BMJ Open

Sensitivity and Specificity of Breast Cancer ICD-9-CM Codes in Three Italian Administrative Healthcare Databases

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020627
Article Type:	Research
Date Submitted by the Author:	14-Nov-2017
Complete List of Authors:	Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service; Azienda Ospedaliera di Perugia, Centro Regionale Sangue Serraino, Diego; IRCCS CRO AVIANO, Epidemiology Montedori, Alessandro; Regional Health Authority of Umbria, Fusco, Mario; ASL NA3 Sud, Registro Tumori Regione Campania Giovannini, Gianni; Regional Health Authority of Umbria, Health Planning Office Casucci, Paola; Regional Health Authority of Umbria, Health ICT Service Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Granata, Annalisa; ASL NA3 Sud, Registro Tumori Regione Campania De Giorgi, Marcello; Regional Health Authority of Umbria, , Health ICT Service Collarile, Paolo; Centro di Riferimento Oncologico Aviano, SOC Epidemiologia Oncologica Chiari, Rita; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Foglietta, Jennifer; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Vitale, Maria Francesca; Registro Tumori Regione Campania , ASL NA3 Sud Stracci, Fabrizio; University of Perugia, Public Health Department Orlandi, Walter; Regional Health Authority of Umbria Bidoli, Ettore; IRCCS CRO AVIANO, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Oncology
Keywords:	sensitivity and specificity, administrative database, validity, breast cancer

Sensitivity and Specificity of Breast Cancer ICD-9-CM Codes in Three Italian Administrative Healthcare Databases

Iosief Abraha[1 2], Diego Serraino[3], Alessandro Montedori[1], Mario Fusco[4], Gianni Giovannini[1], Paola Casucci[5], Francesco Cozzolino[1], Massimiliano Orso[1], Annalisa Granata[4], Marcello De Giorgi[5], Paolo Collarile[6], Rita Chiari[7], Jennifer Foglietta[7], Maria Francesca Vitale[4], Fabrizio Stracci[8], Walter Orlandi[9], Ettore Bidoli[3] for the D.I.V.O. Group

Author affiliations:

1. *Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy*
2. *Centro Regionale Sangue, Azienda Ospedaliera di Perugia, Perugia, Italy*
3. *Cancer Epidemiology Unit, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
4. *Registro Tumori Regione Campania, ASL Napoli 3 Sud, Brusciano, Italy*
5. *Health ICT Service, Regional Health Authority of Umbria, Perugia, Italy*
6. *SOC Epidemiologia Oncologica, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
7. *Dipartimento di Oncologia, Azienda Ospedaliera Perugia, Perugia, Italy*
8. *Public Health Department, University of Perugia, Perugia, Italy*
9. *Regional Health Authority of Umbria, Perugia, Italy*

Corresponding Author:
Alessandro Montedori
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni 61
06124 Perugia, Italy
amontedori@regione.umbria.it

Abstract

Objectives To assess the accuracy of *International Classification of Diseases 9th Revision Clinical Modification* (ICD-9-CM) codes in identifying patients diagnosed with incident carcinoma in situ and invasive breast cancer in three Italian administrative databases.

Design A diagnostic accuracy study comparing ICD-9-CM codes for carcinoma in situ (233.0) and for invasive breast cancer (174.x) with medical chart (as a reference standard). Case definition: (a) presence of a primary nodular lesion in the breast and (b) cytological or histological documentation of cancer from a primary or metastatic site.

Setting Administrative databases from Umbria Region, ASL 3 Napoli Sud (NA), and Friuli Venezia Giulia (FVG) Region

Participants Women with breast carcinoma in situ or invasive breast cancer diagnosed (in primary position) between 2012 and 2014.

Outcome measures Sensitivity and specificity for codes 233.0 and 174.x.

Results For invasive breast cancer the sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. Specificities were 90% (95% CI 82% to 95%) for Umbria, 91% (95% CI 83% to 96%) for NA and 91% (95% CI 84% to 96%) for FVG.

For carcinoma in situ the sensitivities were 100% (95% CI 93% to 100%) for Umbria, 100% (95% CI 95% to 100%) for NA, and 100% (95%CI 96% to 100%) for FVG. Specificities were 98% (95% CI 93% to 100%) for Umbria, 86% (95% CI 78% to 92%) for NA, and 90% (95% CI 82% to 95%) for FVG.

Conclusions The present study showed that administrative healthcare databases from Umbria, NA and FVG can be used to identify the majority of women with newly diagnosed carcinoma of the breast. The proposed case definition is a powerful tool to perform research on large populations of newly diagnosed breast cancer patients.

Strengths and limitations of this study

This study is the first to have validated *International Classification of Diseases-9th Revision – Clinical Modification* (ICD-9-CM) codes for incident breast cancer cases in three large computerized administrative databases in Italy using the same case definition.

Case ascertainment was based on the presence of a primary nodular lesion in the breast documented by imaging and a cytological or histological documentation of cancer from a primary or metastatic site.

This study followed 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.

Validation studies of administrative data are related to that context and are not generalizable to other settings.

The sample size of women with carcinoma in situ was limited due to the low prevalence of disease.

Introduction

The use of administrative databases is increasingly growing in various healthcare settings worldwide. These databases anonymously store data about residents regarding the healthcare assistance they receive including hospital admission, demographic data and disease treatment. Usually, the diagnosis of the disease is associated with a specific code from the *International Classification of Diseases, 9th 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 networking of individual patient data from administrative databases and other sources such as outpatient data and prescription data allows monitoring population health status, performing outcome research [2-4] and exploring a wide range of significant public health questions [2].

In administrative databases, while non-clinical data such as demographic or prescription data are highly accurate [5 6], the accuracy of diagnoses and procedures needs to be determined [6 7]. Typically, the assessment of accuracy consists in confirming the reliability of information within the databases with the corresponding clinical records of patients [5]. To reach this goal, the content of administrative healthcare databases needs to be appropriately validated.

In Italy, all the Regional Health Authorities maintain large healthcare information systems containing patient data from all hospital and operative sources. These databases have the potential to address important issues in post-marketing 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 realized since they have not been widely validated[12].

Breast cancer is the most commonly diagnosed neoplasm in women worldwide, as well as in Italy[13]. Variation in the epidemiology of breast cancers [14], potential heterogeneity in treatment (pharmacological or surgical) and potential clinical[15 16] and economic outcomes[17-19] can all be

evaluated using validated administrative databases. Hence, assessing the accuracy of Italian administrative databases in identifying women with this oncological disease is relevant for the scientific community, the governments, as well as the industry.

As reported in our protocol[20], the objective of the present study was to evaluate the accuracy of the ICD-9-CM codes related to breast cancer in correctly identifying the respective diseases using three large Italian administrative healthcare databases.

Methods

Setting and data source

Administrative databases

The target administrative databases were represented by the Umbria Region (890,000 residents), Local Health Unit 3 of NA (1,170,000 residents) and the FVG Region (1,227,000 residents). The corresponding operative units, the Regional Health Authority of Umbria (for Umbria Region), the Registro Tumori Regione Campania (for Napoli Sud Local Health Unit), and the Centro di Riferimento Oncologico Aviano (for Friuli Venezia Giulia Region), conducted the same validation process.

In Italy, regional and local healthcare administrative databases routinely collect data 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 5 secondary discharge diagnoses and the principal, and up to 5 secondary surgical or pharmacological treatments and diagnostic procedures. 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, health care is covered almost entirely by the Italian National Health System, therefore most residents' significant healthcare information can be found within the healthcare databases.

Source population

The source population was represented by permanent residents aged 18 or above of Umbria Region, Local Health Unit 3 of NA and FVG Region. Any resident that has been discharged from hospital with a diagnosis of breast cancer was considered. Residents that have been hospitalised outside the regional territory of competence were excluded from analysis.

Case and control selection and sampling method

In each administrative database, patients with the first occurrence of diagnosis of breast cancer between 1st January 2012 and 31st December 2014 were identified using the following ICD-9-CM codes (index test) located in primary position: (a) 233.0 for carcinoma in situ of the breast and (b) 174.x for invasive breast cancer. Estimated prevalent cases, that is, cases with the same diagnosis (ICD-9-CM codes in any position) in the five years (2007-2011) before the period of interest, were excluded.

For controls, within the same period, non-cases, i.e. 94 female patients having in primary position a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9) or carcinoma in situ of the breast (ICD-9 233.0), were randomly selected .

Subsequently, for each of the above reported groups of ICD-9-CM codes, random samples of cases and non-cases were selected from each administrative database.

Chart abstraction and case ascertainment

The medical charts of the randomly selected samples of cases and non-cases were obtained from hospitals for the validation task (considered as the reference standard). From each medical chart, the following information were retrieved: clinical chart number, hospital and ward, date of birth, sex, dates of hospital admission and discharge, signs and symptoms, any diagnostic procedures that contributed to the diagnosis of the cancer, any pharmacological or surgical interventions that were provided for the treatment of the cancer.

Within each unit, two reviewers received training on data abstraction and performed an initial consensus chart review, independently examining the same number of medical charts (n=20). The inter-rater agreement regarding the presence or absence of breast cancer among the pairs of reviewers within each unit was near perfect (κ statistics > 0.91).

Case ascertainment of cancer within medical charts was based on (a) the presence of a primary nodular lesion in the breast documented by imaging and (b) the cytological or histological documentation of cancer from a primary or metastatic site.

Following the consensus review, data abstraction were completed independently. To ensure consistency among all the reviewers, cases with uncertainty were discussed and resolved through the involvement of an oncologist (RC).

Validation criteria

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

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

Statistical analysis

We calculated that a sample of 130 charts of cases was necessary to obtain an expected sensitivity of 80% with a precision of 10% and a power of 80%. For specificity calculation, we randomly selected non-cases, that is, records without the ICD-9 codes of interest from administrative database. For controls, we calculated a sample of 94 charts of non-cases were necessary to obtain an expected specificity of 90% with a precision of 10% and a power of 80% according to binomial exact calculation[20].

Sensitivity and specificity with relative 95% confidence intervals were calculated separately for each ICD-9-CM code by constructing 2 x 2 tables.

When missing medical charts occurred we performed a formal sensitivity analysis based on a worst case scenario in which the missing cases was considered as false positive and the missing-non cases were considered false negative.

Results

Invasive breast cancer

After excluding the estimated prevalent cases from the diagnosis of invasive breast cancer in the primary position of hospital discharges, the cases were 2,686, 2,044 and 2,107 from Umbria, NA and FVG respectively. Subsequently each team randomly sampled 130 cases of which the corresponding medical charts were requested for evaluation. Two and four medical charts were not available from Umbria and NA respectively. **Figure 1** displays the identification of cases from the three operative units. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The most common ICD-9-CM subgroup was the code 174.4, that is upper-outer quadrant breast cancer, accounting for 45% of cases for Umbria, 35% for FVG, and 34% for NA. The mean age ranged between 61 and 66 years. The majority of the cases were identified in surgical departments with a percentage ranging from 84% to 94%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units. The sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. The specificities were 90% for Umbria, and 91% for NA and FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

In terms of misclassification, overall there were 28 cases that were considered false positives. The reasons for this misclassification were: histological documentation missing in the medical chart (6 in Umbria, 8 in NA, and 8 in FVG) and negative histology for invasive breast cancer (4 in Umbria, 1 in NA, and 1 in FVG). However, of these false positive cases with negative histology for invasive breast cancer, three were positive for breast carcinoma in situ diagnosis. Conversely, there were 8

non-cases that were judged false negatives: two were possible breast cancer diagnosis and six were metastatic breast diagnoses (**Table 2**).

The 12 missing medical of the cases affected specificity that reduced from 90% to 80% (73% to 87%) but the difference was not statistically significant.

Breast carcinoma in situ

The incident cases of carcinoma in situ of the breast were 67 from Umbria, 95 from NA, and 137 from FVG, from which 50, 95 and 108 were randomly selected and the corresponding medical charts were requested for assessment (**Figure 1)**. Seven charts from NA were not available. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than carcinoma in situ of the breast (ICD-9 233.0), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The mean age ranged between 57 (NA) and 60 years (FVG). Most of the cases were identified in surgical departments with a percentage ranging from 92% to 100%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units.

After reviewing the medical records, 100% (48/48) of the patients with carcinoma in situ of the breast from Umbria, 100% (73/73) from NA, and 100% (97/97) from FVG were correctly identified by the administrative databases . The specificities were 98% for Umbria, 86% for NA, and 90% for FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

Table 2 describes the reasons for misclassification of cases and controls. Overall, there were 28 cases that were judged false positives. The reasons were histological documentation missing in the medical charts (8 in NA and 2 in FVG), and negative histology for carcinoma in situ of the breast (2 in Umbria, 7 in NA and 9 in FVG). None of the controls resulted a false negative.

The sensitivity analysis showed that specificity for NA codes reduced to 81% (95% CI 73% to 88%) due to the seven charts of missing cases but the difference was not however statistically significant.

Discussion

We developed a case definition of breast cancer based on the presence of a primary nodular lesion in the breast documented with imaging and the cytological or histological documentation of cancer from a primary or metastatic site and the performance of the model was evaluated in terms of sensitivities and specificities for the three administrative databases. After revising the medical charts, the results showed that both codes (233.0 and 174.x) performed well in identifying women with breast carcinoma in situ and invasive breast cancer respectively. This means that at least 98% of invasive breast cancer cases and 100% of carcinoma in situ cases were confirmed as true positives.

Previously, researchers have assessed the accuracy of breast cancer diagnosis in administrative databases using different algorithms and in most cases using registry data as a reference standard[21]. 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[22]. The study found a sensitivity of 77% for breast cancer[22] and the lower value of sensitivity compared to our results can be attributed to the fact that in Baldi et al. the validity of the algorithm for each cancer site was assessed by individual matching between cases in hospital discharge and the Piedmont Cancer Registry or because authors were interested in high values of PPV. Using the Surveillance, Epidemiology, and End Results (SEER) database as a reference standard, Freeman et al. developed an approach for identifying incident breast cancer cases based on a logistic regression model, which contained variables that indicate the presence of breast cancer-related diagnoses and procedures in three sources of claims data: hospital inpatient stays, hospital outpatient services, and physician

services. The ROC curve showed that the model achieved over 90% of sensitivity and specificity albeit a lower positive predictive value[23]. Similarly, using the Medicare database, Setoguchi et al. developed a case definition based on diagnoses and procedure codes and compared it with SEER database in Pennsylvania. The authors obtained a sensitivity and specificity for identifying breast cancer cases of 83% and 99% respectively[24]. Using the Cancer Registry data as reference standard, Kemp et al.[25] evaluated Australian administrative and self-reported datasets to identify cases of invasive breast cancer and used several combinations of diagnoses and procedures obtaining the highest sensitivity and PPV (both 86%) when a flag of 'diagnosis of invasive breast cancer' was used. A systematic review of administrative databases that validated breast cancer is currently being completed and will provide a complete account of validation of administrative databases worldwide[26].

Strength and limitation

Strengths of our study include complete transparency based on pre-publication of a protocol, the use of detailed and explicit eligibility criteria, and the use of duplicate and independent processes for medical chart review and abstraction following 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 [27-29]. In addition, we used as a required element for validation the presence of a histological or cytological documentation. Unlike several studies that used Cancer Registries to validate breast cancer codes, in the present study medical records were reviewed directly to evaluate the accuracy of potential cases obtained from the administrative database. Generally, medical charts are considered the gold standard for the diagnosis of a disease.

We acknowledge some limitations in our study. The overall number of carcinoma in situ cases identified in the three units during the period of interest was below the calculated sample size and we are unsure whether this limitation can affect the results of sensitivity and specificity for the

breast carcinoma in situ ICD-9-CM code. In addition, despite the success of validation processes of administrative database, any conclusion that stems from these validated database could not be generalized in other settings.

As declared in our protocol we favoured the estimation of sensitivity rather than predictive values since predictive values are dependent on the prevalence of the disease. However, to comply with the STARD guidelines we provide absolute numbers for true or false case and non-cases from which it is possible to obtain predictive values (**Table 3)**.

Conclusion

In summary, the present study has demonstrated that administrative healthcare databases from Umbria, NA and FVG can be used to identify women with newly diagnosed invasive or in situ carcinoma of the breast. The proposed case definition in the present study provides a powerful tool to perform outcome research on breast cancer based on a population of three million residents. Potential implication of this proposal includes the extension of this case definition to other Italian regional healthcare databases and the combination with other sources of data (such as prescription database) to conduct efficiently pharmacoepidemiological studies that may complement randomised trials.

Footnotes

Contributors: AM, IA, MF, and DS conceived the original idea of the study. IA, DG, AM, MF, EB, GG, FC, MO, and WO designed the study. PCa, MDG, PCo, AG, MFV, VC identified the cohort using administrative database with the supervision of WO, EB, DS, MF, AM and FS. IA, FC, MO, AG, PC, VC, MFV, and JF undertook the data abstraction with the supervision of AM, GG, WO, FS, MF, EB, RC and DS. IA, RC, AM and JF performed case ascertainment. IA, AM, FC, EB, MF and MO performed the analysis. DS, GG, PCa AG, MDG, PCo, RC, JF, MFV, FS, EB and WO helped in the interpretation of the data.
The initial draft of the manuscript was prepared by IA, AM, FC and MO. DS, EB, GG, PCa, AG, MDG, PCo, RC, JF, MFV, FS, and WO revised critically the manuscript for important intellectual content. All the authors read and approved the final manuscript. AM, MF and EB are the guarantors of the data for the respective operative units.

Funding This study was developed within the D.I.V.O. project (*Realizzazione di un **D**atabase **I**nterregionale **V**alidato per l'**O**ncologia 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.

Data sharing statement No additional data are available.

Figure Caption

Figure 1. Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined
Figure 2. Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases.

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. *Annual review of public health* 2011;32:91-108. doi: 10.1146/annurev-publhealth-031210-100700 [published Online First: 2011/01/12]
3. Jick SS. Fresh evidence confirms links between newer contraceptive pills and higher risk of venous thromboembolism. *Bmj* 2015;350:h2422. doi: 10.1136/bmj.h2422 [published Online First: 2015/05/28]
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.
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(3):205-11. [published Online First: 2001/10/05]
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(7405):18-22. doi: 10.1136/bmj.327.7405.18
9. Trifiro G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Studies in health technology and informatics* 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. doi: 10.1186/1471-2458-13-15
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 services research* 2013;13(1):393. doi: 10.1186/1472-6963-13-393
12. Abraha I, Montedori A, Eusebi P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. *Pharmacoepidemiology and Drug Safety* 2012;21:400-00.
13. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer Journal international du cancer* 2015;136(5):E359-86. doi: 10.1002/ijc.29210 [published Online First: 2014/09/16]
14. 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. *European journal of cancer care* 2015;24(3):367-75. doi: 10.1111/ecc.12277 [published Online First: 2014/12/19]
15. Lambertini M, Anserini P, Fontana V, et al. The PREgnancy and FERtility (PREFER) study: an Italian multicenter prospective cohort study on fertility preservation and pregnancy issues in young breast cancer patients. *BMC cancer* 2017;17(1):346. doi: 10.1186/s12885-017-3348-8 [published Online First: 2017/05/21]
16. De Placido S, De Angelis C, Giuliano M, et al. Imaging tests in staging and surveillance of non-metastatic breast cancer: changes in routine clinical practice and cost implications. *Br J Cancer* 2017;116(6):821-27. doi: 10.1038/bjc.2017.24 [published Online First: 2017/02/09]
17. 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 research and treatment* 2013;139(2):469-76. doi: 10.1007/s10549-013-2543-9 [published Online First: 2013/04/30]

18. Konski A. Clinical and economic outcomes analyses of women developing breast cancer in a managed care organization. *American journal of clinical oncology* 2005;28(1):51-7. [published Online First: 2005/02/03]
19. 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(1):E11-7. doi: 10.9778/cmajo.20130026 [published Online First: 2014/08/01]
20. 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(3):e010547. doi: 10.1136/bmjopen-2015-010547
21. 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(3):e010409. doi: 10.1136/bmjopen-2015-010409
22. Baldi I, Vicari P, Di Cuonzo D, et al. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. *Journal of clinical epidemiology* 2008;61(4):373-9.
23. Freeman JL, Zhang D, Freeman DH, et al. An approach to identifying incident breast cancer cases using Medicare claims data. *Journal of clinical epidemiology* 2000;53(6):605-14. doi: 10.1016/s0895-4356(99)00173-0
24. Setoguchi S, Solomon DH, Glynn RJ, et al. Agreement of diagnosis and its date for hematologic malignancies and solid tumors between medicare claims and cancer registry data. *Cancer causes & control : CCC* 2007;18(5):561-9. doi: 10.1007/s10552-007-0131-1 [published Online First: 2007/04/21]
25. Kemp A, Preen DB, Saunders C, et al. Ascertaining invasive breast cancer cases; the validity of administrative and self-reported data sources in Australia. *BMC medical research methodology* 2013;13:17. doi: 10.1186/1471-2288-13-17 [published Online First: 2013/02/13]
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
27. 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. *Journal of clinical epidemiology* 2011;64(8):821-9. doi: 10.1016/j.jclinepi.2010.10.006
28. 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 services research* 2006;6:77. doi: 10.1186/1472-6963-6-77
29. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Annals of internal medicine* 2003;138(1):40-4.

Table 1 Characteristics of patients with breast cancer who were identified in the three administrative healthcare databases

Characteristics	Unit 1 (Umbria)	Unit 2 (Friuli Venezia Giulia)	Unit 3 (Asl Napoli 3 Sud)
Invasive carcinoma			
Incident cases (N medical chart reviewed)	128	130	126
ICD-9 code			
174.0	-	-	1 (1)
174.1	16 (13)	6 (5)	10 (8)
174.2	4 (3)	14 (11)	8 (6)
174.3	6 (5)	9 (7)	5 (4)
174.4	57 (45)	45 (35)	43 (34)
174.5	6 (5)	13 (10)	5 (4)
174.6	-	-	-
174.7	-	-	-
174.8	7 (5)	34 (26)	38 (30)
174.9	32 (25)	9 (7)	16 (13)
Admission to department			
Medical	20 (16)	8 (6)	11 (9)
Surgical	108 (84)	122 (94)	115 (91)
Age, N (%)			
< 40	9 (7)	1 (1)	6 (5)
40 - 59	40 (31)	45 (35)	56 (44)
≥ 60	79 (62)	84 (65)	64 (51)
Instrumental diagnosis			
Breast ultrasound	39 (30)	5 (4)	88 (70)
Mammography	54 (42)	7 (5)	60 (48)
CT scan (breast)	11 (8)	2 (2)	1 (1)
MRI (breast)	3 (2)	8 (6)	17 (14)
Surgical procedures			
Mastectomy	28 (22)	35 (27)	29 (23)
Quadrantectomy	79 (62)	54 (42)	73 (58)
Hystological documentation			
Needle aspiration	32	-	34
Needle biopsy	27	5	40
Nodule (after surgical intervention)	115	112	117
Carcinoma in situ			
Incident cases (N medical chart reviewed)	50	108	88
ICD-9 code			
233.0	50 (100)	108 (100)	88 (100)
Admission to department			
Medical	-	-	7 (8)
Surgical	50 (100)	108 (100)	81 (92)

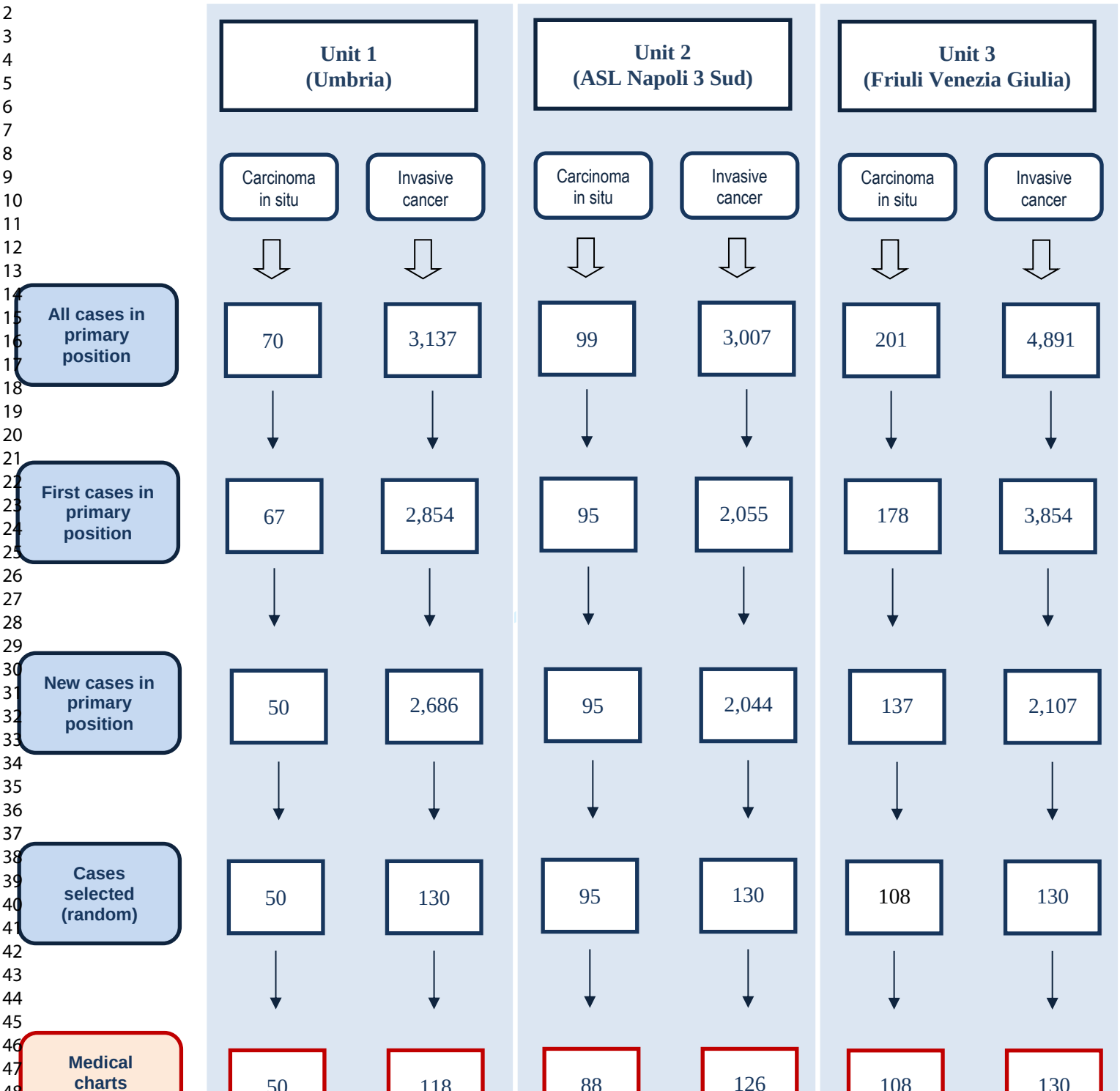
Age, N (%)			
< 40	2 (4)	1 (1)	3 (3)
40 - 59	27 (54)	52 (48)	50 (57)
> 60	21 (42)	55 (51)	35 (40)
Instrumental diagnosis			
Breast ultrasound	3 (6)	30 (28)	65 (74)
Mammography	6 (12)	39 (36)	37 (42)
CT scan (breast)	-	-	-
MRI (breast)	2 (4)	4 (4)	22 (25)
Surgical procedures			
Mastectomy	18 (36)	15 (14)	13 (15)
Quadrantectomy	32 (64)	55 (52)	47 (53)
Hystological documentation			
Needle aspiration	-	1 (1)	16 (18)
Needle biopsy	8 (16)	4 (4)	53 (60)
Nodule (after surgical intervention)	50 (100)	94 (87)	77 (88)

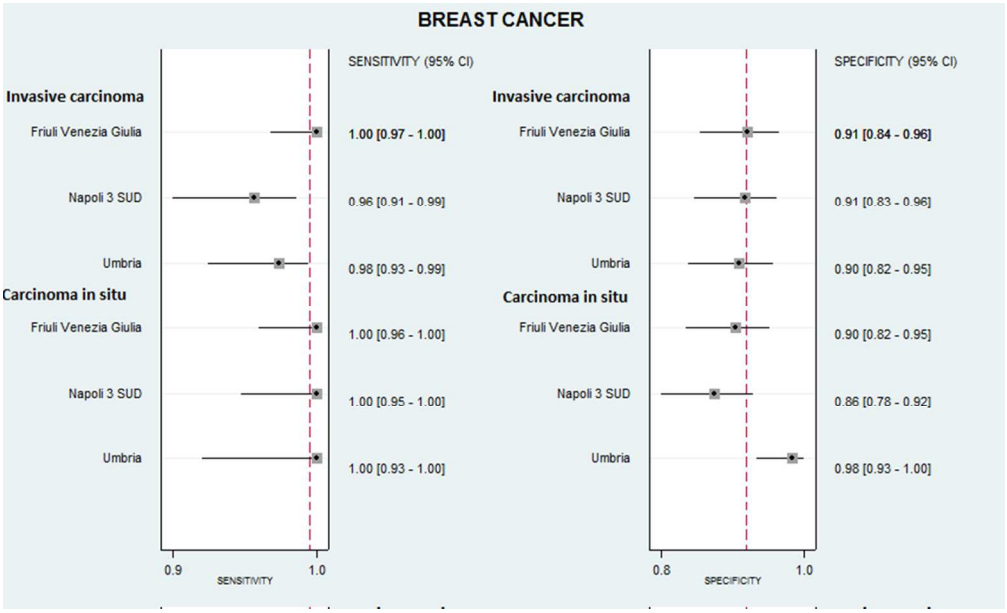
Table 2. Reason for misclassification of cases and controls

	Invasive breast cancer			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (Friuli Venezia Giulia)**	**Unit 3 (ASL 3 Napoli Sud)**
	False positives			
1	histological examination missing	6	8	8
2	negative histology	4	1	1
	a) carcinoma in situ	*2*	*1*	
Total		10	9	9
	False negative			
1	possible breast cancer relapse	1	-	1
2	metastatic breast cancer	2	-	4
Total		3	-	5
	Breast carcinoma in situ			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (Friuli Venezia Giulia)**	**Unit 3 (ASL 3 Napoli Sud)**
	False positives			
1	histological examination missing	-	2	8
2	negative histology	2	9	7
Total		2	11	15
	False negative			
1	Possible carcinoma in situ	-	-	-

Table 3. Cross tabulation of the index test (ICD-9-CM code) results by the results of the reference standard (medical chart)

Type of brest cancer (ICD-9-CM)	Operative unit	TP	FP	TN	FN
Invasive cancer (174.x)	Unit 1 (Umbria)	118	10	90	3
	Unit 2 (Friuli Venezia Giulia)	121	9	94	0
	Unit 3 (ASL 3 Napoli)	117	9	89	5
Carcinoma in situ (233.0)	Unit 1 (Umbria)	48	2	93	0
	Unit 2 (Friuli Venezia Giulia)	97	11	94	0
	Unit 3 (ASL 3 Napoli Sud)	73	15	94	0


Unit 1 (Umbria)
Unit 2 (ASL Napoli 3 Sud)
Unit 3 (Friuli Venezia Giulia)
Carcinoma in situ
Invasive cancer
All cases in primary position
70
3,137
99
3,007
201
4,891
First cases in primary position
67
2,854
95
2,055
178
3,854
New cases in primary position
50
2,686
95
2,044
137
2,107
Cases selected (random)
50
130
95
130
108
130
Medical charts identified
50
118
88
126
108
130



Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	7
Participants	**6**	Eligibility criteria	6-8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	6-7
	8	Where and when potentially eligible participants were identified (setting, location and dates)	6-7
	9	Whether participants formed a consecutive, random or convenience series	7
Test methods	**10a**	Index test, in sufficient detail to allow replication	7
	10b	Reference standard, in sufficient detail to allow replication	7
	11	Rationale for choosing the reference standard (if alternatives exist)	7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	na
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	na
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8
	15	How indeterminate index test or reference standard results were handled	na
	16	How missing data on the index test and reference standard were handled	8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	na
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	Page 9 and Table 1
	21a	Distribution of severity of disease in those with the target condition	na
	21b	Distribution of alternative diagnoses in those without the target condition	na
	22	Time interval and any clinical interventions between index test and reference standard	na
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 3
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Page 9-10, Figure 2
	25	Any adverse events from performing the index test or the reference standard	na
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12-13
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	na
	29	Where the full study protocol can be accessed	5
	30	Sources of funding and other support; role of funders	14

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020627.R1
Article Type:	Research
Date Submitted by the Author:	10-Jan-2018
Complete List of Authors:	Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service; Azienda Ospedaliera di Perugia, Centro Regionale Sangue Serraino, Diego; IRCCS CRO AVIANO, Epidemiology Montedori, Alessandro; Regional Health Authority of Umbria, Fusco, Mario; ASL NA3 Sud, Registro Tumori Regione Campania Giovannini, Gianni; Regional Health Authority of Umbria, Health Planning Office Casucci, Paola; Regional Health Authority of Umbria, Health ICT Service Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Granata, Annalisa; ASL NA3 Sud, Registro Tumori Regione Campania De Giorgi, Marcello; Regional Health Authority of Umbria, , Health ICT Service Collarile, Paolo; Centro di Riferimento Oncologico Aviano, SOC Epidemiologia Oncologica Chiari, Rita; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Foglietta, Jennifer; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Vitale, Maria Francesca; Registro Tumori Regione Campania , ASL NA3 Sud Stracci, Fabrizio; University of Perugia, Public Health Department Orlandi, Walter; Regional Health Authority of Umbria Bidoli, Ettore; IRCCS CRO AVIANO, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Oncology
Keywords:	sensitivity and specificity, administrative database, validity, breast cancer

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Iosief Abraha[1 2], Diego Serraino[3], Alessandro Montedori[1], Mario Fusco[4], Gianni Giovannini[1], Paola Casucci[5], Francesco Cozzolino[1], Massimiliano Orso[1], Annalisa Granata[4], Marcello De Giorgi[5], Paolo Collarile[6], Rita Chiari[7], Jennifer Foglietta[7], Maria Francesca Vitale[4], Fabrizio Stracci[8], Walter Orlandi[9], Ettore Bidoli[3] for the D.I.V.O. Group

Author affiliations:

1. *Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy*
2. *Centro Regionale Sangue, Azienda Ospedaliera di Perugia, Perugia, Italy*
3. *Cancer Epidemiology Unit, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
4. *Registro Tumori Regione Campania, ASL Napoli 3 Sud, Brusciano, Italy*
5. *Health ICT Service, Regional Health Authority of Umbria, Perugia, Italy*
6. *SOC Epidemiologia Oncologica, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
7. *Dipartimento di Oncologia, Azienda Ospedaliera Perugia, Perugia, Italy*
8. *Public Health Department, University of Perugia, Perugia, Italy*
9. *Regional Health Authority of Umbria, Perugia, Italy*

Corresponding Author:
Alessandro Montedori
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni 61
06124 Perugia, Italy
amontedori@regione.umbria.it

Abstract

Objectives To assess the accuracy of *International Classification of Diseases 9th Revision Clinical Modification* (ICD-9-CM) codes in identifying patients diagnosed with incident carcinoma in situ and invasive breast cancer in three Italian administrative databases.

Design A diagnostic accuracy study comparing ICD-9-CM codes for carcinoma in situ (233.0) and for invasive breast cancer (174.x) with medical chart (as a reference standard). Case definition: (a) presence of a primary nodular lesion in the breast and (b) cytological or histological documentation of cancer from a primary or metastatic site.

Setting Administrative databases from Umbria Region, ASL Napoli 3 Sud (NA), and Friuli Venezia Giulia (FVG) Region

Participants Women with breast carcinoma in situ (n = 246) or invasive breast cancer (n = 384) diagnosed (in primary position) between 2012 and 2014.

Outcome measures Sensitivity and specificity for codes 233.0 and 174.x.

Results For invasive breast cancer the sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. Specificities were 90% (95% CI 82% to 95%) for Umbria, 91% (95% CI 83% to 96%) for NA and 91% (95% CI 84% to 96%) for FVG.

For carcinoma in situ the sensitivities were 100% (95% CI 93% to 100%) for Umbria, 100% (95% CI 95% to 100%) for NA, and 100% (95%CI 96% to 100%) for FVG. Specificities were 98% (95% CI 93% to 100%) for Umbria, 86% (95% CI 78% to 92%) for NA, and 90% (95% CI 82% to 95%) for FVG.

Conclusions Administrative healthcare databases from Umbria, NA and FVG can be used to identify with newly diagnosed carcinoma of the breast. The proposed case definition is a powerful tool to perform research on large populations of newly diagnosed breast cancer patients.

Strengths and limitations of this study

- This study is the first to have validated *International Classification of Diseases-9th Revision – Clinical Modification* (ICD-9-CM) codes for incident breast cancer cases in three large computerized administrative databases in Italy using the same case definition.
- Case ascertainment was based on the presence of a primary nodular lesion in the breast documented by imaging and a cytological or histological documentation of cancer from a primary or metastatic site.
- This study followed 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.
- Validation studies of administrative data are related to that context and are not generalizable to other settings.
- The sample size of women with carcinoma in situ was limited due to the low prevalence of disease.

Introduction

The use of administrative databases is increasingly growing in various healthcare settings worldwide. These databases anonymously store data about residents regarding the healthcare assistance they receive including hospital admission, demographic data and disease treatment. Usually, the diagnosis of the disease is associated with a specific code from the *International Classification of Diseases, 9th 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 networking of individual patient data from administrative databases and other sources such as outpatient data and prescription data allows monitoring population health status, performing outcome research [2-4] and exploring a wide range of significant public health questions [2]. In administrative databases, while non-clinical data such as demographic or prescription data are highly accurate [5 6], the accuracy of diagnoses and procedures needs to be determined [6 7]. Typically, the assessment of accuracy consists in confirming the reliability of information within the databases with the corresponding clinical records of patients [5]. To reach this goal, the content of administrative healthcare databases needs to be appropriately validated.

In Italy, all the Regional Health Authorities maintain large healthcare information systems containing patient data from all hospital and operative sources. These databases have the potential to address important issues in post-marketing 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 realized since they have not been widely validated[12].

Breast cancer is the most commonly diagnosed neoplasm in women worldwide, as well as in Italy[13]. Variation in the epidemiology of breast cancers [14], potential heterogeneity in treatment (pharmacological or surgical) and potential clinical[15 16] and economic outcomes[17-19] can all be evaluated using validated administrative databases. Hence, assessing the accuracy of Italian

administrative databases in identifying women with this oncological disease is relevant for the scientific community, the governments, as well as the industry.

As reported in our protocol[20], the objective of the present study was to evaluate the accuracy of the ICD-9-CM codes related to breast cancer in correctly identifying the respective diseases using three large Italian administrative healthcare databases.

Methods

Setting and data source

Administrative databases

The target administrative databases were represented by the Umbria Region (890,000 residents), Local Health Unit 3 of NA (1,170,000 residents) and the FVG Region (1,227,000 residents). The corresponding operative units, the Regional Health Authority of Umbria (for Umbria Region), the Registro Tumori Regione Campania (for Napoli Sud Local Health Unit), and the Centro di Riferimento Oncologico Aviano (for Friuli Venezia Giulia Region), conducted the same validation process.

In Italy, regional and local healthcare administrative databases routinely collect data 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 5 secondary discharge diagnoses and the principal, and up to 5 secondary surgical or pharmacological treatments and diagnostic procedures. 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, health care is covered almost entirely by the Italian National Health System, therefore most residents' significant healthcare information can be found within the healthcare databases.

Source population

The source population was represented by permanent residents aged 18 or above of Umbria Region, Local Health Unit 3 of NA and FVG Region. Any resident that has been discharged from hospital with a diagnosis of breast cancer was considered. Residents that have been hospitalised outside the regional territory of competence were excluded from analysis.

Case and control selection and sampling method

In each administrative database, patients with the first occurrence of diagnosis of breast cancer between 1st January 2012 and 31st December 2014 were identified using the following ICD-9-CM codes (index test) located in primary position: (a) 233.0 for carcinoma in situ of the breast and (b) 174.x for invasive breast cancer. Estimated prevalent cases, that is, cases with the same diagnosis (ICD-9-CM codes in any position) in the five years (2007-2011) before the period of interest, were excluded.

For controls, within the same period, non-cases, i.e. 94 female patients having in primary position a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9) or carcinoma in situ of the breast (ICD-9 233.0), were randomly selected .

Subsequently, for each of the above reported groups of ICD-9-CM codes, random samples of cases and non-cases were selected from each administrative database.

Chart abstraction and case ascertainment

The medical charts of the randomly selected samples of cases and non-cases were obtained from hospitals for the validation task (considered as the reference standard). From each medical chart, the following information were retrieved: clinical chart number, hospital and ward, date of birth, sex, dates of hospital admission and discharge, signs and symptoms, any diagnostic procedures that contributed to the diagnosis of the cancer, any pharmacological or surgical interventions that were provided for the treatment of the cancer.

Within each unit, two reviewers received training on data abstraction and performed an initial consensus chart review, independently examining the same number of medical charts (n=20). The inter-rater agreement regarding the presence or absence of breast cancer among the pairs of reviewers within each unit was near perfect (κ statistics > 0.91).

Case ascertainment of cancer within medical charts was based on (a) the presence of a primary nodular lesion in the breast documented by imaging and (b) the cytological or histological documentation of cancer from a primary or metastatic site.

Following the consensus review, data abstraction were completed independently. To ensure consistency among all the reviewers, cases with uncertainty were discussed and resolved through the involvement of an oncologist (RC).

Validation criteria

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

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

Statistical analysis

We calculated that a sample of 130 charts of cases was necessary to obtain an expected sensitivity of 80% with a 95% confidence interval (CI) of 72% - 86% according to binomial exact calculation[21]. For specificity calculation, we randomly selected non-cases, that is, records without the ICD-9 codes of interest from administrative database. For controls, we calculated a sample of 94 charts of non-cases was necessary to obtain an expected specificity of 90% with a 95% CI of 83% - 95% according to binomial exact calculation[21].

Sensitivity and specificity with relative 95% confidence intervals were calculated separately for each ICD-9-CM code by constructing 2 x 2 tables.

When missing medical charts occurred we performed a formal sensitivity analysis based on a worst case scenario in which the missing cases was considered as false positive and the missing-non cases were considered false negative.

Results

Invasive breast cancer

After excluding the estimated prevalent cases from the diagnosis of invasive breast cancer in the primary position of hospital discharges, the cases were 2,686, 2,044 and 2,107 from Umbria, NA and FVG respectively. Subsequently each team randomly sampled 130 cases of which the corresponding medical charts were requested for evaluation. Two and four medical charts were not available from Umbria and NA respectively. **Figure 1** displays the identification of cases from the three operative units. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The most common ICD-9-CM subgroup was the code 174.4, that is upper-outer quadrant breast cancer, accounting for 45% of cases for Umbria, 34% for NA, and 35% for FVG. The mean age ranged between 61 and 66 years. The majority of the cases were identified in surgical departments with a percentage ranging from 84% to 94%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units. The sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. The specificities were 90% for Umbria, and 91% for NA and FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

In terms of misclassification, overall there were 28 cases that were considered false positives. The reasons for this misclassification were: histological documentation missing in the medical chart (6 in Umbria, 8 in NA, and 8 in FVG) and negative histology for invasive breast cancer (4 in Umbria, 1 in NA, and 1 in FVG). However, of these false positive cases with negative histology for invasive breast cancer, three were positive for breast carcinoma in situ diagnosis. Conversely, there were 8

non-cases that were judged false negatives: two were possible breast cancer diagnosis and six were metastatic breast diagnoses (**Table 2**).

Overall there were six missing charts: two in Umbria and four in NA. Worst case scenario in the sensitivity analysis showed that specificity was affected marginally: it changed from 90% to 88% for Umbria and from 91% to 87% for NA. The differences between the ordinary results and the worst case scenario analysis were not statistically significant.

Breast carcinoma in situ

The incident cases of carcinoma in situ of the breast were 50 from Umbria, 95 from NA, and 137 from FVG, from which 50, 95 and 108 were randomly selected and the corresponding medical charts were requested for assessment (**Figure 1)**. Seven charts from NA were not available. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than carcinoma in situ of the breast (ICD-9 233.0), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The mean age ranged between 57 (NA) and 60 years (FVG). Most of the cases were identified in surgical departments with a percentage ranging from 92% to 100%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units.

After reviewing the medical records, 100% (48/48) of the patients with carcinoma in situ of the breast from Umbria, 100% (73/73) from NA, and 100% (97/97) from FVG were correctly identified by the administrative databases . The specificities were 98% for Umbria, 86% for NA, and 90% for FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

Table 2 describes the reasons for misclassification of cases and controls. Overall, there were 28 cases that were judged false positives. The reasons were histological documentation missing in the

medical charts (8 in NA and 2 in FVG), and negative histology for carcinoma in situ of the breast (2 in Umbria, 7 in NA and 9 in FVG). None of the controls resulted a false negative.

The sensitivity analysis showed that specificity for NA codes reduced to 81% (95% CI 73% to 88%) due to the seven charts of missing cases but the difference was not however statistically significant.

Discussion

We developed a case definition of breast cancer based on the presence of a primary nodular lesion in the breast documented with imaging and the cytological or histological documentation of cancer from a primary or metastatic site and the performance of the model was evaluated in terms of sensitivities and specificities for the three administrative databases. After revising the medical charts, the results showed that both codes (233.0 and 174.x) performed well in identifying women with breast carcinoma in situ and invasive breast cancer respectively. This means that at least 98% of invasive breast cancer cases and 100% of carcinoma in situ cases were confirmed as true positives.

Previously, researchers have assessed the accuracy of breast cancer diagnosis in administrative databases using different algorithms and in most cases using registry data as a reference standard[22]. 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[23]. The study found a sensitivity of 77% for breast cancer[23] and the lower value of sensitivity compared to our results can be attributed to the fact that in Baldi et al. the validity of the algorithm for each cancer site was assessed by individual matching between cases in hospital discharge and the Piedmont Cancer Registry or because authors were interested in high values of PPV. Using the Surveillance, Epidemiology, and End Results (SEER) database as a reference standard, Freeman et al. developed an approach for identifying incident breast cancer cases based on a logistic regression model, which

contained variables that indicate the presence of breast cancer-related diagnoses and procedures in three sources of claims data: hospital inpatient stays, hospital outpatient services, and physician services. The ROC curve showed that the model achieved over 90% of sensitivity and specificity albeit a lower positive predictive value[24]. Similarly, using the Medicare database, Setoguchi et al. developed a case definition based on diagnoses and procedure codes and compared it with SEER database in Pennsylvania. The authors obtained a sensitivity and specificity for identifying breast cancer cases of 83% and 99% respectively[25]. Using the Cancer Registry data as reference standard, Kemp et al.[26] evaluated Australian administrative and self-reported datasets to identify cases of invasive breast cancer and used several combinations of diagnoses and procedures obtaining the highest sensitivity and PPV (both 86%) when a flag of 'diagnosis of invasive breast cancer' was used. A systematic review of administrative databases that validated breast cancer is currently being completed and will provide a complete account of validation of administrative databases worldwide[22].

Strength and limitation

Strengths of our study include complete transparency based on pre-publication of a protocol, the use of detailed and explicit eligibility criteria, and the use of duplicate and independent processes for medical chart review and abstraction following 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 [27-29]. In addition, we used as a required element for validation the presence of a histological or cytological documentation. Unlike several studies that used Cancer Registries to validate breast cancer codes, in the present study medical records were reviewed directly to evaluate the accuracy of potential cases obtained from the administrative database. Generally, medical charts are considered the gold standard for the diagnosis of a disease.

We acknowledge some limitations in our study. The overall number of carcinoma in situ cases identified in the three units during the period of interest was below the calculated sample size and we are unsure whether this limitation can affect the results of sensitivity and specificity for the breast carcinoma in situ ICD-9-CM code. The proportion between carcinoma in situ and the overall breast cancer diagnoses across the three units varied between 2% to 6% and resulted lower than in other settings[30]. These discrepancies could be related to the inclusion criteria or different types of managements that the two diseases required. Future research will clarify these issues.

In addition, we had a higher false positive rate than false negative rate. The number of false positives is due to our stringent case ascertainment criteria, i.e. the presence in the clinical chart of both imaging and histological documentation of breast cancer within the same medical chart. Twenty-two false positives cases for the invasive breast cancer and ten false positives cases for the carcinoma in situ were due to histological documentation missing in the medical chart (**Table 2**). This does not necessarily mean that the subjects were without the diagnosis of cancer. These cases had other elements in the medical chart such as imaging, chemotherapy or radiotherapy, that could confirm the presence of the disease. Should we have used broader case ascertainment criteria we could have obtained a lower false positive rate. Estimates of specificity for invasive carcinoma in our three databases were high ranging from 90% to 91%. For carcinoma in situ specificity was acceptable for Napoli 3 Sud (86%) and high for FVG and Umbria (90% and 98% respectively). For future assessments such as epidemiological studies or post-marketing surveillance studies it is more important that our databases have a higher sensitivity than specificity[31].

In terms of generalizability, despite the success of validation processes of administrative database, any conclusion that stems from these validated database could not be generalized in other settings. As declared in our protocol we favoured the estimation of sensitivity and specificity rather than predictive values since predictive values are dependent on the prevalence of the disease. However, to comply with the STARD guidelines we provide absolute numbers for true or false case and non-cases from which it is possible to obtain predictive values (**Table 3)**.

Conclusion

In summary, the present study has demonstrated that administrative healthcare databases from Umbria, NA and FVG can be used to identify women with newly diagnosed invasive or in situ carcinoma of the breast. The proposed case definition in the present study provides a powerful tool to perform outcome research on breast cancer based on a population of three million residents. Potential implication of this proposal includes the extension of this case definition to other Italian regional healthcare databases and the combination with other sources of data (such as prescription database) to conduct efficiently pharmacoepidemiological studies that may complement randomised trials.

Footnotes

Contributors: AM, IA, MF, and DS conceived the original idea of the study. IA, DG, AM, MF, EB, GG, FC, MO, and WO designed the study. PCa, MDG, PCo, AG, MFV, VC identified the cohort using administrative database with the supervision of WO, EB, DS, MF, AM and FS. IA, FC, MO, AG, PC, VC, MFV, and JF undertook the data abstraction with the supervision of AM, GG, WO, FS, MF, EB, RC and DS. IA, RC, AM and JF performed case ascertainment. IA, AM, FC, EB, MF, PE and MO performed the analysis. DS, GG, PCa AG, MDG, PCo, RC, JF, MFV, FS, EB and WO helped in the interpretation of the data.
The initial draft of the manuscript was prepared by IA, AM, FC and MO. DS, EB, GG, PCa, AG, MDG, PCo, RC, JF, MFV, FS, and WO revised critically the manuscript for important intellectual content. All the authors read and approved the final manuscript. AM, MF and EB are the guarantors of the data for the respective operative units.

Funding This study was developed within the D.I.V.O. project (*Realizzazione di un **D**atabase **I**nterregionale **V**alidato per l'**O**ncologia 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.

Data sharing statement No additional data are available.

Figure Caption

Figure 1. Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined
Figure 2. Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases.

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. Annual review of 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.
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**(3):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**(7405):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. Studies in health technology and informatics 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 services research 2013;**13**(1):393.
12. Abraha I, Montedori A, Eusebi P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. Pharmacoepidemiology and Drug Safety 2012;**21**:400-00.
13. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer Journal international du cancer 2015;**136**(5):E359-86.
14. 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. European journal of cancer care 2015;**24**(3):367-75.
15. Lambertini M, Anserini P, Fontana V, et al. The PREgnancy and FERtility (PREFER) study: an Italian multicenter prospective cohort study on fertility preservation and pregnancy issues in young breast cancer patients. BMC cancer 2017;**17**(1):346.
16. De Placido S, De Angelis C, Giuliano M, et al. Imaging tests in staging and surveillance of non-metastatic breast cancer: changes in routine clinical practice and cost implications. Br J Cancer 2017;**116**(6):821-27.
17. 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 research and treatment 2013;**139**(2):469-76.
18. Konski A. Clinical and economic outcomes analyses of women developing breast cancer in a managed care organization. American journal of clinical oncology 2005;**28**(1):51-7.
19. 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**(1):E11-7.
20. 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**(3):e010547.

21. Wilson EB. Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 1927;**22**:209-12.
22. 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**(3):e010409.
23. Baldi I, Vicari P, Di Cuonzo D, et al. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. Journal of clinical epidemiology 2008;**61**(4):373-9.
24. Freeman JL, Zhang D, Freeman DH, et al. An approach to identifying incident breast cancer cases using Medicare claims data. Journal of clinical epidemiology 2000;**53**(6):605-14.
25. Setoguchi S, Solomon DH, Glynn RJ, et al. Agreement of diagnosis and its date for hematologic malignancies and solid tumors between medicare claims and cancer registry data. Cancer causes & control : CCC 2007;**18**(5):561-9.
26. Kemp A, Preen DB, Saunders C, et al. Ascertaining invasive breast cancer cases; the validity of administrative and self-reported data sources in Australia. BMC medical research methodology 2013;**13**:17.
27. 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. Journal of clinical epidemiology 2011;**64**(8):821-9.
28. 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 services research 2006;**6**:77.
29. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Annals of internal medicine 2003;**138**(1):40-4.
30. Yuan Y, Li M, Yang J, et al. Factors related to breast cancer detection mode and time to diagnosis in Alberta, Canada: a population-based retrospective cohort study. BMC health services research 2016;**16**(65).
31. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. Journal of clinical epidemiology 2012;**65**(3):343-49.e2.

Table 1 Characteristics of patients with breast cancer who were identified in the three administrative healthcare databases

Characteristics	Unit 1 (Umbria)	Unit 2 (Asl Napoli 3 Sud)	Unit 3 (Friuli Venezia Giulia)
Invasive carcinoma			
Incident cases (N medical chart reviewed)	128	126	130
ICD-9 code			
174.0	-	1 (1)	-
174.1	16 (13)	10 (8)	6 (5)
174.2	4 (3)	8 (6)	14 (11)
174.3	6 (5)	5 (4)	9 (7)
174.4	57 (45)	43 (34)	45 (35)
174.5	6 (5)	5 (4)	13 (10)
174.6	-	-	-
174.7	-	-	-
174.8	7 (5)	38 (30)	34 (26)
174.9	32 (25)	16 (13)	9 (7)
Admission to department			
Medical	20 (16)	11 (9)	8 (6)
Surgical	108 (84)	115 (91)	122 (94)
Age, N (%)			
< 40	9 (7)	6 (5)	1 (1)
40 - 59	40 (31)	56 (44)	45 (35)
≥ 60	79 (62)	64 (51)	84 (65)
Instrumental diagnosis			
Breast ultrasound	39 (30)	88 (70)	5 (4)
Mammography	54 (42)	60 (48)	7 (5)
CT scan (breast)	11 (8)	1 (1)	2 (2)
MRI (breast)	3 (2)	17 (14)	8 (6)
Surgical procedures			
Mastectomy	28 (22)	29 (23)	35 (27)
Quadrantectomy	79 (62)	73 (58)	54 (42)
Hystological documentation			
Needle aspiration	32	34	-
Needle biopsy	27	40	5
Nodule (after surgical intervention)	115	117	112
Carcinoma in situ			
Incident cases (N medical chart reviewed)	50	88	108
ICD-9 code			

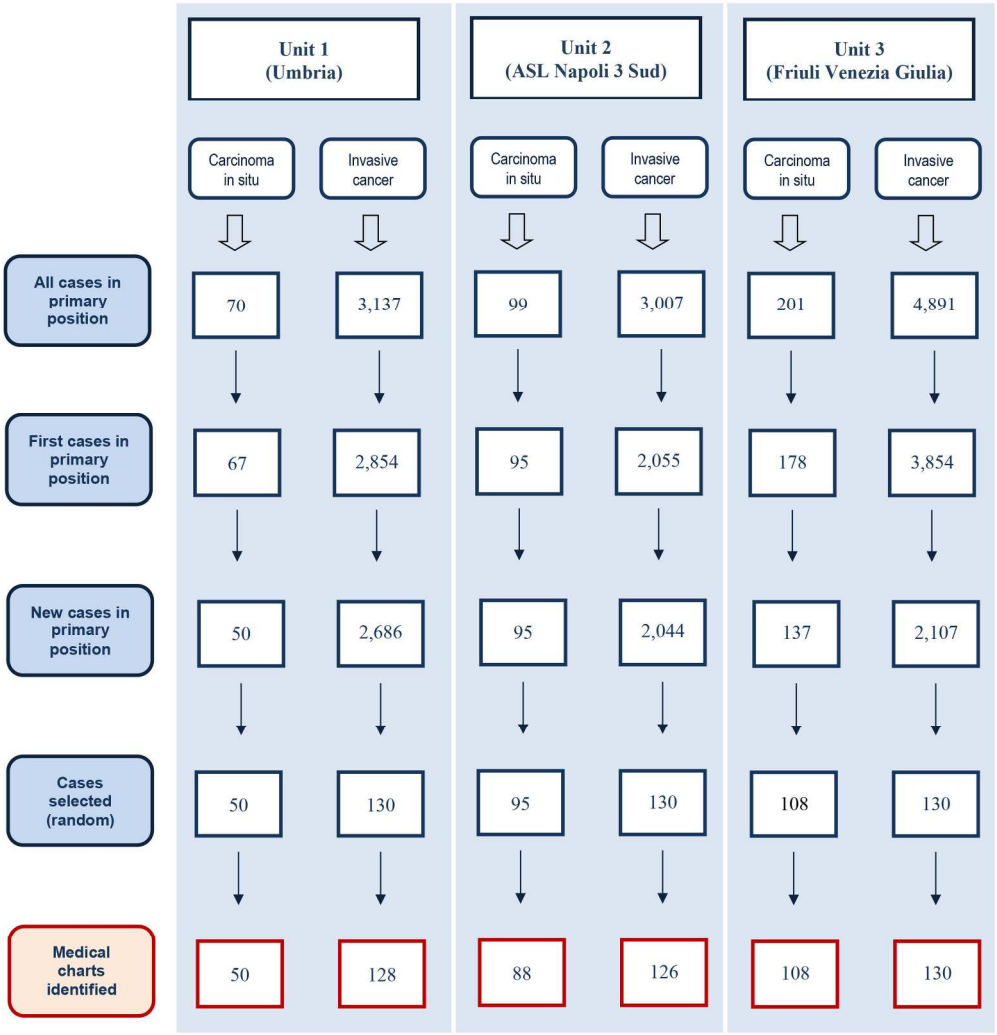
233.0	50 (100)	88 (100)	108 (100)
Admission to department			
Medical	-	7 (8)	-
Surgical	50 (100)	81 (92)	108 (100)
Age, N (%)			
< 40	2 (4)	3 (3)	1 (1)
40 - 59	27 (54)	50 (57)	52 (48)
> 60	21 (42)	35 (40)	55 (51)
Instrumental diagnosis			
Breast ultrasound	3 (6)	65 (74)	30 (28)
Mammography	6 (12)	37 (42)	39 (36)
CT scan (breast)	-	-	-
MRI (breast)	2 (4)	22 (25)	4 (4)
Surgical procedures			
Mastectomy	18 (36)	13 (15)	15 (14)
Quadrantectomy	32 (64)	47 (53)	55 (52)
Hystological documentation			
Needle aspiration	-	16 (18)	1 (1)
Needle biopsy	8 (16)	53 (60)	4 (4)
Nodule (after surgical intervention)	50 (100)	77 (88)	94 (87)

Table 2. Reason for misclassification of cases and controls

	Invasive breast cancer			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	6	8	8
2	negative histology	4	1	1
	a) carcinoma in situ	*2*		*1*
Total		10	9	9
	False negative			
1	possible breast cancer relapse	1	1	-
2	metastatic breast cancer	2	4	-
Total		3	5	-
	Breast carcinoma in situ			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	-	8	2
2	negative histology	2	7	9
Total		2	15	11
	False negative			
1	Possible carcinoma in situ	-	-	-

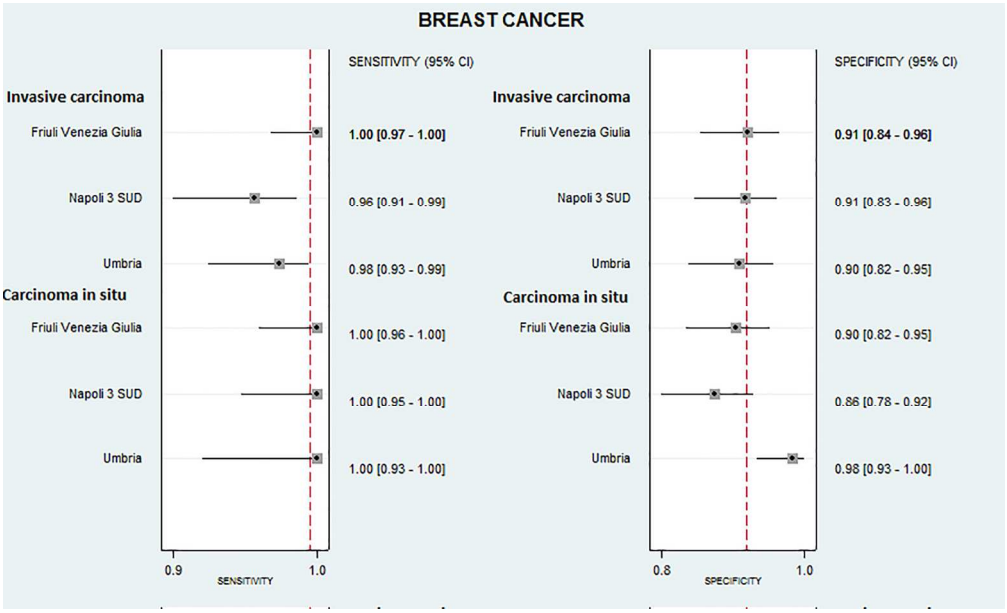
Table 3. Cross tabulation of the index test (ICD-9-CM code) results by the results of the reference standard (medical chart)

Type of brest cancer (ICD-9-CM)	Operative unit	TP	FP	TN	FN
Invasive cancer (174.x)	Unit 1 (Umbria)	118	10	90	3
	Unit 2 (ASL Napoli 3 Sud)	117	9	89	5
	Unit 3 (Friuli Venezia Giulia)	121	9	94	0
Carcinoma in situ (233.0)	Unit 1 (Umbria)	48	2	93	0
	Unit 2 (ASL Napoli 3 Sud)	73	15	94	0
	Unit 3 (Friuli Venezia Giulia)	97	11	94	0



Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined

201x210mm (300 x 300 DPI)



Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases

173x104mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	7
Participants	**6**	Eligibility criteria	6-8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	6-7
	8	Where and when potentially eligible participants were identified (setting, location and dates)	6-7
	9	Whether participants formed a consecutive, random or convenience series	7
Test methods	**10a**	Index test, in sufficient detail to allow replication	7
	10b	Reference standard, in sufficient detail to allow replication	7
	11	Rationale for choosing the reference standard (if alternatives exist)	7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	na
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	na
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8
	15	How indeterminate index test or reference standard results were handled	na
	16	How missing data on the index test and reference standard were handled	8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	na
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	Page 9 and Table 1
	21a	Distribution of severity of disease in those with the target condition	na
	21b	Distribution of alternative diagnoses in those without the target condition	na
	22	Time interval and any clinical interventions between index test and reference standard	na
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 3
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Page 9-10, Figure 2
	25	Any adverse events from performing the index test or the reference standard	na
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12-13
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	na
	29	Where the full study protocol can be accessed	5
	30	Sources of funding and other support; role of funders	14

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020627.R2
Article Type:	Research
Date Submitted by the Author:	06-Mar-2018
Complete List of Authors:	Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service; Agenzia Nazionale per i Servizi Sanitari Regionali (Age.Na.S.), Innovation and Development Serraino, Diego; IRCCS CRO AVIANO, Epidemiology Montedori, Alessandro; Regional Health Authority of Umbria, Fusco, Mario; ASL NA3 Sud, Registro Tumori Regione Campania Giovannini, Gianni; Regional Health Authority of Umbria, Health Planning Office Casucci, Paola; Regional Health Authority of Umbria, Health ICT Service Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Granata, Annalisa; ASL NA3 Sud, Registro Tumori Regione Campania De Giorgi, Marcello; Regional Health Authority of Umbria, , Health ICT Service Collarile, Paolo; Centro di Riferimento Oncologico Aviano, SOC Epidemiologia Oncologica Chiari, Rita; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Foglietta, Jennifer; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Vitale, Maria Francesca; Registro Tumori Regione Campania , ASL NA3 Sud Stracci, Fabrizio; University of Perugia, Public Health Department Orlandi, Walter; Regional Health Authority of Umbria Bidoli, Ettore; IRCCS CRO AVIANO, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Oncology
Keywords:	sensitivity and specificity, administrative database, validity, breast cancer

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Iosief Abraha[1 2], Diego Serraino[3], Alessandro Montedori[1], Mario Fusco[4], Gianni Giovannini[1], Paola Casucci[5], Francesco Cozzolino[1], Massimiliano Orso[1], Annalisa Granata[4], Marcello De Giorgi[5], Paolo Collarile[6], Rita Chiari[7], Jennifer Foglietta[7], Maria Francesca Vitale[4], Fabrizio Stracci[8], Walter Orlandi[9], Ettore Bidoli[3] for the D.I.V.O. Group

Author affiliations:

1. *Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy*
2. *Innovation and development, Agenzia Nazionale per i Servizi Sanitari Regionali (Age.Na.S.), Rome, Italy*
3. *Cancer Epidemiology Unit, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
4. *Registro Tumori Regione Campania, ASL Napoli 3 Sud, Brusciano, Italy*
5. *Health ICT Service, Regional Health Authority of Umbria, Perugia, Italy*
6. *SOC Epidemiologia Oncologica, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
7. *Dipartimento di Oncologia, Azienda Ospedaliera Perugia, Perugia, Italy*
8. *Public Health Department, University of Perugia, Perugia, Italy*
9. *Regional Health Authority of Umbria, Perugia, Italy*

Corresponding Author:
Alessandro Montedori
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni 61
06124 Perugia, Italy
amontedori@regione.umbria.it

Abstract

Objectives To assess the accuracy of *International Classification of Diseases 9th Revision Clinical Modification* (ICD-9-CM) codes in identifying patients diagnosed with incident carcinoma in situ and invasive breast cancer in three Italian administrative databases.

Design A diagnostic accuracy study comparing ICD-9-CM codes for carcinoma in situ (233.0) and for invasive breast cancer (174.x) with medical chart (as a reference standard). Case definition: (a) presence of a primary nodular lesion in the breast and (b) cytological or histological documentation of cancer from a primary or metastatic site.

Setting Administrative databases from Umbria Region, ASL Napoli 3 Sud (NA), and Friuli Venezia Giulia (FVG) Region

Participants Women with breast carcinoma in situ (n = 246) or invasive breast cancer (n = 384) diagnosed (in primary position) between 2012 and 2014.

Outcome measures Sensitivity and specificity for codes 233.0 and 174.x.

Results For invasive breast cancer the sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. Specificities were 90% (95% CI 82% to 95%) for Umbria, 91% (95% CI 83% to 96%) for NA and 91% (95% CI 84% to 96%) for FVG.

For carcinoma in situ the sensitivities were 100% (95% CI 93% to 100%) for Umbria, 100% (95% CI 95% to 100%) for NA, and 100% (95%CI 96% to 100%) for FVG. Specificities were 98% (95% CI 93% to 100%) for Umbria, 86% (95% CI 78% to 92%) for NA, and 90% (95% CI 82% to 95%) for FVG.

Conclusions Administrative healthcare databases from Umbria, NA and FVG are accurate in identifying hospitalized news cases of carcinoma of the breast. The proposed case definition is a powerful tool to perform research on large populations of newly diagnosed breast cancer patients.

Strengths and limitations of this study

- This study is the first to have validated *International Classification of Diseases-9th Revision – Clinical Modification* (ICD-9-CM) codes for incident breast cancer cases in three large computerized administrative databases in Italy using the same case definition.
- Case ascertainment was based on the presence of a primary nodular lesion in the breast documented by imaging and a cytological or histological documentation of cancer from a primary or metastatic site.
- This study followed 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.
- The validated ICD-9 codes for non-invasive and invasive cancer are limited to inpatient setting.
- The sample size of women with carcinoma in situ was limited due to the low prevalence of disease.

Introduction

The use of administrative databases is increasingly growing in various healthcare settings worldwide. These databases anonymously store data about residents regarding the healthcare assistance they receive including hospital admission, demographic data and disease treatment. Usually, the diagnosis of the disease is associated with a specific code from the *International Classification of Diseases, 9th 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 networking of individual patient data from administrative databases and other sources such as outpatient data and prescription data allows monitoring population health status, performing outcome research [2-4] and exploring a wide range of significant public health questions [2]. In administrative databases, while non-clinical data such as demographic or prescription data are highly accurate [5 6], the accuracy of diagnoses and procedures needs to be determined [6 7]. Typically, the assessment of accuracy consists in confirming the reliability of information within the databases with the corresponding clinical records of patients [5]. To reach this goal, the content of administrative healthcare databases needs to be appropriately validated.

In Italy, all the Regional Health Authorities maintain large healthcare information systems containing patient data from all hospital and operative sources. These databases have the potential to address important issues in post-marketing 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 realized since they have not been widely validated[12].

Breast cancer is the most commonly diagnosed neoplasm in women worldwide, as well as in Italy[13]. Variation in the epidemiology of breast cancers [14], potential heterogeneity in treatment (pharmacological or surgical) and potential clinical[15 16] and economic outcomes[17-19] can all be evaluated using validated administrative databases. Hence, assessing the accuracy of Italian

administrative databases in identifying women with this oncological disease is relevant for the scientific community, the governments, as well as the industry.

As reported in our protocol[20], the objective of the present study was to evaluate the accuracy of the ICD-9-CM codes related to breast cancer in correctly identifying the respective diseases using three large Italian administrative healthcare databases.

Methods

Setting and data source

Administrative databases

The target administrative databases were represented by the Umbria Region (890,000 residents), Local Health Unit 3 of NA (1,170,000 residents) and the FVG Region (1,227,000 residents). The corresponding operative units, the Regional Health Authority of Umbria (for Umbria Region), the Registro Tumori Regione Campania (for Napoli Sud Local Health Unit), and the Centro di Riferimento Oncologico Aviano (for Friuli Venezia Giulia Region), conducted the same validation process.

In Italy, regional and local healthcare administrative databases routinely collect data 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 5 secondary discharge diagnoses and the principal, and up to 5 secondary surgical or pharmacological treatments and diagnostic procedures. 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, health care is covered almost entirely by the Italian National Health System, therefore most residents' significant healthcare information can be found within the healthcare databases.

Source population

The source population was represented by permanent residents aged 18 or above of Umbria Region, Local Health Unit 3 of NA and FVG Region. Any resident that has been discharged from hospital with a diagnosis of breast cancer was considered. Residents that have been hospitalised outside the regional territory of competence were excluded from analysis.

Patient and Public Involvement

Patients were not directly involved. This was a retrospective study based on the consultation of medical charts. Ethical approval for the present study was obtained from the Ethics Committee of the Umbria Region Health Authority (CEAS).

Case and control selection and sampling method

In each administrative database, patients with the first occurrence of diagnosis of breast cancer between 1st January 2012 and 31st December 2014 were identified using the following ICD-9-CM codes (index test) located in primary position: (a) 233.0 for carcinoma in situ of the breast and (b) 174.x for invasive breast cancer. Estimated prevalent cases, that is, cases with the same diagnosis (ICD-9-CM codes in any position) in the five years (2007-2011) before the period of interest, were excluded.

For controls, within the same period, non-cases, i.e. 94 female patients having in primary position a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9) or carcinoma in situ of the breast (ICD-9 233.0), were randomly selected .

Subsequently, for each of the above reported groups of ICD-9-CM codes, random samples of cases and non-cases were selected from each administrative database.

Chart abstraction and case ascertainment

The medical charts of the randomly selected samples of cases and non-cases were obtained from hospitals for the validation task (considered as the reference standard). From each medical chart, the following information were retrieved: clinical chart number, hospital and ward, date of birth, sex, dates of hospital admission and discharge, signs and symptoms, any diagnostic procedures that contributed to the diagnosis of the cancer, any pharmacological or surgical interventions that were provided for the treatment of the cancer.

Within each unit, two reviewers received training on data abstraction and performed an initial consensus chart review, independently examining the same number of medical charts (n=20). The

inter-rater agreement regarding the presence or absence of breast cancer among the pairs of reviewers within each unit was near perfect (κ statistics > 0.91).

Case ascertainment of cancer within medical charts was based on (a) the presence of a primary nodular lesion in the breast documented by imaging and (b) the cytological or histological documentation of cancer from a primary or metastatic site.

Following the consensus review, data abstraction were completed independently. To ensure consistency among all the reviewers, cases with uncertainty were discussed and resolved through the involvement of an oncologist (RC).

Validation criteria

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

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

Statistical analysis

We calculated that a sample of 130 charts of cases was necessary to obtain an expected sensitivity of 80% with a 95% confidence interval (CI) of 72% - 86% according to binomial exact calculation[21]. For specificity calculation, we randomly selected non-cases, that is, records without the ICD-9 codes of interest from administrative database. For controls, we calculated a sample of 94 charts of non-cases was necessary to obtain an expected specificity of 90% with a 95% CI of 83% - 95% according to binomial exact calculation[21].

Sensitivity and specificity with relative 95% confidence intervals were calculated separately for each ICD-9-CM code by constructing 2 x 2 tables.

When missing medical charts occurred we performed a formal sensitivity analysis based on a worst case scenario in which the missing cases was considered as false positive and the missing-non cases were considered false negative.

Results

Invasive breast cancer

After excluding the estimated prevalent cases from the diagnosis of invasive breast cancer in the primary position of hospital discharges, the cases were 2,686, 2,044 and 2,107 from Umbria, NA and FVG respectively. Subsequently each team randomly sampled 130 cases of which the corresponding medical charts were requested for evaluation. Two and four medical charts were not available from Umbria and NA respectively. **Figure 1** displays the identification of cases from the three operative units. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The most common ICD-9-CM subgroup was the code 174.4, that is upper-outer quadrant breast cancer, accounting for 45% of cases for Umbria, 34% for NA, and 35% for FVG. The mean age ranged between 61 and 66 years. The majority of the cases were identified in surgical departments with a percentage ranging from 84% to 94%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units. The sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. The specificities were 90% for Umbria, and 91% for NA and FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

In terms of misclassification, overall there were 28 cases that were considered false positives. The reasons for this misclassification were: histological documentation missing in the medical chart (6 in Umbria, 8 in NA, and 8 in FVG) and negative histology for invasive breast cancer (4 in Umbria, 1 in NA, and 1 in FVG). However, of these false positive cases with negative histology for invasive breast cancer, three were positive for breast carcinoma in situ diagnosis. Conversely, there were 8

non-cases that were judged false negatives: two were possible breast cancer diagnosis and six were metastatic breast diagnoses (**Table 2**).

Overall there were six missing charts: two in Umbria and four in NA. Worst case scenario in the sensitivity analysis showed that specificity was affected marginally: it changed from 90% to 88% for Umbria and from 91% to 87% for NA. The differences between the ordinary results and the worst case scenario analysis were not statistically significant.

Breast carcinoma in situ

The incident cases of carcinoma in situ of the breast were 50 from Umbria, 95 from NA, and 137 from FVG, from which 50, 95 and 108 were randomly selected and the corresponding medical charts were requested for assessment (**Figure 1)**. Seven charts from NA were not available. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than carcinoma in situ of the breast (ICD-9 233.0), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The mean age ranged between 57 (NA) and 60 years (FVG). Most of the cases were identified in surgical departments with a percentage ranging from 92% to 100%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units.

After reviewing the medical records, 100% (48/48) of the patients with carcinoma in situ of the breast from Umbria, 100% (73/73) from NA, and 100% (97/97) from FVG were correctly identified by the administrative databases . The specificities were 98% for Umbria, 86% for NA, and 90% for FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

Table 2 describes the reasons for misclassification of cases and controls. Overall, there were 28 cases that were judged false positives. The reasons were histological documentation missing in the

medical charts (8 in NA and 2 in FVG), and negative histology for carcinoma in situ of the breast (2 in Umbria, 7 in NA and 9 in FVG). None of the controls resulted a false negative.

The sensitivity analysis showed that specificity for NA codes reduced to 81% (95% CI 73% to 88%) due to the seven charts of missing cases but the difference was not however statistically significant.

Discussion

We developed a case definition of breast cancer based on the presence of a primary nodular lesion in the breast documented with imaging and the cytological or histological documentation of cancer from a primary or metastatic site and the performance of the model was evaluated in terms of sensitivities and specificities for the three administrative databases. After revising the medical charts, the results showed that both codes (233.0 and 174.x) performed well in identifying new cases of hospitalized women with breast carcinoma in situ and invasive breast cancer respectively. Previously, researchers have assessed the accuracy of breast cancer diagnosis in administrative databases using different algorithms and in most cases using registry data as a reference standard[22]. 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[23]. The study found a sensitivity of 77% for breast cancer[23] and the lower value of sensitivity compared to our results can be attributed to the fact that in Baldi et al. the validity of the algorithm for each cancer site was assessed by individual matching between cases in hospital discharge and the Piedmont Cancer Registry or because authors were interested in high values of PPV. Using the Surveillance, Epidemiology, and End Results (SEER) database as a reference standard, Freeman et al. developed an approach for identifying incident breast cancer cases based on a logistic regression model, which contained variables that indicate the presence of breast cancer-related diagnoses and procedures in three sources of claims data: hospital inpatient stays, hospital outpatient services, and physician

services. The ROC curve showed that the model achieved over 90% of sensitivity and specificity albeit a lower positive predictive value[24]. Similarly, using the Medicare database, Setoguchi et al. developed a case definition based on diagnoses and procedure codes and compared it with SEER database in Pennsylvania. The authors obtained a sensitivity and specificity for identifying breast cancer cases of 83% and 99% respectively[25]. Using the Cancer Registry data as reference standard, Kemp et al.[26] evaluated Australian administrative and self-reported datasets to identify cases of invasive breast cancer and used several combinations of diagnoses and procedures obtaining the highest sensitivity and PPV (both 86%) when a flag of 'diagnosis of invasive breast cancer' was used. A systematic review of administrative databases that validated breast cancer is currently being completed and will provide a complete account of validation of administrative databases worldwide[22].

Strength and limitation

Strengths of our study include complete transparency based on pre-publication of a protocol, the use of detailed and explicit eligibility criteria, and the use of duplicate and independent processes for medical chart review and abstraction following 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 [27-29]. In addition, we used as a required element for validation the presence of a histological or cytological documentation. Unlike several studies that used Cancer Registries to validate breast cancer codes, in the present study medical records were reviewed directly to evaluate the accuracy of potential cases obtained from the administrative database. Generally, medical charts are considered the gold standard for the diagnosis of a disease. Cancer Registries are considered also the gold standard as they produce data over decades and increasingly include data sources[30], which allow complete registration of cases treated in an outpatient setting. This is relevant for the exclusion of prevalent cases and multiple primaries as well as for particular cancer sites or patients (e.g. oldest old) but at a higher cost[31].

Conversely, the information of administrative healthcare databases is generally limited to the information obtained from the Hospital Discharge Register that contain the primary and secondary diagnoses, the surgical or invasive diagnostic approaches performed as well as chemotherapy or radiotherapy. When adequately validated, these databases can provide important data such length of stay and related costs[31], important outcomes such as adverse events[32 33]as well as variations in health care resource utilization[34]. Indeed, administrative databases are readily available for the whole of Italy, whereas cancer registry data are not. Thus, our proposed validation method using a well-defined case definition can be a good alternative in settings in which cancer registries are not available in Italy. Our study confirms that hospital discharge data can be used for some specific purposes (e.g. identification of breast cancer cases treated at a given hospital in a study on caseload). For other aims we would recommend further refinement, even if the validity of the cancer coding is valid (e.g. to provide reliable estimates of breast cancer incidence hospital discharge data should be available for many years and possibly complete anatomic pathology archives should be linked to it too).

We acknowledge specific limitations to our study. The overall number of carcinoma in situ cases identified in the three units during the period of interest was below the calculated sample size and we are unsure whether this limitation can affect the results of sensitivity and specificity for the breast carcinoma in situ ICD-9-CM code. In addition, our assessment was limited to hospitalized patients and does not consider new cases of cancer who had the diagnoses in day hospital or day surgery. Although these cases are limited (e.g., 16 carcinoma in situ cases diagnosed in day surgery or day hospital in Umbria across the three years and 3.8% invasive breast cancer diagnosed in day surgery or day hospital in Umbria across the three years) further research can be addressed the validity of ICD-9 codes in outpatient setting.

In addition, we had a higher false positive rate than false negative rate. The number of false positives is due to our stringent case ascertainment criteria, i.e. the presence in the clinical chart of

both imaging and histological documentation of breast cancer within the same medical chart. Twenty-two false positives cases for the invasive breast cancer and ten false positives cases for the carcinoma in situ were due to histological documentation missing in the medical chart (**Table 2**). This does not necessarily mean that the subjects were without the diagnosis of cancer. These cases had other elements in the medical chart such as imaging, chemotherapy or radiotherapy, that could confirm the presence of the disease. Should we have used broader case ascertainment criteria we could have obtained a lower false positive rate. Estimates of specificity for invasive carcinoma in our three databases were high ranging from 90% to 91%. For carcinoma in situ specificity was acceptable for Napoli 3 Sud (86%) and high for FVG and Umbria (90% and 98% respectively).

In terms of generalizability, despite the success of validation processes of administrative database, any conclusion that stems from these validated database could not be generalized in other settings. As declared in our protocol we favoured the estimation of sensitivity and specificity rather than predictive values since predictive values are dependent on the prevalence of the disease. However, to comply with the STARD guidelines we provide absolute numbers for true or false case and non-cases from which it is possible to obtain predictive values (**Table 3)**.

Conclusion

In summary, the present study has demonstrated that administrative healthcare databases from Umbria, NA and FVG can be used to identify hospitalized women with newly diagnosed invasive or in situ carcinoma of the breast. The proposed case definition in the present study provides a powerful tool to perform outcome research on breast cancer based on a population of three million residents. Potential implication of this proposal includes the extension of this case definition to other Italian regional healthcare databases and the combination with other sources of data (such as prescription database) to conduct efficiently quality of care, health care research and pharmacoepidemiological studies that may complement randomised trials.

Footnotes

Contributors: AM, IA, MF, and DS conceived the original idea of the study. IA, DG, AM, MF, EB, GG, FC, MO, and WO designed the study. PCa, MDG, PCo, AG, MFV, VC identified the cohort using administrative database with the supervision of WO, EB, DS, MF, AM and FS. IA, FC, MO, AG, PC, VC, MFV, and JF undertook the data abstraction with the supervision of AM, GG, WO, FS, MF, EB, RC and DS. IA, RC, AM and JF performed case ascertainment. IA, AM, FC, EB, MF, PE and MO performed the analysis. DS, GG, PCa AG, MDG, PCo, RC, JF, MFV, FS, EB and WO helped in the interpretation of the data.
The initial draft of the manuscript was prepared by IA, AM, FC and MO. DS, EB, GG, PCa, AG, MDG, PCo, RC, JF, MFV, FS, and WO revised critically the manuscript for important intellectual content. All the authors read and approved the final manuscript. AM, MF and EB are the guarantors of the data for the respective operative units.

Funding This study was developed within the D.I.V.O. project (*Realizzazione di un **D**atabase **I**nterregionale **V**alidato per l'**O**ncologia 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.

Data sharing statement No additional data are available.

Figure Caption

Figure 1. Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined
Figure 2. Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases.

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. *Annual review of public health* 2011;32:91-108. doi: 10.1146/annurev-publhealth-031210-100700 [published Online First: 2011/01/12]
3. Jick SS. Fresh evidence confirms links between newer contraceptive pills and higher risk of venous thromboembolism. *Bmj* 2015;350:h2422. doi: 10.1136/bmj.h2422 [published Online First: 2015/05/28]
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.
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(3):205-11. [published Online First: 2001/10/05]
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(7405):18-22. doi: 10.1136/bmj.327.7405.18
9. Trifiro G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Studies in health technology and informatics* 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. doi: 10.1186/1471-2458-13-15
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 services research* 2013;13(1):393. doi: 10.1186/1472-6963-13-393
12. Abraha I, Montedori A, Eusebi P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. *Pharmacoepidemiology and Drug Safety* 2012;21:400-00.
13. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer Journal international du cancer* 2015;136(5):E359-86. doi: 10.1002/ijc.29210 [published Online First: 2014/09/16]
14. 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. *European journal of cancer care* 2015;24(3):367-75. doi: 10.1111/ecc.12277 [published Online First: 2014/12/19]
15. Lambertini M, Anserini P, Fontana V, et al. The PREgnancy and FERtility (PREFER) study: an Italian multicenter prospective cohort study on fertility preservation and pregnancy issues in young breast cancer patients. *BMC cancer* 2017;17(1):346. doi: 10.1186/s12885-017-3348-8 [published Online First: 2017/05/21]
16. De Placido S, De Angelis C, Giuliano M, et al. Imaging tests in staging and surveillance of non-metastatic breast cancer: changes in routine clinical practice and cost implications. *Br J Cancer* 2017;116(6):821-27. doi: 10.1038/bjc.2017.24 [published Online First: 2017/02/09]
17. 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 research and treatment* 2013;139(2):469-76. doi: 10.1007/s10549-013-2543-9 [published Online First: 2013/04/30]

18. Konski A. Clinical and economic outcomes analyses of women developing breast cancer in a managed care organization. *American journal of clinical oncology* 2005;28(1):51-7. [published Online First: 2005/02/03]
19. 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(1):E11-7. doi: 10.9778/cmajo.20130026 [published Online First: 2014/08/01]
20. 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(3):e010547. doi: 10.1136/bmjopen-2015-010547
21. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927;22:209-12.
22. 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(3):e010409. doi: 10.1136/bmjopen-2015-010409
23. Baldi I, Vicari P, Di Cuonzo D, et al. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. *Journal of clinical epidemiology* 2008;61(4):373-9.
24. Freeman JL, Zhang D, Freeman DH, et al. An approach to identifying incident breast cancer cases using Medicare claims data. *Journal of clinical epidemiology* 2000;53(6):605-14. doi: 10.1016/s0895-4356(99)00173-0
25. Setoguchi S, Solomon DH, Glynn RJ, et al. Agreement of diagnosis and its date for hematologic malignancies and solid tumors between medicare claims and cancer registry data. *Cancer causes & control : CCC* 2007;18(5):561-9. doi: 10.1007/s10552-007-0131-1 [published Online First: 2007/04/21]
26. Kemp A, Preen DB, Saunders C, et al. Ascertaining invasive breast cancer cases; the validity of administrative and self-reported data sources in Australia. *BMC medical research methodology* 2013;13:17. doi: 10.1186/1471-2288-13-17 [published Online First: 2013/02/13]
27. 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. *Journal of clinical epidemiology* 2011;64(8):821-9. doi: 10.1016/j.jclinepi.2010.10.006
28. 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 services research* 2006;6:77. doi: 10.1186/1472-6963-6-77
29. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Annals of internal medicine* 2003;138(1):40-4.
30. Gliklich RE, Dreyer NA, Leavy MB. Registries for Evaluating Patient Outcomes: A User's Guide [Internet]. . 2014
31. Rice HE, Englum BR, Gulack BC, et al. Use of Patient Registries and Administrative Datasets for the Study of Pediatric Cancer. *Pediatric blood & cancer* 2015;62(9):1495-500. doi: 10.1002/pbc.25506
32. Rodrigo-Rincon I, Martin-Vizcaino MP, Tirapu-Leon B, et al. Validity of the clinical and administrative databases in detecting post-operative adverse events. *International Journal for Quality in Health Care* 2015;27(4):267-75. doi: 10.1093/intqhc/mzv039
33. Rashid N, Koh HA, Baca HC, et al. Economic burden related to chemotherapy-related adverse events in patients with metastatic breast cancer in an integrated health care system. *Breast Cancer : Targets and Therapy* 2016;8:173-81. doi: 10.2147/BCTT.S105618
34. Mahar AL, Coburn NG, Kagedan DJ, et al. Regional variation in the management of metastatic gastric cancer in Ontario. *Current Oncology* 2016;23(4):250-57. doi: 10.3747/co.23.3123

Table 1 Characteristics of patients with breast cancer who were identified in the three administrative healthcare databases

Characteristics	Unit 1 (Umbria)	Unit 2 (Asl Napoli 3 Sud)	Unit 3 (Friuli Venezia Giulia)
Invasive carcinoma			
Incident cases (N medical chart reviewed)	128	126	130
ICD-9 code			
174.0 nipple and areola	-	1 (1)	-
174.1 central portion	16 (13)	10 (8)	6 (5)
174.2 upper-inner quadrant	4 (3)	8 (6)	14 (11)
174.3 lower-inner quadrant	6 (5)	5 (4)	9 (7)
174.4 upper-outer quadrant	57 (45)	43 (34)	45 (35)
174.5 lower-outer quadrant	6 (5)	5 (4)	13 (10)
174.6 axillary tail	-	-	-
174.8 other specified sites of the female breast	-	-	-
174.9 breast female, unspecified	7 (5)	38 (30)	34 (26)
174.0 nipple and areola	32 (25)	16 (13)	9 (7)
Admission to department			
Medical	20 (16)	11 (9)	8 (6)
Surgical	108 (84)	115 (91)	122 (94)
Age, N (%)			
< 40	9 (7)	6 (5)	1 (1)
40 - 59	40 (31)	56 (44)	45 (35)
≥ 60	79 (62)	64 (51)	84 (65)
Instrumental diagnosis			
Breast ultrasound	39 (30)	88 (70)	5 (4)
Mammography	54 (42)	60 (48)	7 (5)
CT scan (breast)	11 (8)	1 (1)	2 (2)
MRI (breast)	3 (2)	17 (14)	8 (6)
Surgical procedures			
Mastectomy	28 (22)	29 (23)	35 (27)
Quadrantectomy	79 (62)	73 (58)	54 (42)
Hystological documentation			
Needle aspiration	32	34	-
Needle biopsy	27	40	5
Nodule (after surgical intervention)	115	117	112
Carcinoma in situ			
Incident cases (N medical chart reviewed)	50	88	108
ICD-9 code			
233.0	50 (100)	88 (100)	108 (100)
Admission to department			
Medical	-	7 (8)	-
Surgical	50 (100)	81 (92)	108 (100)
Age, N (%)			

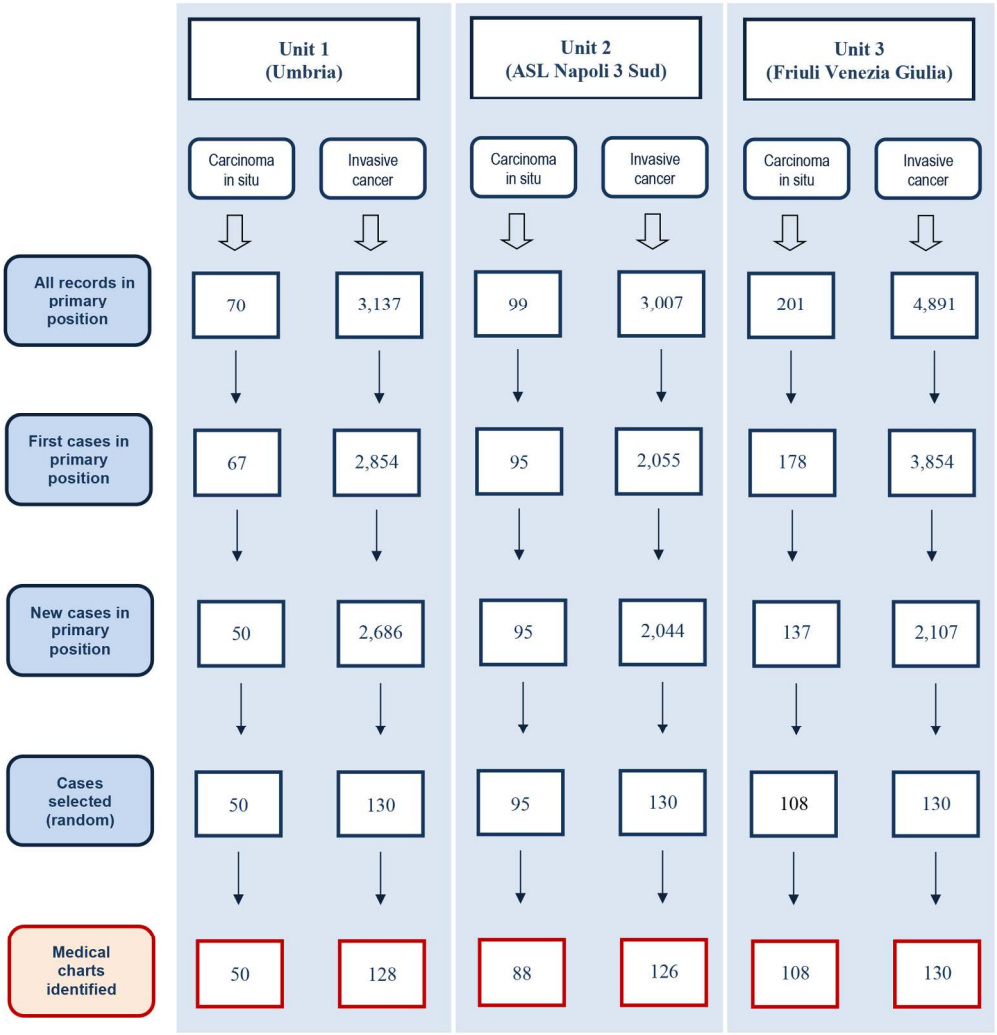
< 40	2 (4)	3 (3)	1 (1)
40 - 59	27 (54)	50 (57)	52 (48)
> 60	21 (42)	35 (40)	55 (51)
Instrumental diagnosis			
Breast ultrasound	3 (6)	65 (74)	30 (28)
Mammography	6 (12)	37 (42)	39 (36)
CT scan (breast)	-	-	-
MRI (breast)	2 (4)	22 (25)	4 (4)
Surgical procedures			
Mastectomy	18 (36)	13 (15)	15 (14)
Quadrantectomy	32 (64)	47 (53)	55 (52)
Hystological documentation			
Needle aspiration	-	16 (18)	1 (1)
Needle biopsy	8 (16)	53 (60)	4 (4)
Nodule (after surgical intervention)	50 (100)	77 (88)	94 (87)

Table 2. Reason for misclassification of cases and controls

	Invasive breast cancer			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	6	8	8
2	negative histology	4	1	1
	a) carcinoma in situ	*2*		*1*
Total		10	9	9
	False negative			
1	possible breast cancer relapse	1	1	-
2	metastatic breast cancer	2	4	-
Total		3	5	-
	Breast carcinoma in situ			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	-	8	2
2	negative histology	2	7	9
Total		2	15	11
	False negative			
1	Possible carcinoma in situ	-	-	-

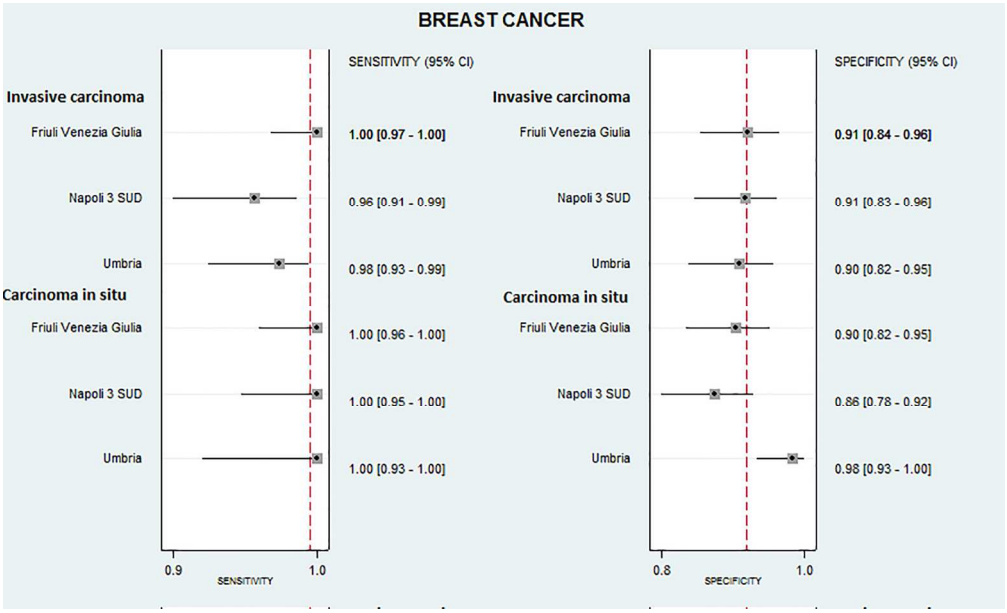
Table 3. Cross tabulation of the index test (ICD-9-CM code) results by the results of the reference standard (medical chart)

Type of brest cancer (ICD-9-CM)	Operative unit	TP	FP	TN	FN
Invasive cancer (174.x)	Unit 1 (Umbria)	118	10	90	3
	Unit 2 (ASL Napoli 3 Sud)	117	9	89	5
	Unit 3 (Friuli Venezia Giulia)	121	9	94	0
Carcinoma in situ (233.0)	Unit 1 (Umbria)	48	2	93	0
	Unit 2 (ASL Napoli 3 Sud)	73	15	94	0
	Unit 3 (Friuli Venezia Giulia)	97	11	94	0



Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined

173x181mm (300 x 300 DPI)



Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases

173x104mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	7
Participants	**6**	Eligibility criteria	6-8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	6-7
	8	Where and when potentially eligible participants were identified (setting, location and dates)	6-7
	9	Whether participants formed a consecutive, random or convenience series	7
Test methods	**10a**	Index test, in sufficient detail to allow replication	7
	10b	Reference standard, in sufficient detail to allow replication	7
	11	Rationale for choosing the reference standard (if alternatives exist)	7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	na
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	na
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8
	15	How indeterminate index test or reference standard results were handled	na
	16	How missing data on the index test and reference standard were handled	8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	na
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	Page 9 and Table 1
	21a	Distribution of severity of disease in those with the target condition	na
	21b	Distribution of alternative diagnoses in those without the target condition	na
	22	Time interval and any clinical interventions between index test and reference standard	na
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 3
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Page 9-10, Figure 2
	25	Any adverse events from performing the index test or the reference standard	na
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12-13
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	na
	29	Where the full study protocol can be accessed	5
	30	Sources of funding and other support; role of funders	14

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020627.R3
Article Type:	Research
Date Submitted by the Author:	25-Apr-2018
Complete List of Authors:	Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service; Agenzia Nazionale per i Servizi Sanitari Regionali (Age.Na.S.), Innovation and Development Serraino, Diego; IRCCS CRO AVIANO, Epidemiology Montedori, Alessandro; Regional Health Authority of Umbria, Fusco, Mario; ASL NA3 Sud, Registro Tumori Regione Campania Giovannini, Gianni; Regional Health Authority of Umbria, Health Planning Office Casucci, Paola; Regional Health Authority of Umbria, Health ICT Service Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Granata, Annalisa; ASL NA3 Sud, Registro Tumori Regione Campania De Giorgi, Marcello; Regional Health Authority of Umbria, , Health ICT Service Collarile, Paolo; Centro di Riferimento Oncologico Aviano, SOC Epidemiologia Oncologica Chiari, Rita; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Foglietta, Jennifer; Azienda Ospedaliera Perugia, Dipartimento di Oncologia Vitale, Maria Francesca; Registro Tumori Regione Campania , ASL NA3 Sud Stracci, Fabrizio; University of Perugia, Public Health Department Orlandi, Walter; Regional Health Authority of Umbria Bidoli, Ettore; IRCCS CRO AVIANO, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Oncology
Keywords:	sensitivity and specificity, administrative database, validity, breast cancer

Sensitivity and specificity of breast cancer ICD-9-CM codes in three Italian administrative healthcare databases: a diagnostic accuracy study

Iosief Abraha[1 2], Diego Serraino[3], Alessandro Montedori[1], Mario Fusco[4], Gianni Giovannini[1], Paola Casucci[5], Francesco Cozzolino[1], Massimiliano Orso[1], Annalisa Granata[4], Marcello De Giorgi[5], Paolo Collarile[6], Rita Chiari[7], Jennifer Foglietta[7], Maria Francesca Vitale[4], Fabrizio Stracci[8], Walter Orlandi[9], Ettore Bidoli[3] for the D.I.V.O. Group

Author affiliations:

1. *Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy*
2. *Innovation and development, Agenzia Nazionale per i Servizi Sanitari Regionali (Age.Na.S.), Rome, Italy*
3. *Cancer Epidemiology Unit, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
4. *Registro Tumori Regione Campania, ASL Napoli 3 Sud, Brusciano, Italy*
5. *Health ICT Service, Regional Health Authority of Umbria, Perugia, Italy*
6. *SOC Epidemiologia Oncologica, Centro di Riferimento Oncologico Aviano, Aviano, Italy*
7. *Dipartimento di Oncologia, Azienda Ospedaliera Perugia, Perugia, Italy*
8. *Public Health Department, University of Perugia, Perugia, Italy*
9. *Regional Health Authority of Umbria, Perugia, Italy*

Corresponding Author:
Alessandro Montedori
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni 61
06124 Perugia, Italy
amontedori@regione.umbria.it

Abstract

Objectives To assess the accuracy of *International Classification of Diseases 9th Revision Clinical Modification* (ICD-9-CM) codes in identifying patients diagnosed with incident carcinoma in situ and invasive breast cancer in three Italian administrative databases.

Design A diagnostic accuracy study comparing ICD-9-CM codes for carcinoma in situ (233.0) and for invasive breast cancer (174.x) with medical chart (as a reference standard). Case definition: (a) presence of a primary nodular lesion in the breast and (b) cytological or histological documentation of cancer from a primary or metastatic site.

Setting Administrative databases from Umbria Region, ASL Napoli 3 Sud (NA), and Friuli Venezia Giulia (FVG) Region

Participants Women with breast carcinoma in situ (n = 246) or invasive breast cancer (n = 384) diagnosed (in primary position) between 2012 and 2014.

Outcome measures Sensitivity and specificity for codes 233.0 and 174.x.

Results For invasive breast cancer the sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. Specificities were 90% (95% CI 82% to 95%) for Umbria, 91% (95% CI 83% to 96%) for NA and 91% (95% CI 84% to 96%) for FVG.

For carcinoma in situ the sensitivities were 100% (95% CI 93% to 100%) for Umbria, 100% (95% CI 95% to 100%) for NA, and 100% (95%CI 96% to 100%) for FVG. Specificities were 98% (95% CI 93% to 100%) for Umbria, 86% (95% CI 78% to 92%) for NA, and 90% (95% CI 82% to 95%) for FVG.

Conclusions Administrative healthcare databases from Umbria, NA and FVG are accurate in identifying hospitalized news cases of carcinoma of the breast. The proposed case definition is a powerful tool to perform research on large populations of newly diagnosed breast cancer patients.

Strengths and limitations of this study

- This study is the first to have validated *International Classification of Diseases-9th Revision – Clinical Modification* (ICD-9-CM) codes for incident breast cancer cases in three large computerized administrative databases in Italy using the same case definition.
- Case ascertainment was based on the presence of a primary nodular lesion in the breast documented by imaging and a cytological or histological documentation of cancer from a primary or metastatic site.
- This study followed 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.
- The validated ICD-9 codes for non-invasive and invasive cancer are limited to inpatient setting.
- The sample size of women with carcinoma in situ was limited due to the low prevalence of disease.

Introduction

The use of administrative databases is increasingly growing in various healthcare settings worldwide. These databases anonymously store data about residents regarding the healthcare assistance they receive including hospital admission, demographic data and disease treatment. Usually, the diagnosis of the disease is associated with a specific code from the *International Classification of Diseases, 9th 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 networking of individual patient data from administrative databases and other sources such as outpatient data and prescription data allows monitoring population health status, performing outcome research [2-4] and exploring a wide range of significant public health questions [2]. In administrative databases, while non-clinical data such as demographic or prescription data are highly accurate [5 6], the accuracy of diagnoses and procedures needs to be determined [6 7]. Typically, the assessment of accuracy consists in confirming the reliability of information within the databases with the corresponding clinical records of patients [5]. To reach this goal, the content of administrative healthcare databases needs to be appropriately validated.

In Italy, all the Regional Health Authorities maintain large healthcare information systems containing patient data from all hospital and operative sources. These databases have the potential to address important issues in post-marketing 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 realized since they have not been widely validated[12].

Breast cancer is the most commonly diagnosed neoplasm in women worldwide, as well as in Italy[13]. Variation in the epidemiology of breast cancers [14], potential heterogeneity in treatment (pharmacological or surgical) and potential clinical[15 16] and economic outcomes[17-19] can all be evaluated using validated administrative databases. Hence, assessing the accuracy of Italian

administrative databases in identifying women with this oncological disease is relevant for the scientific community, the governments, as well as the industry.

As reported in our protocol[20], the objective of the present study was to evaluate the accuracy of the ICD-9-CM codes related to breast cancer in correctly identifying the respective diseases using three large Italian administrative healthcare databases.

Methods

Setting and data source

Administrative databases

The target administrative databases were represented by the Umbria Region (890,000 residents), Local Health Unit 3 of NA (1,170,000 residents) and the FVG Region (1,227,000 residents). The corresponding operative units, the Regional Health Authority of Umbria (for Umbria Region), the Registro Tumori Regione Campania (for Napoli Sud Local Health Unit), and the Centro di Riferimento Oncologico Aviano (for Friuli Venezia Giulia Region), conducted the same validation process.

In Italy, regional and local healthcare administrative databases routinely collect data 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 5 secondary discharge diagnoses and the principal, and up to 5 secondary surgical or pharmacological treatments and diagnostic procedures. 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, health care is covered almost entirely by the Italian National Health System, therefore most residents' significant healthcare information can be found within the healthcare databases.

Source population

The source population was represented by permanent residents aged 18 or above of Umbria Region, Local Health Unit 3 of NA and FVG Region. Any resident that has been discharged from hospital with a diagnosis of breast cancer was considered. Residents that have been hospitalised outside the regional territory of competence were excluded from analysis.

Patient and Public Involvement

Patients were not directly involved. This was a retrospective study based on the consultation of medical charts. Ethical approval for the present study was obtained from the Ethics Committee of the Umbria Region Health Authority (CEAS).

Case and control selection and sampling method

In each administrative database, patients with the first occurrence of diagnosis of breast cancer between 1st January 2012 and 31st December 2014 were identified using the following ICD-9-CM codes (index test) located in primary position: (a) 233.0 for carcinoma in situ of the breast and (b) 174.x for invasive breast cancer. Estimated prevalent cases, that is, cases with the same diagnosis (ICD-9-CM codes in any position) in the five years (2007-2011) before the period of interest, were excluded.

For controls, within the same period, non-cases, i.e. 94 female patients having in primary position a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9) or carcinoma in situ of the breast (ICD-9 233.0), were randomly selected .

Subsequently, for each of the above reported groups of ICD-9-CM codes, random samples of cases and non-cases were selected from each administrative database.

Chart abstraction and case ascertainment

The medical charts of the randomly selected samples of cases and non-cases were obtained from hospitals for the validation task (considered as the reference standard). From each medical chart, the following information were retrieved: clinical chart number, hospital and ward, date of birth, sex, dates of hospital admission and discharge, signs and symptoms, any diagnostic procedures that contributed to the diagnosis of the cancer, any pharmacological or surgical interventions that were provided for the treatment of the cancer.

Within each unit, two reviewers received training on data abstraction and performed an initial consensus chart review, independently examining the same number of medical charts (n=20). The

inter-rater agreement regarding the presence or absence of breast cancer among the pairs of reviewers within each unit was near perfect (κ statistics > 0.91).

Case ascertainment of cancer within medical charts was based on (a) the presence of a primary nodular lesion in the breast documented by imaging and (b) the cytological or histological documentation of cancer from a primary or metastatic site.

Following the consensus review, data abstraction were completed independently. To ensure consistency among all the reviewers, cases with uncertainty were discussed and resolved through the involvement of an oncologist (RC).

Validation criteria

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

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

Statistical analysis

We calculated that a sample of 130 charts of cases was necessary to obtain an expected sensitivity of 80% with a 95% confidence interval (CI) of 72% - 86% according to binomial exact calculation[21]. For specificity calculation, we randomly selected non-cases, that is, records without the ICD-9 codes of interest from administrative database. For controls, we calculated a sample of 94 charts of non-cases was necessary to obtain an expected specificity of 90% with a 95% CI of 83% - 95% according to binomial exact calculation[21].

Sensitivity and specificity with relative 95% confidence intervals were calculated separately for each ICD-9-CM code by constructing 2 x 2 tables.

When missing medical charts occurred we performed a formal sensitivity analysis based on a worst case scenario in which the missing cases was considered as false positive and the missing-non cases were considered false negative.

Results

Invasive breast cancer

After excluding the estimated prevalent cases from the diagnosis of invasive breast cancer in the primary position of hospital discharges, the cases were 2,686, 2,044 and 2,107 from Umbria, NA and FVG respectively. Subsequently each team randomly sampled 130 cases of which the corresponding medical charts were requested for evaluation. Two and four medical charts were not available from Umbria and NA respectively. **Figure 1** displays the identification of cases from the three operative units. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than invasive breast cancer (ICD-9 174.0-174.9), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The most common ICD-9-CM subgroup was the code 174.4, that is upper-outer quadrant breast cancer, accounting for 45% of cases for Umbria, 34% for NA, and 35% for FVG. The mean age ranged between 61 and 66 years. The majority of the cases were identified in surgical departments with a percentage ranging from 84% to 94%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units. The sensitivities were 98% (95% CI 93% to 99%) for Umbria, 96% (95% CI 91% to 99%) for NA, and 100% (95% CI 97% to 100%) for FVG. The specificities were 90% for Umbria, and 91% for NA and FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

In terms of misclassification, overall there were 28 cases that were considered false positives. The reasons for this misclassification were: histological documentation missing in the medical chart (6 in Umbria, 8 in NA, and 8 in FVG) and negative histology for invasive breast cancer (4 in Umbria, 1 in NA, and 1 in FVG). However, of these false positive cases with negative histology for invasive breast cancer, three were positive for breast carcinoma in situ diagnosis. Conversely, there were 8

non-cases that were judged false negatives: two were possible breast cancer diagnosis and six were metastatic breast diagnoses (**Table 2**).

Overall there were six missing charts: two in Umbria and four in NA. Worst case scenario in the sensitivity analysis showed that specificity was affected marginally: it changed from 90% to 88% for Umbria and from 91% to 87% for NA. The differences between the ordinary results and the worst case scenario analysis were not statistically significant.

Breast carcinoma in situ

The incident cases of carcinoma in situ of the breast were 50 from Umbria, 95 from NA, and 137 from FVG, from which 50, 95 and 108 were randomly selected and the corresponding medical charts were requested for assessment (**Figure 1)**. Seven charts from NA were not available. For the non-cases, i.e. female patients having a diagnosis of cancer (ICD-9 140-239) other than carcinoma in situ of the breast (ICD-9 233.0), each unit randomly selected 94 medical charts. One medical chart of non-cases from Umbria was missing.

The mean age ranged between 57 (NA) and 60 years (FVG). Most of the cases were identified in surgical departments with a percentage ranging from 92% to 100%. The types of surgical intervention were similar across the three operative units with quandrantectomy being the most reported surgical intervention. **Table 1** describes the basic characteristics of the incident cases across the three units.

After reviewing the medical records, 100% (48/48) of the patients with carcinoma in situ of the breast from Umbria, 100% (73/73) from NA, and 100% (97/97) from FVG were correctly identified by the administrative databases . The specificities were 98% for Umbria, 86% for NA, and 90% for FVG. Accuracy results with their confidence intervals are displayed in **Figure 2**.

Table 2 describes the reasons for misclassification of cases and controls. Overall, there were 28 cases that were judged false positives. The reasons were histological documentation missing in the

medical charts (8 in NA and 2 in FVG), and negative histology for carcinoma in situ of the breast (2 in Umbria, 7 in NA and 9 in FVG). None of the controls resulted a false negative.

The sensitivity analysis showed that specificity for NA codes reduced to 81% (95% CI 73% to 88%) due to the seven charts of missing cases but the difference was not however statistically significant.

Discussion

We developed a case definition of breast cancer based on the presence of a primary nodular lesion in the breast documented with imaging and the cytological or histological documentation of cancer from a primary or metastatic site and the performance of the model was evaluated in terms of sensitivities and specificities for the three administrative databases. After revising the medical charts, the results showed that both codes (233.0 and 174.x) performed well in identifying new cases of hospitalized women with breast carcinoma in situ and invasive breast cancer respectively. Previously, researchers have assessed the accuracy of breast cancer diagnosis in administrative databases using different algorithms and in most cases using registry data as a reference standard[22]. 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[23]. The study found a sensitivity of 77% for breast cancer[23] and the lower value of sensitivity compared to our results can be attributed to the fact that in Baldi et al. the validity of the algorithm for each cancer site was assessed by individual matching between cases in hospital discharge and the Piedmont Cancer Registry or because authors were interested in high values of PPV. Using the Surveillance, Epidemiology, and End Results (SEER) database as a reference standard, Freeman et al. developed an approach for identifying incident breast cancer cases based on a logistic regression model, which contained variables that indicate the presence of breast cancer-related diagnoses and procedures in three sources of claims data: hospital inpatient stays, hospital outpatient services, and physician

services. The ROC curve showed that the model achieved over 90% of sensitivity and specificity albeit a lower positive predictive value[24]. Similarly, using the Medicare database, Setoguchi et al. developed a case definition based on diagnoses and procedure codes and compared it with SEER database in Pennsylvania. The authors obtained a sensitivity and specificity for identifying breast cancer cases of 83% and 99% respectively[25]. Using the Cancer Registry data as reference standard, Kemp et al.[26] evaluated Australian administrative and self-reported datasets to identify cases of invasive breast cancer and used several combinations of diagnoses and procedures obtaining the highest sensitivity and PPV (both 86%) when a flag of 'diagnosis of invasive breast cancer' was used. A systematic review of administrative databases that validated breast cancer is currently being completed and will provide a complete account of validation of administrative databases worldwide[22].

Strength and limitation

Strengths of our study include complete transparency based on pre-publication of a protocol, the use of detailed and explicit eligibility criteria, and the use of duplicate and independent processes for medical chart review and abstraction following 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 [27-29]. In addition, we used as a required element for validation the presence of a histological or cytological documentation. Unlike several studies that used Cancer Registries to validate breast cancer codes, in the present study medical records were reviewed directly to evaluate the accuracy of potential cases obtained from the administrative database. Generally, medical charts are considered the gold standard for the diagnosis of a disease. Cancer Registries are considered also the gold standard as they produce data over decades and increasingly include data sources[30], which allow complete registration of cases treated in an outpatient setting. This is relevant for the exclusion of prevalent cases and multiple primaries as well as for particular cancer sites or patients (e.g. oldest old) but at a higher cost[31].

Conversely, the information of administrative healthcare databases is generally limited to the information obtained from the Hospital Discharge Register that contain the primary and secondary diagnoses, the surgical or invasive diagnostic approaches performed as well as chemotherapy or radiotherapy. When adequately validated, these databases can provide important data such length of stay and related costs[31], important outcomes such as adverse events[32 33]as well as variations in health care resource utilization[34]. Indeed, administrative databases are readily available for the whole of Italy, whereas cancer registry data are not. Thus, our proposed validation method using a well-defined case definition can be a good alternative in settings in which cancer registries are not available in Italy[20 35 36]. Our study confirms that hospital discharge data can be used for some specific purposes (e.g. identification of breast cancer cases treated at a given hospital in a study on caseload). For other aims we would recommend further refinement, even if the validity of the cancer coding is valid (e.g. to provide reliable estimates of breast cancer incidence hospital discharge data should be available for many years and possibly complete anatomic pathology archives should be linked to it too).

We acknowledge specific limitations to our study. The overall number of carcinoma in situ cases identified in the three units during the period of interest was below the calculated sample size and we are unsure whether this limitation can affect the results of sensitivity and specificity for the breast carcinoma in situ ICD-9-CM code. Indeed, within the figure of the overall breast cancer diagnoses, carcinoma in situ cases diagnosed within hospitals are underrepresented and any future epidemiological assessment will need to take it into account trying possibly to clarify the reason.

In addition, our assessment was limited to hospitalized patients and does not consider new cases of cancer who had the diagnoses in day hospital or day surgery. Although these cases are limited (e.g., 16 carcinoma in situ cases diagnosed in day surgery or day hospital in Umbria across the three years and 3.8% invasive breast cancer diagnosed in day surgery or day hospital in Umbria across the three years) further research can be addressed the validity of ICD-9 codes in outpatient setting.

In addition, we had a higher false positive rate than false negative rate. The number of false positives is due to our stringent case ascertainment criteria, i.e. the presence in the clinical chart of both imaging and histological documentation of breast cancer within the same medical chart. Twenty-two false positives cases for the invasive breast cancer and ten false positives cases for the carcinoma in situ were due to histological documentation missing in the medical chart (**Table 2**). This does not necessarily mean that the subjects were without the diagnosis of cancer. These cases had other elements in the medical chart such as imaging, chemotherapy or radiotherapy, that could confirm the presence of the disease. Should we have used broader case ascertainment criteria we could have obtained a lower false positive rate. Estimates of specificity for invasive carcinoma in our three databases were high ranging from 90% to 91%. For carcinoma in situ specificity was acceptable for Napoli 3 Sud (86%) and high for FVG and Umbria (90% and 98% respectively).
In terms of generalizability, despite the success of validation processes of administrative database, any conclusion that stems from these validated database could not be generalized in other settings. As declared in our protocol we favoured the estimation of sensitivity and specificity rather than predictive values since predictive values are dependent on the prevalence of the disease. However, to comply with the STARD guidelines we provide absolute numbers for true or false case and non-cases from which it is possible to obtain predictive values (**Table 3)**.

Conclusion

In summary, the present study has demonstrated that administrative healthcare databases from Umbria, NA and FVG can be used to identify hospitalized women with newly diagnosed invasive or in situ carcinoma of the breast. The proposed case definition in the present study provides a powerful tool to perform outcome research on breast cancer based on a population of three million residents. Potential implication of this proposal includes the extension of this case definition to other Italian regional healthcare databases and the combination with other sources of data (such as

prescription database) to conduct efficiently quality of care, health care research and pharmacoepidemiological studies that may complement randomised trials.

Footnotes

Contributors: AM, IA, MF, and DS conceived the original idea of the study. IA, AM, MF, EB, GG, FC, MO, and WO designed the study. PCa, MDG, PCo, AG, MFV, identified the cohort using administrative database with the supervision of WO, EB, DS, MFV, AM and FS. IA, FC, MO, AG, PCa, MFV, and JF undertook the data abstraction with the supervision of AM, GG, WO, FS, MF, EB, RC and DS. IA, RC, AM and JF performed case ascertainment. IA, AM, FC, EB, MFV, and MO performed the analysis. DS, GG, PCa, AG, MDG, PCo, RC, JF, MFV, FS, EB and WO helped in the interpretation of the data.
The initial draft of the manuscript was prepared by IA, AM, FC and MO. DS, EB, GG, PCa, AG, MDG, PCo, RC, JF, MFV, FS, and WO revised critically the manuscript for important intellectual content. All the authors read and approved the final manuscript. AM, MF and EB are the guarantors of the data for the respective operative units.

Funding This study was developed within the D.I.V.O. project (*Realizzazione di un **D**atabase **I**nterregionale **V**alidato per l'**O**ncologia 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.

Data sharing statement No additional data are available.

Figure Caption

Figure 1. Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined
Figure 2. Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases.

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. *Annual review of public health* 2011;32:91-108. doi: 10.1146/annurev-publhealth-031210-100700 [published Online First: 2011/01/12]
3. Jick SS. Fresh evidence confirms links between newer contraceptive pills and higher risk of venous thromboembolism. *Bmj* 2015;350:h2422. doi: 10.1136/bmj.h2422 [published Online First: 2015/05/28]
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.
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(3):205-11. [published Online First: 2001/10/05]
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(7405):18-22. doi: 10.1136/bmj.327.7405.18
9. Trifiro G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Studies in health technology and informatics* 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. doi: 10.1186/1471-2458-13-15
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 services research* 2013;13(1):393. doi: 10.1186/1472-6963-13-393
12. Abraha I, Montedori A, Eusebi P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. *Pharmacoepidemiology and Drug Safety* 2012;21:400-00.
13. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer Journal international du cancer* 2015;136(5):E359-86. doi: 10.1002/ijc.29210 [published Online First: 2014/09/16]
14. 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. *European journal of cancer care* 2015;24(3):367-75. doi: 10.1111/ecc.12277 [published Online First: 2014/12/19]
15. Lambertini M, Anserini P, Fontana V, et al. The PREgnancy and FERtility (PREFER) study: an Italian multicenter prospective cohort study on fertility preservation and pregnancy issues in young breast cancer patients. *BMC cancer* 2017;17(1):346. doi: 10.1186/s12885-017-3348-8 [published Online First: 2017/05/21]
16. De Placido S, De Angelis C, Giuliano M, et al. Imaging tests in staging and surveillance of non-metastatic breast cancer: changes in routine clinical practice and cost implications. *Br J Cancer* 2017;116(6):821-27. doi: 10.1038/bjc.2017.24 [published Online First: 2017/02/09]
17. 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 research and treatment* 2013;139(2):469-76. doi: 10.1007/s10549-013-2543-9 [published Online First: 2013/04/30]

18. Konski A. Clinical and economic outcomes analyses of women developing breast cancer in a managed care organization. *American journal of clinical oncology* 2005;28(1):51-7. [published Online First: 2005/02/03]
19. 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(1):E11-7. doi: 10.9778/cmajo.20130026 [published Online First: 2014/08/01]
20. 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(3):e010547. doi: 10.1136/bmjopen-2015-010547
21. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927;22:209-12.
22. 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(3):e010409. doi: 10.1136/bmjopen-2015-010409
23. Baldi I, Vicari P, Di Cuonzo D, et al. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. *Journal of clinical epidemiology* 2008;61(4):373-9.
24. Freeman JL, Zhang D, Freeman DH, et al. An approach to identifying incident breast cancer cases using Medicare claims data. *Journal of clinical epidemiology* 2000;53(6):605-14. doi: 10.1016/s0895-4356(99)00173-0
25. Setoguchi S, Solomon DH, Glynn RJ, et al. Agreement of diagnosis and its date for hematologic malignancies and solid tumors between medicare claims and cancer registry data. *Cancer causes & control : CCC* 2007;18(5):561-9. doi: 10.1007/s10552-007-0131-1 [published Online First: 2007/04/21]
26. Kemp A, Preen DB, Saunders C, et al. Ascertaining invasive breast cancer cases; the validity of administrative and self-reported data sources in Australia. *BMC medical research methodology* 2013;13:17. doi: 10.1186/1471-2288-13-17 [published Online First: 2013/02/13]
27. 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. *Journal of clinical epidemiology* 2011;64(8):821-9. doi: 10.1016/j.jclinepi.2010.10.006
28. 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 services research* 2006;6:77. doi: 10.1186/1472-6963-6-77
29. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Annals of internal medicine* 2003;138(1):40-4.
30. Gliklich RE, Dreyer NA, Leavy MB. Registries for Evaluating Patient Outcomes: A User's Guide [Internet]. . 2014
31. Rice HE, Englum BR, Gulack BC, et al. Use of Patient Registries and Administrative Datasets for the Study of Pediatric Cancer. *Pediatric blood & cancer* 2015;62(9):1495-500. doi: 10.1002/pbc.25506
32. Rodrigo-Rincon I, Martin-Vizcaino MP, Tirapu-Leon B, et al. Validity of the clinical and administrative databases in detecting post-operative adverse events. *International Journal for Quality in Health Care* 2015;27(4):267-75. doi: 10.1093/intqhc/mzv039
33. Rashid N, Koh HA, Baca HC, et al. Economic burden related to chemotherapy-related adverse events in patients with metastatic breast cancer in an integrated health care system. *Breast Cancer : Targets and Therapy* 2016;8:173-81. doi: 10.2147/BCTT.S105618
34. Mahar AL, Coburn NG, Kagedan DJ, et al. Regional variation in the management of metastatic gastric cancer in Ontario. *Current Oncology* 2016;23(4):250-57. doi: 10.3747/co.23.3123
35. Orso M, Serraino D, Abraha I, et al. Validating malignant melanoma ICD-9-CM codes in Umbria, ASL Napoli 3 Sud and Friuli Venezia Giulia administrative healthcare databases: a diagnostic accuracy study. *BMJ Open* 2018;8(4):e020631. doi: 10.1136/bmjopen-2017-020631 [published Online First: 2018/04/22]

36. Montedori A, Bidoli E, Serraino D, et al. Accuracy of lung cancer ICD-9-CM codes in Umbria, Napoli 3 Sud and Friuli Venezia Giulia administrative healthcare databases: a diagnostic accuracy study. *BMJ Open* 2018

Table 1 Characteristics of patients with breast cancer who were identified in the three administrative healthcare databases

Characteristics	Unit 1 (Umbria)	Unit 2 (Asl Napoli 3 Sud)	Unit 3 (Friuli Venezia Giulia)
Invasive carcinoma			
Incident cases (N medical chart reviewed)	128	126	130
ICD-9 code			
174.0 nipple and areola	-	1 (1)	-
174.1 central portion	16 (13)	10 (8)	6 (5)
174.2 upper-inner quadrant	4 (3)	8 (6)	14 (11)
174.3 lower-inner quadrant	6 (5)	5 (4)	9 (7)
174.4 upper-outer quadrant	57 (45)	43 (34)	45 (35)
174.5 lower-outer quadrant	6 (5)	5 (4)	13 (10)
174.6 axillary tail	-	-	-
174.8 other specified sites of the female breast	-	-	-
174.9 breast female, unspecified	7 (5)	38 (30)	34 (26)
174.0 nipple and areola	32 (25)	16 (13)	9 (7)
Admission to department			
Medical	20 (16)	11 (9)	8 (6)
Surgical	108 (84)	115 (91)	122 (94)
Age, N (%)			
< 40	9 (7)	6 (5)	1 (1)
40 - 59	40 (31)	56 (44)	45 (35)
≥ 60	79 (62)	64 (51)	84 (65)
Instrumental diagnosis			
Breast ultrasound	39 (30)	88 (70)	5 (4)
Mammography	54 (42)	60 (48)	7 (5)
CT scan (breast)	11 (8)	1 (1)	2 (2)
MRI (breast)	3 (2)	17 (14)	8 (6)
Surgical procedures			
Mastectomy	28 (22)	29 (23)	35 (27)
Quadrantectomy	79 (62)	73 (58)	54 (42)
Hystological documentation			
Needle aspiration	32	34	-
Needle biopsy	27	40	5
Nodule (after surgical intervention)	115	117	112
Carcinoma in situ			
Incident cases (N medical chart reviewed)	50	88	108
ICD-9 code			
233.0	50 (100)	88 (100)	108 (100)
Admission to department			
Medical	-	7 (8)	-
Surgical	50 (100)	81 (92)	108 (100)
Age, N (%)			

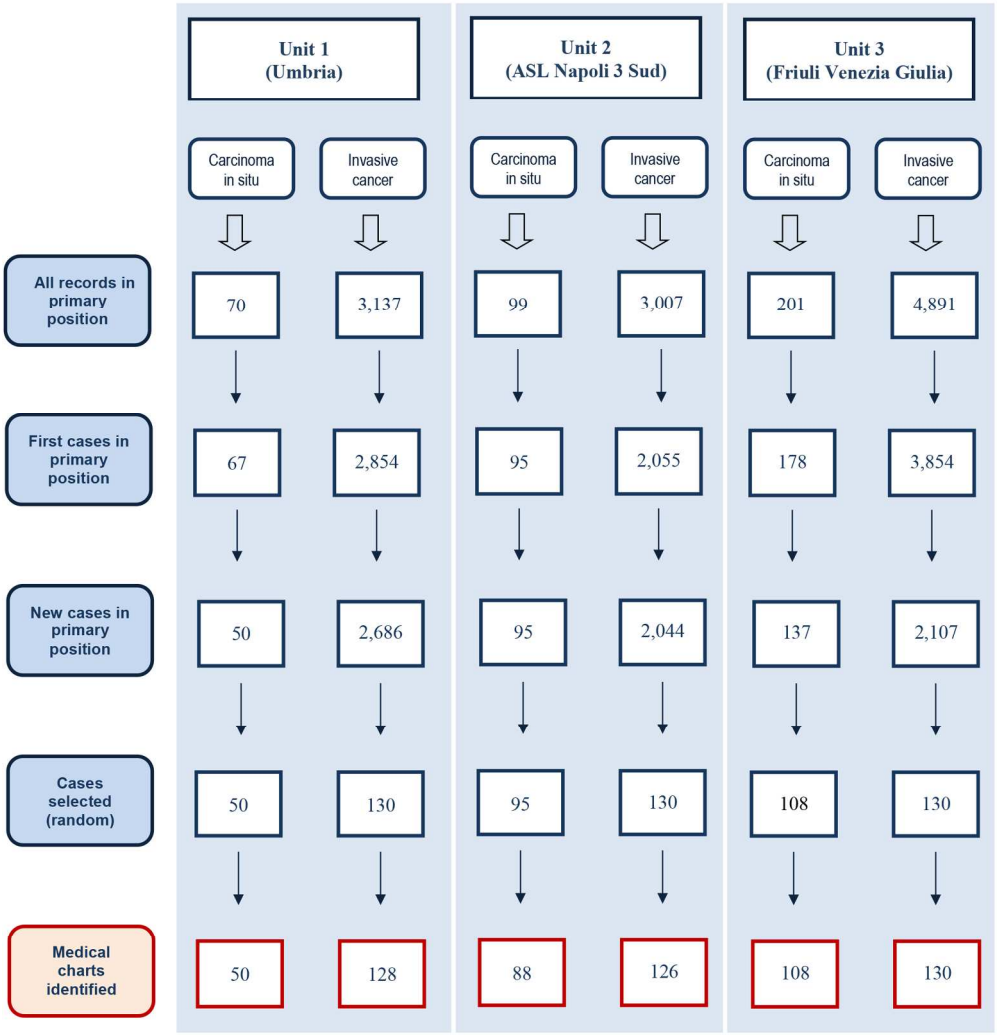
< 40	2 (4)	3 (3)	1 (1)
40 - 59	27 (54)	50 (57)	52 (48)
> 60	21 (42)	35 (40)	55 (51)
Instrumental diagnosis			
Breast ultrasound	3 (6)	65 (74)	30 (28)
Mammography	6 (12)	37 (42)	39 (36)
CT scan (breast)	-	-	-
MRI (breast)	2 (4)	22 (25)	4 (4)
Surgical procedures			
Mastectomy	18 (36)	13 (15)	15 (14)
Quadrantectomy	32 (64)	47 (53)	55 (52)
Hystological documentation			
Needle aspiration	-	16 (18)	1 (1)
Needle biopsy	8 (16)	53 (60)	4 (4)
Nodule (after surgical intervention)	50 (100)	77 (88)	94 (87)

Table 2. Reason for misclassification of cases and controls

	Invasive breast cancer			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	6	8	8
2	negative histology	4	1	1
	a) carcinoma in situ	*2*		*1*
Total		10	9	9
	False negative			
1	possible breast cancer relapse	1	1	-
2	metastatic breast cancer	2	4	-
Total		3	5	-
	Breast carcinoma in situ			
	Type of misclassification	**Unit 1 (Umbria)**	**Unit 2 (ASL Napoli 3 Sud)**	**Unit 3 (Friuli Venezia Giulia)**
	False positives			
1	histological examination missing	-	8	2
2	negative histology	2	7	9
Total		2	15	11
	False negative			
1	Possible carcinoma in situ	-	-	-

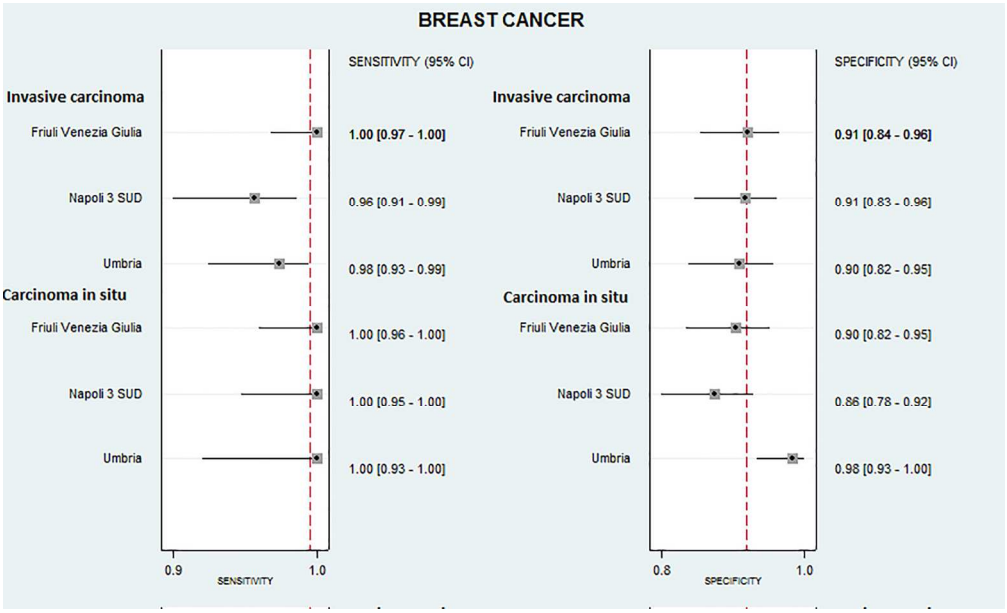
Table 3. Cross tabulation of the index test (ICD-9-CM code) results by the results of the reference standard (medical chart)

Type of brest cancer (ICD-9-CM)	Operative unit	TP	FP	TN	FN
Invasive cancer (174.x)	Unit 1 (Umbria)	118	10	90	3
	Unit 2 (ASL Napoli 3 Sud)	117	9	89	5
	Unit 3 (Friuli Venezia Giulia)	121	9	94	0
Carcinoma in situ (233.0)	Unit 1 (Umbria)	48	2	93	0
	Unit 2 (ASL Napoli 3 Sud)	73	15	94	0
	Unit 3 (Friuli Venezia Giulia)	97	11	94	0



Flow-chart of incident cases identification using the administrative databases and the corresponding charts identified and examined

173x181mm (300 x 300 DPI)



Sensitivity and specificity results for ICD-9-CM codes related to breast carcinoma in situ and invasive breast cancer for the three administrative databases

173x104mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	7
Participants	**6**	Eligibility criteria	6-8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	6-7
	8	Where and when potentially eligible participants were identified (setting, location and dates)	6-7
	9	Whether participants formed a consecutive, random or convenience series	7
Test methods	**10a**	Index test, in sufficient detail to allow replication	7
	10b	Reference standard, in sufficient detail to allow replication	7
	11	Rationale for choosing the reference standard (if alternatives exist)	7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	na
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	na
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8
	15	How indeterminate index test or reference standard results were handled	na
	16	How missing data on the index test and reference standard were handled	8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	na
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	Page 9 and Table 1
	21a	Distribution of severity of disease in those with the target condition	na
	21b	Distribution of alternative diagnoses in those without the target condition	na
	22	Time interval and any clinical interventions between index test and reference standard	na
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 3
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Page 9-10, Figure 2
	25	Any adverse events from performing the index test or the reference standard	na
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12-13
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	na
	29	Where the full study protocol can be accessed	5
	30	Sources of funding and other support; role of funders	14

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.