How accurate are diagnoses recorded in administrative databases: A systematic review

Maria Cassandre Médor BSc (student), Amanda Vandyk RN, PhD
Faculty of Health Sciences, University of Ottawa

Background

Diagnostic codes (such as ICD codes) recorded in administrative databases are often used in clinical research to identify participant diagnoses.

De Coster and al. (2006) specify 13 priorities for research on administrative databases, including the need for studies assessing diagnostic congruence with medical records.

Preliminary evidence suggests that at least 2 ICD codes are necessary to accurately identify individuals with confirmed chronic conditions (Goldberg and al., 2013)

Unfortunately, researchers continue to use single ICD codes to identify study participants (Krueger and al., 2011)

Furthermore, the methods used to assign ICD codes may also jeopardize the reliability and validity of the diagnoses documented.

More work is needed to understand the strengths and limitations of using ICD codes when to identify participant diagnoses

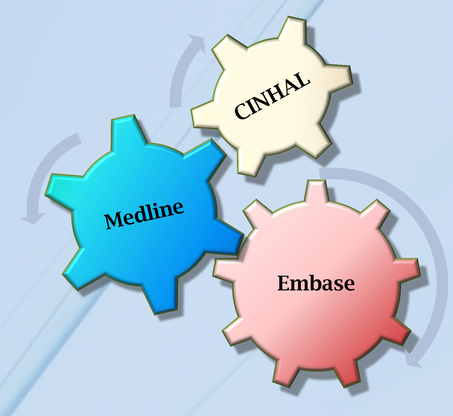
This systematic review was based on the Joanne Briggs Methodology

Objective

To assess the accuracy of diagnostic information captured in administrative systems, compared to most responsible physician diagnosis, for individuals with chronic and complex health conditions.

Search Strategy

- A comprehensive three-step search of online health care literature published within the most relevant databases



- Subject headings and keywords were identified and verified using the thesaurus feature in each database
- References of all included articles were hand-searched
- A library scientist at the University of Ottawa helped design the search strategies in an effort to avoid missing pertinent literature

Methods

Two researchers independently examined each study obtained during this process and only the articles approved by both were kept.

Inclusion Criteria
Identified using PICO tool (Stone, 2002)

Population	:	Individuals with a chronic illness
Intervention	:	Identification of DX using admin. codes and physician record
Context	:	All tertiary health care settings
Outcome	:	Concordance between admin. assigned and physician documented DX
Study Design	:	All quantitative study designs
Language	:	English or French

Exclusion Criteria

Study Design	:	Qualitative studies, abstracts, books, and conference proceedings
Language	:	Other than English or French

Data Extraction
Structured abstract designed for this purpose:

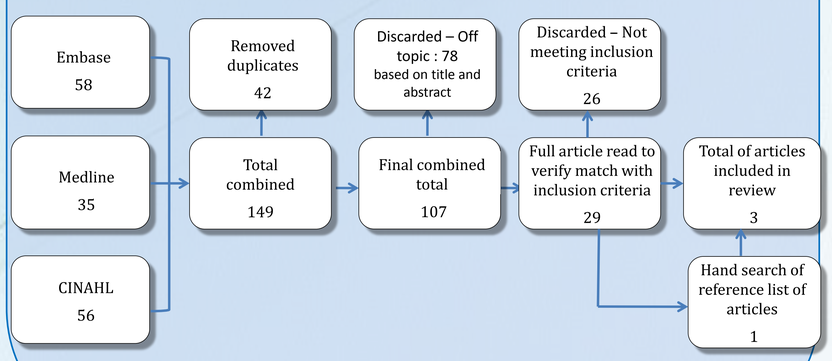
1. Study characteristics
2. Illness
3. Number of participants with DX in administrative database
4. Number of participants with DX in medical record
5. Number of participants with DX in both admin. and medical records

Search strategy using Keywords in Embase Classic + Embase (1946 to present)

Line nº	Subject Headings and Keywords	Results
1	Chronic diseases or Chronic disease*.mp or Chronic illness*.mp	201320
2	International classification of disease.mp or ''International classification of diseases''	8618
3	1 and 2	237
4	Emergencies or Emergency Service, Hospital or Emergenc*. mp or Urgent.mp	435636
5	3 and 4	10
6	Patient Readmission	16889
7	(''Frequent users'' or ''returning patients'' or ''multiple contact'').mp	1823
8	6 or 7	18698
9	5 and 8	6
10	12 or 17	15
11	Physician diagnosis.mp or Physician record.mp or Medical record.mp or Medical Records	193553
12	9 and 12	27

Embase total included = 58

Article selection process - Search decision tree

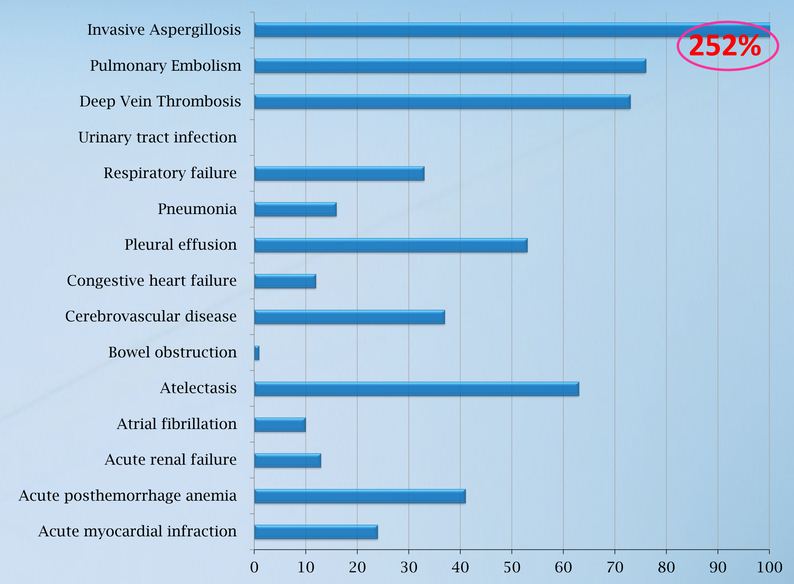


Preliminary Results

Comparison of diagnoses recorded in administrative databases and medical records (N = 3 studies)

Diagnosis	Diagnosis identified in medical record N	Diagnosis identified by ICD code N	ICD diagnosis congruence +/- N (%)
Acute myocardial infraction	25	19	- 6 (24)
Acute posthemorrhage anemia	29	44	+ 15 (41)
Acute renal failure	45	39	- 6 (13)
Atrial fibrillation	113	102	- 11 (10)
Atelectasis	112	42	- 70 (63)
Bowel obstruction	77	76	- 1 (1)
Cerebrovascular disease	54	34	- 20 (37)
Congestive heart failure	129	113	- 16 (12)
Pleural effusion	135	63	- 72 (53)
Pneumonia	105	88	- 17 (16)
Respiratory failure	45	30	- 15 (33)
Urinary tract infection	51	51	0 (0)
Deep Vein Thrombosis	232	72	- 170 (73)
Pulmonary Embolism	95	23	- 72 (76)
Invasive Aspergillosis	19	67	+ 48 (252)

Accuracy of the diagnoses captured in administrative databases compared to medical records



References

- Chang, D. C., Burwell, L. A., Lyon, G. M., Pappas, P. G., Chiller, T. M., Wannemuehler, K. A., . . . & Park, B. J. (2008). Comparison of the use of administrative data and an active system for surveillance of invasive aspergillosis. *Infection Control & Hospital Epidemiology, 29*(1), 25-30.
- De Coster, C., Quan, H., Finlayson, A., Gao, M., Halfon, P., Humphries, K. H., . . . & Ghali, W. A. (2006). Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: Report from an international consortium. *BMC Health Services Research, 6*, 77.
- Goldberg, D. S., Lewis, J. D., Halpern, S. D., Weiner, M. G., & Lo Re, V.,3rd. (2013). Validation of a coding algorithm to identify patients with hepatocellular carcinoma in an administrative database. *Pharmacoepidemiology & Drug Safety, 22*(1), 103-107.
- Krueger, K. P., Armstrong, E. P., & Langley, P. C. (2001). The accuracy of asthma and respiratory disease diagnostic codes in a managed care medical claims database. *Disease Management, 4*(4), 155-161.
- Quan, H., Parsons, G. A., & Ghali, W. A. (2004). Assessing accuracy of diagnosis-type indicators for flagging complications in administrative data. *Journal of Clinical Epidemiology*, 57(4), 366-372.
- Zhan, C., Battles, J., Chiang, Y., & Hunt, D. (2007). The validity of ICD-9-CM codes in identifying postoperative deep vein thrombosis and pulmonary embolism. *Joint Commission Journal on Quality & Patient Safety, 33*(6), 326-331.

Acknowledgements

This research received partial support from the University of Ottawa's Undergraduate Research Opportunity Program. We express great gratitude for this opportunity and the funding to make this project possible.