or atrial flutter	427.31 in any position or 427.32 among the first three codes	Review of medical records including ECGs	PPV 96% (96/100)
Shireman (2004) ¹³	Medicare patients discharged from acute-care hospitals; 750 with primary or secondary diagnosis of AF from each US state 1998–9	Prevalent AF	427.31 in primary or secondary position	Review of inpatient records for index AF hospitalization	PPV 71.1 % (27,674/38,924)

Citation	Population and Time Period	Outcome(s) Studied	Algorithm* ICD-9 code(s)	Validation/Adjudication Procedure and Definition	Validation Statistics
Whittle (1997) ¹⁴	Medicare patients age 65 discharged from 5 small hospitals; no open heart surgery 1993–4	AF present during hospitalization	427.31	Medical record review to confirm whether AF occurred during hospitalization	PPV 85% (242/322)
Yuan (1998) ¹⁵	Medicare patients age 65 at one teaching hospital; exclusions: stroke or venous thrombosis in the prior year; cancer; race unknown or “other” 1985	Prevalent AF	427.3 in one of five diagnosis fields	Hospital discharge records containing up to 27 diagnosis fields ^{††}	SN 87.7% ^{**} SP 100.0% PPV 100% ^{**} NPV 98.6%

* Unless specified, the position of ICD-9 codes was not described.

** Not included in summary statistics

[†] People whose baseline study ECG showed AF were excluded, but the study lacked other information about AF history baseline needed to properly account for pre-existing AF in their incidence calculations.

^{††} Validation statistics also cited by Gage.²³

[§] Psaty et al. also conducted a study of people who appeared to have had incident AF during follow-up based on hospital discharge diagnoses. Based on review of medical records, 98.6% (209/212) truly had AF. This is not actually the PPV for incident (vs. prevalent) AF because people who on review were found to have had AF pre-dating study entry were excluded (removed from both numerator and denominator). Information was not presented about the PPV for incident AF, specifically.

^{†††} Gold standard uses only hospital database of diagnosis codes, rather than medical record review. Their results shed light on how the number of diagnosis codes in a database impacts sensitivity but are not informative about the actual validity of AF diagnosis codes.

Abbreviations: AF, atrial fibrillation; CABG, coronary artery bypass grafting; ECG, electrocardiogram; ICD, International Classification of Diseases; MI, myocardial infarction; NPV, negative predictive value; PPV, positive predictive value; SN, sensitivity; SP, specificity.

Table 2

Articles that identified AF from outpatient data

Citation	Population and Time Period	Outcome	Algorithm* ICD-9 code(s) (inpatient or outpatient unless otherwise specified)	Validation/Adjudication Procedure and Definition	Validation Statistics
Borzecki (2004) ¹⁶	V/A outpatients with 2 clinic visits at least 6 months apart; 100 with hypertension and 20 without from each of 10 sites 1998–9	Prevalent AF	427.3: 1 code in 1 year Examined impact of varying # of codes required and # of years of claims data used	Review of outpatient medical record; AF considered confirmed if mentioned in chart	1 claims diagnosis in 1 yr: SN 80%, SP 99%, PPV 84% 2 claims diagnoses in 1 yr: SN 67%, SP 99% 1 diagnosis in 2 yrs: SN 86%, SP 97% 2 diagnoses in 2 yrs: SN 74%, SP 99%
Brophy (2004) ¹⁸	Patients with 1 encounter at VA Boston Healthcare System and 1 ECG showing AF in the electronic ECG database 1997–2001	Prevalent AF	427.3, 427.31	ECG showing AF in electronic database	SN 77.8% (2619/3366)
Dublin (2006) ^{19,†}	HMO members; average age 73, 38% male, and 73% with treated hypertension 1989–98	Prevalent AF or atrial flutter	427.31 or 427.32 in any position from an inpatient, outpatient or ED encounter	Medical record review	SN 95% (236/248) SP 99% (244/247)
Glazer (2007) ²⁰	HMO enrollees without a previous ICD-9 code for AF 2001–2	Incident AF or atrial flutter	427.31 or 427.32	Medical record review	PPV 76.8% (1105/1438)
Go (2000) ^{17,†}	HMO members (Kaiser Permanente Northern California) 1996–7	Prevalent AF	Validation study: > 1 outpatient 427.31 but no electronic ECG showing AF	ECG showing AF in medical record	PPV 78% (39/50)
Go (2001) ²¹	HMO members (Kaiser Permanente Northern California) 20 years old, 1996–7	Prevalent AF	Validation study: 1 outpatient 427.31 but no electronic ECG showing AF [§]	ECG showing AF in medical record	PPV 56% ^{**} (28/50)

* Unless otherwise specified, the position of ICD-9 codes was not described.

** Not included in summary statistics

† Validation statistics also cited by Go.²⁴‡ Validation statistics also cited by Glazer.²⁰

§ Validation performed on a population that was unlikely to have prevalent AF (1 outpatient code 427.31 but no electronic ECG showing AF)

Abbreviations: AF, atrial fibrillation; ECG, electrocardiogram; HMO, health maintenance organization; ICD, International Classification of Diseases; MI, myocardial infarction; PPV, positive predictive value; SN, sensitivity; SP, specificity; VA, Veterans Administration.