BMJ Open

Validity of Peptic Ulcer Disease and Upper Gastrointestinal Bleeding Diagnoses in Administrative Databases: A Systematic Review Protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011776
Article Type:	Protocol
Date Submitted by the Author:	04-Mar-2016
Complete List of Authors:	Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service Chiatti, Carlos; Italian National Research Centre on Aging (INRCA), Italy Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Rimland, Joseph; Italian National Research Center on Aging (INRCA), Geriatrics and Geriatric Emergency Care Luchetta, Maria Laura; Azienda USL Umbria 1, General Medicine Ambrosio, Giuseppe; University of Perugia School of Medicine, Cardiology; Ospedale S. Maria della Misericordia, Medical Administration Montedori, Alessandro; Regional Health Authority of Umbria
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Gastroenterology and hepatology, Public health, Epidemiology
Keywords:	administrative database, ICD-9, ICD-10, peptic ulcer, gastrointestinal bleeding, sensitivity and specificity

Validity of Peptic Ulcer Disease and Upper Gastrointestinal Bleeding Diagnoses in Administrative Databases: A Systematic Review Protocol

Iosief Abraha, Carlos Chiatti, Francesco Cozzolino, Massimiliano Orso, Joseph M Rimland, Maria Laura Luchetta , Giuseppe Ambrosio, Alessandro Montedori,

Author affiliations:
Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy
Iosief Abraha
Alessandro Montedori
Francesco Cozzolino
Massimiliano Orso

Scientific Directorate, Italian National Research Center on Aging, Ancona, Italy
Carlos Chiatti

Geriatrics and Geriatric Emergency Care, Italian National Research Center on Aging, Ancona, Italy
Joseph M Rimland

Azienda USL Umbria 1, General Medicine, Perugia, Italy
Maria Laura Luchetta

University of Perugia School of Medicine, Cardiology, Perugia, Italy
Giuseppe Ambrosio

Correspondence to:
Dr. Iosief Abraha
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni, 61
06124 Perugia (Italy)
tel +39 075 504 5251
cell. +39349 077 0910
fax +39 075 504 5569
e-mail: iosief_a@yahoo.it
iabraha@regione.umbria.it

Abstract

Introduction Administrative healthcare databases are useful to investigate the epidemiology, health outcomes, quality indicators and healthcare utilization concerning peptic ulcers and gastrointestinal bleeding, but the databases need to be validated in order to be a reliable source for research. The aim of this protocol is to perform the first systematic review of studies reporting the validation of *International Classification of Diseases 9th Revision* and *10th version* (ICD-9; ICD-10) codes for peptic ulcer and upper gastrointestinal bleeding diagnoses.

Methods and analysis MEDLINE, EMBASE, Web of Science and the Cochrane Library databases will be searched, using appropriate search strategies. We will include validation studies that used administrative data to identify peptic ulcer disease and upper gastrointestinal bleeding diagnoses or studies that evaluated the validity of peptic ulcer and upper gastrointestinal bleeding codes in administrative data. The following inclusion criteria will be used: (a) the presence of a reference standard case definition for the diseases of interest; (b) the presence of at least one test measure (e.g., sensitivity, etc.); and (c) the use of an administrative database as a source of data. Pairs of reviewers will independently abstract data using standardized forms and will evaluate quality using the checklist of the Standards for Reporting of Diagnostic accuracy (STARD) criteria. This systematic review protocol has been produced in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) 2015 statement.

Ethics and dissemination Ethics approval is not required given that this is a protocol for a systematic review. We will submit results of this study to a peer-reviewed journal for publication. The results will serve as a guide for researchers validating administrative healthcare databases to determine appropriate case definitions for peptic ulcer disease and upper gastrointestinal bleeding, as well as to perform outcome research using administrative healthcare databases of these conditions.

Protocol registration number PROSPERO 2015 CRD42015029216

Strengths and limitations of this study

- Validation of *International Classification of Diseases 9th Revision* and *10th reversion* (ICD-9; ICD-10) diagnosis codes for peptic ulcer disease and upper gastrointestinal bleeding using administrative healthcare databases can contribute to health outcome research.
- This review will be the first to systematically identify and evaluate primary studies that validated the accuracy of ICD-9 and ICD-10 codes for peptic ulcer disease and upper gastrointestinal bleeding in administrative healthcare databases.
- The results from this systematic review will serve as a guide to determine appropriate case definitions for peptic ulcer and upper gastrointestinal bleeding.

Introduction

Non-variceal upper gastrointestinal bleeding (UGIB) is associated with significant morbidity and mortality. It has an incidence rate from 48 to 160 cases per 100,000 per year, and greater incidences in men and older people [1]. Although UGIB and peptic ulcer bleeding are diminishing in the general population, hospitalization rates from ulcer complications are growing in older populations [2]. The most frequent risk factors for non-variceal UGIB comprise H. pylori infection, and the use of NSAIDs/aspirin, and other antiplatelet and anticoagulant medications. (Up to 67% of cases of UGIB are caused by peptic ulcer disease (PUD) [1].) Both H pylori infection and NSAIDs are independent risk factors for PUD and UGIB [3].

Health authorities generate and maintain large administrative healthcare databases that typically contain information and data regarding health resource utilization (e.g., hospitalizations, outpatient care, drug prescriptions) and vital statistics[4]. For research, one of the advantages of administrative databases is that they passively collect data at a population level with longitudinal follow-up, making their results easily generalizable. In addition, they are considered to be cost-effective compared to primary data collection[5, 6]. The main disadvantage of these databases is that they are generated for administrative purposes, such as billing, and as a repository for patient hospital records, and not for research, hence, the diagnostic codes for specific disorders must be validated according to an accepted "gold standard" reference diagnosis [7-11].

In the gastrointestinal field, administrative healthcare databases have been used to estimate the epidemiology of peptic ulcer disease [12] and upper gastrointestinal bleeding[13], to assess drug related gastrointestinal outcomes[14-16], to conduct active drug surveillance [17] and health service quality evaluation [18, 19].

The current *International Classification of Diseases, 9th Revision,* (ICD-9) codes for peptic ulcer disease and upper gastrointestinal bleeding are: 531.0 - 531.7, 531.9 for gastric ulcers and hemorrhage, 532.0 - 532.7, 532.9 for duodenal ulcers and hemorrhage, 533.0 - 533.7, 533.9 for

peptic ulcers and hemorrhage, 534.0 - 534.7, 534.9 for gastrojejunal ulcers and hemorrhage, 578.0, 578.1, 578.9 for gastrointestinal hemorrhage. The *International Classification of Diseases, 10th Revision,* (ICD-10) codes are K25 for gastric ulcers and hemorrhage, K26 for duodenal ulcers and hemorrhage, K27 for peptic ulcers and hemorrhage and K28 for gastrojejunal ulcers and hemorrhage and K92.0, K92.1 and K92.8 for gastrointestinal hemorrhage. The latest diagnostic criteria for upper gastrointestinal ulcers are based on: (i) upper endoscopy (ii) testing for H. pylori (breath test, biopsy, stool antigen). Various claim-based algorithms have been employed for case identification of UGIB, such as medical chart review [20] and endoscopy reports [21].

In the medical literature, at the present time, data on the validity of diagnostic codes for peptic ulcer disease and upper gastrointestinal bleeding have not been investigated. With the current protocol, we plan to systematically evaluate validation studies of diagnostic codes corresponding to these gastrointestinal conditions in administrative databases.

Research question

The principal research question is: "what is the accuracy of ICD-9 or ICD-10 codes, for peptic ulcer disease and upper gastrointestinal bleeding, to correctly identify the corresponding diseases in administrative databases?". The target populations are patients with a diagnosis of peptic ulcer disease or upper gastrointestinal bleeding, the index tests for the principal question are ICD-9 or ICD-10 codes for peptic ulcer disease and upper gastrointestinal bleeding. The index test will be ICD-9 or ICD-10 codes in administrative data and the reference standard will be medical charts or validated electronic health records. Our primary outcome is the accuracy, in terms of sensitivity, specificity, positive and negative predictive values, of ICD-9 and ICD-10 administrative data codes to discriminate cases of peptic ulcer disease or upper gastrointestinal bleeding.

Methods

Literature search

Published peer-reviewed articles will be identified through comprehensive searches of MEDLINE, EMBASE, Web of Science and the Cochrane Library from their inception. We will use a search strategy that we developed based on the combination of: (a) keywords and MeSH terms to identify records regarding peptic ulcer disease and upper gastrointestinal bleeding; (b) terms to identify studies likely to contain validity or accuracy measures; and (c) a search strategy, based on the combination of terms used by Benchimol et al. [22] and the Mini-Sentinel's program [23, 24], which is designed to accurately identify studies that use healthcare administrative databases. The search strategy is available as supplementary material (**Appendix 1**). Relevant reference lists of key articles will be hand searched in order to retrieve additional articles. Pertinent articles that cited the article of interest, identified through the preceding search strategy, will be sought through the "Cited-By" tools in PubMed and Google Scholar. Two independent reviewers will screen titles and abstracts for eligibility. Discussion will be used to resolve discrepancies.

This review protocol has been prepared according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis Protocols (PRISMA-P) 2015 Statement[25] and the results will be presented following the PRISMA flow diagram (**Figure**) [26]. This protocol has also been published in the PROSPERO International Prospective Register of systematic reviews with registration number CRD42015029216 (http://www.crd.york.ac.uk/PROSPERO).

Inclusion criteria

Full-texts of eligible peer-reviewed articles, without limits in publication date, and published in English, that used administrative data to validate the ICD-9 or ICD-10 codes for peptic ulcer disease or upper gastrointestinal bleeding, will be obtained. For each study, the following inclusion criteria will be applied: (a) the presence of a reference standard case definition for peptic ulcer disease and

upper gastrointestinal bleeding; (b) the presence of at least one test measure (e.g., sensitivity, positive predictive values, etc.); (c) the data source was from an administrative database (i.e., a database in which data is routinely and passively collected without an a priori research question); and (d) the study database was from a representative sample of the general population. Studies that used electronic health records (EHRs, i.e., digital records which commonly include clinical information, prescription records, and radiological and laboratory data) to validate our target disease will also be included [27, 28]. Studies that employed databases, that were not truly administrative (e.g. disease registries, epidemiology surveillance systems, etc.), will be excluded.

Selection process

During the initial stage, titles and abstracts will be screened to identify potentially eligible studies. Subsequently, full texts of articles will be obtained and evaluated to determine if they meet the inclusion and exclusion criteria. We will perform data abstraction with standardized data collection forms, that will be tested on a sample of eligible articles beforehand. Title and abstract screening, full-text screening and data abstraction will be carried out, independently, and in duplicate, by two review authors. Any discrepancies will be resolved by consensus, and where necessary, by involving a third review author. Calibration exercises will be performed at each step of the process.

Data extraction

Data extraction will include the following information:

(a) the details of the included study (including title, year and journal of publication, country of origin, and sources of funding; the first author will be used as the study ID);

(b) the disease of interest (peptic ulcer or upper gastrointestinal bleeding);

(c) the target population from which the administrative data were collected;

(d) the type of administrative database used (e.g., hospitalization discharge data), outpatient records (e.g., physician billing claims) etc.;

(e) the ICD-9 or ICD-10 code used;

(f) external validation;

(g) use of training and testing cohorts;

(h) the reference standard used to determine the validity of the diagnostic code (e.g., medical chart review, patient self-reports, disease registry, etc.,);

(i) the characteristic of the test used to determine the validity of the diagnostic code or algorithm (e.g., sensitivity, specificity, positive predictive values (PPVs) and negative predictive values (NPVs), area under the receiver operating characteristic curve, likelihood ratios, and kappa statistics);

(j) any funding source and conflict of interest.

Quality assessment

The design and method of the included primary studies will be assessed using a checklist developed by Benchimol et al.[22], based on the criteria published by the Standards for Reporting of Diagnostic accuracy (STARD) initiative for the accurate reporting of studies using diagnostic studies[29]. The checklist is provided in **Appendix 2**. The presence of potential biases within the studies will be reported descriptively.

No subgroup analysis or publication bias assessment are anticipated.

Analysis

For each algorithm, we will abstract the validation statistics provided in the included studies. Validation statistics may include sensitivity, specificity, PPV, and NPV. We will calculate 95% confidence intervals (95% CI) when they are not reported in the articles. Where sufficient data are available we will calculate PPV and NPV. Where possible, validation statistics will be aggregated and stratified by administrative data source (outpatient vs. inpatient data), type of ICD code (ICD-9 or ICD-10), type of disease (duodenal ulcer vs gastric ulcer), and country of origin.

Ethics and dissemination

Approval from an ethics committee is not required, since this review protocol will use publicly available data without directly involving human participants. An outline of the protocol has been published in the PROSPERO International Prospective Register of Systematic Reviews in 2015, registration number CRD42015029216. The results of the review will summarize the studies validating diagnostic codes that identify peptic ulcer disease and upper gastrointestinal bleeding in administrative data. In addition, the results will serve as a guide to identify appropriate case definitions and algorithms of peptic ulcer disease and upper gastrointestinal bleeding for researchers validating administrative healthcare databases, as well as for outcome research that uses administrative healthcare databases on these conditions. Findings of the review will be presented at relevant scientific conferences and disseminated through publication in a peer-reviewed journal.

Footnotes

Contributors IA, JMR, FC, MO and AM conceived the study. JMR, IA, MLL, FC, MO, CC, GA, and AM were responsible for designing the protocol. IA, AM, MO, JMR and FC drafted the protocol manuscript. JMR, IA, FC, and MO developed the search strategy. JMR, IA, MLL, FC, MO, CC, GA, and AM critically revised the successive versions of the manuscript and approved the final version.

Funding This review protocol was funded by the Regional Health Authority of Umbria. The study funder was not involved in the study design or the writing of the protocol.

Competing interests None.

Reference

1. Tielleman, T., D. Bujanda, and B. Cryer, *Epidemiology and Risk Factors for Upper Gastrointestinal Bleeding.* Gastrointest Endosc Clin N Am, 2015. **25**(3): p. 415-28.
2. Lanas, A., et al., *The changing face of hospitalisation due to gastrointestinal bleeding and perforation.* Aliment Pharmacol Ther, 2011. **33**(5): p. 585-91.
3. Papatheodoridis, G.V., S. Sougioultzis, and A.J. Archimandritis, *Effects of Helicobacter pylori and nonsteroidal anti-inflammatory drugs on peptic ulcer disease: a systematic review.* Clin Gastroenterol Hepatol, 2006. **4**(2): p. 130-42.
4. Jutte, D.P., L.L. Roos, and M.D. Brownell, *Administrative record linkage as a tool for public health research.* Annu Rev Public Health, 2011. **32**: p. 91-108.
5. Schneeweiss, S. and J. Avorn, *A review of uses of health care utilization databases for epidemiologic research on therapeutics.* J Clin Epidemiol, 2005. **58**(4): p. 323-37.
6. Motheral, B.R. and K.A. Fairman, *The use of claims databases for outcomes research: rationale, challenges, and strategies.* Clinical Therapeutics, 1997. **19**(2): p. 346-366.
7. West, S.L., B.L. Strom, and C. Poole, *Validity of Pharmacoepidemiologic Drug and Diagnosis Data*. Pharmacoepidemiology. 2007: John Wiley & Sons, Ltd. 709-765.
8. Abraha, I., et al., *Validity of Breast, Lung and Colorectal Cancer Diagnoses in Administrative Databases: A Systematic Review Protocol.* BMJ Open, 2016. **Accept (10-Feb-2016)**: p. 010409.
9. Abraha, I., et al., *The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review.* Pharmacoepidemiology and Drug Safety, 2012. **21**: p. 400-400.
10. Abraha, I., 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. **Accept (22-Feb-2016)**: p. 010547.
11. Pisa, F., et al., *Accuracy of International Classification of Diseases, 9th Revision, Clinical Modification codes for upper gastrointestinal complications varied by position and age: a validation study in a cohort of nonsteroidal anti-inflammatory drugs users in Friuli Venezia Giulia, Italy.* Pharmacoepidemiol Drug Saf, 2013. **22**(11): p. 1195-204.
12. Thorsen, K., et al., *Epidemiology of perforated peptic ulcer: age- and gender-adjusted analysis of incidence and mortality.* World J Gastroenterol, 2013. **19**(3): p. 347-54.
13. Abougergi, M.S., A.C. Travis, and J.R. Saltzman, *The in-hospital mortality rate for upper GI hemorrhage has decreased over 2 decades in the United States: a nationwide analysis.* Gastrointest Endosc, 2015. **81**(4): p. 882-8.e1.
14. Chang, H.Y., et al., *Risk of gastrointestinal bleeding associated with oral anticoagulants: population based retrospective cohort study.* Bmj, 2015. **350**: p. h1585.
15. De Berardis, G., et al., *Association of aspirin use with major bleeding in patients with and without diabetes.* JAMA, 2012. **307**(21): p. 2286-94.
16. Jun, M., et al., *The association between kidney function and major bleeding in older adults with atrial fibrillation starting warfarin treatment: population based observational study.* BMJ, 2015. **350**: p. h246.
17. Coloma, P.M., et al., *Electronic healthcare databases for active drug safety surveillance: is there enough leverage?* Pharmacoepidemiol Drug Saf, 2012. **21**(6): p. 611-21.
18. Mattke, S., et al., *Evaluating the role of patient sample definitions for quality indicators sensitive to nurse staffing patterns.* Med Care, 2004. **42**(2 Suppl): p. II21-33.
19. Needleman, J., et al., *Nurse-staffing levels and the quality of care in hospitals.* N Engl J Med, 2002. **346**(22): p. 1715-22.
20. Wahl, P.M., et al., *Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population.* Pharmacoepidemiol Drug Saf, 2010. **19**(6): p. 596-603.
21. Quan, S., et al., *Upper-gastrointestinal bleeding secondary to peptic ulcer disease: incidence and outcomes.* World J Gastroenterol, 2014. **20**(46): p. 17568-77.

22. Benchimol, E.I., et al., *Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data.* J Clin Epidemiol, 2011. **64**(8): p. 821-9.
23. Carnahan, R.M. and K.G. Moores, *Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned.* Pharmacoepidemiol Drug Saf, 2012. **21 Suppl 1**: p. 82-9.
24. McPheeters, M.L., et al., *Methods for systematic reviews of administrative database studies capturing health outcomes of interest.* Vaccine, 2013. **31 Suppl 10**: p. K2-6.
25. Liberati, A., et al., *The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration.* Bmj, 2009. **339**: p. b2700.
26. Shamseer, L., et al., *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation.* Bmj, 2015. **349**: p. g7647.
27. Chubak, J., G. Pocobelli, and N.S. Weiss, *Tradeoffs between accuracy measures for electronic health care data algorithms.* J Clin Epidemiol, 2012. **65**(3): p. 343-349.e2.
28. Dean, B.B., et al., *Review: Use of Electronic Medical Records for Health Outcomes Research: A Literature Review.* Medical Care Research and Review, 2009. **66**(6): p. 611-638.
29. Bossuyt, P.M., et al., *Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative.* Ann Intern Med, 2003. **138**(1): p. 40-4.

Figure . Study screening process (PRISMA Flow Diagram)

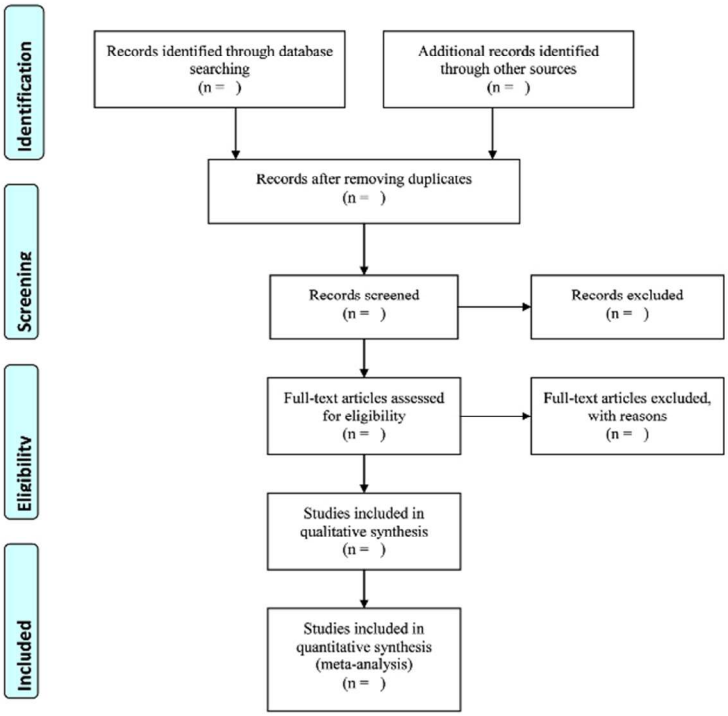

Figure 1. Study screening process
61x86mm (300 x 300 DPI)

Appendix 1

MEDLINE (via Pubmed) search strategy

1. (health administrative) OR (administrative data) OR (administrative database) OR (claim administrative) OR (International Classification of Diseases) OR "International Classification of Diseases"[Mesh] OR ICD-9-CM OR ICD-10 OR "Database Management Systems"[Mesh] OR "Medical Records Systems, Computerized"[Mesh] OR "CPT" OR "Current procedural terminology"[Mesh]
2. (factual databases) OR (geographic information systems) OR (national practitioner data bank) OR (insurance database)
3. #1 OR #2
4. sensitivity or "Sensitivity and Specificity"[Mesh]
5. specificity[Title/Abstract]
6. (positive predictive value) OR (negative predictive value) OR (likelihood ratio) OR (receiver operating characteristic) OR kappa
7. ((case or cases) AND (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*))
8. Algorithm OR "Algorithm"[Mesh]
9. #4 OR #5 OR #6 OR #7 OR #8
10. (stomach ulcer*) OR ("Stomach Ulcer"[Mesh]) OR (gastr* ulcer*)
11. (duodenal ulcer*) OR ("Duodenal Ulcer"[Mesh]) OR (curling* ulcer*)
12. (peptic ulcer*) OR ("Peptic Ulcer"[Mesh]) OR (marginal ulcer*)
13. (ulcer bleed*) OR ("Peptic Ulcer Hemorrhage"[MESH]) OR (ulcer hemorrhag*) OR (ulcer haemorrhag*) OR (ulcer perforat*)
14. (gastrointestinal bleed*) OR ("Gastrointestinal Hemorrhage"[Mesh]) OR (gastrointestinal hemorrhag*) OR (gastrointestinal haemorrhag*)
15. #10 OR #11 OR #12 OR #13 OR #14
16. #3 AND #9 AND #15

EMBASE search strategy (via embase.com)

1. health NEAR/3 administrative OR administrative NEAR/3 data OR administrative NEAR/3 database OR claim NEAR/3 administrative OR (International Classification of Diseases) OR 'International Classification of Diseases'/exp OR ICD-9-CM OR ICD-10 OR 'Database Management Systems'/exp OR 'Medical Records Systems, Computerized'/exp OR 'CPT' OR 'Current procedural terminology'/exp
2. database:ab,ti OR (('practitioner'/exp OR practitioner) AND data AND bank) OR (('practitioner'/exp OR practitioner) AND ('database'/exp OR database)) OR ('insurance' AND ('database'/exp OR database))
3. #1 OR #2
4. 'sensitivity and specificity'/exp OR 'sensitivity and specificity'
5. specificity:ab,ti
6. 'predictive value of tests'/exp OR 'predictive value of tests'
7. (positive:ab,ti AND predictive:ab,ti AND value:ab,ti) OR (negative:ab,ti AND predictive:ab,ti AND value:ab,ti) OR (likelyhood:ab,ti AND ratio:ab,ti) OR (receiver:ab,ti AND operating:ab,ti AND characteristic:ab,ti) OR kappa:ab,ti
8. case NEAR/1 (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*)
9. 'algorithms'/exp OR algorithm
10. #4 OR #5 OR #6 OR #7 OR #8 OR #9
11. 'stomach'/exp OR 'stomach ulcer'/exp OR (stomach NEAR/3 ulcer*):ab,ti OR (gastr* NEAR/3 ulcer*):ab,ti
12. 'duodenal'/exp OR 'duodenal ulcer'/exp OR (duodenal NEAR/3 ulcer*):ab,ti OR (curling* NEAR/3 ulcer*):ab,ti
13. 'peptic'/exp OR 'peptic ulcer'/exp OR (peptic NEAR/3 ulcer*):ab,ti OR (marginal NEAR/3 ulcer*):ab,ti
14. 'ulcer'/exp OR 'ulcer bleed'/exp OR (ulcer NEAR/3 bleed*) OR (ulcer NEAR/3 hemorrhag*):ab,ti OR (ulcer NEAR/3 haemorrhag*):ab,ti OR (ulcer NEAR/3 perforat*):ab,ti
15. 'gastrointestinal'/exp OR 'gastrointestinal bleed'/exp OR (gastrointestinal NEAR/3 bleed*):ab,ti OR (gastrointestinal NEAR/3 hemorrhag*):ab,ti OR (gastrointestinal NEAR/3 haemorrhag*):ab,ti
16. #11 OR #12 OR #13 OR #14 OR #15
17. #3 AND #10 AND #16

Web of Science search strategy

1. (health NEAR/3 administrative) OR (administrative NEAR/3 data) OR (administrative NEAR/3 database) OR (claim NEAR/3 administrative) OR (International Classification of Diseases) OR ICD-9-CM OR ICD-10 OR (Database Management Systems) OR ("Medical Records Systems" NEAR/2 Computerized) OR "CPT" OR (Current procedural terminology)
2. (factual databases) OR (geographic information systems) OR (national practitioner data bank) OR (insurance database)
3. #1 OR #2
4. sensitivity or "Sensitivity and Specificity"
5. specificity
6. (positive predictive value) OR (negative predictive value) OR (likelihood ratio) OR (receiver operating characteristic) OR kappa
7. ((case or cases) AND (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*))
8. algorithm
9. #4 OR #5 OR #6 OR #7 OR #8
10. (stomach NEAR/3 ulcer*) OR (gastr* NEAR/3 ulcer*)
11. (duodenal NEAR/3 ulcer*) OR (curling* NEAR/3 ulcer*)
12. (peptic NEAR/3 ulcer*) OR (marginal NEAR/3 ulcer*)
13. (ulcer NEAR/3 bleed*) OR (ulcer NEAR/3 hemorrhag*) OR (ulcer NEAR/3 haemorrhag*) OR (ulcer NEAR/3 perforat*)
14. (gastrointestinal NEAR/3 bleed*) OR (gastrointestinal NEAR/3 hemorrhag*) OR (ulcer NEAR/3 haemorrhag*)
15. #10 OR #11 OR #12 OR #13 OR #14
16. #3 AND #9 AND #15

The Cochrane Library

1. (health near/3 administrative) or (administrative near/3 data) or (administrative near/3 database) or (claim near/3 administrative) or (International Classification of Diseases) or [mh "International Classification of Diseases"] or ICD-9-CM or ICD-10 or [mh "Database Management Systems"] or [mh "Medical Records Systems, Computerized"] or "CPT" or [mh "Current procedural terminology"]
2. (factual databases) or (geographic information systems) or (national practitioner data bank) or (insurance database)
3. #1 or #2
4. sensitivity or [mh "Sensitivity and Specificity"]
5. specificity:ti,ab,kw
6. (positive predictive value) or (negative predictive value) or (likelihood ratio) or (receiver operating characteristic) or kappa
7. ((case or cases) and (verificat* or valid* or identif* or definition* or define* or evaluat*))
8. Algorithm or [mh "Algorithm"]
9. #4 or #5 or #6 or #7 or #8
10. [mh "Stomach Ulcer"] or (stomach near/3 ulcer*) or (gastr* near/3 ulcer*)
11. [mh "Duodenal Ulcer"] or (duodenal near/3 ulcer*) or (curling* near/3 ulcer*)
12. [mh "Peptic Ulcer"] or (peptic near/3 ulcer*) or (marginal near/3 ulcer*)
13. [mh "Peptic Ulcer Hemorrhage"]) or (ulcer near/3 bleed*) or (ulcer near/3 hemorrhag*) or (ulcer near/3 haemorrhag*) or (ulcer near/3 perforat*)
14. [mh "Gastrointestinal Hemorrhage"] or (gastrointestinal near/3 bleed*) or (gastrointestinal near/3 hemorrhag*) or (gastrointestinal near/3 haemorrhag*)
15. #10 or #11 or #12 or #13 or #14
16. #3 and #9 and #15

Appendix 2

Checklist of reporting criteria for studies validating health administrative data algorithms (developed by Benchimol et al., based on the criteria published by the Standards for Reporting of Diagnostic accuracy (STARD) initiative for the accurate reporting of studies using diagnostic studies.

	YES	NO	UNCERTAIN	NOT APPLICABLE
TITLE, KEYWORDS, ABSTRACT				
Identify article as study of assessing diagnostic accuracy				
Identify article as study of administrative data				
INTRODUCTION:				
State disease identification & validation one of goals of study				
METHODS:				
Participants in validation cohort:				
Describe validation cohort (Cohort of patients to which reference standard was applied)				
• Age				
• Disease				
• Severity				
• Location/Jurisdiction				
Describe recruitment procedure of validation cohort				
• Inclusion criteria				
• Exclusion criteria				
Describe patient sampling (random, consecutive, all, etc.)				
Describe data collection				
• Who identified patients and did selection adhere to patient recruitment criteria				
• Who collected data				
• *A priori* data collection form				
• Disease classification				
• Split sample (i.e. re-validation using a separate cohort) a) Training set b) Testing set				
Test Methods:				
Describe number, training and expertise of persons reading reference standard				
If >1 person reading reference standard, quote measure of consistency (e.g. kappa)				

Blinding of interpreters of reference standard to results of classification by administrative data e.g. Chart abstractor blinded to how that chart was coded				
Statistical Methods:				
Describe methods of calculating/comparing diagnostic accuracy				
RESULTS:				
Participants:				
Report when study done, start/end dates of enrollment				
Describe number of people who satisfied inclusion/exclusion criteria				
Study flow diagram				
Test results:				
Report distribution of disease severity				
Report cross-tabulation of index tests by results of reference standard				
Estimates:				
Report at least 4 estimates of diagnostic accuracy				
Diagnostic Accuracy Measures Reported:				
• Sensitivity				
• Spec				
• PPV				
• NPV				
• Likelihood ratios				
• Kappa				
• Area under the ROC curve / c-statistic				
• Accuracy/agreement				
• Other (specify)				
Report accuracy for subgroups (e.g. age, geography, different sex, etc.)				
If PPV/NPV reported, ratio of cases/controls of validation cohort approximate prevalence of condition in the population				
Report 95% confidence intervals for each diagnostic measure				
DISCUSSION:				
Discuss the applicability of the validation findings				

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	Page 1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	Page 2: Trial registration number PROSPERO 2015 CRD42015029216
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	Page 1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	Page 9
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	At this stage there are no relevant amendments to perform
Support:			
Sources	5a	Indicate sources of financial or other support for the review	Page 9 *(Regional Health Authority of Umbria, Italy)*
Sponsor	5b	Provide name for the review funder and/or sponsor	Page 9 (*Regional Health Authority of Umbria, Italy.)*
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	Page 9
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	Page 4 and 5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Page 5 Research question
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting,	Page 5, 6:

		time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	Page 6.
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Appendix 1 in Supplemental file
Study records:			Page 7:
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	Page 7:.
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	Page 7:
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Page 7:
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	Page 7/8 .
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	Page 5
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Not applicable. The present review will apply the STARD criteria.
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	No cumulative evidence will be presented.
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication	Not applicable

		bias across studies, selective reporting within studies)	
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	The present review will apply the STARD criteria. Page 8.

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Validity of Peptic Ulcer Disease and Upper Gastrointestinal Bleeding Diagnoses in Administrative Databases: A Systematic Review Protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011776.R1
Article Type:	Protocol
Date Submitted by the Author:	16-Jun-2016
Complete List of Authors:	Montedori, Alessandro; Regional Health Authority of Umbria Abraha, Iosief; Regional Health Authority of Umbria, Health Planning Service Chiatti, Carlos; Italian National Research Centre on Aging (INRCA), Italy Cozzolino, Francesco; Regional Health Authority of Umbria, Orso, Massimiliano; Regional Health Authority of Umbria, Health Planning Service of Perugia Luchetta, Maria Laura; Azienda USL Umbria 1, General Medicine Rimland, Joseph; Italian National Research Center on Aging (INRCA), Geriatrics and Geriatric Emergency Care Ambrosio, Giuseppe; University of Perugia School of Medicine, Cardiology; Ospedale S. Maria della Misericordia, Medical Administration
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Gastroenterology and hepatology, Public health, Epidemiology
Keywords:	peptic ulcer, gastrointestinal haemorrhage, administrative database, sensitivity, accuracy, validity

Validity of Peptic Ulcer Disease and Upper Gastrointestinal Bleeding Diagnoses in Administrative Databases: A Systematic Review Protocol

Alessandro Montedori, Iosief Abraha, Carlos Chiatti, Francesco Cozzolino, Massimiliano Orso, Maria Laura Luchetta, Joseph M Rimland, Giuseppe Ambrosio,

Author affiliations:
Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy
Iosief Abraha
Alessandro Montedori
Francesco Cozzolino
Massimiliano Orso

Scientific Directorate, Italian National Research Center on Aging, Ancona, Italy
Carlos Chiatti

Geriatrics and Geriatric Emergency Care, Italian National Research Center on Aging, Ancona, Italy
Joseph M Rimland

Azienda USL Umbria 1, General Medicine, Perugia, Italy
Maria Laura Luchetta

University of Perugia School of Medicine, Cardiology, Perugia, Italy
Giuseppe Ambrosio

Correspondence to:
Dr. Iosief Abraha
Health Planning Service
Regional Health Authority of Umbria
Via Mario Angeloni, 61
06124 Perugia (Italy)
tel +39 075 504 5251
cell. +39349 077 0910
fax +39 075 504 5569
e-mail: iosief_a@yahoo.it
iabraha@regione.umbria.it

Abstract

Introduction Administrative healthcare databases are useful to investigate the epidemiology, health outcomes, quality indicators and healthcare utilization concerning peptic ulcers and gastrointestinal bleeding, but the databases need to be validated in order to be a reliable source for research. The aim of this protocol is to perform the first systematic review of studies reporting the validation of *International Classification of Diseases 9th Revision* and *10th version* (ICD-9; ICD-10) codes for peptic ulcer and upper gastrointestinal bleeding diagnoses.

Methods and analysis MEDLINE, EMBASE, Web of Science and the Cochrane Library databases will be searched, using appropriate search strategies. We will include validation studies that used administrative data to identify peptic ulcer disease and upper gastrointestinal bleeding diagnoses or studies that evaluated the validity of peptic ulcer and upper gastrointestinal bleeding codes in administrative data. The following inclusion criteria will be used: (a) the presence of a reference standard case definition for the diseases of interest; (b) the presence of at least one test measure (e.g., sensitivity, etc.); and (c) the use of an administrative database as a source of data. Pairs of reviewers will independently abstract data using standardized forms and will evaluate quality using the checklist of the Standards for Reporting of Diagnostic accuracy (STARD) criteria. This systematic review protocol has been produced in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) 2015 statement.

Ethics and dissemination Ethics approval is not required given that this is a protocol for a systematic review. We will submit results of this study to a peer-reviewed journal for publication. The results will serve as a guide for researchers validating administrative healthcare databases to determine appropriate case definitions for peptic ulcer disease and upper gastrointestinal bleeding, as well as to perform outcome research using administrative healthcare databases of these conditions.

Protocol registration number PROSPERO 2015 CRD42015029216

Strengths and limitations of this study

- Validation of *International Classification of Diseases 9th Revision* and *10th reversion* (ICD-9; ICD-10) diagnosis codes for peptic ulcer disease and upper gastrointestinal bleeding using administrative healthcare databases can contribute to health outcome research.
- This review will be the first to systematically identify and evaluate primary studies that validated the accuracy of ICD-9 and ICD-10 codes for peptic ulcer disease and upper gastrointestinal bleeding in administrative healthcare databases.
- The results from this systematic review will serve as a guide to determine appropriate case definitions for peptic ulcer and upper gastrointestinal bleeding.
- The main limitation is that validated diagnosis codes or algorithms are context-specific, and may not be generalizable to other settings.

Introduction

Non-variceal upper gastrointestinal bleeding (UGIB) is associated with significant morbidity and mortality. It has an incidence rate from 48 to 160 cases per 100,000 per year, and greater incidences in men and older people [1 2]. Although UGIB and peptic ulcer bleeding are diminishing in the general population, hospitalization rates from ulcer complications are growing in older populations [3]. The most frequent risk factors for non-variceal UGIB comprise H. pylori infection, and the use of NSAIDs/aspirin, and other antiplatelet and anticoagulant medications. (Up to 67% of cases of UGIB are caused by peptic ulcer disease (PUD) [1].) Both H pylori infection and NSAIDs are independent risk factors for PUD and UGIB [4].

Health authorities generate and maintain large administrative healthcare databases that typically contain information and data regarding health resource utilization (e.g., hospitalizations, outpatient care, drug prescriptions) and vital statistics[5]. For research, one of the advantages of administrative databases is that they passively collect data at a population level with longitudinal follow-up, making their results easily generalizable. In addition, they are considered to be cost-effective compared to primary data collection[6 7]. The main disadvantage of these databases is that they are generated for administrative purposes, such as billing, and as a repository for patient hospital records, and not for research, hence, the diagnostic codes for specific disorders must be validated according to an accepted "gold standard" reference diagnosis [8-14].

In the gastrointestinal field, administrative healthcare databases have been used to estimate the epidemiology of peptic ulcer disease [15] and upper gastrointestinal bleeding[16], to assess drug related gastrointestinal outcomes[17-19], to conduct active drug surveillance [20] and health service quality evaluation [21 22].

Current administrative databases use the *International Classification of Diseases, 9th Revision,* (ICD-9) or 10th Revision (ICD-10) codes or for peptic ulcer disease and upper gastrointestinal bleeding. Validation of diagnostic codes is of particular interest to national healthcare authorities to

perform surveillance of medical products and epidemiological studies of diseases. For example, the US Food and Drug Administration has sponsored a pilot project, Mini-Sentinel, with the aim of performing active surveillance to improve safety signals that emerge for newly released medical products. To implement this work, the program needed to identify algorithms used to detect a number of health outcomes of interest using administrative data sources and identify the performance characteristics of these algorithms[23]. The Mini-Sentinel program produced a series of systematic reviews of validated methods and case definitions, to identify various diseases or health outcomes in administrative data, including cardio-cerebrovascular diseases [24-28] and other conditions [29-33]. For the purpose of establishing best practices in the use of administrative data for health research and surveillance, the Canadian Rheumatology Network conducted a systematic review of studies reporting on the validity of diagnostic codes to identify cardiovascular diseases[34-36]. Likewise, the Regional Health Authority of Umbria, is interested in the validity of administrative data diagnoses and in identifying case definitions and the algorithms developed for different diseases, including cancer (breast, lung and colorectal)[9 11], Chronic Obstructive Pulmonary Disease[13] (Rimland, BMJ Open. 2016 Jun 1;6(6):e011777) and non-variceal upper gastrointestinal bleeding, which is the focus of this article.

In the medical literature, at the present time, the validity and performance of algorithms employing diagnostic codes for peptic ulcer disease and upper gastrointestinal bleeding have not been systematically investigated. With the current protocol, we plan to systematically evaluate validation studies of diagnostic codes corresponding to these gastrointestinal conditions in administrative databases.

Methods

Literature search

Published peer-reviewed articles will be identified through comprehensive searches of MEDLINE, EMBASE, Web of Science and the Cochrane Library from their inception. We will use a search strategy that we developed based on the combination of: (a) keywords and MeSH terms to identify records regarding peptic ulcer disease and upper gastrointestinal bleeding; (b) terms to identify studies likely to contain validity or accuracy measures; and (c) a search strategy, based on the combination of terms used by Benchimol et al. [37] and the Mini-Sentinel program [38 39], which is designed to accurately identify studies that use healthcare administrative databases. The search strategy is available as supplementary material (**Supplementary Appendix 1**). Relevant reference lists of key articles will be hand searched in order to retrieve additional articles. Pertinent articles that cited the article of interest, identified through the preceding search strategy, will be sought through the "Cited-By" tools in PubMed and Google Scholar. Two independent reviewers will screen titles and abstracts for eligibility. Discussion will be used to resolve discrepancies.

This review protocol has been prepared according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis Protocols (PRISMA-P) 2015 Statement[40] and the results will be presented following the PRISMA flow diagram (**Figure**) [41]. This protocol has also been published in the PROSPERO International Prospective Register of systematic reviews with registration number CRD42015029216 (http://www.crd.york.ac.uk/PROSPERO).

Inclusion criteria

Type of studies

We will consider any type of diagnostic (cross-sectional, retrospective or prospective) cohort study, without limits in publication date, and published in English, for inclusion.

Population

The target populations will include patients of any age and sex with peptic ulcer or gastrointestinal haemorrhage. Since there are substantial differences between in-hospital and outpatient upper gastrointestinal bleeders in terms of both clinical risk profile and treatment patterns [42] we will considered two types of cohorts with bleeding: (a) patients who have been admitted to a hospital due to non-variceal upper gastrointestinal bleeding caused by peptic ulcer; and (b) outpatients who have been visited for peptic ulcer or gastrointestinal bleeding.

Index test

Studies that validated diagnostic codes or algorithms related to ICD-9 or ICD-10 for peptic ulcer disease or upper gastrointestinal bleeding will be considered. The current ICD-9 codes for peptic ulcer disease and upper gastrointestinal bleeding are: 531.0 - 531.7, 531.9 for gastric ulcers and haemorrhage, 532.0 - 532.7, 532.9 for duodenal ulcers and haemorrhage, 533.0 - 533.7, 533.9 for peptic ulcers and haemorrhage, 534.0 - 534.7, 534.9 for gastrojejunal ulcers and haemorrhage, and 578.0, 578.1, 578.9 for gastrointestinal haemorrhage. The ICD-10 codes are K25 for gastric ulcers and haemorrhage, K26 for duodenal ulcers and haemorrhage, K27 for peptic ulcers and haemorrhage and K28 for gastrojejunal ulcers and haemorrhage and K92.0, K92.1 and K92.8 for gastrointestinal haemorrhage. Detailed descriptions of each ICD code are reported in **Supplementary Appendix 2** of the Supplemental file .

Reference standard

Studies will be considered in which the diagnoses of target diseases were confirmed through review of medical charts, medical notes ,or electronic health records. Confirmed peptic ulcers will include cases of active gastric or duodenal ulcers, or gastroduodenal perforation, as confirmed by surgery, endoscopy, X-ray, or autopsy. Confirmed upper gastrointestinal bleeding will include cases of haemorrhage from gastric or duodenal ulcers, haemorrhagic gastritis, duodenitis, or gastroduodenal perforation, confirmed by surgery, endoscopy, X-ray, or autopsy.

Outcome

Studies that reported the accuracy of administrative data codes to discriminate cases of peptic ulcer disease or upper gastrointestinal bleeding, at least in terms of sensitivity or positive predictive values will be eligible for inclusion.

Selection process

During the initial stage, titles and abstracts will be screened to identify potentially eligible studies. Subsequently, full texts of articles will be obtained and evaluated to determine if they meet the inclusion and exclusion criteria. We will perform data abstraction with standardized data collection forms, that will be tested on a sample of eligible articles beforehand. Title and abstract screening, full-text screening and data abstraction will be carried out, independently, and in duplicate, by two review authors. Any discrepancies will be resolved by consensus, and where necessary, by involving a third review author. Calibration exercises will be performed at each step of the process.

Data extraction

Data extraction will include the following information:

(a) the details of the included study (including title, year and journal of publication, country of origin, and sources of funding; the first author will be used as the study ID);

(b) the disease of interest (peptic ulcer or upper gastrointestinal bleeding);

(c) the target population from which the administrative data were collected;

(d) the type of administrative database used (e.g., hospitalization discharge data), outpatient records (e.g., physician billing claims) etc.;

(e) the ICD-9 or ICD-10 code used;

(f) external validation;

(g) use of training and testing cohorts;

(h) the reference standard used to determine the validity of the diagnostic code (e.g., medical chart review, patient self-reports, disease registry, etc.,);

(i) the characteristic of the test used to determine the validity of the diagnostic code or algorithm (e.g., sensitivity, specificity, positive predictive values (PPVs) and negative predictive values (NPVs), area under the receiver operating characteristic curve, likelihood ratios, and kappa statistics);

(j) any conflict of interest.

Quality assessment

The design and method of the included primary studies will be assessed using a checklist developed by Benchimol et al.[37], based on the criteria published by the Standards for Reporting of Diagnostic accuracy (STARD) initiative for the accurate reporting of studies using diagnostic studies[43]. The checklist is provided in **Supplementary Appendix 3**. The presence of potential biases within the studies will be reported descriptively.

No subgroup analysis or publication bias assessment are anticipated.

Analysis

For each algorithm, we will abstract the validation statistics provided in the included studies. Validation statistics may include sensitivity, specificity, PPV, and NPV. We will calculate 95% confidence intervals (95% CI) when they are not reported in the articles. Where sufficient and homogeneous data are available we will derive summary estimates of sensitivity and specificity and their 95% CIs data using a bivariate model[44]. Data will be meta-analysed using a random-effects model so that sensitivity and specificity are assumed to vary across studies. Separate meta-analyses will be provided based on the administrative data source (outpatient vs. inpatient data), type of ICD code (ICD-9 or ICD-10), and type of disease (ulcer or haemorrhage). We will perform subgroup analyses according to timing of publication and ICD code assessed to examine whether accuracy data have changed overtime.

In addition, summary receiver operating characteristic (ROC) curves will be constructed and pooled

estimates of LR+, LR− and diagnostic odds ratio will be calculated. Heterogeneity will be assessed by visual inspection of forest plots and ROC plots, as well as regression analysis suggested by Reitsma [44]. Where there is important heterogeneity, we will not pool the data.

Publication bias will not evaluated, as the common tests available (Begg, Egger and Deeks tests) provide different results and thus are not interchangeable.[45]

Ethics and dissemination

Approval from an ethics committee is not required, since this review protocol will use publicly available data without directly involving human participants. An outline of the protocol has been published in the PROSPERO International Prospective Register of Systematic Reviews in 2015, registration number CRD42015029216. The results of the review will summarize the studies validating diagnostic codes that identify peptic ulcer disease and upper gastrointestinal bleeding in administrative data. In addition, the results will serve as a guide to identify appropriate case definitions and algorithms of peptic ulcer disease and upper gastrointestinal bleeding for researchers validating administrative healthcare databases, as well as for outcome research that uses administrative healthcare databases on these conditions. Findings of the review will be presented at relevant scientific conferences and disseminated through publication in a peer-reviewed journal.

Footnotes

Contributors IA, JMR, FC, MO and AM conceived the study. JMR, IA, MLL, FC, MO, CC, GA, and AM were responsible for designing the protocol. IA, GA, AM, MO, JMR and FC drafted the protocol manuscript. JMR, IA, FC, and MO developed the search strategy. JMR, IA, MLL, FC, MO, CC, GA, and AM critically revised the successive versions of the manuscript and approved the final version.

Funding This review protocol was funded by the Regional Health Authority of Umbria. The study funder was not involved in the study design or the writing of the protocol.

Competing interests None declared.

Reference

1. Tielleman T, Bujanda D, Cryer B. Epidemiology and Risk Factors for Upper Gastrointestinal Bleeding. Gastrointest Endosc Clin N Am 2015;**25**:415-28.
2. Kurien M, Lobo AJ. Acute upper gastrointestinal bleeding. Clin Med (Lond) 2015;**15**:481-5.
3. Lanas A, Garcia-Rodriguez LA, Polo-Tomas M, Ponce M, Quintero E, Perez-Aisa MA, et al. The changing face of hospitalisation due to gastrointestinal bleeding and perforation. Aliment Pharmacol Ther 2011;**33**:585-91.
4. Papatheodoridis GV, Sougioultzis S, Archimandritis AJ. Effects of Helicobacter pylori and nonsteroidal anti-inflammatory drugs on peptic ulcer disease: a systematic review. Clin Gastroenterol Hepatol 2006;**4**:130-42.
5. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;**32**:91-108.
6. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;**58**:323-37.
7. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. Clinical Therapeutics 1997;**19**:346-66.
8. West SL, Strom BL, Poole C. *Validity of Pharmacoepidemiologic Drug and Diagnosis Data*: John Wiley & Sons, Ltd, 2007.
9. Abraha I, Giovannini G, Serraino D, Fusco M, Montedori A. Validity of Breast, Lung and Colorectal Cancer Diagnoses in Administrative Databases: A Systematic Review Protocol. BMJ Open 2016;**Accept (10-Feb-2016)**:010409.
10. Abraha I, Montedori A, Eusebi P, Orso M, Cozzolino F, De Florio R, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. Pharmacoepidemiology and Drug Safety 2012;**21**:400-00.
11. Abraha I, Serraino D, Giovannini G, Stracci F, Casucci P, Alessandrini A, 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;**Accept (22-Feb-2016)**:010547.
12. Pisa F, Castellsague J, Drigo D, Riera-Guardia N, Giangreco M, Rosolen V, et al. Accuracy of International Classification of Diseases, 9th Revision, Clinical Modification codes for upper gastrointestinal complications varied by position and age: a validation study in a cohort of nonsteroidal anti-inflammatory drugs users in Friuli Venezia Giulia, Italy. Pharmacoepidemiol Drug Saf 2013;**22**:1195-204.
13. Rimland JM, Abraha I, Luchetta ML, Cozzolino F, Orso M, Cherubini A, et al. Validation of chronic obstructive pulmonary disease (COPD) diagnoses in healthcare databases: a systematic review protocol. BMJ Open 2016;**6**:e011777.
14. Abraha I, Orso M, Grilli P, Cozzolino F, Eusebi P, Casucci P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. International Journal of Statistics in Medical Research 2014;**3**:309-20.
15. Thorsen K, Soreide JA, Kvaloy JT, Glomsaker T, Soreide K. Epidemiology of perforated peptic ulcer: age- and gender-adjusted analysis of incidence and mortality. World J Gastroenterol 2013;**19**:347-54.
16. Abougergi MS, Travis AC, Saltzman JR. The in-hospital mortality rate for upper GI haemorrhage has decreased over 2 decades in the United States: a nationwide analysis. Gastrointest Endosc 2015;**81**:882-8.e1.
17. Chang HY, Zhou M, Tang W, Alexander GC, Singh S. Risk of gastrointestinal bleeding associated with oral anticoagulants: population based retrospective cohort study. BMJ 2015;**350**:h1585.
18. De Berardis G, Lucisano G, D'Ettorre A, Pellegrini F, Lepore V, Tognoni G, et al. Association of aspirin use with major bleeding in patients with and without diabetes. JAMA 2012;**307**:2286-94.
19. Jun M, James MT, Manns BJ, Quinn RR, Ravani P, Tonelli M, et al. The association between kidney function and major bleeding in older adults with atrial fibrillation starting warfarin treatment: population based observational study. BMJ 2015;**350**:h246.

20. Coloma PM, Trifiro G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? Pharmacoepidemiol Drug Saf 2012;**21**:611-21.
21. Mattke S, Needleman J, Buerhaus P, Stewart M, Zelevinsky K. Evaluating the role of patient sample definitions for quality indicators sensitive to nurse staffing patterns. Med Care 2004;**42**:II21-33.
22. Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. N Engl J Med 2002;**346**:1715-22.
23. Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:1-8.
24. Andrade SE, Harrold LR, Tjia J, Cutrona SL, Saczynski JS, Dodd KS, et al. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:100-28.
25. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. A systematic review of validated methods for identifying atrial fibrillation using administrative data. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:141-7.
26. Saczynski JS, Andrade SE, Harrold LR, Tjia J, Cutrona SL, Dodd KS, et al. A systematic review of validated methods for identifying heart failure using administrative data. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:129-40.
27. Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:154-62.
28. Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying ventricular arrhythmias using administrative and claims data. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:148-53.
29. Chung CP, Rohan P, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. Vaccine 2013;**31 Suppl 10**:K41-61.
30. Idowu RT, Carnahan R, Sathe NA, McPheeters ML. A systematic review of validated methods to capture myopericarditis using administrative or claims data. Vaccine 2013;**31 Suppl 10**:K34-40.
31. Moores KG, Sathe NA. A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data. Vaccine 2013;**31 Suppl 10**:K62-73.
32. Williams SE, Carnahan R, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying transverse myelitis using administrative or claims data. Vaccine 2013;**31 Suppl 10**:K83-7.
33. Williams SE, Carnahan R, McPheeters ML. A systematic review of validated methods for identifying uveitis using administrative or claims data. Vaccine 2013;**31 Suppl 10**:K88-97.
34. McCormick N, Bhole V, Lacaille D, Avina-Zubieta JA. Validity of Diagnostic Codes for Acute Stroke in Administrative Databases: A Systematic Review. PLoS One 2015;**10**:e0135834.
35. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. PLoS One 2014;**9**:e104519.
36. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. PLoS One 2014;**9**:e92286.
37. Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. J Clin Epidemiol 2011;**64**:821-9.
38. Carnahan RM, Moores KG. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned. Pharmacoepidemiol Drug Saf 2012;**21 Suppl 1**:82-9.
39. McPheeters ML, Sathe NA, Jerome RN, Carnahan RM. Methods for systematic reviews of administrative database studies capturing health outcomes of interest. Vaccine 2013;**31 Suppl 10**:K2-6.
40. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;**339**:b2700.

41. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;**349**:g7647.
42. Marmo R, Koch M, Cipolletta L, Bianco MA, Grossi E, Rotondano G. Predicting mortality in patients with in-hospital nonvariceal upper GI bleeding: a prospective, multicenter database study. Gastrointest Endosc 2014;**79**:741-49.e1.
43. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003;**138**:40-4.
44. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;**58**:982-90.
45. van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC Medical Research Methodology 2014;**14**:1-11.

Figure . Study screening process (PRISMA Flow Diagram)

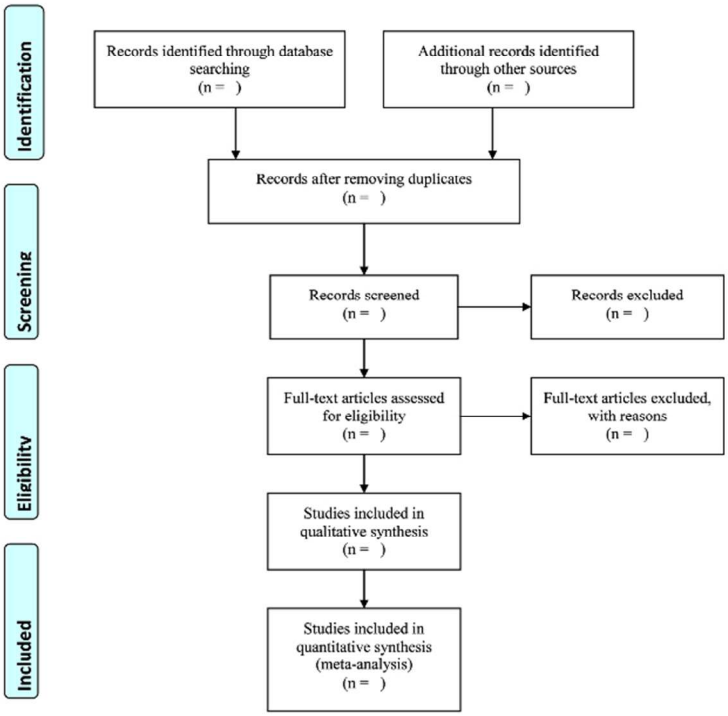

Figure 1. Study screening process
61x86mm (300 x 300 DPI)

Appendix 1

MEDLINE (via Pubmed) search strategy

1. (health administrative) OR (administrative data) OR (administrative database) OR (claim administrative) OR (International Classification of Diseases) OR "International Classification of Diseases"[Mesh] OR ICD-9-CM OR ICD-10 OR "Database Management Systems"[Mesh] OR "Medical Records Systems, Computerized"[Mesh] OR "CPT" OR "Current procedural terminology"[Mesh]
2. (factual databases) OR (geographic information systems) OR (national practitioner data bank) OR (insurance database)
3. #1 OR #2
4. sensitivity or "Sensitivity and Specificity"[Mesh]
5. specificity[Title/Abstract]
6. (positive predictive value) OR (negative predictive value) OR (likelihood ratio) OR (receiver operating characteristic) OR kappa
7. ((case or cases) AND (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*))
8. Algorithm OR "Algorithm"[Mesh]
9. #4 OR #5 OR #6 OR #7 OR #8
10. (stomach ulcer*) OR ("Stomach Ulcer"[Mesh]) OR (gastr* ulcer*)
11. (duodenal ulcer*) OR ("Duodenal Ulcer"[Mesh]) OR (curling* ulcer*)
12. (peptic ulcer*) OR ("Peptic Ulcer"[Mesh]) OR (marginal ulcer*)
13. (ulcer bleed*) OR ("Peptic Ulcer Hemorrhage"[MESH]) OR (ulcer hemorrhag*) OR (ulcer haemorrhag*) OR (ulcer perforat*)
14. (gastrointestinal bleed*) OR ("Gastrointestinal Hemorrhage"[Mesh]) OR (gastrointestinal hemorrhag*) OR (gastrointestinal haemorrhag*)
15. #10 OR #11 OR #12 OR #13 OR #14
16. #3 AND #9 AND #15

EMBASE search strategy (via embase.com)

1. health NEAR/3 administrative OR administrative NEAR/3 data OR administrative NEAR/3 database OR claim NEAR/3 administrative OR (International Classification of Diseases) OR 'International Classification of Diseases'/exp OR ICD-9-CM OR ICD-10 OR 'Database Management Systems'/exp OR 'Medical Records Systems, Computerized'/exp OR 'CPT' OR 'Current procedural terminology'/exp
2. database:ab,ti OR (('practitioner'/exp OR practitioner) AND data AND bank) OR (('practitioner'/exp OR practitioner) AND ('database'/exp OR database)) OR ('insurance' AND ('database'/exp OR database))
3. #1 OR #2
4. 'sensitivity and specificity'/exp OR 'sensitivity and specificity'
5. specificity:ab,ti
6. 'predictive value of tests'/exp OR 'predictive value of tests'
7. (positive:ab,ti AND predictive:ab,ti AND value:ab,ti) OR (negative:ab,ti AND predictive:ab,ti AND value:ab,ti) OR (likelyhood:ab,ti AND ratio:ab,ti) OR (receiver:ab,ti AND operating:ab,ti AND characteristic:ab,ti) OR kappa:ab,ti
8. case NEAR/1 (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*)
9. 'algorithms'/exp OR algorithm
10. #4 OR #5 OR #6 OR #7 OR #8 OR #9
11. 'stomach'/exp OR 'stomach ulcer'/exp OR (stomach NEAR/3 ulcer*):ab,ti OR (gastr* NEAR/3 ulcer*):ab,ti
12. 'duodenal'/exp OR 'duodenal ulcer'/exp OR (duodenal NEAR/3 ulcer*):ab,ti OR (curling* NEAR/3 ulcer*):ab,ti
13. 'peptic'/exp OR 'peptic ulcer'/exp OR (peptic NEAR/3 ulcer*):ab,ti OR (marginal NEAR/3 ulcer*):ab,ti
14. 'ulcer'/exp OR 'ulcer bleed'/exp OR (ulcer NEAR/3 bleed*) OR (ulcer NEAR/3 hemorrhag*):ab,ti OR (ulcer NEAR/3 haemorrhag*):ab,ti OR (ulcer NEAR/3 perforat*):ab,ti
15. 'gastrointestinal'/exp OR 'gastrointestinal bleed'/exp OR (gastrointestinal NEAR/3 bleed*):ab,ti OR (gastrointestinal NEAR/3 hemorrhag*):ab,ti OR (gastrointestinal NEAR/3 haemorrhag*):ab,ti
16. #11 OR #12 OR #13 OR #14 OR #15
17. #3 AND #10 AND #16

Web of Science search strategy

1. (health NEAR/3 administrative) OR (administrative NEAR/3 data) OR (administrative NEAR/3 database) OR (claim NEAR/3 administrative) OR (International Classification of Diseases) OR ICD-9-CM OR ICD-10 OR (Database Management Systems) OR ("Medical Records Systems" NEAR/2 Computerized) OR "CPT" OR (Current procedural terminology)
2. (factual databases) OR (geographic information systems) OR (national practitioner data bank) OR (insurance database)
3. #1 OR #2
4. sensitivity or "Sensitivity and Specificity"
5. specificity
6. (positive predictive value) OR (negative predictive value) OR (likelihood ratio) OR (receiver operating characteristic) OR kappa
7. ((case or cases) AND (verificat* OR valid* OR identif* OR definition* OR define* OR evaluat*))
8. algorithm
9. #4 OR #5 OR #6 OR #7 OR #8
10. (stomach NEAR/3 ulcer*) OR (gastr* NEAR/3 ulcer*)
11. (duodenal NEAR/3 ulcer*) OR (curling* NEAR/3 ulcer*)
12. (peptic NEAR/3 ulcer*) OR (marginal NEAR/3 ulcer*)
13. (ulcer NEAR/3 bleed*) OR (ulcer NEAR/3 hemorrhag*) OR (ulcer NEAR/3 haemorrhag*) OR (ulcer NEAR/3 perforat*)
14. (gastrointestinal NEAR/3 bleed*) OR (gastrointestinal NEAR/3 hemorrhag*) OR (ulcer NEAR/3 haemorrhag*)
15. #10 OR #11 OR #12 OR #13 OR #14
16. #3 AND #9 AND #15

The Cochrane Library

1. (health near/3 administrative) or (administrative near/3 data) or (administrative near/3 database) or (claim near/3 administrative) or (International Classification of Diseases) or [mh "International Classification of Diseases"] or ICD-9-CM or ICD-10 or [mh "Database Management Systems"] or [mh "Medical Records Systems, Computerized"] or "CPT" or [mh "Current procedural terminology"]
2. (factual databases) or (geographic information systems) or (national practitioner data bank) or (insurance database)
3. #1 or #2
4. sensitivity or [mh "Sensitivity and Specificity"]
5. specificity:ti,ab,kw
6. (positive predictive value) or (negative predictive value) or (likelihood ratio) or (receiver operating characteristic) or kappa
7. ((case or cases) and (verificat* or valid* or identif* or definition* or define* or evaluat*))
8. Algorithm or [mh "Algorithm"]
9. #4 or #5 or #6 or #7 or #8
10. [mh "Stomach Ulcer"] or (stomach near/3 ulcer*) or (gastr* near/3 ulcer*)
11. [mh "Duodenal Ulcer"] or (duodenal near/3 ulcer*) or (curling* near/3 ulcer*)
12. [mh "Peptic Ulcer"] or (peptic near/3 ulcer*) or (marginal near/3 ulcer*)
13. [mh "Peptic Ulcer Hemorrhage"]) or (ulcer near/3 bleed*) or (ulcer near/3 hemorrhag*) or (ulcer near/3 haemorrhag*) or (ulcer near/3 perforat*)
14. [mh "Gastrointestinal Hemorrhage"] or (gastrointestinal near/3 bleed*) or (gastrointestinal near/3 hemorrhag*) or (gastrointestinal near/3 haemorrhag*)
15. #10 or #11 or #12 or #13 or #14
16. #3 and #9 and #15

Appendix 2 – List with descriptions of ICD-9 and ICD-10 codes for gastrointestinal ulcer and haemorrhage.

ICD-9	Description	ICD-10	Description
531	**Gastric ulcer**	**K25**	**Gastric ulcer**
531.00	Acute gastric ulcer with haemorrhage without obstruction	K25.0	Acute gastric ulcer with hemorrhage
531.01	Acute gastric ulcer with hemorrhage with obstruction	K25.0	Acute gastric ulcer with haemorrhage
531.10	Acute gastric ulcer with perforation without obstruction	K25.1	Acute gastric ulcer with perforation
531.11	Acute gastric ulcer with perforation with obstruction	K25.1	Acute gastric ulcer with perforation
531.20	Acute gastric ulcer with haemorrhage and perforation without obstruction	K25.2	Acute gastric ulcer with both haemorrhage and perforation
531.21	Acute gastric ulcer with haemorrhage and perforation with obstruction	K25.2	Acute gastric ulcer with both haemorrhage and perforation
531.30	Acute gastric ulcer without haemorrhage or perforation without obstruction	K25.3	Acute gastric ulcer without haemorrhage or perforation
531.31	Acute gastric ulcer without haemorrhage or perforation with obstruction	K25.3	Acute gastric ulcer without haemorrhage or perforation
531.40	Chronic or unspecified gastric ulcer with haemorrhage without obstruction	K25.4	Chronic or unspecified gastric ulcer with haemorrhage
531.41	Chronic or unspecified gastric ulcer with haemorrhage with obstruction	K25.4	Chronic or unspecified gastric ulcer with haemorrhage
531.50	Chronic or unspecified gastric ulcer with perforation without obstruction	K25.5	Chronic or unspecified gastric ulcer with perforation
531.51	Chronic or unspecified gastric ulcer with perforation with obstruction	K25.5	Chronic or unspecified gastric ulcer with perforation
531.60	Chronic or unspecified gastric ulcer with haemorrhage and perforation without obstruction	K25.6	Chronic or unspecified gastric ulcer with both haemorrhage and perforation
531.61	Chronic or unspecified gastric ulcer with haemorrhage and perforation with obstruction	K25.6	Chronic or unspecified gastric ulcer with both haemorrhage and perforation

531.70	Chronic gastric ulcer without haemorrhage or perforation without obstruction	K25.7	Chronic gastric ulcer without haemorrhage or perforation
531.71	Chronic gastric ulcer without haemorrhage or perforation with obstruction	K25.7	Chronic gastric ulcer without haemorrhage or perforation
531.90	Gastric ulcer unspecified as acute or chronic without haemorrhage or perforation without obstruction	K25.9	Gastric ulcer, unspecified as acute or chronic, without haemorrhage or perforation
531.91	Gastric ulcer unspecified as acute or chronic without haemorrhage or perforation with obstruction	K25.9	Gastric ulcer, unspecified as acute or chronic, without haemorrhage or perforation
532	**Duodenal Ulcer**	**K26**	**Duodenal Ulcer**
532.00	Acute duodenal ulcer with haemorrhage without obstruction	K26.0	Acute duodenal ulcer with haemorrhage
532.01	Acute duodenal ulcer with haemorrhage with obstruction	K26.0	Acute duodenal ulcer with haemorrhage
532.10	Acute duodenal ulcer with perforation without obstruction	K26.1	Acute duodenal ulcer with perforation
532.11	Acute duodenal ulcer with perforation with obstruction	K26.1	Acute duodenal ulcer with perforation
532.20	Acute duodenal ulcer with haemorrhage and perforation without obstruction	K26.2	Acute duodenal ulcer with both haemorrhage and perforation
532.21	Acute duodenal ulcer with haemorrhage and perforation with obstruction	K26.2	Acute duodenal ulcer with both haemorrhage and perforation
532.30	Acute duodenal ulcer without haemorrhage or perforation without obstruction	K26.3	Acute duodenal ulcer without haemorrhage or perforation
532.31	Acute duodenal ulcer without haemorrhage or perforation with obstruction	K26.3	Acute duodenal ulcer without haemorrhage or perforation
532.40	Chronic or unspecified duodenal ulcer with haemorrhage without obstruction	K26.4	Chronic or unspecified duodenal ulcer with haemorrhage
532.41	Chronic or unspecified duodenal ulcer with haemorrhage with obstruction	K26.4	Chronic or unspecified duodenal ulcer with haemorrhage
532.50	Chronic or unspecified duodenal ulcer with perforation without obstruction	K26.5	Chronic or unspecified duodenal ulcer with perforation
532.51	Chronic or unspecified duodenal ulcer with perforation with obstruction	K26.5	Chronic or unspecified duodenal ulcer with perforation

532.60	Chronic or unspecified duodenal ulcer with haemorrhage and perforation without obstruction	K26.6	Chronic or unspecified duodenal ulcer with both haemorrhage and perforation
532.61	Chronic or unspecified duodenal ulcer with haemorrhage and perforation with obstruction	K26.6	Chronic or unspecified duodenal ulcer with both haemorrhage and perforation
532.70	Chronic duodenal ulcer without haemorrhage or perforation without obstruction	K26.7	Chronic duodenal ulcer without haemorrhage or perforation
532.71	Chronic duodenal ulcer without haemorrhage or perforation with obstruction	K26.7	Chronic duodenal ulcer without haemorrhage or perforation
532.90	Duodenal ulcer unspecified as acute or chronic without haemorrhage or perforation without obstruction	K26.9	Duodenal ulcer, unspecified as acute or chronic, without haemorrhage or perforation
532.91	Duodenal ulcer unspecified as acute or chronic without haemorrhage or perforation with obstruction	K26.9	Duodenal ulcer, unspecified as acute or chronic, without haemorrhage or perforation
533	**Peptic ulcer, site unspecified**	**K27**	**Peptic ulcer, site unspecified**
533.00	Acute peptic ulcer of unspecified site with haemorrhage without obstruction	K27.0	Acute peptic ulcer, site unspecified, with haemorrhage
533.01	Acute peptic ulcer of unspecified site with haemorrhage with obstruction	K27.0	Acute peptic ulcer, site unspecified, with haemorrhage
533.10	Acute peptic ulcer of unspecified site with perforation without obstruction	K27.1	Acute peptic ulcer, site unspecified, with perforation
533.11	Acute peptic ulcer of unspecified site with perforation with obstruction	K27.1	Acute peptic ulcer, site unspecified, with perforation
533.20	Acute peptic ulcer of unspecified site with haemorrhage and perforation without obstruction	K27.2	Acute peptic ulcer, site unspecified, with both haemorrhage and perforation
533.21	Acute peptic ulcer of unspecified site with haemorrhage and perforation with obstruction	K27.2	Acute peptic ulcer, site unspecified, with both haemorrhage and perforation
533.30	Acute peptic ulcer of unspecified site without haemorrhage and perforation without obstruction	K27.3	Acute peptic ulcer, site unspecified, without haemorrhage or perforation
533.31	Acute peptic ulcer of unspecified site without haemorrhage and perforation with obstruction	K27.3	Acute peptic ulcer, site unspecified, without haemorrhage or perforation
533.40	Chronic or unspecified peptic ulcer of unspecified site with haemorrhage without obstruction	K27.4	Chronic or unspecified peptic ulcer, site unspecified, with haemorrhage

533.41	Chronic or unspecified peptic ulcer of unspecified site with haemorrhage with obstruction	K27.4	Chronic or unspecified peptic ulcer, site unspecified, with haemorrhage
533.50	Chronic or unspecified peptic ulcer of unspecified site with perforation without obstruction	K27.5	Chronic or unspecified peptic ulcer, site unspecified, with perforation
533.51	Chronic or unspecified peptic ulcer of unspecified site with perforation with obstruction	K27.5	Chronic or unspecified peptic ulcer, site unspecified, with perforation
533.60	Chronic or unspecified peptic ulcer of unspecified site with haemorrhage and perforation without obstruction	K27.6	Chronic or unspecified peptic ulcer, site unspecified, with both haemorrhage and perforation
533.61	Chronic or unspecified peptic ulcer of unspecified site with haemorrhage and perforation with obstruction	K27.6	Chronic or unspecified peptic ulcer, site unspecified, with both haemorrhage and perforation
533.70	Chronic peptic ulcer of unspecified site without haemorrhage or perforation without obstruction	K27.7	Chronic peptic ulcer, site unspecified, without haemorrhage or perforation
533.71	Chronic peptic ulcer of unspecified site without haemorrhage or perforation with obstruction	K27.7	Chronic peptic ulcer, site unspecified, without haemorrhage or perforation
533.90	Peptic ulcer of unspecified site unspecified as acute or chronic without haemorrhage or perforation without obstruction	K27.9	Peptic ulcer, site unspecified, unspecified as acute or chronic, without haemorrhage or perforation
533.91	Peptic ulcer of unspecified site unspecified as acute or chronic without haemorrhage or perforation with obstruction	K27.9	Peptic ulcer, site unspecified, unspecified as acute or chronic, without haemorrhage or perforation
534	**Gastrojejunal ulcer**	**K28**	**Gastrojejunal ulcer**
534.00	Acute gastrojejunal ulcer with haemorrhage without obstruction	K28.0	Acute gastrojejunal ulcer with haemorrhage
534.01	Acute gastrojejunal ulcer with haemorrhage with obstruction	K28.0	Acute gastrojejunal ulcer with haemorrhage
534.10	Acute gastrojejunal ulcer with perforation without obstruction	K28.1	Acute gastrojejunal ulcer with perforation
534.11	Acute gastrojejunal ulcer with perforation with obstruction	K28.1	Acute gastrojejunal ulcer with perforation
534.20	Acute gastrojejunal ulcer with haemorrhage and perforation without obstruction	K28.2	Acute gastrojejunal ulcer with both haemorrhage and perforation
534.21	Acute gastrojejunal ulcer with haemorrhage and perforation with	K28.2	Acute gastrojejunal ulcer with both haemorrhage and perforation

	obstruction		
534.30	Acute gastrojejunal ulcer without haemorrhage or perforation without obstruction	K28.3	Acute gastrojejunal ulcer without haemorrhage or perforation
534.31	Acute gastrojejunal ulcer without haemorrhage or perforation with obstruction	K28.3	Acute gastrojejunal ulcer without haemorrhage or perforation
534.40	Chronic or unspecified gastrojejunal ulcer with haemorrhage without obstruction	K28.4	Chronic or unspecified gastrojejunal ulcer with haemorrhage
534.41	Chronic or unspecified gastrojejunal ulcer with haemorrhage with obstruction	K28.4	Chronic or unspecified gastrojejunal ulcer with haemorrhage
534.50	Chronic or unspecified gastrojejunal ulcer with perforation without obstruction	K28.5	Chronic or unspecified gastrojejunal ulcer with perforation
534.51	Chronic or unspecified gastrojejunal ulcer with perforation with obstruction	K28.5	Chronic or unspecified gastrojejunal ulcer with perforation
534.60	Chronic or unspecified gastrojejunal ulcer with haemorrhage and perforation without obstruction	K28.6	Chronic or unspecified gastrojejunal ulcer with both haemorrhage and perforation
534.61	Chronic or unspecified gastrojejunal ulcer with haemorrhage and perforation with obstruction	K28.6	Chronic or unspecified gastrojejunal ulcer with both haemorrhage and perforation
534.70	Chronic gastrojejunal ulcer without haemorrhage or perforation without obstruction	K28.7	Chronic gastrojejunal ulcer without haemorrhage or perforation
534.71	Chronic gastrojejunal ulcer without haemorrhage or perforation with obstruction	K28.7	Chronic gastrojejunal ulcer without haemorrhage or perforation
534.90	Gastrojejunal ulcer unspecified as acute or chronic without haemorrhage or perforation without obstruction	K28.9	Gastrojejunal ulcer, unspecified as acute or chronic, without haemorrhage or perforation
534.91	Gastrojejunal ulcer unspecified as acute or chronic without haemorrhage or perforation with obstruction	K28.9	Gastrojejunal ulcer, unspecified as acute or chronic, without haemorrhage or perforation
578	**Gastrointestinal haemorrhage**	**K92**	**Other diseases of digestive system**
578.0	Hematemesis	K92.0	Hematemesis
578.1	Blood in stool	K92.1	Melena
578.9	Haemorrhage of gastrointestinal tract unspecified	K92.2	Gastrointestinal haemorrhage , unspecified

Appendix 3

Checklist of reporting criteria for studies validating health administrative data algorithms (developed by Benchimol et al., based on the criteria published by the Standards for Reporting of Diagnostic accuracy (STARD) initiative for the accurate reporting of studies using diagnostic studies.

	YES	NO	UNCERTAIN	NOT APPLICABLE
TITLE, KEYWORDS, ABSTRACT				
Identify article as study of assessing diagnostic accuracy				
Identify article as study of administrative data				
INTRODUCTION:				
State disease identification & validation one of goals of study				
METHODS:				
Participants in validation cohort:				
Describe validation cohort (Cohort of patients to which reference standard was applied)				
• Age				
• Disease				
• Severity				
• Location/Jurisdiction				
Describe recruitment procedure of validation cohort				
• Inclusion criteria				
• Exclusion criteria				
Describe patient sampling (random, consecutive, all, etc.)				
Describe data collection				
• Who identified patients and did selection adhere to patient recruitment criteria				
• Who collected data				
• *A priori* data collection form				
• Disease classification				
• Split sample (i.e. re-validation using a separate cohort) a) Training set b) Testing set				
Test Methods:				
Describe number, training and expertise of persons reading reference standard				
If >1 person reading reference standard, quote measure of consistency (e.g. kappa)				

Blinding of interpreters of reference standard to results of classification by administrative data e.g. Chart abstractor blinded to how that chart was coded				
Statistical Methods:				
Describe methods of calculating/comparing diagnostic accuracy				
RESULTS:				
Participants:				
Report when study done, start/end dates of enrollment				
Describe number of people who satisfied inclusion/exclusion criteria				
Study flow diagram				
Test results:				
Report distribution of disease severity				
Report cross-tabulation of index tests by results of reference standard				
Estimates:				
Report at least 4 estimates of diagnostic accuracy				
Diagnostic Accuracy Measures Reported:				
• Sensitivity				
• Spec				
• PPV				
• NPV				
• Likelihood ratios				
• Kappa				
• Area under the ROC curve / c-statistic				
• Accuracy/agreement				
• Other (specify)				
Report accuracy for subgroups (e.g. age, geography, different sex, etc.)				
If PPV/NPV reported, ratio of cases/controls of validation cohort approximate prevalence of condition in the population				
Report 95% confidence intervals for each diagnostic measure				
DISCUSSION:				
Discuss the applicability of the validation findings				

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	Page 1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	Page 2: Trial registration number PROSPERO 2015 CRD42015029216
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	Page 1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	Page 11
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	At this stage there are no relevant amendments to perform
Support:			
Sources	5a	Indicate sources of financial or other support for the review	Page 11 *(Regional Health Authority of Umbria, Italy)*
Sponsor	5b	Provide name for the review funder and/or sponsor	Page 11 (*Regional Health Authority of Umbria, Italy.)*
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	Page 11
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	Page 4 and 5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Page 5
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting,	Pages 6-8:

		time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	Page 6.
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Appendix 1 in Supplemental file
Study records:			Pages 8, 9
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	Pages 8, 9
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	Pages 8, 9
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Pages 8, 9
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	Pages 8, 9.
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	Pages 7, 8
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Page 9. The present review will apply the STARD criteria.
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	Pages 9-10.
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	Pages 9-10.
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	Pages 9-10.
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	Pages 9-10.
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication	Page 10

		bias across studies, selective reporting within studies)	
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	The present review will apply the STARD criteria. Page 9.

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.