

BMJ Open Validity of ICD-9-CM codes for breast, lung and colorectal cancers in three Italian administrative healthcare databases: a diagnostic accuracy study protocol

Iosief Abraha,¹ Diego Serraino,² Gianni Giovannini,¹ Fabrizio Stracci,³ Paola Casucci,⁴ Giuliana Alessandrini,⁴ Ettore Bidoli,² Rita Chiari,⁵ Roberto Ciocchi,⁶ Marcello De Giorgi,⁴ David Franchini,⁴ Maria Francesca Vitale,⁷ Mario Fusco,⁷ Alessandro Montedori¹

To cite: 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. doi:10.1136/bmjopen-2015-010547

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-010547>).

Received 16 November 2015
Revised 31 January 2016
Accepted 22 February 2016



CrossMark

For numbered affiliations see end of article.

Correspondence to

Dr Iosief Abraha;
iosief_a@yahoo.it;
abraha@regione.umbria.it

ABSTRACT

Introduction: Administrative healthcare databases are useful tools to study healthcare outcomes and to monitor the health status of a population. Patients with cancer can be identified through disease-specific codes, prescriptions and physician claims, but prior validation is required to achieve an accurate case definition. The objective of this protocol is to assess the accuracy of International Classification of Diseases Ninth Revision—Clinical Modification (ICD-9-CM) codes for breast, lung and colorectal cancers in identifying patients diagnosed with the relative disease in three Italian administrative databases.

Methods and analysis: Data from the administrative databases of Umbria Region (910 000 residents), Local Health Unit 3 of Napoli (1 170 000 residents) and Friuli-Venezia Giulia Region (1 227 000 residents) will be considered. In each administrative database, patients with the first occurrence of diagnosis of breast, lung or colorectal cancer between 2012 and 2014 will be identified using the following groups of ICD-9-CM codes in primary position: (1) 233.0 and (2) 174.x for breast cancer; (3) 162.x for lung cancer; (4) 153.x for colon cancer and (5) 154.0–154.1 and 154.8 for rectal cancer. Only incident cases will be considered, that is, excluding cases that have the same diagnosis in the 5 years (2007–2011) before the period of interest. A random sample of cases and non-cases will be selected from each administrative database and the corresponding medical charts will be assessed for validation by pairs of trained, independent reviewers. Case ascertainment within the medical charts will be based on (1) the presence of a primary nodular lesion in the breast, lung or colon–rectum, documented with imaging or endoscopy and (2) a cytological or histological documentation of cancer from a primary or metastatic site. Sensitivity and specificity with 95% CIs will be calculated.

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

Strengths and limitations of this study

- The study will evaluate the validity of the International Classification of Diseases–Ninth Revision—Clinical Modification (ICD-9-CM) codes for breast, lung and colorectal cancers in three large Italian administrative databases.
- The strength of this study is that it will use a medical chart review to ascertain cases of cancer diseases.
- Once these administrative databases are validated for breast, lung and colorectal cancer diseases, they can be used for outcome research including pharmacoepidemiology, health service research and quality of care research.
- This study will be the first to validate ICD-9-CM codes of three cancers in three large administrative databases in Italy.
- Validation studies of administrative data are related to that context and are not generalisable to other settings.

INTRODUCTION

As computer technology continues to advance, administrative databases are increasingly growing in numerous healthcare settings worldwide. These databases anonymously store data about patients regarding the healthcare assistance they received, including birth, death or disease treatment. Usually, the diagnosis of the disease is associated with a specific code from the International Classification of Diseases, Ninth Revision (ICD-9) or 10th Revision (ICD-10) edition. The ICD is designed to map health conditions to corresponding generic categories together with specific variations.¹

The merging of individual patient data from administrative databases with other sources (eg, prescription and laboratory data) allows one to investigate a wide range of relevant and often unique public health questions,² monitor population health status over time and perform population-based pharmacoepidemiological research.²⁻⁴

To constitute a reliable source for research studies, adequate validation of administrative healthcare databases is mandatory. While non-clinical information in healthcare databases, such as demographic and prescription data, are highly accurate,^{5 6} the validity of registered diagnoses and procedures is variable.^{6 7} Determining the accuracy of the latter two categories of clinical information is vitally important to all potential users and involves confirming the consistency of information within the databases with the corresponding clinical records of patients.⁵

In Italy, all the Regional Health Authorities maintain large healthcare information systems containing patient data from all hospital and territorial sources. These databases have the potential to address important issues in postmarketing surveillance,^{8 9} epidemiology,¹⁰ quality performance and health services research.¹¹ However, there is a concern that their considerable potential as a source of reliable healthcare information has not been realised since they have not been widely validated. A systematic review of ICD-9 code validation in Italian administrative databases¹² reported that only a few regional databases have been validated for a limited number of ICD-9 codes of diseases including stroke,^{13 14} gastrointestinal bleeding,¹⁵ thrombocytopenia,¹⁶ epilepsy,¹⁷ infections,¹⁸ chronic obstructive pulmonary disease,^{19 20} Guillain-Barré syndrome²¹ and cancers.^{22 23} In addition, the use of these databases was scarce, as only six administrative databases served as sources for published research articles based on the validated ICD-9 codes. Hence, it is imperative that Regional Health Authorities systematically validate their databases for critical diseases to productively use the information they contain.

Breast, colorectal and lung cancers are the most commonly diagnosed neoplasms worldwide, as well as in Italy.²⁴ Consequently, they generate interest in the scientific community and industry as targets for the development of new drugs and for governments, given that they are an important cause of public health and economic burden. For example, variation in the epidemiology of breast,²⁵ colorectal^{26 27} and lung²⁸ cancers, treatment (pharmacological or surgical) administered to patients suffering from these cancers and potential clinical and economic outcomes²⁹⁻³¹ can all be evaluated using validated administrative databases.

The objective of the present protocol is to evaluate the accuracy of the ICD-9-CM codes related to breast, lung and colorectal cancers in correctly identifying the respective diseases using three large Italian administrative healthcare databases.

METHODS

Setting and data source

Administrative databases

Starting from the early 1990s, local and regional Italian healthcare administrative databases have collected information from all patient medical records from public and private hospitals including demographics, hospital admission and discharge dates, vital statistics, the admitting hospital department, the principal diagnosis and a maximum of five secondary discharge diagnoses, and the principal and five secondary, surgical and diagnostic procedures. In addition, these databases contain the records of all drug prescriptions listed in the National Drug Formulary and the basic characteristics of patients' physicians. 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.

The target administrative databases for the present study will be from the Umbria Region (910 000 residents), Local Health Unit 3 of Napoli (1 170 000 residents) and the Friuli-Venezia Giulia Region (1 227 000 residents). For each database, the corresponding Unit (Regional Health Authority of Umbria for Umbria Region, Registro Tumori Regione Campania for Local Health Unit 3 of Napoli and Centro di Riferimento Oncologico Aviano for Friuli-Venezia Giulia Region) will conduct the same validation process.

Source population

The source population will be represented by permanent residents aged 18 years or above of Umbria Region, Local Health Unit 3 of Napoli and the Friuli-Venezia Giulia Region. Any resident who has been discharged from hospital with a diagnosis of breast, lung or colorectal cancer will be considered. Residents who have been hospitalised outside the regional territory of competence will be excluded from analysis due to the difficulty in obtaining the medical charts.

Case selection and sampling method

In each administrative database, patients with the first occurrence of diagnosis of breast, lung or colorectal cancer between 2012 and 2014 will be identified using the following groups of ICD-9-CM codes located in primary position: (1) 233.0 and (2) 174.x for breast cancer; (3) 162.x for lung cancer; (4) 153.x for colon cancer and (5) 154.0–154.1 and 154.8 for rectal cancer. Only incident cases will be considered, that is, excluding cases with the same diagnosis (ICD-9-CM codes in any position) in the 5 years (2007–2011) before the period of interest. Subsequently, for each of the above reported groups of ICD-9-CM codes, a random sample of cases

will be selected from each administrative database. Table 1 displays the description of the ICD-9-CM codes for each of the cancer diseases of interest.

Chart abstraction and case ascertainment

The corresponding medical charts of the randomly selected sample cases will be obtained from hospitals for validation purposes. From each medical chart, the following information will be retrieved: initials of the patient, date of birth, sex, dates of hospital admission and discharge, any diagnostic procedure that contributed to the diagnosis of the cancer, any pharmacological or surgical intervention that was provided for the treatment of the cancer.

Table 1 Description of the ICD-9-CM codes related to breast, lung and colorectal cancers

Condition	ICD-9-CM diagnosis code
Breast	Carcinoma in situ: 233.0
	Malignant neoplasm of the female breast
	174.0 nipple and areola
	174.1 central portion
	174.2 upper-inner quadrant
	174.3 lower-inner quadrant
	174.4 upper-outer quadrant
	174.5 lower-outer quadrant
	174.6 axillary tail
	174.8 other specified sites of the female breast
Lung	174.9 breast female, unspecified
	Malignant neoplasm of the trachea, bronchus and lung
	162.0 Trachea
	162.2 main bronchus
	162.3 upper lobe, bronchus or lung
	162.4 middle lobe, bronchus or lung
	162.5 lower lobe, bronchus or lung
	162.8 other parts of the bronchus or lung
Colorectal	162.9 bronchus and lung, unspecified
	Malignant neoplasm of the colon
	153.0 hepatic flexure
	153.1 transverse colon
	153.2 descending colon
	153.3 sigmoid colon
	153.4 caecum
	153.5 appendix
	153.6 ascending colon
	153.7 splenic flexure
	153.8 other specified sites of the large intestine
	153.9 colon, unspecified
	Malignant neoplasm of the rectum and rectosigmoid junction
	154.0 rectosigmoid junction
	154.1 rectum
	154.8 other

ICD-9-CM, International Classification of Diseases Ninth Revision
—Clinical Modification.

Within each unit, two reviewers will receive training on data abstraction. An initial consensus chart review will be performed with each reviewer independently examining the same number of medical charts ($n=20$). The inter-rater agreement regarding the presence or absence of breast, lung or colorectal cancer among the pairs of reviewers within each unit will be calculated using the κ statistics. This process will be repeated until the strength of agreement among the pairs of reviewers will be near perfect (κ statistics between 0.81 and 1.00). Any discrepancies will be discussed and resolved through third party involvement (RC).

Case ascertainment of cancer within a medical chart will be based on (1) the presence of a primary nodular lesion in the breast, lung or colon-rectum, documented with imaging or endoscopy and (2) the cytological or histological documentation of cancer from a primary or metastatic site.

Following consensus review, data abstraction will be completed independently. To ensure consistency among all the reviewers, cases with uncertainty will be discussed and resolved through third party involvement (RC).

Validation criteria

For non-invasive breast cancer, we will consider the ICD-9-CM code 233.0 valid when there is evidence of a breast nodule documented with imaging (eg, mammography) and a histological diagnosis of ductal or lobular breast carcinoma in situ (pTis).

For invasive breast cancer, we will consider the ICD-9-CM codes 174.x valid when there is evidence of a breast nodule documented with imaging (eg, mammography) and a cytological or histological diagnosis from a primary or metastatic site positive for ductal or lobular adenocarcinoma.

For lung cancer, we will consider the ICD-9-CM codes 162.x valid when there is evidence of a pulmonary nodule documented with imaging (eg, CT scan) and a cytological or histological diagnosis from a primary or metastatic site positive for either small cell lung cancer (microcitoma) or non-small cell lung cancer (NSCLC).

For colon cancer, we will consider the ICD-9-CM codes 153.x valid when there is evidence of a neoplastic lesion within the colon, documented with endoscopy (eg, colonoscopy) or imaging (eg, barium enema), and a histological diagnosis from a primary or metastatic site positive for adenocarcinoma, squamous cell carcinoma or neuroendocrine carcinoma.

For rectal cancer, we will consider the ICD-9-CM codes 154.0–154.1 and 154.8 valid when there is evidence of a neoplastic lesion in the rectosigmoid junction or the rectum, documented with endoscopy (eg, colonoscopy) or imaging (eg, barium enema), and a histological diagnosis from a primary or metastatic site positive for adenocarcinoma or squamous cell carcinoma.

Statistical analysis

We calculated that a sample of 130 charts of cases will be necessary to obtain an expected sensitivity of 80%

with a precision of 10% and a power of 80%. For specificity calculation, we will randomly select non-cases, that is, records without the ICD-9-codes of interest from an administrative database. The corresponding medical charts will be retrieved and evaluated. We calculated that a sample of 94 charts of non-cases will be retrieved to obtain an expected specificity of 90% with a precision of 10% and a power of 80%. Overall, each unit will evaluate 1120 charts.

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, cancer cases classified as positive by both the administrative database and medical record review) and all cases deemed positive by the medical chart review. Specificity expresses the proportion of 'true negatives' (ie, cases without cancer identified by both the administrative database and medical record review), and with all cases deemed negative by the medical chart review. For both sensitivity and specificity, 95% CIs will be calculated.

Reporting

Complete, transparent and accurate reporting is essential in diagnostic accuracy studies because it allows readers to assess internal validity as well as to evaluate the generalisability and applicability of results.³² To ensure quality reporting, any reporting or publication of the results from this 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.^{32–34}

DISCUSSION

In this protocol, we present the approach we will use to analyse the validity of ICD-9-CM codes for breast, lung and colorectal cancers in administrative databases representing northern, central and southern Italy.

Administrative databases constitute a valid alternative to situations in which randomised trials are not able to provide the required evidence for practical or economic reasons. In addition, despite epidemiological studies on cancer being frequently based on cancer registries,^{35–37} administrative databases can add a further value especially on pharmacoepidemiology^{3 12 38} and health services research.^{39 40}

Accurate identification of cancer cases using the ICD-9-CM codes may contribute to monitoring cancer trends and to proposing interventions to ameliorate cancer care. In 2008, an Italian study developed and validated an algorithm using a regional administrative database to determine incident cases of breast, lung and colorectal cancers and found a sensitivity of 76.7%, 80.8% and 72.4%, respectively, for the three cancers.²² This study will add value to the knowledge of the three cancer diseases given that it covers different areas of Italy.

Ethics and dissemination

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

Author affiliations

- ¹Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy
- ²Epidemiology and Biostatistic Unit, Centro di Riferimento Oncologico Aviano, Aviano, Italy
- ³Public Health Department, University of Perugia, Perugia, Italy
- ⁴Health ICT Service, Regional Health Authority of Umbria, Perugia, Italy
- ⁵Dipartimento di Oncologia, Azienda Ospedaliera Perugia, Perugia, Italy
- ⁶Department of Digestive Surgery and Liver Unit, University of Perugia, Perugia, Italy
- ⁷Registro Tumori Regione Campania, ASL NA3 Sud, Brusciano, Italy

Contributors AM, IA, MF and DS conceived the study and all authors were responsible for designing the protocol. IA and AM drafted the protocol manuscript. IA, DS, GG, FS, PC, GA, EB, RCh, RCi, MDG, DF, MFV, MF and AM critically revised the successive versions of the manuscript and approved the final version.

Funding This study protocol was developed within the D.I.V.O. project (Realizzazione di un Database Interregionale Validato per l'Oncologia quale strumento di valutazione di impatto e di appropriatezza delle attività di prevenzione primaria e secondaria in ambito oncologico) supported by funding from the National Centre for Disease Prevention and Control (CCM 2014), Ministry of Health, Italy. The study funder was not involved in the study design or the writing of the protocol.

Competing interests None declared.

Ethics approval Regional Committee Ethics of Umbria (CEAS).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All raw data will be available from the corresponding author.

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

1. World Health Organization. *International statistical classification of diseases and health related problems, 10th revision*. Geneva: WHO, 1992.
2. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health* 2011;32:91–108.
3. Jick SS. Fresh evidence confirms links between newer contraceptive pills and higher risk of venous thromboembolism. *BMJ* 2015;350:h2422.
4. García Rodríguez LA, Pérez-Gutthann S, Jick S. *The UK general practice research database. Pharmacoepidemiology*. John Wiley & Sons, Ltd, 2002:375–85.
5. 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, 2008:495–517.
6. West SL, Ritchey ME, Poole C. *Validity of pharmacoepidemiologic drug and diagnosis data. Pharmacoepidemiology*. Wiley-Blackwell, 2012:757–94.
7. Campbell SE, Campbell MK, Grimshaw JM, et al. A systematic review of discharge coding accuracy. *J Public Health Med* 2001;23:205–11.
8. 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.
9. 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.
10. 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.
 11. 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.
 12. Abraha I, Orso M, Grilli P, *et al.* The current state of validation of administrative healthcare databases in Italy: a systematic review. *J Stat Med Res* 2014;3:309–20.
 13. Leone MA, Capponi A, Varrasi C, *et al.* Accuracy of the ICD-9 codes for identifying TIA and stroke in an Italian automated database. *Neurol Sci* 2004;25:281–8.
 14. Rinaldi R, Vignatelli L, Galeotti M, *et al.* Accuracy of ICD-9 codes in identifying ischemic stroke in the general hospital of Lugo di Romagna (Italy). *Neurol Sci* 2003;24:65–9.
 15. Cattaruzzi C, Troncon MG, Agostinis L, *et al.* Positive predictive value of ICD-9th codes for upper gastrointestinal bleeding and perforation in the Sistema Informativo Sanitario Regionale database. *J Clin Epidemiol* 1999;52:499–502.
 16. 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.
 17. 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.
 18. Tinelli M, Mannino S, Lucchi S, *et al.* Healthcare-acquired infections in rehabilitation units of the Lombardy Region, Italy. *Infection* 2011;39:353–8.
 19. 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.
 20. 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.
 21. Bogliun G, Beghi E. Validity of hospital discharge diagnoses for public health surveillance of the Guillain-Barré syndrome. *Neurol Sci* 2002;23:113–17.
 22. 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.
 23. 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.
 24. Ferlay J, Soerjomataram I, Dikshit R, *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.
 25. Yuan Y, Li M, Yang J, *et al.* Using administrative data to estimate time to breast cancer diagnosis and percent of screen-detected breast cancers—a validation study in Alberta, Canada. *Eur J Cancer Care* 2015;24:367–75.
 26. Deshpande AD, Schootman M, Mayer A. Development of a claims-based algorithm to identify colorectal cancer recurrence. *Ann Epidemiol* 2015;25:297–300.
 27. Quantin C, Benzenine E, Hägi M, *et al.* Estimation of national colorectal-cancer incidence using claims databases. *J Cancer Epidemiol* 2012;2012:298369.
 28. Abdulmalak C, Cottenet J, Beltramo G, *et al.* Haemoptysis in adults: a 5-year study using the French nationwide hospital administrative database. *Eur Respir J* 2015;46:503–11.
 29. Dehal A, Abbas A, Johna S. Comorbidity and outcomes after surgery among women with breast cancer: analysis of nationwide in-patient sample database. *Breast Cancer Res Treat* 2013;139:469–76.
 30. Konski A. Clinical and economic outcomes analyses of women developing breast cancer in a managed care organization. *Am J Clin Oncol* 2005;28:51–7.
 31. Mittmann N, Liu N, Porter J, *et al.* Utilization and costs of home care for patients with colorectal cancer: a population-based study. *CMAJ Open* 2014;2:E11–17.
 32. 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.
 33. 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.
 34. 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.
 35. Guzzinati S, Buzzoni C, De Angelis R, *et al.* Cancer prevalence in Italy: an analysis of geographic variability. *Cancer Causes Control* 2012;23:1497–510.
 36. Nicita C, Buzzoni C, Chellini E, *et al.* [A comparative analysis between regional mesothelioma registries and cancer registries: results of the ReNaM-AIRTUM project]. *Epidemiol Prev* 2014;38:191–9.
 37. Zorzi M, Mangone L, Sassatelli R, *et al.* Screening for colorectal cancer in Italy: 2011–2012 survey. *Epidemiol Prev* 2015;39(3 Suppl 1):115–25.
 38. Nordstrom BL, Whyte JL, Stolar M, *et al.* Identification of metastatic cancer in claims data. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 2):21–8.
 39. Franchi C, Lucca U, Tettamanti M, *et al.* Cholinesterase inhibitor use in Alzheimer's disease: the EPIFARM-Elderly Project. *Pharmacoepidemiol Drug Saf* 2011;20:497–505.
 40. Rosato R, Sacerdote C, Pagano E, *et al.* Appropriateness of early breast cancer management in relation to patient and hospital characteristics: a population based study in Northern Italy. *Breast Cancer Res Treat* 2009;117:349–56.