BMC
Medical Informatics & Decision Making

RESEARCH ARTICLE   **Open Access**

Developing model-based algorithms to identify screening colonoscopies using administrative health databases

Maida J Sewitch[1,2,3,4*], Mengzhu Jiang[2], Lawrence Joseph[2,4], Robert J Hilsden[5] and Alain Bitton[1,3]

Abstract

Background: Algorithms to identify screening colonoscopies in administrative databases would be useful for monitoring colorectal cancer (CRC) screening uptake, tracking health resource utilization, and quality assurance. Previously developed algorithms based on expert opinion were insufficiently accurate. The purpose of this study was to develop and evaluate the accuracy of model-based algorithms to identify screening colonoscopies in health administrative databases.

Methods: Patients aged 50-75 were recruited from endoscopy units in Montreal, Quebec, and Calgary, Alberta. Physician billing records and hospitalization data were obtained for each patient from the provincial administrative health databases. Indication for colonoscopy was derived using Bayesian latent class analysis informed by endoscopist and patient questionnaire responses. Two modeling methods were used to fit the data, multivariate logistic regression and recursive partitioning. The accuracies of these models were assessed.

Results: 689 patients from Montreal and 541 from Calgary participated (January to March 2007). The latent class model identified 554 screening exams. Multivariate logistic regression predictions yielded an area under the curve of 0.786. Recursive partitioning using the latent outcome had sensitivity and specificity of 84.5% (95% CI: 81.5-87.5) and 63.3% (95% CI: 59.7-67.0), respectively.

Conclusions: Model-based algorithms using administrative data failed to identify screening colonoscopies with sufficient accuracy. Nevertheless, the approach of constructing a latent reference standard against which model-based algorithms were evaluated may be useful for validating administrative data in other contexts where there lacks a gold standard.

Background

Administrative records are frequently used in health research. While some diagnoses and procedures are recorded with reasonable accuracy [1,2], others are prone to misclassification [3,4]. Indications for medical procedures are particularly challenging to derive from administrative health data because of the lack of indication codes, and, therefore, require automated data algorithms [5-7]. Studies validating administrative data algorithms have typically used medical chart review as the gold standard [7-13]. However, information in medical charts may be inaccurate for reasons including the variable quality of record keeping and record extraction. Moreover, with Canadian average wait-times from referral to performance of any gastrointestinal procedure of 155 days and to screening colonoscopy of 201 days [14], the medical chart may not reflect symptoms at the time the colonoscopy is performed. In this study, we undertook the challenge of evaluating model-based algorithms in the absence of a gold standard measure for the indication of colonoscopy.

Colorectal cancer (CRC) screening is recommended worldwide for asymptomatic persons aged 50 to 75 [15-18]. Many industrialized countries have committed to or already implemented population-based CRC screening programs using modalities such as fecal occult blood test, fecal immunochemical test, flexible sigmoidoscopy,

* Correspondence: maida.sewitch@mcgill.ca
[1]Department of Medicine, McGill University, Montreal, Quebec, Canada
[2]Division of Clinical Epidemiology, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada
Full list of author information is available at the end of the article

and colonoscopy [19]. Colonoscopy is the only exam that enables visualization and removal of precancerous and cancerous lesions throughout the entire colon, and recent guidelines promote its role as a first-line screening modality [15,16]. Colonoscopy utilization has increased dramatically owing to its use in CRC screening [20,21]. In addition to its use in screening, colonoscopy is also performed for surveillance for bowel diseases, diagnostics for large bowel symptoms, and follow-up for positive results by other CRC screening modalities. A screening colonoscopy is defined as one performed in asymptomatic individuals for the early detection of CRC or the detection and removal of precancerous lesions [22].

Population screening initiatives are accompanied by increasing interest in undertaking large-scale colonoscopic screening studies using existing administrative health databases. However, most such databases either do not have a screening colonoscopy procedure code or the code is underused since the primary purpose is remuneration [21,23]. Methods to distinguish screening and non-screening colonoscopies would enable monitoring of CRC screening uptake, tracking of health resource utilization, and estimation of cost-effectiveness; they may also be used for quality assessment as a key quality indicator is the adenoma detection rate in screening colonoscopies [22,24,25]. Automated data screening colonoscopy algorithms developed in previous studies had sensitivities ranging between 29% and 84% and specificities ranging between 58% and 93%; none of the algorithms had both high sensitivity and specificity [7,12,13], which led to the conclusion that administrative data cannot reliably be used to distinguish between colonoscopy indications [7,26]. However, these prior studies relied on expert opinion regarding which diagnostic and procedural codes to use and in what order. In contrast, model-based algorithms use the data to help determine which variables to include and their weights, and thus have the potential to be more accurate.

A validated database algorithm would be convenient and efficient for researchers and administrators. The objectives of this study were to develop model-based algorithms to identify screening colonoscopies in health administrative databases, to evaluate their accuracy against a latent reference standard, and to compare their accuracies to that from a previously developed algorithm based on expert opinion.

Methods
Study design
A retrospective cohort study was conducted of endoscopists and a convenience sample of their patients about to undergo colonoscopy between January and March 2007 in two Canadian cities (Montreal, Quebec and Calgary, Alberta). Participating institutions were those where colonoscopy was performed and billed to the provincial

health insurance plans (Montreal: McGill University Health Centre, Sir Mortimer B. Davis Jewish General Hospital, St. Mary's Hospital Centre, Centre hospitalier de l'Université de Montréal, Fleury Hospital, Maisonneuve-Rosemont Hospital; Calgary: Foothills Medical Centre, Peter Lougheed Centre). At the time of the study, the provincial CRC screening program in Alberta relied on family physicians to offer fecal occult blood tests to patients aged 50 to 75 but no such program existed in Quebec, where 'screening' occurred opportunistically at the individual physician's discretion. In both provinces, patients and/or family physicians could opt for colonoscopy as the initial screening exam.

Data collection
The research assistant assessed endoscopists and patients for eligibility. Eligible endoscopists received remuneration for colonoscopy from the provincial health insurance board. Immediately after each colonoscopy, the endoscopist completed a questionnaire on the colonoscopy indication; the screening indication was defined as 'performed in asymptomatic people at average-risk for developing colorectal cancer, or in people with a family history of colorectal cancer'. It is unknown whether the endoscopist based the indication on the colonoscopy referral, communication with the patient, or something else. Eligible patients were aged 50 to 75 years; those without provincial health insurance plan coverage in the prior year or unable to give consent were excluded. The research assistant approached patients prior to colonoscopy, explained the study, obtained consent, and administered the patient questionnaire. Patient perceived indication was defined in two ways: 1) non-screening if patient reported that the reason for colonoscopy was to follow-up for a previous test or problem, and is screening otherwise; and 2) non-screening if patient reported specific lower abdominal symptoms and personal history of gastrointestinal (GI) condition, and is screening otherwise. These patient indications, and their agreement with endoscopist indication, have been described in detail elsewhere [27].

We obtained provincial administrative health data on participating patients for the five years prior to the index colonoscopy as follows: Physician billing records from the Régie de l'assurance maladie du Québec (RAMQ) and Alberta Health and Wellness provided data on patient age and sex, all medical acts (RAMQ billing codes in Quebec, and Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP) codes in Alberta). The Maintenance et Exploitation des Données pour l'Étude de la Clientèle Hospitalière (MED-Echo) and the Canadian Institute for Health Information (CIHI) provided information on hospitalizations (The International Classification of Diseases ICD-9 and ICD-10 codes) and

surgeries (ICD-9, CCP, and Canadian Classification of Health Interventions (CCI) codes). Data were linked using unique patient health insurance numbers. Prior to study inception, ethics approval was obtained from the Institutional Review Board in the Faculty of Medicine at McGill University and the local research ethics boards.

Statistical analyses

Since there is no gold standard method for identifying screening colonoscopies, a Bayesian latent class model for diagnostic testing was used to provide the probability that any given colonoscopy was for screening purposes [28], based on endoscopist indication and the two patient indications. Flat or non-informative prior distributions were used for the two patient indications, which were assumed to be conditionally independent. To examine the robustness of this assumption, a model including a dependence between these two indications was also fitted to the same data [29]. For the endoscopist screening indication, a beta(10.67, 1.06) density, 97% of which covers the range from 70% to 100%, was used for both sensitivity and specificity. A beta(6, 7.6) density, 95% of which covers the range from 20-70% was used for the prevalence of screening. These priors were based on expert opinion, and covered the ranges of all plausible values with relatively flat density. Latent class modeling was carried out using WinBUGS software (MRC Biostatistics Unit, Cambridge).

The predicted probabilities for screening from the latent class model, based on posterior medians, were dichotomized into screening and non-screening using a cut-off of 50% probability. The dichotomized latent class indication was then used as the outcome variable for multivariate logistic regression and recursive partitioning. For comparison purposes, we also fitted models using endoscopist indication alone as the outcome. Predictor variables entered into the models were: age, sex, procedure codes for previous procedures (colonoscopy, polypectomy, sigmoidoscopy, and double contrast barium enema (DCBE)) in the past 4 years, diagnostic codes for risk factors (inflammatory bowel disease (IBD), colorectal polyp, and CRC) in the past 5 years, diagnostic codes for symptoms (rectal bleeding, anemia, diarrhea, vomiting, and weight loss) in the past year, diagnostic codes for hospitalization for large bowel diseases in the past 5 years, and procedure codes for large bowel surgeries in the past 5 years (Additional file 1). The variables selected were based on practice guidelines [15,16,18], published studies [7,12,13], and expert opinion.

The Bayesian information criterion (BIC) was used to select the multivariate logistic regression model that best predicts the screening indication [30]. Model discrimination was assessed by the area under the curve (AUC) of the receiver operator characteristic curve [31]. The accuracy of classification trees generated by the recursive partitioning model was assessed against the latent class predictions and endoscopist indication; sensitivities, specificities, positive and negative predicative values (PPV, NPV) were computed. Multivariate and recursive modeling were performed using R [32].

We also applied an algorithm based on expert-opinion previously developed by El-Serag et al., which defines screening colonoscopies as the absence of ICD-9 codes for 28 symptoms or conditions and no colonoscopy in the past 4 years [12]. Sensitivities, specificities, PPVs, and NPVs were estimated by comparing the algorithm classification to latent class predictions and to endoscopist indication.

Results

Participant characteristics

A total of 1,411 patients were approached, of which 1,230 (87.2%) were eligible and agreed to participate, 689 (56.0%) from Montreal and 541 (44.0%) from Calgary. In Montreal and Calgary, 52 and 0 eligible patients approached refused study participation, respectively. The average age was 60 and 48.5% of participants were male (Table 1). Endoscopists reported screening as the reason for colonoscopy in 46.8% of colonoscopies, whereas patient indications 1 (patient perceived reason) and 2 (based on patient reported symptoms and GI history) were screening in 51.0% and 38.9% of the colonoscopies, respectively. The frequency of occurrence of diagnostic and procedure codes of interest in patient administrative health records are presented in Table 2.

Table 1 Patient characteristics (N = 1,230)

Patient characteristic	N	%
Age (mean, sd)	60.1 (6.9)	
Male sex	597	48.5
Patient reported symptoms[a]	573	46.6
Patient reported gastrointestinal conditions[b]	350	28.5
Patient reported positive FOBT[c] in the past 12 months	71	5.8
Endoscopist indication = screening	576	46.8
Patient indication 1[d] = screening	627	51.0
Patient indication 2[e] = screening	478	38.9

[a] Symptoms include rectal bleeding, lower abdominal pain, unintentional weight loss, and change in bowel habits in the past 6 months, as well as anemia in the past 12 months.
[b] Gastrointestinal conditions include colorectal cancer, colorectal polyps, inflammatory bowel disease, and previous bowel surgery.
[c] Fecal occult blood test.
[d] Patient indication 1 is non-screening when patient perceived reason for the colonoscopy was to follow up on a problem or a previous test, and is screening otherwise.
[e] Patient indication 2 is non-screening when patient reports specific symptoms, personal history of gastrointestinal conditions, or recent positive FOBT, and is screening otherwise.

Table 2 Frequency of occurrence of diagnostic and procedure codes in provincial administrative databases (N = 1,230)

Diagnostic or procedure codes	N[a]	%
Procedures in the past 4 years		
Colonoscopy	267	21.7
Sigmoidoscopy	76	6.2
Polypectomy	114	9.3
Double contrast barium enema	65	5.3
Symptoms in the past year		
Rectal bleeding	95	7.7
Diarrhea	50	4.1
Vomiting	7	0.6
Weight loss	8	0.7
Anemia	86	7
Gastrointestinal conditions in the past 5 years		
Colorectal cancer	65	5.3
Colorectal polyps	205	16.7
Inflammatory bowel disease	48	3.9
Hospitalizations in the past 5 years		
Large bowel diseases	32	2.6
Large bowel surgery	26	2.1

[a] Number of patients whose health administrative records contain the relevant diagnostic or procedure code(s). Specific codes are listed in Additional file 1.

Table 3 Odds ratio estimates for the multivariate logistic regression models selected by BIC[a]

Variable	Latent class indication	Endoscopist indication
	OR[b] (95% CI)	OR (95% CI)
Colonoscopy in the past 4 years	0.16 (0.10-0.23)	0.18 (0.12-0.25)
Sigmoidoscopy in the past 4 years	0.29 (0.17-0.45)	0.28 (0.16-0.43)
Polypectomy in the past 4 years	0.24 (0.12-0.42)	0.13 (0.05-0.28)
DCBE[c] in the past 4 years	0.19 (0.11-0.31)	0.19 (0.11-0.31)
Rectal bleeding in the past year	0.15 (0.09-0.25)	0.12 (0.07-0.20)
Diarrhea in the past year	0.20 (0.11-0.34)	0.14 (0.07-0.27)
Anemia in the past year	0.15 (0.09-0.24)	0.19 (0.12-0.30)
IBD[d] in the past 5 years	0.06 (0.02-0.23)	0.09 (0.03-0.24)

[a] Bayesian information criterion.
[b] Odds ratio adjusted for all 7 other variables in the model.
[c] Double contrast barium enema.
[d] Inflammatory bowel disease.

Model-based algorithms

The latent class model predicted 554 (45.0%) screening exams. The Kappa statistic for its agreement with endoscopist indication was 0.794 (95% CI: 0.760-0.828). Allowing conditional dependence between the two patient indications yielded virtually identical results (data not shown).

Using the latent class indication as the outcome, the multivariate logistic regression model that included age, sex, and all administrative data variables yielded an AUC of 0.786 (95% CI: 0.760-0.812) when the logistic model predictions were compared to the latent class indication. The best model selected by BIC had an AUC of 0.754 (95% CI: 0.726-0.782) and included 8 administrative data variables (Table 3). In comparison, multivariate logistic regression using the endoscopist indication alone as the outcome provided an AUC of 0.791 (95% CI: 0.765-0.816) for the full model and 0.761 (95% CI: 0.734-0.788) for the best model selected by BIC. The same 8 variables were selected by the BIC, with similar odds ratio estimates (Table 3).

Recursive partitioning using the latent class indication as the outcome used 7 variables, yielding the classification tree in Figure 1. The sensitivity and specificity, when comparing the classification tree to the latent class indication as the reference standard, were 84.5% and 63.3% respectively (Table 4). Recursive partitioning using the endoscopist indication as the outcome yielded a similar classification tree but used only 6 variables in the following order:: colonoscopy in the past 4 years, rectal bleeding in the past year, DCBE in the past 4 years, diarrhea in the past year, anemia in the past year, and IBD in the past 5 years. Sensitivity was 85.1% and specificity was 62.2% (Table 4).

Expert opinion algorithm

The algorithm developed by El-Serag et al. was applied to our data. The algorithm identified 395 (32.1%) colonoscopies as screening. The sensitivity and specificity were 49.3% and 82.0%, respectively, compared to the latent class indication, and 49.3% and 83.0% compared to the endoscopist indication (Table 4).

Discussion

We evaluated model-based algorithms in the absence of a gold standard measure of the outcome for determining the colonoscopy indication (screening vs. non-screening). We tackled this problem by constructing a latent reference standard and then using it to develop and evaluate logistic regression and recursive partitioning models of administrative data variables. Both modeling approaches have been used in previous studies to identify the indications for of medical procedures [6,10]. The latent class predictions were quite accurate when the various tests all agreed on the indication, i.e. when all tests together indicated either positive or negative for screening. However, when one or more tests disagreed with the others, there was higher variability and less certainty about the inputs. Overall, the stability of the logistic regression model was very good, as evidenced by the robustness of our analyses
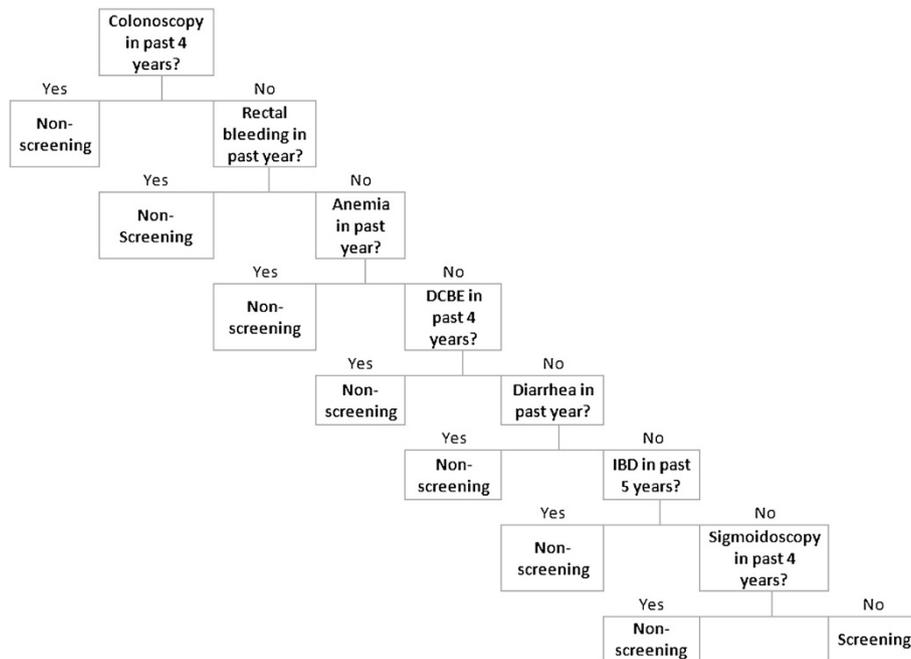
Figure 1 Classification tree for colonoscopy indication generated by recursive partitioning model using latent class predictions as the outcome. Colonoscopy exams were classified as screening or non-screening based on the presence or absence of diagramed diagnostic or procedure codes in patient administrative health records. DCBE: double contrast barium enema. IBD: inflammatory bowel disease.

(using a second latent class model that gave very similar predictions).

From multivariate logistic models, the AUC was 0.786 and 0.791 for the full models fitted with latent and endoscopist indications as the outcomes, respectively. The greater than 20% chance of mistakenly ranking a non-screening exam as more likely to be screening than a screening exam is not sufficiently accurate for most research purposes. The recursive models also did not achieve sufficient accuracy, despite their propensity to overfit data. Compared to the expert opinion algorithm, the recursive models had higher sensitivity but lower specificity; this occurred largely because the El-Serag

algorithm defined screening as the absence of 28 diagnosis and colonoscopy procedure in the past 4 years, whereas recursive partitioning chose only the most discriminating variables. However, direct comparisons between the model-based and the El-Serag algorithms should be done with caution, as the models have the advantage of being validated on the same data upon which they were constructed.

Since we had two datasets available, the prudent approach would have been to construct the models using one dataset and validate them in another. However, our intention was to show that even under optimal circumstances – using all the data to generate the models and

Table 4 Accuracy measures for recursive partitioning and expert opinion algorithms

Algorithm	Reference standard	Sensitivity% (95% CI)	Specificity% (95% CI)	PPV [a]% (95% CI)	NPV [b]% (95% CI)
Recursive partitioning with latent class outcome	Latent class indication	84.5	63.3	65.4	83.3
		(81.5-87.5)	(59.7-67.0)	(61.9-68.9)	(80.0-86.5)
Recursive partitioning with endoscopist outcome	Endoscopist indication	85.1	62.2	66.5	82.6
		(82.2-88.0)	(58.5-66.0)	(63.1-70.0)	(79.2-85.9)
El-Serag	Latent class indication	49.2	82.0	69.1	66.4
		(45.1-53.4)	(79.1-84.9)	(64.6-73.7)	(63.3-69.6)
El-Serag	Endoscopist indication	49.3	83.0	71.9	65.0
		(45.2-53.4)	(80.2-85.9)	(67.5-76.3)	(61.8-68.3)

[a] Positive predicative value.
[b] Negative predictive value.

evaluating with the same dataset – model accuracies were less than satisfactory. Not pruning the classification tree, which tends to overfit the data in the recursive model [33,34], also did not result in more accurate predictions.

Both the latent class and the endoscopist indications produced very similar results in most cases, since there was relatively high agreement between them. The latent class predictions were likely driven more by the endoscopist indication than patient indications, given the informative prior distribution used for the endoscopist indication, while uninformative priors were used for patient indications. The assumption that physicians know the true indication at least 70% of the time seems conservative.

The poor performance of algorithms, whether model-based or expert-opinion based, may be due to imperfect accuracies of administrative codes for the predictor variables we used [4] or to the variability in physician clinical and billing practices [35]. The polypectomy procedure code in Quebec, for example, underestimates the number of polypectomies by 15% [1]. Since the primary purpose of health administrative data collection is remuneration, diagnostic codes may be poorly recorded [35], leading to misclassification. Model selection in the multivariate logistic regression analysis retained all 4 procedure variables (colonoscopy, sigmoidoscopy, polypectomy, DCBE in the past 4 years) as important predictors of indication, while it retained only 4 of 8 diagnostic variables from physician billing data. Procedure codes may have been better predictors than some diagnostic codes because they were recorded with greater accuracy due to the need for remuneration [35]. CRC diagnosis and colorectal polyps were not identified as useful predictors by either logistic regression model selection procedures or recursive partitioning, possibly due to overlapping information with other variables or poor accuracy. Hospitalizations for large bowel disease and large bowel surgeries were also not selected as important predictors, possibly due to the small numbers of patients whose records contained these codes.

Algorithms are typically needed to accurately identify cases of IBD because of problems with misclassification in administrative data [9,36]. We did not use such an algorithm for IBD as the purpose was not to correctly identify IBD but to evaluate the utility of administrative codes (including those for IBD) in predicting colonoscopy indication. Despite the relatively low frequency of occurrence of IBD codes, the variable emerged as an important predictor in all models.

Conclusions
In conclusion, model-based screening colonoscopy algorithms for administrative databases were insufficiently accurate to be used for most research purposes. However the novel approach that we have employed, constructing a latent reference standard against which model-based algorithms were evaluated, may be useful for validating administrative data in other contexts where gold standards are unavailable.

Additional file

> **Additional file 1: Diagnostic and procedure codes in health administrative databases.**

Authors' contributions
MJS conceived of the study, participated in the study design and coordination, and helped draft the manuscript. MJ participated in the statistical analysis and drafting of the manuscript. LJ participated in the statistical analysis and in the drafting of the manuscript. RJH participated in the study design and with acquisition of the data. AB participated in the study design and with acquisition of the data. All authors contributed to the interpretation of the findings, and read and approved the final manuscript.

Author details
[1]Department of Medicine, McGill University, Montreal, Quebec, Canada. [2]Division of Clinical Epidemiology, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada. [3]Division of Gastroenterology, McGill University Health Centre, Montreal, Quebec, Canada. [4]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada. [5]Departments of Medicine/Community Health Sciences, University of Calgary, Calgary, Alberta, Canada.

References
1. Wyse JM, Joseph L, Barkun AN, Sewitch MJ: **Accuracy of administrative claims data for polypectomy.** *CMAJ* 2011, **183**(11):E743–E747.
2. Lee DS, Donovan L, Austin PC, Gong Y, Liu PP, Rouleau JL, Tu JV: **Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research.** *Med Care* 2005, **43**(2):182–188.
3. Januel JM, Luthi JC, Quan H, Borst F, Taffe P, Ghali WA, Burnand B: **Improved accuracy of co-morbidity coding over time after the introduction of ICD-10 administrative data.** *BMC Health Serv Res* 2011, **11**:194.
4. Wilchesky M, Tamblyn RM, Huang A: **Validation of diagnostic codes within medical services claims.** *J Clin Epidemiol* 2004, **57**(2):131–141.
5. Randolph WM, Mahnken JD, Goodwin JS, Freeman JL: **Using Medicare data to estimate the prevalence of breast cancer screening in older women: comparison of different methods to identify screening mammograms.** *Health Serv Res* 2002, **37**(6):1643–1657.
6. Gregory KD, Korst LM, Gornbein JA, Platt LD: **Using administrative data to identify indications for elective primary cesarean delivery.** *Health Serv Res* 2002, **37**(5):1387–1401.
7. Fisher DA, Grubber JM, Castor JM, Coffman CJ: **Ascertainment of colonoscopy indication using administrative data.** *Dig Dis Sci* 2010, **55**(6):1721–1725.
8. Myers RP, Shaheen AA, Fong A, Wan AF, Swain MG, Hilsden RJ, Sutherland L, Quan H: **Validation of coding algorithms for the identification of patients with primary biliary cirrhosis using administrative data.** *Can J Gastroenterol* 2010, **24**(3):175–182.
9. Benchimol EI, Guttmann A, Griffiths AM, Rabeneck L, Mack DR, Brill H, Howard J, Guan J, To T: Increasing incidence of paediatric inflammatory

bowel disease in Ontario, Canada: evidence from health administrative data. *Gut* 2009, **58**(11):1490–1497.

10. Richardson P, Henderson L, Davila JA, Kramer JR, Fitton CP, Chen GJ, El-Serag HB: **Surveillance for hepatocellular carcinoma: development and validation of an algorithm to classify tests in administrative and laboratory data.** *Dig Dis Sci* 2010, **55**(11):3241–3251.

11. Korst LM, Gregory KD, Gornbein JA: **Elective primary caesarean delivery: accuracy of administrative data.** *Paediatric perinatal epidemiol* 2004, **18**(2):112–119.

12. El-Serag HB, Petersen L, Hampel H, Richardson P, Cooper G: **The use of screening colonoscopy for patients cared for by the Department of Veterans Affairs.** *Arch Intern Med* 2006, **166**(20):2202–2208.

13. Haque R, Chiu V, Mehta KR, Geiger AM: **An automated data algorithm to distinguish screening and diagnostic colorectal cancer endoscopy exams.** *J Nat Cancer Inst Mono* 2005, **35**:116–118.

14. Leddin D, Bridges RJ, Morgan DG, Fallone C, Render C, Plourde V, Gray J, Switzer C, McHattie J, Singh H, et al: **Survey of access to gastroenterology in Canada: the SAGE wait times program.** *Can J Gastroenterol* 2010, **24**(1):20–25.

15. Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, Dash C, Giardiello FM, Glick S, Levin TR, et al: **Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology.** *CA Cancer J Clin* 2008, **58**(3):130–160.

16. Rex DK, Johnson DA, Anderson JC, Schoenfeld PS, Burke CA, Inadomi JM: **American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected].** *Am J Gastroenterol* 2009, **104**(3):739–750.

17. Segnan N, Patnick J, von Karsa L: *European guidelines for quality assurance in colorectal cancer screening and diagnosis.* Luxembourg: Publications Office of the European Union; 2010.

18. Leddin DJ, Enns R, Hilsden R, Plourde V, Rabeneck L, Sadowski DC, Signh H: **Canadian Association of Gastroenterology position statement on screening individuals at average risk for developing colorectal cancer: 2010.** *Can J Gastroenterol* 2010, **24**(12):705–714.

19. **Inventory of Colorectal Cancer Screening Activities in ICSN Countries.** [http://appliedresearch.cancer.gov/icsn/colorectal/screening.html (Accessed 4 April 2013)].

20. Harewood GC, Lieberman DA: **Colonoscopy practice patterns since introduction of medicare coverage for average-risk screening.** *Clin Gastroenterol Hepatol* 2004, **2**(1):72–77.

21. White A, Vernon SW, Franzini L, Du XL: **Racial and ethnic disparities in colorectal cancer screening persisted despite expansion of Medicare's screening reimbursement.** *Cancer Epidem Biomar* 2011, **20**(5):811–817.

22. Rex DK, Petrini JL, Baron TH, Chak A, Cohen J, Deal SE, Hoffman B, Jacobson BC, Mergener K, Petersen BT, et al: **Quality indicators for colonoscopy.** *Am J Gastroenterol* 2006, **101**(4):873–885.

23. Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L: **Association of colonoscopy and death from colorectal cancer.** *Ann Intern Med* 2009, **150**(1):1–8.

24. Fletcher RH, Nadel MR, Allen JI, Dominitz JA, Faigel DO, Johnson DA, Lane DS, Lieberman D, Pope JB, Potter MB, et al: **The quality of colonoscopy services–responsibilities of referring clinicians: a consensus statement of the Quality Assurance Task Group, National Colorectal Cancer Roundtable.** *J Gen Intern Med* 2010, **25**(11):1230–1234.

25. Rex DK, Bond JH, Winawer S, Levin TR, Burt RW, Johnson DA, Kirk LM, Litlin S, Lieberman DA, Waye JD, et al: **Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the U.S. Multi-Society Task Force on Colorectal Cancer.** *Am J Gastroenterol* 2002, **97**(6):1296–1308.

26. Lieberman D: **Pitfalls of using administrative data for research.** *Dig Dis Sci* 2010, **55**(6):1506–1508.

27. Sewitch MJ, Stein D, Joseph L, Bitton A, Hilsden RJ, Rabeneck L, Paszat L, Tinmouth J, Cooper MA: **Comparing patient and endoscopist perceptions of the colonoscopy indication.** *Can J Gastroenterol* 2010, **24**(11):656–660.

28. Joseph L, Gyorkos TW, Coupal L: **Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard.** *Am J Epidemiol* 1995, **141**(3):263–272.

29. Dendukuri N, Joseph L: **Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests.** *Biometrics* 2001, **57**(1):158–167.

30. Kass RE, Raftery AE: **Bayes Factors.** *J Am Stat Assoc* 1995, **90**:773–795.

31. Hanley JA, McNeil BJ: **The meaning and use of the area under a receiver operating characteristic (ROC) curve.** *Radiology* 1982, **143**(1):29–36.

32. R Development Core Team: *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2005.

33. Breiman L, Friedman JH, Olshen RA, Stone CJ: *Classification and Regression Trees.* Belmont CA: Wadsworth International Group; 1984.

34. Strobl C, Malley J, Tutz G: **An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests.** *Psychol Methods* 2009, **14**(4):323–348.

35. Sewitch MJ, Hilsden R, Joseph L, Rabeneck L, Paszat L, Bitton A, Cooper M-A: **Qualitative study of physician perspectives on classifying screening and non-screening colonoscopy using administrative health data: adding practice does not make perfect.** *Can J Gastroenterol* 2012, **26**:889–893.

36. Bernstein CN, Blanchard JF, Rawsthorne P, Wajda A: **Epidemiology of Crohn's disease and ulcerative colitis in a central Canadian province: a population-based study.** *Am J Epidemiol* 1999, **149**(10):916–924.