PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Validity of ICD-9-CM Codes for Breast, Lung, and Colorectal Cancers in Three Italian Administrative Healthcare Databases: A Diagnostic Accuracy Study Protocol
AUTHORS	Abraha, Iosief; Serraino, Diego; Giovannini, Gianni; Stracci, Fabrizio; Casucci, Paola; Alessandrini, Giuliana; Bidoli, Ettore; Chiari, Rita; Cirocchi, Roberto; De Giorgi, Marcello; Franchini, David; Vitale, Maria Francesca; Fusco, Mario; Montedori, Alessandro

VERSION 1 - REVIEW

REVIEWER	Susan Jick Boston Collaborative Drug Surveillance Program, Boston University School of Public Health USA
REVIEW RETURNED	02-Dec-2015

GENERAL COMMENTS	This purports to be a protocol for a study to evaluate the validity of breast, lung and colon cancer ICD 9 codes in 3 Italian administrative databases. The Introduction is well written as is the description of the administrative databases, however, the rest of the Methods section of this protocol is very confusing and has no logical flow. This should be a very straight forward study yet the authors have managed to make it complicated and hard to follow. It is not acceptable as written. 1. There is no logical flow to the Methods section. After the description of the databases the authors should then describe how the cases to be validated will be selected. Then describe the various types of validation (comparison of ICD 9 codes to information in charts, and looking for other codes in the electronic record to support the ICD 9 diagnosis.). 2. I think the authors are making this protocol more complicated than it needs to be while at the same time they provide little information on how they will actually validate the cases. For example, they need to describe the kinds of information they will abstract from the medical charts and then say something like: For cancer x if the following diagnoses are recorded in the chart we will consider the ICD 9 code to be valid. Etc. I cannot tell from the description how the authors will actually decide if a cancer code is valid or not. 3. Page 6 study design – I do not understand what they are saying. What is the reference standard about? And what is the point of the systematic review? I thought this was a study to validate breast, lung and colon cancer. What are these other diagnoses they refer to? This is another example of making this protocol more complicated than it needs to be. If they are describing how they will define true cases then this should not be in the study design section. 4. The headings of the protocol are not accurate. For example,

	"Study Design" includes unhelpful information about I'm not sure what. It does not seem to describe the actual validation study being proposed. Are the authors describing how they will derive an objective definition of the 3 cancers under study? This is not part of the study design. 5. The Case ascertainment section seems to describe what kinds of information will be abstracted from the chart review. This is not case ascertainment. It is chart abstraction. The authors also say, in this section, that agreement between reviewers will be assessed. Agreement on what? This is not described. 6. Perhaps the heading "Index test" (I'm not sure what that means?) should have the heading "case ascertainment" since it describes how cases to be validated will be identified. 7. The heading "Reference standards" seems to describe validation criteria to be abstracted from the record. But what criteria will be used to validate a cancer diagnosis? This is not stated. See comment 2 above. 8. The heading "Algorithm development" should be changed to "Validation algorithm". I think this section describes the algorithm for evaluating information from the electronic data – not from chart review. This should be clarified. 9. Table 1: This table is too vague. What are "Any procedures"? Or "Any secondary ICD – codes"? Do the medical charts really contain these codes? I would expect the codes to be in the electronic record. How will the 2 data sources be compared? 10. The information under the heading "Case ascertainment and sampling" should be included in a section with the ICD 9 codes that will be used to identify all cases. Once all of these cases have been identified the authors will sample from among all of them to determine which will be validated using the following technique…. As written there is no logic to the presentation of the methods.

REVIEWER	Giovannino Ciccone, MD, Unit of Clinical Epidemiology AOU Città della Salute e della Scienza, Turin and CPO Piemonte, Italy
REVIEW RETURNED	26-Dec-2015

GENERAL COMMENTS	1. The specific study objectives are clearly defined, but it is not clear the primary purpose of this validation study. The authors mention different potential uses of the results (incidence and prevalence estimation, quality of care evaluation or clinical research), but the methods they describe may not be equally adequate for all these possible uses (see below for details). 2. The abstract reflects some of the protocol limitations, especially in the Methods and analysis section. 3. The study design should be better described and, possibly, improved. Using the administrativedatabases of three Italian areas, covering a population about 3.3 millions of residents, the authors will select all hospital discharge records with ICD9-CM diagnosis codes of breast, lung or colorectal cancer in a three year period (which calendar years?). A random sample (how large?) of these records will be extracted and the corresponding medical charts will be used as reference standard to validate the ICD9-CM codes for each cancer. The same design and methods will be applied in the three study areas. The author stated that with this study design they would be able to estimate both sensitivity and specificity of the cancer diagnosis codes. However, with this design it seems that they can only estimate the proportion of false positive cases (i.e. discharge

	records with a cancer diagnosis code not confirmed through the medical charts) and the positive predictive value of each code (the proportion of confirmed diagnosis on the total cases initially identified through the ICD9-CM codes). Therefore, neither sensitivity, because they cannot measure the number of false negative cases (i.e. true cancer cases not identified through the administrative databases), nor specificity (the proportion of true negative cases) could be estimated with this study design. To properly estimate sensitivity and specificity of cancer diagnosis codes the optimal reference is a population cancer registry covering the same population of the administrative databases. However, the proposed study design may be more acceptable to validate the diagnosis codes of cohorts of cancer patients analysed for quality of care evaluations. 4. Methods should be better specified to allow replication (for example, it is not clear if the cases will be validated according to their previous cancer history, giving separate estimates for incident and prevalent cases), but it is possible that these details are available in the full protocol. 7. The statistical methods are adequately described, but the sample size calculation is completely lacking. As previously explained (see point 3) some of the cells of the 2x2 tables cannot be filled with the current study design.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Susan Jick
Boston Collaborative Drug Surveillance Program, Boston University School of Public Health

Please leave your comments for the authors below
This purports to be a protocol for a study to evaluate the validity of breast, lung and colon cancer ICD 9 codes in 3 Italian administrative databases. The Introduction is well written as is the description of the administrative databases, however, the rest of the Methods section of this protocol is very confusing and has no logical flow. This should be a very straight forward study yet the authors have managed to make it complicated and hard to follow. It is not acceptable as written.
**We thank prof. Jick for providing helpful comments. We have amended the Methods section. Details are reported below.
1. There is no logical flow to the Methods section. After the description of the databases the authors should then describe how the cases to be validated will be selected. Then describe the various types of validation (comparison of ICD 9 codes to information in charts, and looking for other codes in the electronic record to support the ICD 9 diagnosis.).
**We have now completely revised the methods section.
We now have followed the following logic:
(a) we first described the characteristics of Italian administrative databases and then introduced our 3 databases of interest. Here we have also underlined that the three operational units will use the same process of validation (Administrative databases).
(b) after introducing the source population , we described the case selection – these are incident cases with breast, lung or colorectal cancer diagnosed between 2012 – 2014 (the cases with same diagnosis between 2007 – 2011 will be excluded; subsequently we described the sampling method. Please note that for clarification we grouped the codes according the disease of interest; this will also clarify the sample of the cases that will be investigated for validation (Case selection and sampling method).
(c) subsequently we described medical chart review, the data abstraction, the core information needed for ascertaining the cases – on which the agreement between reviewers will be measured

(Chart abstraction and the and case ascertainment)
(d) finally, we described the validation process (Validation criteria)
Please note also that the selection, sampling, and ascertainment of non-cases necessary for specificity calculation is briefly reported in the “statistical analysis” section.
2. I think the authors are making this protocol more complicated than it needs to be while at the same time they provide little information on how they will actually validate the cases. For example, they need to describe the kinds of information they will abstract from the medical charts and then say something like:
For cancer x if the following diagnoses are recorded in the chart we will consider the ICD 9 code to be valid. Etc. I cannot tell from the description how the authors will actually decide if a cancer code is valid or not.
** we are grateful for this suggestion. We have made changes as follows: “For non-invasive breast cancer, we will consider the ICD-9-CM code 233.0 valid, when there is evidence of a breast nodule documented with imaging (e.g., mammography) and a histological diagnosis of ductal or lobular breast carcinoma in situ (pTis) etc. …”
3. Page 6 study design – I do not understand what they are saying. What is the reference standard about? And what is the point of the systematic review? I thought this was a study to validate breast, lung and colon cancer. What are these other diagnoses they refer to? This is another example of making this protocol more complicated than it needs to be. If they are describing how they will define true cases then this should not be in the study design section.
** The reference standard was meant as gold standard (medical chart in our case) against which to measure the validity (or accuracy) of an index text(an ICD code in our case). However, as reported above we have changed the paragraphs in the methods section and abandoned the use of index test and reference standard to avoid confusion. We have deleted the sentences related to the performance of systematic review.
4. The headings of the protocol are not accurate. For example, “Study Design” includes unhelpful information about I’m not sure what. It does not seem to describe the actual validation study being proposed. Are the authors describing how they will derive an objective definition of the 3 cancers under study? This is not part of the study design.
** The heading as well as the content of the methods have now been changed to avoid confusion.

5. The Case ascertainment section seems to describe what kinds of information will be abstracted from the chart review. This is not case ascertainment. It is chart abstraction. The authors also say, in this section, that agreement between reviewers will be assessed. Agreement on what? This is not described.
** Case ascertainment and Chart abstraction have now been revised. Please see above.

6. Perhaps the heading “Index test” (I’m not sure what that means?) should have the heading “case ascertainment” since it describes how cases to be validated will be identified.
** Please see above

7. The heading “Reference standards” seems to describe validation criteria to be abstracted from the record. But what criteria will be used to validate a cancer diagnosis? This is not stated. See comment 2 above.
** Please see above regarding reference standard. As for the validation criteria we used the following items: “Case ascertainment of cancer within medical chart will be based on (a) the presence of a primary nodular lesion in the breast, lung or colon-rectum, documented with imaging or endoscopy and (b) the cytological or histological documentation of cancer from a primary or metastatic site.”

8. The heading “Algorithm development” should be changed to “Validation algorithm”. I think this section describes the algorithm for evaluating information from the electronic data – not from chart review. This should be clarified.

**We want to highlight In this amended protocol that we aimed to focus on validating the ICD-9-CM codes (only for incident cases) and, after considering the number charts we need to evaluate for validation, we are uncertain whether we will have enough financial support to perform the validation of the algorithms using procedure or therapeutic codes. Consequently, we deleted the part related to algorithms.

9. Table 1: This table is too vague. What are “Any procedures”? Or “Any secondary ICD – codes”? Do the medical charts really contain these codes? I would expect the codes to be in the electronic record. How will the 2 data sources be compared?
** Now the table reports only the ICD-9CM codes description.

10. The information under the heading “Case ascertainment and sampling” should be included in a section with the ICD 9 codes that will be used to identify all cases. Once all of these cases have been identified the authors will sample from among all of them to determine which will be validated using the following technique…. As written there is no logic to the presentation of the methods.
** As reported above all the section are now amended.
Many thanks again for your helpful suggestions.
Iosief Abraha

Reviewer: 2

Giovannino Ciccone, MD, Unit of Clinical Epidemiology
AOU Città della Salute e della Scienza, Turin and CPO Piemonte, Italy

Please leave your comments for the authors below
1. The specific study objectives are clearly defined, but it is not clear the primary purpose of this validation study. The authors mention different potential uses of the results (incidence and prevalence estimation, quality of care evaluation or clinical research), but the methods they describe may not be equally adequate for all these possible uses (see below for details).
**We thank Dr. Ciccone for reviewing the protocol and for providing helpful comments.

2. The abstract reflects some of the protocol limitations, especially in the Methods and analysis section.
** After revising all the Methods section of the full protocol now the abstract has been modified as follows: “Data from the administrative databases of Umbria Region (910,000 residents), Local Health Unit 3 of Napoli (1,170,000 residents), Friuli-Venezia Giulia Region (1,227,000 residents) will be considered. In each administrative database, patients with the first occurrence of diagnosis of breast, lung or colorectal cancer between 2012 and 2014 will be identified using the following groups of ICD-9-CM codes in primary position: (a) 233.0 and (b) 174.x for breast cancer; (c) 162.x for lung cancer; (d) 153.x for colon cancer, and (e) 154.0 - 154.1 and 154.8 for rectal cancer. Only incident cases will be considered, that is, excluding cases that have the same diagnosis in the five years (2007-2011) before the period of interest. A random sample of cases and non-cases will be selected from each administrative database and the corresponding medical charts will be assessed for validation by pairs of trained, independent reviewers. Case ascertainment within the medical charts will be based on (a) the presence of a primary nodular lesion in the breast, lung or colon-rectum, documented with imaging or endoscopy and (b) a cytological or histological documentation of cancer from a primary or metastatic site. Sensitivity and specificity with 95% confidence intervals will be calculated.”

3. The study design should be better described and, possibly, improved. Using the administrative databases of three Italian areas, covering a population about 3.3 millions of residents, the authors will select all hospital discharge records with ICD9-CM diagnosis codes of breast, lung or colorectal

cancer in a three year period (which calendar years?). A random sample (how large?) of these records will be extracted and the corresponding medical charts will be used as reference standard to validate the ICD9-CM codes for each cancer. The same design and methods will be applied in the three study areas. The author stated that with this study design they would be able to estimate both sensitivity and specificity of the cancer diagnosis codes. However, with this design it seems that they can only estimate the proportion of false positive cases (i.e. discharge records with a cancer diagnosis code not confirmed through the medical charts) and the positive predictive value of each code (the proportion of confirmed diagnosis on the total cases initially identified through the ICD9-CM codes). Therefore, neither sensitivity, because they cannot measure the number of false negative cases (i.e. true cancer cases not identified through the administrative databases), nor specificity (the proportion of true negative cases) could be estimated with this study design. To properly estimate sensitivity and specificity of cancer diagnosis codes the optimal reference is a population cancer registry covering the same population of the administrative databases. However, the proposed study design may be more acceptable to validate the diagnosis codes of cohorts of cancer patients analysed for quality of care evaluations.
** We have now reported the sample size calculation. Regarding the sensitivity and specificity we think that the Methods were not adequately presented and this may have originated a misinterpretation. But still we believe that it is possible to determine the accuracy (sensitivity and specificity) of ICD-9 codes using the medical chart as the gold standard.
We agree that validated cancer registries are important for epidemiologic studies, but we think that administrative database can add value in terms of pharmacoepidemiology and other research such as on healthcare services. Indeed, in the 3 regions in which this study will be performed exist cancer registries and the present study may definitely contribute to the knowledge of the three cancer diseases. We made some changes in the Discussion section.

4. Methods should be better specified to allow replication (for example, it is not clear if the cases will be validated according to their previous cancer history, giving separate estimates for incident and prevalent cases), but it is possible that these details are available in the full protocol.
**This is an important point that we have taken into account. We will consider only incident cases that is excluding the cases for whom there is the same cancer diagnosis in the 5 years before the period of interest (2012-2014). This is in the Case selection and sampling method: “”In each administrative database, patients with the first occurrence of diagnosis of breast, lung or colorectal cancer between 2012 and 2014 will be identified using the following groups of ICD-9-CM codes located in primary position: (a) 233.0 and (b) 174.x for breast cancer; (c) 162.x for lung cancer; (d) 153.x for colon cancer, and (e) 154.0 - 154.1 and 154.8 for rectal cancer. Only incident cases will be considered, that is, excluding cases with the same diagnosis (ICD-9-CM codes in any position) in the five years (2007-2011) before the period of interest.”

7. The statistical methods are adequately described, but the sample size calculation is completely lacking. As previously explained (see point 3) some of the cells of the 2x2 tables cannot be filled with the current study design.
**The sample size calculation is now provided and we have amended also the paragraph in order to clarify the true positives or true negatives. We admit that the previous version did not clarify the cases and non-case. Cases are as described above (response to point 6) and non-case are described in the “statistical analysis” section as follows: For specificity calculation, we will randomly select non-cases, that is, records without the ICD-9-codes of interest from administrative database. The corresponding medical charts will be retrieved and evaluated. We calculated a sample of 94 charts of non-cases will be retrieved to obtain an expected specificity of 90% with a precision of 10% and a power of 80%.”