

BMJ Open Protocol for validating cardiovascular and cerebrovascular ICD-9-CM codes in healthcare administrative databases: the Umbria Data Value Project

Francesco Cozzolino,¹ Iosief Abraha,¹ Massimiliano Orso,¹ Anna Mengoni,² Maria Francesca Cerasa,² Paolo Eusebi,¹ Giuseppe Ambrosio,² Alessandro Montedori¹

To cite: Cozzolino F, Abraha I, Orso M, *et al*. Protocol for validating cardiovascular and cerebrovascular ICD-9-CM codes in healthcare administrative databases: the Umbria Data Value Project. *BMJ Open* 2017;**7**:e013785. doi:10.1136/bmjopen-2016-013785

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-013785>).

Received 6 August 2016
Revised 19 January 2017
Accepted 23 January 2017



CrossMark

¹Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy
²Division of Cardiology, Santa Maria della Misericordia Hospital, University of Perugia School of Medicine, Perugia, Italy

Correspondence to
Dr Iosief Abraha;
iosief_a@yahoo.it

ABSTRACT

Introduction: Administrative healthcare databases can provide a comprehensive assessment of the burden of diseases in terms of major outcomes, such as mortality, hospital readmissions and use of healthcare resources, thus providing answers to a wide spectrum of research questions. However, a crucial issue is the reliability of information gathered. Aim of this protocol is to validate International Classification of Diseases, 9th Revision—Clinical Modification (ICD-9-CM) codes for major cardiovascular diseases, including acute myocardial infarction (AMI), heart failure (HF), atrial fibrillation (AF) and stroke.

Methods and analysis: Data from the centralised administrative database of the entire Umbria Region (910 000 residents, located in Central Italy) will be considered. Patients with a first hospital discharge for AMI, HF, AF or stroke, between 2012 and 2014, will be identified in the administrative database using the following groups of ICD-9-CM codes located in primary position: (1) 410.x for AMI; (2) 427.31 for AF; (3) 428 for HF; (4) 433.x1, 434 (excluding 434.x0), 436 for ischaemic stroke, 430 and 431 for haemorrhagic stroke (subarachnoid haemorrhage and intracerebral haemorrhage). A random sample of cases, and of non-cases, will be selected, and the corresponding medical charts retrieved and reviewed for validation by pairs of trained, independent reviewers. For each condition considered, case adjudication of disease will be based on symptoms, laboratory and diagnostic tests, as available in medical charts. Divergences will be resolved by consensus. Sensitivity and specificity with 95% CIs will be calculated.

Ethics and dissemination: Research protocol has been granted approval by the Regional Ethics Committee. Study results will be disseminated widely through peer-reviewed publications and presentations at national and international conferences.

INTRODUCTION

As computer technology continues to advance, use of digital administrative databases

Strengths and limitations of this study

- The study will evaluate the validity of the International Classification of Diseases, 9th Revision—Clinical Modification (ICD-9-CM) codes for cerebrovascular and cardiovascular diseases (acute myocardial infarction, heart failure, atrial fibrillation and stroke), using the administrative database of Umbria Region.
- The strength of this study is that it will review original source documentation available in medical charts to adjudicate case of cerebrovascular and cardiovascular diseases.
- Once the administrative database will be validated for the considered diseases, it can be used for outcome research including pharmacoepidemiology, health service research and quality of care research.
- Validation studies of administrative data are related to their context, and are not immediately generalisable to other settings.

is increasingly growing in healthcare settings worldwide. These databases anonymously store patients' data about healthcare assistance they received, including birth, death or disease treatment. Usually, diagnosis of the disease is associated with a specific code from the International Classification of Diseases (ICD-9) 9th or 10th revision. The ICD is designed to map health conditions to corresponding generic categories along with specific variations. Merging of individual patient data from administrative databases with other sources (eg, prescription and laboratory data) allows investigating a wide range of relevant and often unique public health questions¹ monitoring population health status over time, and performing population-based pharmacoepidemiological research.¹ These databases, therefore, have the potential to address important

issues in postmarketing surveillance^{2 3} epidemiology⁴ quality performance, and health services research.⁵

However, there is a concern that the power of administrative databases as a source of healthcare information cannot be fully exploited unless they are thoroughly validated.^{6–9} For instance, a systematic review¹⁰ of ICD-9 code validation in Italian administrative databases reported that only a few regional databases have been validated, and just for a limited number of ICD-9 codes of diseases.^{11–18}

While non-clinical information in healthcare databases, such as demographic and prescription data, are highly accurate,^{19 20} reliability of registered diagnoses and procedures is variable.^{20 21} Determining the accuracy of the latter two categories of clinical information is important to all potential users, and involves confirming the consistency of information within the databases with the corresponding clinical records of patients.¹⁹ Hence, it is imperative that health authorities systematically validate their databases for major diseases to productively use the information they contain.

In Western countries, cardiovascular and cerebrovascular diseases are the leading cause of mortality and morbidity, and represent a major social and economic problem. In this context, most of the burden is due to occurrence of acute myocardial infarction (AMI),²² heart failure (HF),^{23 24} atrial fibrillation (AF)²⁵ and stroke.¹⁷ Administrative healthcare databases can provide a comprehensive understanding of the burden of these diseases in terms of important outcomes such as mortality, hospital readmissions and use of other healthcare resources. In addition, such databases can aid in monitoring adherence to essential drug therapies, including the use of evidence-based therapies.

The Regional Health Authority of Umbria has started a research regarding case definitions of diseases^{26–28} as well as validating ICD-9 codes for diseases.²⁹ The objective of the present protocol is to evaluate the accuracy of the ICD-9-CM codes related to AMI, HF, AF, ischaemic and haemorrhagic stroke (intracerebral and subarachnoidal haemorrhage) in the administrative database of the Regional Health Authority of Umbria.

METHODS

Setting and data source

Administrative database

The Regional Administrative Database of Umbria gathers information regarding all hospital admission medical records on all 910 000 residents, including personal demographics, hospital admission and discharge dates, vital status, hospital department, primary and secondary diagnoses and surgical or diagnostic procedures. In addition, the database records all drug prescriptions listed in the National Drug Formulary, and it allows identification of the prescriber. In Italy, all hospital medical doctors, including those who fill discharge diagnoses, are salaried.

The Regional Administrative Database has been used for pharmacoepidemiology and drug-related outcome researches.^{2 30–32} Each resident has a unique national identification code, with which it is possible to link the various types of information, corresponding to each person, within the database. In Italy, healthcare assistance is covered almost entirely by the Italian National Health System (NHS), therefore most residents' significant healthcare information can be found within the healthcare databases.

Source population

Source population will be represented by permanent residents aged 18 or above of Umbria Region. Any resident that has been discharged alive (with exclusion of voluntary discharge and interhospital transfer of patients) from a hospital with a diagnosis of AMI, AF, HF and stroke (ischaemic stroke, intracerebral haemorrhage, subarachnoidal haemorrhage) will be considered. Residents that have been hospitalised outside the regional territory of competence will be excluded from analysis.

Case selection and sampling method

Patients with the first occurrence of below diagnosis located in primary positioning of the administrative database between 2012 and 2014 will be identified using the following groups of ICD-9-CM codes: (1) 410 for AMI; (2) 427.31 for AF and 427.32 for atrial flutter; (3) 428 for HF; (4) 433.x1, 434 (excluding 434.x0), 436 for ischaemic stroke, 430 and 431 for haemorrhagic stroke (subarachnoid haemorrhage and intracerebral haemorrhage). It is worth mentioning that according to the Italian legislation, the principal diagnosis coding needs to be based on the condition identified at the end of hospitalisation, which constitutes the main cause of the need of treatment and/or diagnostic tests, and is mainly responsible for the use of resources.^{33 34} Table 1 displays the description of the ICD-9-CM codes for each disease of interest.

Validation criteria

Criteria used for validation of the considered ICD-9-CM codes will be those derived from international guidelines if available, or systematic reviews published on these topics.

Acute myocardial infarction

For validation of AMI, we will consider the Guidelines of the European Society of Cardiology³⁵ using the relative set of guidelines available at the time of each patient's discharge. To adjudicate the event as an AMI, in addition to troponin release (with at least one peak level greater than the upper limit of reference range), at least one of the following criteria of the reference standard needs to be met:

A. Cardiac symptoms at rest (eg, such as angina, chest pain, chest discomfort, chest heaviness, or substernal

Table 1 Description of ICD-9-CM codes

ICD-9-CM diagnosis code	Description
Acute myocardial infarction	410.0 Anterolateral wall 410.1 Other anterior wall 410.2 Inferolateral wall 410.3 Inferoposterior wall 410.4 Other inferior wall 410.5 Other lateral wall 410.6 True posterior wall infarction 410.7 Subendocardial infarction 410.8 Other specified sites 410.9 Unspecified site
Atrial fibrillation and flutter	427.31 Atrial fibrillation 427.32 Atrial flutter
Heart failure	428.0 Congestive heart failure, unspecified 428.1 Left heart failure 428.2 Systolic heart failure 428.20 Unspecified 428.21 Acute 428.22 Chronic 428.23 Acute on chronic 428.3 Diastolic heart failure 428.30 Unspecified 428.31 Acute 428.32 Chronic 428.33 Acute on chronic 428.4 Combined systolic and diastolic heart failure 428.40 Unspecified 428.41 Acute 428.42 Chronic 428.43 Acute on chronic 428.9 Heart failure, unspecified
Haemorrhagic stroke	430 Subarachnoid haemorrhage 431 Intracerebral haemorrhage
Ischaemic stroke	433 Occlusion and stenosis of precerebral arteries with cerebral infarction 433.01 Occlusion and stenosis of basilar artery 433.11 Occlusion and stenosis of carotid artery 433.21 Occlusion and stenosis of vertebral artery 433.31 Occlusion and stenosis of multiple and bilateral 433.81 Other specified occlusion and stenosis of precerebral artery 433.91 Unspecified occlusion and stenosis of precerebral artery 434 Occlusion of cerebral arteries with cerebral infarction 434.01 Cerebral thrombosis with cerebral infarction 434.11 Cerebral embolism with cerebral infarction 434.91 Cerebral artery occlusion, unspecified with cerebral infarction

chest pain), consistent with coronary artery disease/AMI.

- B. Evidence for a new AMI on the presenting ECG, as demonstrated by any one of these:
1. ST elevation (≥ 1 mm) in two contiguous leads;
 2. new Q waves;
 3. new left bundle branch block.
- C. Presence of myocardial infarction/injury (exclude old myocardial infarction) on any ECG during index hospital stay.

- D. Imaging evidence of new loss of viable myocardium, or new regional wall motion abnormality.
- E. Identification of an intracoronary thrombus by angiography.

Atrial fibrillation

To validate ICD-9 code 427.31 (index test) related to AF, we will require at least one ECG tracing documenting the presence of AF when present in the medical record. However, in the medical literature,³⁶ alternative

criteria have been used to identify cases of AF: (1) any mention of current AF, without a reference ECG; (2) two ECG measurements documenting the presence of AF in the same hospital admission; (3) the presence of at least two ECG documenting the presence of AF in more than one hospital admissions for AF in the same period. Accordingly, we will also record and analyse charts according to these different criteria for subsequent comparison, in order to understand which algorithm gives the best diagnostic yield.

Heart failure

For the validation of the ICD-9-CM 428.x relating to HF, we will consider the European Society of Cardiology Heart Failure Guidelines published at the time of patient's discharge.³⁷ Those guidelines consider the algorithm for the diagnosis of HF in the non-acute setting and in the acute setting. Since the ICD-9 codes do not distinguish between the acute and non-acute setting, we will combine the clinical presentation of both settings.

Diagnosis of HF will be adjudicated when, in addition to the presence of symptoms (such as dyspnoea, orthopnoea), or presence of signs at physical examination (rales, bilateral ankle oedema, increased jugular venous pressure, displaced apical beat), at least one of the following conditions are found in the medical chart:

1. any abnormality in resting ECG (ie, sinoatrial disease, atrioventricular block, or abnormal intraventricular conduction);
2. plasma natriuretic peptides (NPs) concentration (elevated levels of NPs are considered brain natriuretic peptide (BNP) ≥ 35 pg/mL and/or N-terminal pro-brain natriuretic peptide (NT-proBNP) ≥ 125 pg/mL);
3. echocardiography abnormalities (ventricular and atrial volumes and function) attributable to heart failure.

Stroke

For the validation of stroke, we will consider the ICD-9 codes related to stroke valid when both of the following conditions will be present:^{6 38 39} (1) detection of focal lesions by neurological examination; (2) imaging test (CT or MRI).

Neuroimaging will be the main discriminator among the two types of stroke: negative imaging for haemorrhage will classify the case as ischaemic stroke (ICD-9-CM codes: 433.x1, 434.xx (excluding 434.x0), 436, 437.1); neuroimaging that shows the presence of haemorrhagic lesion will classify the case as haemorrhagic stroke (codes 431, 430).

Chart abstraction and case ascertainment

The corresponding medical charts of the randomly selected sample cases for each ICD-9 code will be obtained from hospitals for validation purposes. From each medical chart, the following information will be retrieved: unique identification patient code, date of birth, gender, dates of hospital admission and discharge,

any diagnostic procedure and treatment that contributed to the diagnosis of the disease.

For each target disease, we will abstract data regarding clinical, laboratory and instrumental data, including the date or dates of performance. Specifically, for AMI, we will record any cardiac signs and symptoms, cardiac enzymes (eg, troponin levels), ECG data, other diagnostic instrumental examinations (ie, angiography); for AF, we will record any cardiac signs and symptoms, ECG data or other instrumental examinations (eg, echocardiography); for HF, we will record any cardiac symptoms and signs, plasma NPs concentration, ECG data and echocardiography measurements or other instrumental examinations of interest; for stroke, we will record any neurological signs and symptoms, imaging test (CT or MRI), any other laboratory or instrumental measurement.

Two medical doctors acting as chart reviewers will receive specific training on data abstraction. An initial chart review will be performed, with each reviewer independently examining the same medical charts (n=20). The interreviewer agreement between the two reviewers regarding presence or absence of the diseases considered among the pairs of reviewers will be calculated using κ statistics. This process will be reiterated until the strength of agreement among the pairs of reviewers will be optimal (κ statistics between 0.81 and 1.00). Any discrepancies between the two reviewers will be resolved by consensus and where necessary, a third expert will be involved.

Case adjudication of disease within medical charts will be based on symptoms, laboratory and diagnostic tests for each cerebrovascular and cardiovascular disease considered.

Statistical analysis

For each condition, we anticipate that a sample of 121 charts of cases will be necessary to obtain an expected sensitivity of 80% with a precision of 8% and a power of 80% according to binomial exact calculation. For an expected specificity of 90% (precision of 8% and a power of 80%), we will randomly select from the same administrative database 73 non-cases, that is, medical records belonging to diseases of the circulatory system (ICD-9-CM codes 390–459) but without the ICD-9-codes of interest. Expected accuracy figures were based on the published literature.^{6–9}

Sensitivity and specificity will be analysed separately for each ICD-9-CM code by constructing 2x2 tables. Sensitivity expresses the proportion of 'true positives' (ie, AMI classified as positive by the administrative database and medical record review) and all cases deemed positive by medical chart review. Specificity expresses the proportion of 'true negatives' (ie, AMI identified as negative by the administrative database and medical record review), and with all cases deemed negative by medical chart review. For both sensitivity and specificity, 95% CIs will be calculated. We account for a 10% of missing charts for sensitivity and specificity. Positive and negative predicting values will also be reported along with their 95% CI.

Reporting

Complete, transparent and accurate reporting is essential in research,⁴⁰ because it allows readers to assess internal validity as well as to evaluate the generalisability and applicability of results.⁴¹ To ensure quality and thoroughness of reporting, any dissemination or publication of the results from the present study will follow recommended guidelines based on the criteria published by the Standards for Reporting of Diagnostic accuracy (STARD) initiative for the accurate reporting of investigations of diagnostic studies.^{41–43} The project will be completed within 2 years.

DISCUSSION

Administrative databases constitute a valid instrument to get insights into epidemiology and health management, and to get ‘real-world’ knowledge, of major disease under situations in which randomised trials are not able to provide the required evidence, for practical or economic reasons, or because of selective inclusion criteria. Epidemiological studies are frequently based on administrative claims databases to identify cases of specific diseases, such as AMI, and often contain ICD-9 codes. These codes have the advantage of being widely available and require lower effort and cost than consulting medical charts.²⁰ However, their effective exploitation hinges on the prerequisite that coding does reliably reflect the clinical condition of a given patient.

In this protocol, we present the approach to be used to analyse the validity of ICD-9-CM codes for important cardiovascular diseases, such as AMI, HF, AF, and stroke using the Regional Health database of Umbria.

We seek to verify whether an acceptable level (at least 80% for sensitivity and 90% for specificity) of operating characteristics is in fact achieved for each identified ICD-9-CM codes. In case of inadequate level of validation, algorithms will be considered to explore potential elements to enhance the level of validation. Elements of an algorithm may include the type of the data source, number of years of administrative data, other diagnostic codes, medication use or laboratory data. The algorithms selected for evaluation will be based on the literature review and consultations with clinicians, health services researchers. If the desired accuracy is not reached despite the use of algorithms, results of our research will be used to identify potential areas for improvement for future validation.

DISSEMINATION

Study results will be disseminated widely through peer-reviewed publications and presentations at national and international conferences.

Twitter Follow Paolo Eusebi @paolousebi

Contributors IA, FC, MO, AMo and GA conceived the study. IA, FC, MO, GA, AMe, MFC, PE and AMo were responsible for designing the protocol. IA, FC, MO, GA and AMo drafted the protocol manuscript. IA, FC and MO developed

the search strategy. IA, FC, MO, GA, AMe, MFC, PE and AMo critically revised the successive versions of the manuscript and approved the final version.

Funding This review protocol was funded by the Regional Health Authority of Umbria. [Progetto Data-Value: valorizzazione del dato sanitario regionale per la Ricerca dei Servizi Sanitari (Health Services Research)].

Disclaimer The study funder was not involved in study design or writing of the protocol.

Competing interests None declared.

Ethics approval Ethics approval has been obtained from the Regional Ethics Committee of Umbria (CEAS).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Study results will be disseminated widely through peer-reviewed publications and presentations at national and international conferences.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health* 2011;32:91–108.
- Traversa G, Bianchi C, Da Cas R, *et al.* Cohort study of hepatotoxicity associated with nimesulide and other non-steroidal anti-inflammatory drugs. *BMJ* 2003;327:18–22.
- Trifiro G, Patadia V, Schuemie MJ, *et al.* EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform* 2011;166:25–30.
- Gini R, Francesconi P, Mazzaglia G, *et al.* Chronic disease prevalence from Italian administrative databases in the VALORE project: a validation through comparison of population estimates with general practice databases and national survey. *BMC Public Health* 2013;13:15.
- Colais P, Pinnarelli L, Fusco D, *et al.* The impact of a pay-for-performance system on timing to hip fracture surgery: experience from the Lazio Region (Italy). *BMC Health Serv Res* 2013;13:393.
- McCormick N, Bhole V, Lacaille D, *et al.* Validity of diagnostic codes for acute stroke in administrative databases: a systematic review. *PLoS ONE* 2015;10:e0135834.
- McCormick N, Lacaille D, Bhole V, *et al.* Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS ONE* 2014;9:e104519.
- McCormick N, Lacaille D, Bhole V, *et al.* Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS ONE* 2014;9:e92286.
- Jensen PN, Johnson K, Floyd J, *et al.* A systematic review of validated methods for identifying atrial fibrillation using administrative data. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 1):141–7.
- Abraha I, Orso M, Grilli P, *et al.* The current state of validation of administrative healthcare databases in Italy: a systematic review. *Int J Stat Med Res* 2014;3:309–20.
- Galdarossa M, Vianello F, Tezza F, *et al.* Epidemiology of primary and secondary thrombocytopenia: first analysis of an administrative database in a major Italian institution. *Blood Coagul Fibrinolysis* 2012;23:271–7.
- Franchi C, Giussani G, Messina P, *et al.* Validation of healthcare administrative data for the diagnosis of epilepsy. *J Epidemiol Community Health* 2013;67:1019–24.
- Tinelli M, Mannino S, Lucchi S, *et al.* Healthcare-acquired infections in rehabilitation units of the Lombardy Region, Italy. *Infection* 2002;30:353–8.
- Fano V, D'Ovidio M, del Zio K, *et al.* [The role of the quality of hospital discharge records on the comparative evaluation of outcomes: the example of chronic obstructive pulmonary disease (COPD)]. *Epidemiol Prev* 2011;36:172–9.
- Faustini A, Canova C, Cascini S, *et al.* The reliability of hospital and pharmaceutical data to assess prevalent cases of chronic obstructive pulmonary disease. *COPD* 2012;9:184–96.

16. Bogliun G, Beghi E. Validity of hospital discharge diagnoses for public health surveillance of the Guillain-Barré syndrome. *Neurol Sci* 2002;23:113–17.
17. Baldi I, Vicari P, Di Cuonzo D, *et al*. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. *J Clin Epidemiol* 2008;61:373–9.
18. Schifano P, Papini P, Agabiti N, *et al*. Indicators of breast cancer severity and appropriateness of surgery based on hospital administrative data in the Lazio Region, Italy. *BMC public health* 2006;6:25.
19. Rawson NSB, Shatin D. Assessing the validity of diagnostic data in large administrative healthcare utilization databases. In: Hartzema A, Tilson H, Chan K, eds. *Pharmacoepidemiology and therapeutic risk management*. Harvey Whitney Books, 2008:495–518.
20. West SL, Ritchey ME, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. Wiley-Blackwell, 2012:757–94.
21. Campbell SE, Campbell MK, Grimshaw JM, *et al*. A systematic review of discharge coding accuracy. *J Public Health Med* 2001;23:205–11.
22. World Health Organisation. Global Health Estimates. 2014 Summary tables: Deaths by cause, age and sex, by WHO region, 2000–2012, 2014 2014. http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html (accessed 23 Oct 2015).
23. McMurray JJ, Pfeffer MA. Heart failure. *Lancet* 2005;365:1877–89.
24. Carluccio E, Dini FL, Biagioli P, *et al*. The 'Echo Heart Failure Score': an echocardiographic risk prediction score of mortality in systolic heart failure. *Eur J Heart Fail* 2013;15:868–76.
25. Angeli F, Rebaldi G, Garofoli M, *et al*. Atrial fibrillation and mortality in patients with acute myocardial infarction: a systematic overview and meta-analysis. *Curr Cardiol Rep* 2012;14:601–10.
26. Abraha I, Giovannini G, Serraino D, *et al*. Validity of breast, lung and colorectal cancer diagnoses in administrative databases: a systematic review protocol. *BMJ Open* 2016;6:e010409.
27. Montedori A, Abraha I, Chiatti C, *et al*. Validity of peptic ulcer disease and upper gastrointestinal bleeding diagnoses in administrative databases: a systematic review protocol. *BMJ Open* 2016;6:e011776.
28. Rimland JM, Abraha I, Luchetta ML, *et al*. Validation of chronic obstructive pulmonary disease (COPD) diagnoses in healthcare databases: a systematic review protocol. *BMJ Open* 2016;6:e011777.
29. Abraha I, Serraino D, Giovannini G, *et al*. Validity of ICD-9-CM codes for breast, lung and colorectal cancers in three Italian administrative healthcare databases: a diagnostic accuracy study protocol. *BMJ Open* 2016;6:e010547.
30. Menniti-Ippolito F, Maggini M, Raschetti R, *et al*. Ketorolac use in outpatients and gastrointestinal hospitalization: a comparison with other non-steroidal anti-inflammatory drugs in Italy. *Eur J Clin Pharmacol* 1998;54:393–7.
31. Abraha I, Montedori A, Stracci F, *et al*. Statin compliance in the Umbrian population. *Eur J Clin Pharmacol* 2003;59:659–61.
32. Raschetti R, Maggini M, Da Cas R, *et al*. Time trends in the coprescribing of cisapride and contraindicated drugs in Umbria, Italy. *JAMA* 2001;285:1840–1.
33. Italian Ministerial Decree no. 380 of 27 October 2000. <http://www.privacyit/dm2000380.html>
34. Guidelines for national hospital discharge data coding, approved by the Italian State-Regions Conference www.statoregioni.it/DettaglioDoc.asp?IDDoc=26555&IdProv=8266&tipodoc=2&CONF=CSR (rep. Acts 64 / CSR of 29 April 2010).
35. Thygesen K, Alpert JS, Jaffe AS, *et al*. Third universal definition of myocardial infarction. *Eur Heart J* 2012;33:2551–67.
36. Camm AJ, Lip GY, De Caterina R, *et al*. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *Eur Heart J* 2012;33:2719–47.
37. McMurray JJ, Adamopoulos S, Anker SD, *et al*. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012;33:1787–847.
38. Stroke. *National clinical guideline for diagnosis and initial management of acute stroke and Transient Ischaemic Attack (TIA)*. Royal College of Physicians, UK London, 2008.
39. Sacco RL, Kasner SE, Broderick JP, *et al*. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2013;44:2064–89.
40. Macura A, Abraha I, Kirkham J, *et al*. Selective outcome reporting: telling and detecting true lies. The state of the science. *Intern Emerg Med* 2010;5:151–5.
41. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41–4.
42. Benchimol EI, Manuel DG, To T, *et al*. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol* 2011;64:821–9.
43. De Coster C, Quan H, Finlayson A, *et al*. Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium. *BMC Health Serv Res* 2006;6:77.