Am J Cardiovasc Dis 2012;2(3):184-191
www.AJCD.us /ISSN:2160-200X/AJCD1205003

Original Article

Choices in the use of ICD-9 codes to identify stroke risk factors can affect the apparent population-level risk factor prevalence and distribution of CHADS2 scores

James A Rothendler[1,2], Adam J Rose[1,3], Joel I Reisman[1], Dan R Berlowitz[1,2,3], Lewis E Kazis[1,2]

[1]Center for Health Quality, Outcomes and Economic Research, Health Services Research and Development, Bedford VA Medical Center, Bedford, MA, USA; [2]Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, USA; [3]Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA

Received May 18, 2012; Accepted June 10, 2012; Epub July 25, 2012; Published August 15, 2012

Abstract: While developed for managing individuals with atrial fibrillation, risk stratification schemes for stroke, such as CHADS2, may be useful in population-based studies, including those assessing process of care. We investigated how certain decisions in identifying diagnoses from administrative data affect the apparent prevalence of CHADS2-associated diagnoses and distribution of scores. Two sets of ICD-9 codes (more restrictive/ more inclusive) were defined for each CHADS2-associated diagnosis. For stroke/transient ischemic attack (TIA), the more restrictive set was applied to only inpatient data. We varied the number of years (1-3) in searching for relevant codes, and, except for stroke/TIA, the number of instances (1 vs. 2) that diagnoses were required to appear. The impact of choices on apparent disease prevalence varied by type of choice and condition, but was often substantial. Choices resulting in substantial changes in prevalence also tended to be associated with more substantial effects on the distribution of CHADS2 scores.

Keywords: Stroke, atrial fibrillation, risk stratification, CHADS2, ICD-9-CM codes

Introduction

Atrial fibrillation (AF) is a risk factor for ischemic stroke, but among patients with AF, the level of risk for an individual patient varies according to other clinical characteristics [1-5]. Several risk stratification schemes have been developed for assessing the relative risk of stroke in patients with non-valvular AF [2-8], and such information may help clinicians in making more informed decisions regarding clinical management. While these risk stratification schemes were developed for use in individual patients, they may also be useful in population-based studies [9-14]. For individual patients, chart reviews to obtain prior and current medical conditions are typically used for identifying characteristics associated with increased risks of stroke. However, for assessing relative risks for large cohorts of patients, administrative data is often more feasible.

There is disagreement on which risk stratification scheme provides the most useful information in patients with AF. We have chosen to focus on the CHADS2 because its elements (age and diagnoses) are often available in large administrative databases [6]. (Some risk stratification schemes use data that is less commonly available in such databases, such as systolic blood pressure or heart size measurements from echocardiograms [3-5].) In the CHADS2 scheme, a score is created by adding one point for each of the elements of Congestive heart failure (CHF), Hypertension, Age ≥75 or Diabetes, and two points for prior Stroke or transient ischemic attack (TIA) [6]. In the initial validation of CHADS2 in Medicare beneficiaries, a chart review was used to identify the presence of hypertension, a history of stroke or TIA and, for most of the cohort, diabetes [6]. However, International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9) codes were

used to identify "recent" CHF exacerbations associated with hospitalizations. ICD-9 codes were also used to identify the presence of diabetes for some of the patients. In a subsequent study comparing the CHADS2 to other risk adjustment schemes, any history of CHF was substituted for recent exacerbations [8].

Since the initial validation of CHADS2, a number of population-based studies have defined the elements of the CHADS2, in whole or in part, by the use of ICD-9 diagnosis codes [9-12]. However, in defining the various elements of the CHADS2, there have been differences among studies in the specifications of sets of ICD-9 codes associated with each condition, the number of times that a code was required to be present or the use of inpatient vs. outpatient datasets from which codes are extracted. Furthermore, there has been variation in the time interval used in searching for such codes.

Studies that have focused on validating sets of ICD-9 codes to represent the types of clinical conditions associated with the CHADS2 scheme have reported differences in sensitivity and specificity based on similar types of choices [15-19]. Use of ICD-9 codes to identify stroke has been particularly difficult in regard to identifying code sets that have both high sensitivity and specificity [19-21]. Therefore, among investigators who have used ICD-9-based datasets in defining CHADS2 scores, variation in codes to identify CHADS2 conditions is probably not surprising.

Differences in choices regarding how ICD-9 codes are used to identify the clinical conditions associated with CHADS2 could also impact the distribution of CHADS2 scores at the population level. This, in turn, could affect statistical associations and conclusions. The purpose of this study was to assess how different sets of ICD-9 codes as well as other decisions (such as the number of times a code appears and the time over which the presence of codes is searched) affect the apparent prevalence of CHADS2-associated conditions and the distribution of CHADS2 scores in a population of patients with non-valvular AF.

Methods

Datasets

Department of Veterans Affairs (VA) national datasets encompassing outpatient visits and/or acute inpatient hospitalizations were used to determine the presence of certain ICD-9 codes, as further specified below.

Population

Patients with "non-valvular" AF were identified from among those cared for in the VA during fiscal year (FY) 2007 (October 2006 through September 2007). Diagnoses for AF were based on having two or more inpatient or outpatient ICD-9 codes of 427.31 separated by at least 60 days. Patients were excluded if there was a history of valvular heart disease manifested by mitral stenosis or certain types of prior cardiac valve surgery [22], as indicated by the following ICD-9 diagnostic or procedure codes: 394.0, 394.2, 396.0, 396.1, 396.8, V42.2, V43.3, 35.10 to 35.14 or 35.20 to 35.28. These exclusion diagnoses were based on outpatient or inpatient records in the three years prior to the initial diagnosis of AF identified in FY2007.

ICD-9 definitions for CHADS2 components

A number of reports in the literature that incorporate CHADS2 scores have defined the presence of CHF, hypertension, diabetes or stroke/TIA using sets of ICD-9 codes [9-12]. Other reports, not involved with CHADS2, have studied ICD-9 code sets for these conditions in comparison to chart reviews [15, 16, 18, 19]. "Core" sets of codes often include 428.x for CHF, 401.xx-405.xx for hypertension and 250.x for diabetes. Additional codes for each condition, typically associated with other comorbidities or complications, have been used by some investigators. For example, the set for CHF has included 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.03, 404.13, 404.91, 404.93, that for hypertension has included 437.2, and that for diabetes has included 357.2, 362.0 and 366.41 [10, 19]. While it did not seem likely that use of these additional codes would identify a substantial number of additional patients with the relevant condition, we investigated this issue by specifying "more restrictive" and "more inclusive" code sets, as indicated in **Table 1**.

Compared to the other CHADS2 conditions, the use of ICD-9 codes to identify patients with prior ischemic stroke or TIA is more problematic. Because the CHADS2 scheme includes any prior history of ischemic stroke or TIA, it becomes

Table 1. ICD-9 codes used to define more restrictive and more inclusive sets associated with the CHADS2 conditions. TIA: transient ischemic attack

CHADS2 condition	More restrictive set	More inclusive set
Congestive heart failure	428.x	428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.03, 404.13, 404.91, 404.93
Hypertension	401.x, 402.x, 403.x, 404.x, 405.x	401.x, 402.x, 403.x, 404.x, 405.x, 437.2
Diabetes	250.x	250.x, 357.2, 362.0, 366.41
Stroke and TIA	433.x1,* 434.x1,* 435.x*	433.x1, 434.x1, 435.x, 436, 438.x

*Assessed using only inpatient data

important to search for records indicating prior in addition to acute events. Validation studies in the literature have typically assessed the codes associated with acute strokes associated with hospitalizations, and even in that setting, it has been difficult to identify code sets that have both good sensitivity and specificity [20, 21]. Using one year of inpatient records, Birman-Deych reported use of an ICD-9 code set that achieved a sensitivity of 58% for detecting acute or prior strokes, but patients in the cohort were selected based on a hospitalization for AF [19]. In attempting to identify any history of an ischemic stroke or TIA in a population that includes individuals without hospitalizations within a specified span of data, it would seem important to include outpatient records. However, we are not aware of published studies that have reported on systematic validation of outpatient ICD-9 codes for either acute or prior strokes (or TIAs). While we included outpatient codes in our more inclusive code set, as indicated below, we recognize that the accuracy of such codes is unknown at this time.

For the more restrictive code set for stroke/TIA, we included ICD-9 codes 433.x1 and 434.x1 identified from acute inpatient records, as these codes have been found to have high specificity and positive predictive value [20]. Also included in the restrictive set is 435.x for TIA. We accepted these codes in any position in the list of discharge diagnoses.

For the more inclusive code set, we included codes that appeared in either acute inpatient or in outpatient records. In addition to 433.x1, 434.x1 and 435.x, we included codes 436 and 438.x. Investigators who have defined CHADS2 using ICD-9 codes have often included code 436 ("Acute, but ill-defined, cerebrovascular disease") in identifying stroke [9-12], although it had a positive predictive value of only 48% in VA inpatient datasets [20]. We included ICD-9 code 438.x ("Late effects of cerebrovascular disease") in an effort to detect previous strokes [11, 19]. One potential issue is whether, in the outpatient setting, prior strokes are being coded with 438.x or with one of the codes more typically associated with acute events. Yiannakoulias et al found that in certain outpatient facilities in the province of Alberta, Canada, ICD-9 codes 435.x, 436 and 438.x were the most common of those for cerebrovascular disease [23]. However, there were substantial differences among the facilities in which of these codes predominated. For example, at some facilities, 438.x predominated, while at others, it was 436. While the actual medical conditions were not determined by chart review, the authors suggested that such wide variation in coding was unlikely to be explainable by differences in disease prevalence.

There was a change in the definition of ICD-9 code 436 beginning in October 2004 that was intended to remove the words "stroke" and "cerebrovascular accident" (CVA) from its description and to re-index those terms to other codes [24, 25]. However, it is unclear how this has changed coding practices in either the inpatient or outpatient setting, and some reference documents have continued to include, for certain sub-entries, the terms "stroke" and "CVA" in association with code 436 [26, 27]. Therefore we counted its presence in periods after September 2004 as part of the more inclusive code set. V12.54 is a current code for indicating prior stroke or TIA without residual sequelae, but was not relevant to this study as it was not an official code until Oct. 1, 2007.

In searching for relevant ICD-9 codes for each of the CHADS2 conditions, we assessed the effect

Table 2. Prevalence (%) of Components of CHADS2 by ICD-9 set (restrictive/inclusive), number of diagnoses required and lookback period. TIA: transient ischemic attack

Component Variant	Lookback Period					
	1 year		2 years		3 years	
	2 diagnoses	1 diagnosis	2 diagnoses	1 diagnosis	2 diagnoses	1 diagnosis
Congestive Heart Failure						
More Restrictive	17.2	24.5	21.7	28.2	23.7	30.3
More Inclusive	17.3	24.7	21.9	28.4	23.9	30.5
Hypertension						
More Restrictive	61.6	76.9	73.7	81.6	77.0	83.8
More Inclusive	61.6	76.9	73.7	81.6	77.0	83.8
Diabetes						
More Restrictive	30.2	34.3	32.7	35.5	33.4	36.1
More Inclusive	30.4	34.4	32.8	35.6	33.4	36.2
Stroke and TIA						
More Restrictive	N/A	0.6	N/A	1.0	N/A	1.3
More Inclusive	N/A	9.6	N/A	11.7	N/A	13.1

of looking back 1, 2 and 3 years ("lookback" times) prior to the initial diagnoses of AF in FY2007. Also, for diagnoses involving CHF, hypertension and diabetes, we explored the effect of requiring only 1 vs. 2 diagnoses during each of the relevant "lookback" times.

Statistical methods

We used descriptive analyses to indicate the prevalence of disease as a percent of all individuals in the cohort. In addition, we assessed the distribution of the CHADS2 scores (0, 1, 2, or 3+) as a percent of all scores in the cohort for each of the choices related to ICD-9 code sets, number of diagnoses and years of lookback.

Results

A total of 126,167 individuals met the study criteria for non-valvular AF. Their mean age was 74.0, and 98.4% were male. **Table 2** shows the prevalence of the individual conditions that comprise the CHADS2 as a function of the ICD-9 code set (more restrictive vs. more inclusive), the number diagnoses required (1 vs. 2) and the number of "lookback" years used in assessing for the presence of diagnoses (1, 2 or 3).

For CHF, hypertension and diabetes, the additional codes in the more inclusive set made relatively little difference, as expected, in the calculated disease prevalence. However, the requirement of having at least two diagnoses substantially decreased the apparent prevalence for all three conditions, especially with only a one year lookback, although the effect was greater for CHF and hypertension in comparison to diabetes. For Stroke/TIA, there was a substantial difference in the calculated prevalence of disease between the more restrictive and more inclusive code sets, mainly due to the inclusion in the more inclusive code set of diagnoses drawn from outpatient datasets. A majority of such outpatient codes were for 436, and there was also a substantial prevalence of 438.x and 435.x, whereas there were relatively few 433.x1 and 434.x1.

For all four of the condition categories, the number of years of lookback also had a substantial effect on the calculated disease prevalence. On a relative basis, the difference was most pronounced for stroke/TIA. Diabetes showed the least change as a function of years of lookback, perhaps because diabetes is so often a focus of office visits.

Table 3 shows the distribution of CHADS2 scores as function of ICD-9 criteria, number of diagnoses required and years of database lookback. In this case, the more restrictive and more inclusive categories include the corresponding code sets for each of the four conditions. There were moderate differences in the CHADS2 distribution between the more restrictive and more inclusive categories, mainly due to the prevalence differences for stroke/TIA. The effect of the more restrictive/ more inclusive categories was greater for the higher CHADS2 scores.

Varying the number of required diagnoses and

Table 3. Distribution of CHADS2 (%) scores by ICD-9 set (restrictive/inclusive), number of diagnoses required, and lookback period

CHADS2 Variant	1-Year Lookback Value of CHADS2				2-Year Lookback Value of CHADS2				3-Year Lookback Value of CHADS2			
	0	1	2	3+	0	1	2	3+	0	1	2	3+
More restrictive												
2 diagnoses*	12.1	32.7	36.5	18.7	8.4	27.5	39.8	24.3	7.5	25.9	40.0	26.5
1 diagnosis	6.8	26.2	40.5	26.5	5.5	23.9	40.4	30.2	5.0	22.7	40.0	32.3
More Inclusive												
2 diagnoses*	11.4	30.0	33.7	24.9	7.8	25.0	35.9	31.3	7.0	23.3	35.5	34.2
1 diagnosis	6.4	24.2	37.1	32.3	5.2	21.7	36.2	36.8	4.6	20.5	35.5	39.4

*Except for stroke/transient ischemic attack, which uses 1 diagnosis

the number of years of database lookback both had substantial effects on the distribution of the CHADS2 scores. As expected, there was a decrease in those with scores of 3+ when two diagnoses were required. Similarly, the number of patients with CHADS2 scores of 0 was substantially higher with one year compared to three years of lookback, and there was a corresponding decrease in the percentage of those with scores of 3+ with one compared to three years of lookback. For the more restrictive code sets, use of two diagnoses over one year was associated with 12.1% of patients with CHADS2 score of 0, compared to 5% when one diagnosis was required over three years.

Discussion

Because the more inclusive sets of ICD-9 codes for hypertension, CHF and diabetes differed from the more restrictive sets only by inclusion of codes that indicate certain complications or comorbidities, it is not surprising that there were only small differences in prevalence between these sets. However, for these same conditions, the requirements based on the number of years of lookback and number of diagnoses required both had more substantial effects on disease prevalence. The extent of these differences may vary among different healthcare systems and their associated administrative datasets.

For stroke and TIA, there were also substantial differences in apparent disease prevalence as a function of the number of years of lookback. There were also large relative differences in disease prevalence between more restrictive and more inclusive code sets. For the more inclusive code set, the increased prevalence was mainly due to the inclusion of codes in the outpatient setting, including those more typically associated with acute events. However, it is possible that in many cases of "acute" codes, the intent was to indicate a prior stroke or TIA. To the extent that prior strokes or TIAs were being diagnosed, the relative contribution of such data to the overall prevalence of stroke/TIA might be expected to have a larger effect in VA datasets compared to those of some other healthcare systems. This hypothesis is based on a report that most patients dually eligible for VA care and Medicare receive their initial care for acute strokes in non-VA hospitals, and such acute events would not typically be included in the VA acute hospitalization files [28]. As previously noted, we are not aware of any validation studies in the outpatient setting for using ICD-9 codes to identify stroke/TIA, and chart reviews were not part of this study. Therefore, the relative accuracy of different outpatient codes to identify stroke/TIA is unknown, and further studies in this area are warranted. However, the apparent prevalence of stroke/TIA of 13.1% using three years of data is similar to other studies that have determined the prevalence of these conditions in patients with AF [29, 30].

Consistent with the effect on disease prevalence, choices studied in this report regarding identification of conditions from ICD-9 codes also had a substantial effect on the distribution of the CHADS2 scores. Reports in the literature that have assessed CHADS2 scores based on ICD-9 codes have varied in these parameters, and this may affect comparison among studies [9-12].

Even in our "more inclusive" set of codes for stroke/TIA, we did not include such codes as 433.x0 or 434.x0 since these codes have been found to have poor positive predictive value for

stroke, although they may indicate other manifestations of cerebrovascular disease [20]. We also did not assess variation of disease prevalence or CHADS2 distribution as a function of the position of a particular diagnosis (i.e. primary vs. anywhere in the record). Such an assessment would likely have resulted in further variation in the observed parameters. For those ICD-9 codes identified from outpatient records, we included all outpatient services (including laboratory, radiology, etc) and not just face-to-face clinical encounters. This may also have affected the relative prevalence of these codes.

There are several additional issues that may limit the generalizability of our findings. The VA population in our cohort was 98.4% male, and the VA population tends to have a higher prevalence of comorbidities than non-VA populations. Also, as previously noted, a substantial percentage of patients dually eligible for VA care and Medicare receive initial care for strokes outside the VA. Therefore, the presence of outpatient codes suggestive of stroke/TIA may have had a greater influence on apparent disease prevalence compared to other health care systems. In addition to care for stroke, many VA patients receive at least some routine outpatient care outside the VA, and the VA datasets may therefore not reflect some comorbidities not addressed at VA clinic encounters. Regarding outpatient diagnoses, the VA uses an electronic medical record (EMR) to document care. The EMR's characteristics and the methods by which providers choose ICD-9 codes using this EMR may result in different distributions of codes compared to those found in other healthcare systems. Because of the presence of these issues, the actual numbers in this report derived from VA datasets are likely to be different in other systems of care. Although specific numbers may vary, it seems likely that in other healthcare systems, decisions in identifying CHADS2 conditions using ICD-9 codes can similarly affect calculated disease prevalence and distribution of scores.

Our goal in this report was not to define "optimal" algorithms to identify the presence of the various conditions associated with CHADS2 from administrative data, but rather to demonstrate the variation that can occur in the observed prevalence of these conditions as a function of choices made in their identification. While the CHADS2 often is used on an "n" of 1 basis for the clinical care of patients who might be at increased risk of stroke, the intent of this paper is to show the effect of different choices regarding use of ICD-9 codes in deriving the CHADS2 scores for application in population-based studies and not for individual patients.

Researchers performing population-based studies based on ICD-9 codes identified in administrative datasets should be aware of how choices in specifying components of stroke risk stratification schemes, such as the CHADS2, affect the distribution of risk scores. Further studies would be useful to assess how such choices affect the accuracy of administrative diagnostic codes for different healthcare systems. In particular, validation studies are needed regarding the accuracy of outpatient ICD-9 codes to indicate prior strokes and TIAs.

Acknowledgements

This material is based upon work supported in part by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Health Services Research and Development; the Boston University School of Public Health and a contract from Boehringer Ingelheim, Pharmaceuticals Inc. Dr. Rose is supported by a Career Development Award from the Veterans Affairs Health Services Research and Development Service.

The views expressed in this report are solely those of the authors and do not necessarily represent the position or policy of the Department of Veterans Affairs or the United States Government, Boston University School of Public Health or Boehringer Ingelheim, Pharmaceuticals, Inc.

Address correspondence to: Dr. Lewis E Kazis, Pharmaceutical Effectiveness Program, Center for Health Quality, Outcomes and Economic Research (CHQOER), 200 Springs Rd. (152), Building 70, Bedford, MA 01730-1114, USA. Tel: 781-687-2860; E-mail: lek@bu.edu

References

[1] Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. Neurology 2007; 69: 546-554.

[2] Hughes M and Lip GY. Guideline Development Group, National Clinical Guideline for Management of Atrial Fibrillation in Primary and Secondary Care, National Institute for Health and Clinical Excellence. Stroke and thromboembolism in atrial fibrillation: a systematic review

of stroke risk factors, risk stratification schema and cost effectiveness data. Thromb Haemost 2008; 99: 295-304.
[3] Hart RG and Pearce LA. Current status of stroke risk stratification in patients with atrial fibrillation. Stroke 2009; 40: 2607-2610.
[4] Stroke Risk in Atrial Fibrillation Working Group. Comparison of 12 risk stratification schemes to predict stroke in patients with nonvalvular atrial fibrillation. Stroke; a journal of cerebral circulation 2008; 39: 1901-1910.
[5] Fang MC, Go AS, Chang Y, Borowsky L, Pomernacki NK, Singer DE; ATRIA Study Group. Comparison of risk stratification schemes to predict thromboembolism in people with nonvalvular atrial fibrillation. J Am Coll Cardiol 2008; 51: 810-815.
[6] Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA 2001; 285: 2864-2870.
[7] Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010; 137: 263-272.
[8] Gage BF, van Walraven C, Pearce L, Hart RG, Koudstaal PJ, Boode BS, Petersen P. Selecting patients with atrial fibrillation for anticoagulation: stroke risk stratification in patients taking aspirin. Circulation 2004; 110: 2287-2292.
[9] Choudhry NK, Soumerai SB, Normand SL, Ross-Degnan D, Laupacis A, Anderson GM. Warfarin prescribing in atrial fibrillation: the impact of physician, patient, and hospital characteristics. Am J Med 2006; 119: 607-615.
[10] Lakshminarayan K, Solid CA, Collins AJ, Anderson DC, Herzog CA. Atrial fibrillation and stroke in the general medicare population: a 10-year perspective (1992 to 2002). Stroke 2006; 37: 1969-1974.
[11] Kaatz S, Douketis JD, Zhou H, Gage BF, White RH. Risk of stroke after surgery in patients with and without chronic atrial fibrillation. J Thromb Haemost 2010; 8: 884-890.
[12] Henriksson KM, Farahmand B, Johansson S, Asberg S, Terént A, Edvardsson N. Survival after stroke--the impact of CHADS2 score and atrial fibrillation. Int J Cardiol 2010; 141: 18-23.
[13] Rose AJ, Berlowitz DR, Ash AS, Ozonoff A, Hylek EM, Goldhaber-Fiebert JD. The business case for quality improvement: oral anticoagulation for atrial fibrillation. Circ Cardiovasc Qual Outcomes 2011; 4: 416-424.
[14] Sandhu RK, Bakal JA, Ezekowitz JA, McAlister FA. Risk stratification schemes, anticoagulation use and outcomes: the risk--treatment paradox in patients with newly diagnosed non-valvular atrial fibrillation. Heart 2011; 97: 2046-2050.
[15] Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? J Clin Epidemiol 2000; 53: 343-349.
[16] Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: what's the optimal approach? Am J Med Qual 2004; 19: 201-206.
[17] Miller DR, Safford MM and Pogach LM. Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. Diabetes Care 2004; 27 Suppl 2: B10-21.
[18] Quan H, Khan N, Hemmelgarn BR, Tu K, Chen G, Campbell N, Hill MD, Ghali WA, McAlister FA; Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs. Validation of a case definition to define hypertension using administrative data. Hypertension 2009; 54: 1423-1428.
[19] Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 2005; 43: 480-485.
[20] Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A. Stroke: who's counting what? J Rehabil Res Dev 2001; 38: 281-289.
[21] Benesch C, Witter DM Jr, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. Neurology 1997; 49: 660-664.
[22] Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV, Singer DE. 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.
[23] Yiannakoulias N, Svenson LW, Hill MD, Schopflocher DP, James RC, Wielgosz AT, Noseworthy TW. Regional comparisons of inpatient and outpatient patterns of cerebrovascular disease diagnosis in the province of Alberta. Chronic Dis Can 2003; 24: 9-16.
[24] ICD-9-CM Coordination and Maintenance Committee Meeting, April 1- 2, 2004. Available from: http://www.cdc.gov/nchs/data/icd9/agendaapril04%20revised.pdf.
[25] International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM): ICD-9-CM Tabular List of Diseases (FY05); File DTAB05. 2004; Available from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD9-CM/2004.
[26] International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). ICD-9-CM Files via FTP, 2009. ICD-9-CM Index to Diseases (FY10)/Index to Diseases and Injuries. File: DINDEX10; Available from: ftp://

ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD9-CM/2009

[27] Centers for Medicare and Medicaid Services. Diagnosis and Procedure Codes: Abbreviated and Full Code Titles. Version 27 Full and Abbreviated Code Titles - Effective October 1, 2009. See file: V27LONG_SHORT_DX_110909.xls. Available from: https://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/codes.html

[28] Shen Y, Findley PA, Maney M, Pogach L, Crystal S, Rajan M, Findley TW. Department of Veterans Affairs-Medicare dual beneficiaries with stroke: where do they get care? J Rehabil Res Dev 2008; 45: 43-51.

[29] Rose AJ, Ozonoff A, Henault LE, Hylek EM. Warfarin for atrial fibrillation in community-based practise. J Thromb Haemost 2008; 6: 1647-1654.

[30] ACTIVE Writing Group of the ACTIVE Investigators. Connolly S, Pogue J, Hart R, Pfeffer M, Hohnloser S, Chrolavicius S and Yusuf S. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet 2006; 367: 1903-1912.