Research

Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: a cross-sectional study

Jill Tinmouth MD PhD, Rinku Sutradhar PhD, Ning Liu MSc, Nancy N. Baxter MD PhD, Lawrence Paszat MD MS, Linda Rabeneck MD MPH

Abstract

Background: Colonoscopy is used widely, but its quality is highly variable, which may adversely affect patients. Research and quality-improvement initiatives in a variety of jurisdictions have sought to address this issue, often supported by the use of health administrative data. As these data are generally not collected for these purposes, it is critical to measure their validity before use. The aim of this study was to validate health administrative data definitions for 5 key colonoscopy elements through comparison to the clinical record.

Methods: In a cross-sectional study, we randomly sampled 1968 colonoscopy and noncolonoscopy procedures performed at 23 hospitals and 5 nonhospital endoscopy clinics between April 2008 and March 2009 in Ontario. We compared definitions for 5 key colonoscopy elements (colonoscopy case, colonoscopy setting, colonoscopy completeness, anesthesiologist assistance and polypectomy) derived from the health administrative data to the clinical record. We calculated weighted and unweighted sensitivity, specificity and positive predictive value, adjusted for clustering of patients within physicians, for each definition relative to the reference standard.

Results: We abstracted 1845 records; in 1282 cases (69.5%), colonoscopy was intended or performed. The weighted sensitivity and specificity of colonoscopy case, nonhospital colonoscopy setting and anesthesiologist assistance exceeded 95%. The weighted sensitivity for colonoscopy completeness and polypectomy exceeded 95%, but specificity was less than 90%.

Interpretation: Ontario health administrative data definitions for 5 key colonoscopy data elements performed well, with sensitivity and specificity values acceptable for use in research and quality-improvement initiatives. In jurisdictions where health administrative data are similarly used for research or quality improvement, similar studies could be considered.

Colonoscopy is common, essential for the management of gastrointestinal diseases. The procedure has important risks, including perforation (0.85/1000), bleeding (1.64/1000) and death (0.074/1000).[1] Furthermore, there is considerable variation in the performance and quality of colonoscopy,[2,3] with attendant consequences for the health care system and patients. For example, although colonoscopy can be safely and comfortably performed with moderate sedation administered by the endoscopist, increasingly in some jurisdictions, anesthesiologists are providing deep sedation for the procedure;[4-6] however, there are added costs[6] and potentially complications[7,8] associated with this practice. Patients are adversely affected by poor-quality colonoscopy; such procedures are associated with missed colorectal cancers and cancer-related death.[9]

Initiatives related to colonoscopy practice standards,[10,11] research[12,13] and policy[14,15] have been implemented to improve quality. The use of routinely collected data[16] such as health administrative data for these purposes is highly attractive as they are inexpensive to use and readily available and can be captured uniformly across a health care payer or system. These data have been used to study and measure the quality of

colonoscopy, including completeness,[17] type of setting,[4,18] polypectomy rate/adenoma detection rate,[9,19] anesthesiologist assistance,[4–8] complications[1,8] and missed cancers.[9,18] Health administrative data are also used for funding and accountability and to measure performance and quality of care by government agencies.[20] As these data are often collected for other reasons,[21,22] it is critical to ensure their validity when using them for these purposes. In other jurisdictions, colonoscopy validation studies have generally focused on procedure indication;[23–25] other aspects of colonoscopy have not been widely evaluated. In Ontario, the Institute for Clinical Evaluative Sciences houses health administrative databases containing the health care records for the population of Ontario. These databases have been used extensively in colonoscopy research[1,4,7,17] and quality improvement,[26] but, to date, they have not been validated. The objective of this study was to validate health administrative data definitions for 5 key colonoscopy elements: colonoscopy case, colonoscopy setting, colonoscopy completeness, anesthesiologist assistance and polypectomy.

Methods

Overview

In this multisite chart abstraction study, we created health administrative definitions of 5 colonoscopy data elements: a "colonoscopy case," colonoscopy setting, colonoscopy completeness, anesthesiologist assistance and polypectomy. We compared these definitions to reference standards: clinical data obtained via chart abstraction at 23 hospitals and 5 non-hospital endoscopy clinics in Ontario. For some data elements, we included more than 1 health administrative definition and/or more than 1 reference standard (see below and Table 1 for detailed descriptions). The clinical data largely comprised medical records of complete and incomplete colonoscopy procedures but also included gastroscopy and flexible sigmoidoscopy procedures to allow estimation of the true-negative rate for the case definition of colonoscopy.

Sources of administrative data

We used the Ontario Health Insurance Plan (OHIP) database and the Canadian Institute for Health Information databases housed at the Institute for Clinical Evaluative Sciences. The OHIP database contains physician billing data on inpatient and outpatient visits and procedures including colonoscopy since 1991. The Canadian Institute for Health Information databases comprise diagnosis and procedure (both inpatient [Discharge Abstract Database] and same-day [Same Day Surgery database]) codes for all hospital admissions in Canada since 1988. These databases were linked at the individual level by means of an encrypted version of the provincial health card number.

Hospital and nonhospital endoscopy clinic sites

We randomly selected 23 hospital and 5 nonhospital facilities in Ontario to participate in the study (Figure 2). The hospital sites were divided into 4 strata, and from those that performed more than 200 colonoscopy procedures in the prior year (n = 106/115), we selected a random sample from each stratum as

follows: rural (n = 2), urban and lowest tertile of annual colonoscopy volume (n = 8), urban and middle volume tertile (n = 7), and urban and highest volume tertile (n = 6). With the assistance of endoscopy equipment manufacturers, we identified 34 nonhospital clinics active during the study period. We randomly selected 5 clinics from among those who performed more than 200 colonoscopy procedures in the prior year and indicated willingness to participate in the study (n = 21/34).

Medical record abstraction

We identified all outpatient procedure visits for gastroscopy, flexible sigmoidoscopy and colonoscopy at the 28 facilities from Apr. 1, 2008, to Mar. 31, 2009, using health administrative data (hospitals) or self-reported billing data (clinics). These visits constituted the sampling frame (Figure 2). We used OHIP codes to identify procedures as this approach allowed identification of procedures performed outside of hospitals as well as the stratified sampling for colonoscopy procedures described below. Colonoscopy procedures were identified with the base code (Z555A), which was common to all colonoscopy procedures. From the 144 078 procedures in the sampling frame, we then randomly selected 1968 medical records for abstraction using a stratified sampling strategy.[27] Briefly, the procedures in the sampling frame were divided into 6 strata based on billing data (complete colonoscopy, incomplete colonoscopy to descending colon, incomplete colonoscopy to splenic flexure, incomplete colonoscopy to hepatic flexure, gastroscopy, flexible sigmoidoscopy). We then randomly sampled procedures in unique patients from each stratum, aiming for our intended sample sizes as described below. We deliberately oversampled, expecting missing charts at the time of abstraction. The final sample included complete colonoscopy procedures (n = 794), incomplete colonoscopy procedures (n = 806), gastroscopy procedures (n = 128) and flexible sigmoidoscopy procedures (n = 240). Our sampling strategy ensured adequate inclusion of less common events, such as incomplete colonoscopy, while maintaining a sample size feasible for medical record abstraction.

At the 28 sites, trained medical record abstractors from the Institute for Clinical Evaluative Sciences, blinded to the administrative data, abstracted from 3 prespecified sources (the endoscopist's procedure note, the anesthesiologist's record and the pathologist's record) using a standardized data collection protocol. The abstractors collected the data in a standardized fashion using a customized data collection platform residing on encrypted laptop computers. The data were then transmitted from each site to the Institute for Clinical Evaluative Sciences via a secure virtual private network. If the endoscopist's procedure note could not be found at the sites or if the dates for the procedure found in the medical record at abstraction did not agree with those in the administrative data, the case was excluded.

Cohort creation

We created 3 cohorts for the analyses. The first cohort comprised all successfully abstracted medical records and was used to assess the health administrative data definition for

Table 1: Description of the cohorts, administrative data definitions and reference standards for 5 colonoscopy data elements

Element	Cohort description, size	Administrative data definition	Reference standard
Colonoscopy case	All successfully abstracted charts, *n* = 1845	OHIP codes: Z555A alone or in combination with any of: E740A, E741A, E747A or E705A 14 of the most clinically plausible combinations were evaluated (see Figure 1 for the specific codes included in each definition)	Colonoscopy was performed, or there was intent to perform colonoscopy according to endoscopist's procedure note*
Nonhospital clinic	Charts in which colonoscopy was intended or performed, *n* = 1282	1. OHIP code E649A billed on date of colonoscopy 2. No record in CIHI database overlapping with date of colonoscopy according to OHIP database (i.e., no record of procedure's being done in hospital) 3. OHIP code E649A and no overlapping record in CIHI database	Presence of endoscopist's procedure note in nonhospital facility chart
Anesthesiologist assistance	Charts in which colonoscopy was intended or performed, *n* = 1282	OHIP codes for anesthesia (003C or procedure code with "C" suffix [see supplementary tables, Appendix 1, available at www.cmajopen.ca/content/6/3/E330/suppl/DC1]) billed on date as colonoscopy in same patient	1. Presence of anesthesiologist record in chart regardless of type of sedating agent 2. Use of propofol as sedating agent according to anesthesiologist's record†
Colonoscopy completeness‡	Charts in which procedure billed with colonoscopy codes, and colonoscopy or flexible sigmoidoscopy was intended,§ *n* = 1477 (administrative data definition 1), *n* = 1016 (administrative data definition 2)	1. OHIP code E747A (to cecum) or E705A (to terminal ileum) billed among colonoscopy procedures defined using most sensitive definition (Z555A ± other E codes) 2. OHIP code E747A or E705A billed among colonoscopy procedures defined using most accurate definition (Z555A + E741 ± other E codes)	Colonoscopy "intended" and "complete" according to endoscopist's procedure note
Polypectomy¶	Charts in which colonoscopy was intended or performed,** *n* = 1256 (reference standard 1), *n* = 1252 (reference standard 2)	1. OHIP code Z571A alone 2. OHIP code Z571A, Z570A or E685A 3. OHIP code Z571A, Z570A, E685A or E717A	1. Polyp visualized or polypectomy described according to endoscopist's procedure note 2. Adenoma, advanced adenoma or sessile serrated polyp according to pathologist's report††

Note: CIHI = Canadian Institute for Health Information, OHIP = Ontario Health Insurance Plan.
*Completed by the endoscopist; includes a description of the procedure, including findings.
†Completed by the anesthesiologist; record of anesthetic administered during the procedure.
‡Procedures intended as flexible sigmoidoscopy for which E747A or E705A was billed were classified as false-positive. Procedures intended as flexible sigmoidoscopy for which E747A and E705A were not billed were classified as false-negative.
§Excluding those with prior total colectomy or right hemicolectomy.
¶Because the histologic findings of the polyp are not available in administrative databases, we could not define adenoma using these data.
**Excluding those with missing data for reference standard.
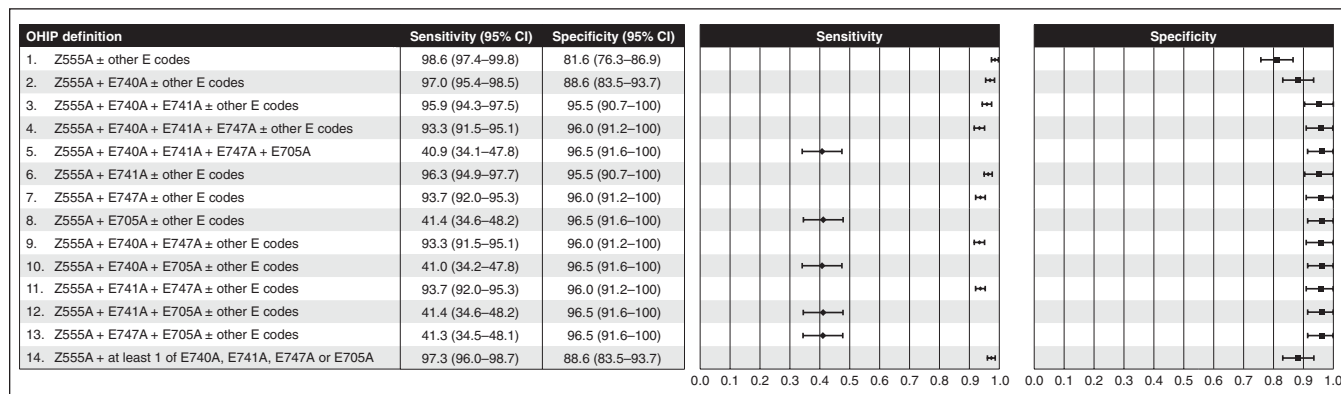††Report on the histologic findings of specimens, such as polyps, obtained at colonoscopy.

OHIP definition	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity	Specificity
1. Z555A ± other E codes	98.6 (97.4–99.8)	81.6 (76.3–86.9)		
2. Z555A + E740A ± other E codes	97.0 (95.4–98.5)	88.6 (83.5–93.7)		
3. Z555A + E740A + E741A ± other E codes	95.9 (94.3–97.5)	95.5 (90.7–100)		
4. Z555A + E740A + E741A + E747A ± other E codes	93.3 (91.5–95.1)	96.0 (91.2–100)		
5. Z555A + E740A + E741A + E747A + E705A	40.9 (34.1–47.8)	96.5 (91.6–100)		
6. Z555A + E741A ± other E codes	96.3 (94.9–97.7)	95.5 (90.7–100)		
7. Z555A + E747A ± other E codes	93.7 (92.0–95.3)	96.0 (91.2–100)		
8. Z555A + E705A ± other E codes	41.4 (34.6–48.2)	96.5 (91.6–100)		
9. Z555A + E740A + E747A ± other E codes	93.3 (91.5–95.1)	96.0 (91.2–100)		
10. Z555A + E740A + E705A ± other E codes	41.0 (34.2–47.8)	96.5 (91.6–100)		
11. Z555A + E741A + E747A ± other E codes	93.7 (92.0–95.3)	96.0 (91.2–100)		
12. Z555A + E741A + E705A ± other E codes	41.4 (34.6–48.2)	96.5 (91.6–100)		
13. Z555A + E747A + E705A ± other E codes	41.3 (34.5–48.1)	96.5 (91.6–100)		
14. Z555A + at least 1 of E740A, E741A, E747A or E705A	97.3 (96.0–98.7)	88.6 (83.5–93.7)		

Figure 1: Weighted sensitivity and specificity of 14 administrative data definitions using Ontario Health Insurance Plan (OHIP) codes of colonoscopy case compared to the reference standard of colonoscopy intended or performed according to the medical record. Note: CI = confidence interval.

colonoscopy case. The second cohort comprised medical records in which colonoscopy or flexible sigmoidoscopy was intended and colonoscopy codes were billed; we used this cohort to evaluate the definition for colonoscopy completeness. Ontario physicians bill colonoscopy per segment reached; this data structure allows measurement of colonoscopy completeness.[17] Although there is a separate code for flexible sigmoidoscopy, anecdotally, colonoscopy codes are often used to bill this procedure as remuneration is better. As this practice may result in misclassification of incomplete colonoscopy procedures when Ontario administrative data are used, we included flexible sigmoidoscopy procedures billed with colonoscopy codes in the cohort. We excluded patients with prior total colectomy or right hemicolectomy, as these patients are typically excluded when administrative data are used to measure colonoscopy completion. Finally, the third cohort comprised medical charts in which colonoscopy was intended or performed and was used to test the remaining 3 colonoscopy data elements (colonoscopy setting, anesthesiologist assistance and polypectomy).

Administrative data definitions and reference standards for colonoscopy data elements

Descriptions of administrative data definitions and reference standards for each data element are given in Table 1.

There were 14 alternative definitions for colonoscopy case because of the structure of OHIP colonoscopy codes: a base code (Z555A) must be used indicating that the scope was inserted to the level of the descending colon. Up to 4 additional "E" codes are then used for every additional segment of colon or terminal ileum visualized. Figure 1 provides a listing of the specific OHIP codes used in each definition. We evaluated colonoscopy completeness using 2 definitions of colonoscopy case: achieving the cecum or terminal ileum among 1) colonoscopy procedures identified using the most sensitive definition and 2) those identified using the most accurate definition, based on the analyses described below.

We compared health administrative data definitions to relevant reference standards obtained via chart abstraction by the chart abstractors. The reference standard for colonoscopy case was the intent to perform or performance of a colonoscopy according to the endoscopist's procedure note in the medical record at the institution where the colonoscopy was performed. The "intention to perform a colonoscopy" criterion allows the distinction of an incomplete colonoscopy from a flexible sigmoidoscopy. There were 2 reference standards for anesthesiologist assistance (presence of anesthesiologist record, indicating an anesthesiologist attended the procedure,
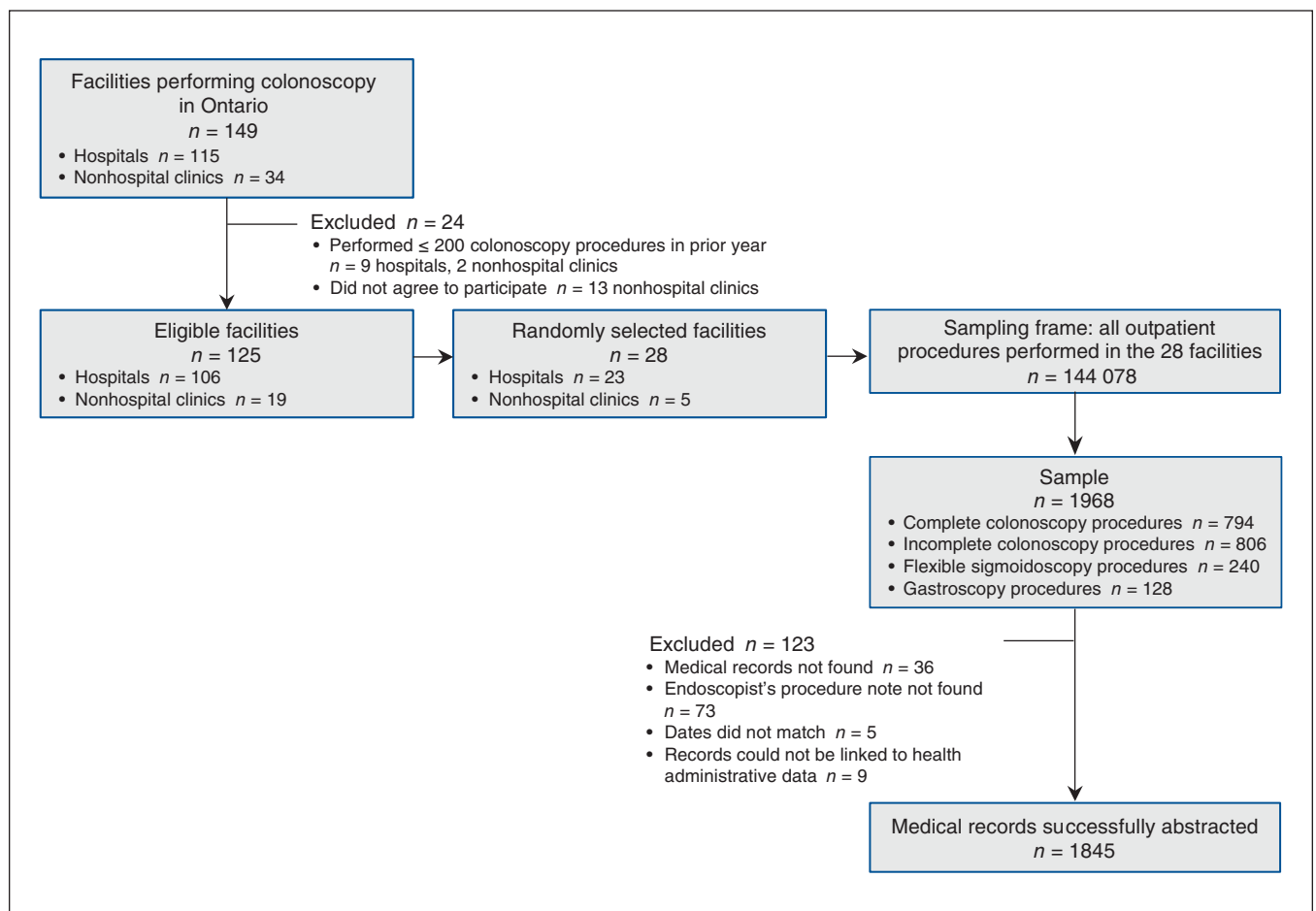

Figure 2: Flow chart showing sampling of facilities and procedures.

and propofol use documented in anesthetic record or endoscopist's procedure note) and for polypectomy (polyp documented by endoscopist and histologic confirmation of adenoma, including advanced adenoma, or sessile serrated adenoma/polyp documented). We included 2 reference standards for these data elements as anesthesiologist assistance and polypectomy are often used as surrogates for the use of propofol and adenomas, respectively. Reference standards were not validated; however, the medical record is generally considered an acceptable reference standard for studies validating health administrative data.[28] The specific OHIP billing codes used and their definitions are given in Supplementary Tables A1–A3, Appendix 1.

Sample size

We performed a sample size calculation a priori for the key data elements of colonoscopy case and colonoscopy completeness. Using methods for 2-sided binomial tests,[29] we estimated that we would need at least 600 complete colonoscopy cases, 600 incomplete colonoscopy cases and 300 noncolonoscopy cases to have over 80% power ($\alpha = 0.05$) to detect an absolute difference of at least a 4% in coding accuracy (assuming a coding accuracy proportion, determined by comparing the medical record data with the administrative data, of 0.85) within each group.

Statistical analysis

We calculated sensitivity and specificity with 95% confidence intervals for each administrative data definition relative to the relevant reference standard. We adjusted the 95% confidence limits for clustering of patients within physicians using the Taylor series expansion method.[30] Because we oversampled incomplete colonoscopy procedures, we performed both unweighted and weighted analyses, where the weights reflected the distribution of procedures in the sampling frame relative to those in the sample. Weighted results are presented unless there was important variation between weighted and unweighted results. For colonoscopy case, we created a receiver operating curve by plotting the sensitivity (on the y-axis) and 1-specificity (on the x-axis) for the 14 definitions. We defined the definition located in the upper left-hand corner of this curve as the most accurate as per the Youden method.[31] We performed all sampling and analyses using SAS version 9.3 (SAS Institute, Inc.).

Ethics approval

Ethics approval was obtained from the Sunnybrook Health Sciences Centre Research Ethics Board as well as from the research ethics boards at the 23 hospitals where chart abstraction was performed.

Results

Among the 1968 randomly selected endoscopy procedures, 1845 charts (93.8%) were successfully abstracted (see Figure 2 for exclusions). Colonoscopy was intended or performed in

1282 (69.5%) of the 1845 cases. About 25% of patients undergoing the included procedures were less than age 50 years, and over half were women; 284 procedures (15.4%) were performed in nonhospital endoscopy clinics (Table 2). Seven to 153 charts were abstracted per hospital, and 31–86 charts were abstracted per clinic.

Table 2: Patient and procedure characteristics for all successfully abstracted charts and for charts in which colonoscopy was intended or performed

Characteristic	No. (%) of charts*	
	All abstracted charts $n = 1845$	Colonoscopy intended or performed $n = 1282$
Age group, yr		
< 50	469 (25.4)	251 (19.6)
50–59	517 (28.0)	395 (30.8)
60–69	430 (23.3)	327 (25.5)
70–74	174 (9.4)	134 (10.4)
> 74	255 (13.8)	175 (13.6)
Sex		
Female	986 (53.4)	709 (55.3)
Male	859 (46.6)	573 (44.7)
Procedure(s) performed†		
Colonoscopy only	1143 (62.0)	1125 (87.8)
Gastroscopy only	45 (2.4)	≤ 5
Flexible sigmoidoscopy only	432 (23.4)	≤ 5
Colonoscopy + gastroscopy	200 (10.8)	151 (11.8)
Flexible sigmoidoscopy + gastroscopy	12 (0.6)	≤ 5
Colonoscopy + other procedure	≤ 5	≤ 5
Flexible sigmoidoscopy + other procedure	≤ 5	0 (0)
Other procedure only	6 (0.3)	0 (0)
Median neighbourhood income quintile‡		
1 (lowest)	300 (16.3)	205 (16.0)
2	331 (17.9)	225 (17.6)
3	337 (18.3)	233 (18.2)
4	393 (21.3)	265 (20.7)
5 (highest)	474 (25.7)	348 (27.1)
Missing	10 (0.5)	6 (0.5)
Setting		
Nonhospital clinic	284 (15.4)	216 (16.8)
Hospital	1561 (84.6)	1066 (83.2)

*To ensure confidentiality, counts of 5 or less are suppressed.
†Based on findings at chart abstraction.
‡Median annual neighborhood household income at the level of enumeration area, obtained from Statistics Canada, was linked to patient postal code.

Colonoscopy case

We evaluated 14 definitions of a colonoscopy case (Figure 1, Figure 3). The most sensitive definition was the base colonoscopy code, Z555A, with or without 1 or more additional E codes. However, this definition was also the least specific. The most accurate definition was the base code plus the code "to hepatic flexure" (E741A), with or without additional E codes. Positive predictive values exceeded 95% for all definitions (Table 3).

Nonhospital clinics

All 3 administrative data definitions for nonhospital clinic setting were found to have sensitivity, specificity and positive predictive values in excess of 95% (Figure 4, Table 3). The first 2 definitions applied criteria using OHIP codes alone or Canadian Institute for Health Information codes alone; when used simultaneously, there was minor loss of sensitivity.

Colonoscopy completeness

The sensitivity and specificity for colonoscopy completeness differed depending on the definition of colonoscopy case and whether they were weighted or unweighted (Figure 4). Regardless of the colonoscopy case definition, the weighted sensitivity exceeded 95% and the weighted specificity was poor (< 80%). All unweighted estimates were about 95% or higher, with 1 exception: the unweighted sensitivity using the most sensitive colonoscopy definition was 70.2%. Positive predictive values exceeded 95% for all definitions (Table 3).

Anesthesiologist assistance

The sensitivity and specificity of the administrative data definition for anesthesiologist-assisted colonoscopy exceeded 95% with either presence of anesthesiologist record or the use of propofol as a sedating agent (Figure 4) as the reference standard. Positive predictive values were less robust (83.8% and 77.2%, respectively) (Table 3).
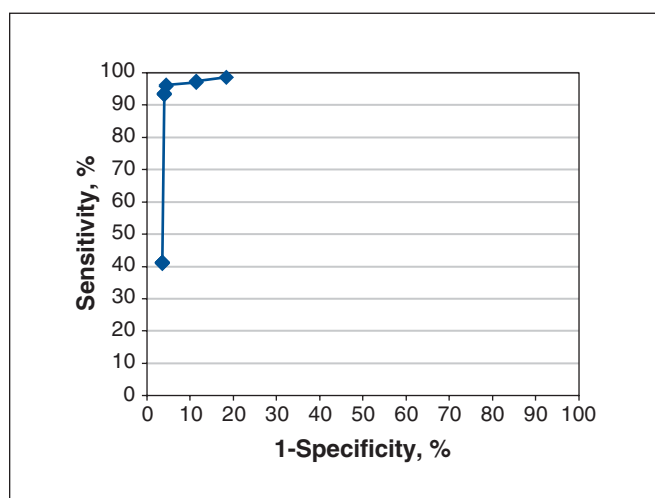

Figure 3: Receiver operating curve of the 14 definitions for colonoscopy case. Upper left point of curve corresponds to Ontario Health Insurance Plan (OHIP) definition 6: OHIP codes Z555A + E741A ± other E codes.

Table 3: Positive predictive values of OHIP definition of colonoscopy, nonhospital setting, colonoscopy completeness, anesthesiologist assistance and polypectomy compared to reference standards

OHIP definition	PPV (95% CI)
Colonoscopy case	
1. Z555A ± other E codes	96.4 (95.2–97.6)
2. Z555A + E740A ± other E codes	97.7 (96.6–98.8)
3. Z555A + E740A + E741A ± other E codes	99.1 (98.0–100)
4. Z555A + E740A + E741A + E747A ± other E codes	99.1 (98.1–100)
5. Z555A + E740A + E741A + E747A + E705A	98.3 (96.0–100)
6. Z555A + E741A ± other E codes	99.1 (98.1–100)
7. Z555A + E747A ± other E codes	99.2 (98.1–100)
8. Z555A + E705A ± other E codes	98.3 (96.0–100)
9. Z555A + E740A + E747A ± other E codes	99.1 (98.1–100)
10. Z555A + E740A + E705A ± other E codes	98.3 (96.0–100)
11. Z555A + E741A + E747A ± other E codes	99.2 (98.1–100)
12. Z555A + E741A + E705A ± other E codes	98.3 (96.0–100)
13. Z555A + E747A + E705A ± other E codes	98.3 (96.0–100)
14. Z555A + 1 or more of E740A, E741A, E747A or E705A	97.7 (96.6–98.8)
Nonhospital setting	
1. E649A billed on date of colonoscopy	100 (100–100)
2. No record in CIHI database overlapping with date of colonoscopy according to OHIP database	100 (100–100)
3. E649A and no overlapping record in CIHI database	100 (100–100)
Colonoscopy completeness, defined using most sensitive colonoscopy definition	
1. Weighted analysis results	99.0 (98.3–99.7)
2. Unweighted analysis results	99.0 (98.3–99.7)
Colonoscopy completeness, defined using most accurate colonoscopy definition	
1. Weighted analysis results	99.0 (98.3–99.7)
2. Unweighted analysis results	99.0 (98.3–99.7)
Anesthesiologist assistance	
1. v. "anesthesiologist's record"	83.8 (71.5–96.0)
2. v. "use of propofol"	77.2 (64.3–90.1)
Polypectomy	
v. "polyp seen or removed"	
1. Z571A alone	99.0 (97.7–100)
2. Z571A, Z570A or E685A	98.7 (97.3–100)
3. Z571A, Z570A, E717A or E685A	79.2 (74.3–84.2)
v. "histology"	
1. Z571A alone	68.1 (62.2–74.1)
2. Z571A, Z570A or E685A	64.1 (58.1–70.2)
3. Z571A, Z570A, E717A or E685A	49.7 (44.4–55.0)

Note: CI = confidence interval, CIHI = Canadian Institute for Health Information, OHIP = Ontario Health Insurance Plan, PPV = positive predictive value.
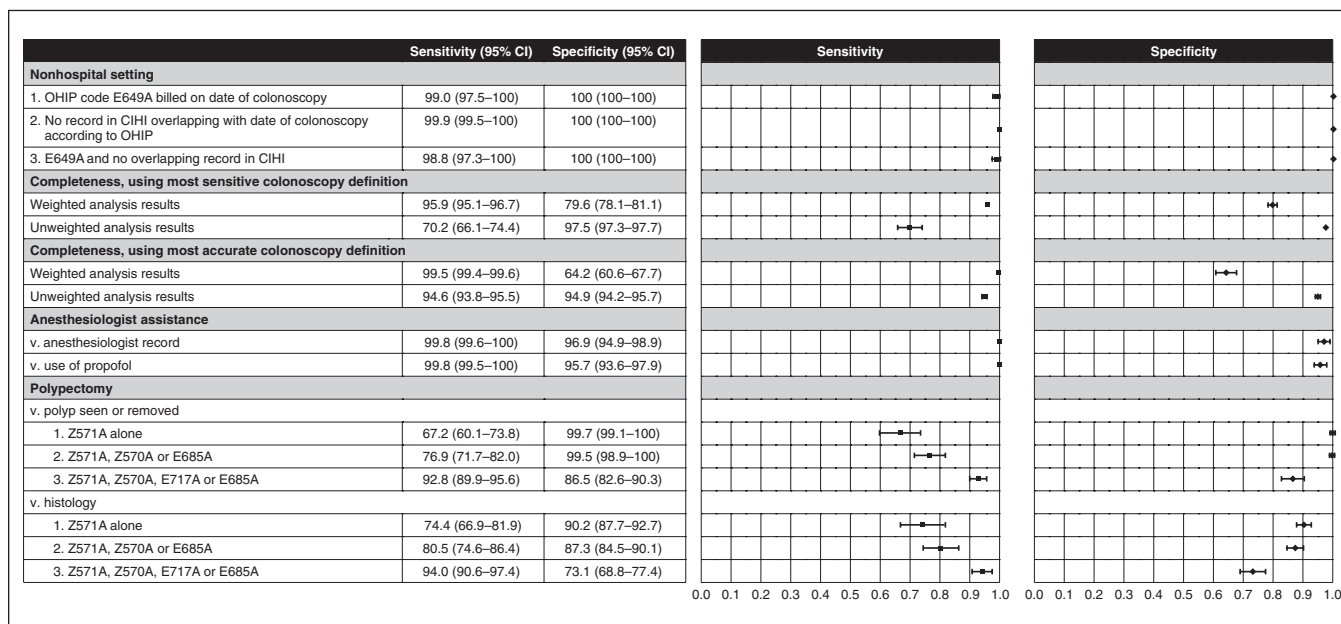
	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity	Specificity
Nonhospital setting				
1. OHIP code E649A billed on date of colonoscopy	99.0 (97.5–100)	100 (100–100)		
2. No record in CIHI overlapping with date of colonoscopy according to OHIP	99.9 (99.5–100)	100 (100–100)		
3. E649A and no overlapping record in CIHI	98.8 (97.3–100)	100 (100–100)		
Completeness, using most sensitive colonoscopy definition				
Weighted analysis results	95.9 (95.1–96.7)	79.6 (78.1–81.1)		
Unweighted analysis results	70.2 (66.1–74.4)	97.5 (97.3–97.7)		
Completeness, using most accurate colonoscopy definition				
Weighted analysis results	99.5 (99.4–99.6)	64.2 (60.6–67.7)		
Unweighted analysis results	94.6 (93.8–95.5)	94.9 (94.2–95.7)		
Anesthesiologist assistance				
v. anesthesiologist record	99.8 (99.6–100)	96.9 (94.9–98.9)		
v. use of propofol	99.8 (99.5–100)	95.7 (93.6–97.9)		
Polypectomy				
v. polyp seen or removed				
1. Z571A alone	67.2 (60.1–73.8)	99.7 (99.1–100)		
2. Z571A, Z570A or E685A	76.9 (71.7–82.0)	99.5 (98.9–100)		
3. Z571A, Z570A, E717A or E685A	92.8 (89.9–95.6)	86.5 (82.6–90.3)		
v. histology				
1. Z571A alone	74.4 (66.9–81.9)	90.2 (87.7–92.7)		
2. Z571A, Z570A or E685A	80.5 (74.6–86.4)	87.3 (84.5–90.1)		
3. Z571A, Z570A, E717A or E685A	94.0 (90.6–97.4)	73.1 (68.8–77.4)		

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0   0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

Figure 4: Sensitivity and specificity of nonhospital setting, colonoscopy completeness, anesthesiologist assistance and polypectomy compared to reference standards. Note: CI = confidence interval, CIHI = Canadian Institute for Health Information, OHIP = Ontario Health Insurance Plan.

Polypectomy

Three administrative data definitions of polypectomy were compared to 2 reference standards, polyp documented by endoscopist and histologic confirmation of adenoma, or sessile serrated adenoma/polyp documented (Figure 4). Using more codes improved sensitivity but worsened specificity. These definitions were more sensitive but less specific when the reference standard was histologic confirmation of adenoma versus polyp seen/removed. Positive predictive values were higher when the reference standard was polyp documented versus histologic confirmation of adenoma (Table 3).

Interpretation

We found that health administrative data definitions of colonoscopy case, colonoscopy setting and anesthesiologist assistance performed well compared to the medical record. The weighted definitions of colonoscopy completeness were sensitive but not specific. The definitions of polypectomy performed less well for the identification of the more clinically relevant reference standard, histologically significant polyps, than they did for polyp documented by endoscopist.

In the current study, the most accurate definition of colonoscopy incorporated codes indicating that the endoscopist reached the hepatic flexure. However, in practice, other definitions may be used depending on context, e.g., for studies on perforation (as the procedure may be aborted if perforation is recognized).[7] Li and colleagues[32] also found that health administrative data identified colonoscopy procedures accurately.

The weighted and unweighted results for colonoscopy completeness differed. The weighted results should be more valid than the unweighted results as long as the distribution of procedures in our sampling frame is representative of the distribution in the underlying population. The weighted results for colonoscopy completeness indicate that the administrative data definitions are sensitive but less specific, which would occur if the endoscopist billed for a complete colonoscopy procedure but, in fact, the procedure was not complete. In a study of 15 168 colonoscopy procedures in which Medicare claims were matched to records in the Clinical Outcomes Research Initiative database, the Medicare data also failed to identify incomplete colonoscopy procedures accurately.[33]

We found that the administrative data definition comprising codes for removal of a polyp 3 mm or larger, for fulguration and for removal of polyps greater than 3 cm was highly specific, but the sensitivity was 76.9%. This specificity value is comparable to that reported in the study of Medicare data[33] and in a study from Quebec;[34] however, the sensitivity values reported in those 2 studies were better (92% and 86%, respectively). In the current study, this administrative data definition performed reasonably well for the clinically relevant reference standard of adenoma or sessile serrated adenoma/polyp, although the false-negative and false-positive rates were 20% and 13%, respectively.

Strengths and limitations

Strengths of our study include chart abstraction by trained, blinded abstractors who used a standardized data collection tool, a large sample that was representative of regional and facility differences, and use of a rigorous sampling strategy that randomly selected institutions and procedures and ensured adequate inclusion of rare events. However, the latter approach, which necessitates the use of weighted cases, also introduces a potential limitation, as inaccuracies may occur if the selected cases are not representative of the underlying sampling frame. Although we sampled procedures from

2008/09, our findings remain relevant, as recent studies on colonoscopy in Ontario used data from this year and earlier years.[7,35] In addition, although some new codes for anesthesiologist assistance and colonoscopy (by colonoscopy indication) have been introduced, none of the codes we validated have been removed, and the structure of the codes remains the same (see Supplementary Table A3, Appendix 1, for an explanation of the changes). We would expect minimal effect on the test characteristics reported above as long as these new codes are included where relevant in the administrative data definitions above. As some of the clinics we approached did not agree to participate, there is a risk of volunteer or selection bias if the nonparticipating clinics were more likely than the participating clinics to code inaccurately, which would make our algorithms less accurate. However, as the nonparticipating clinics would have contributed a small number of colonoscopy procedures to the total sample, any effect would have been small. Cases were excluded from analysis if the medical record could not be located, if the endoscopist procedure note could not be located at the time of chart abstraction, if the dates did not match or if records could not be linked to administrative data, and cases were excluded from the polypectomy analyses if the pathology report could not be located. Given the small proportion of cases in which these records were missing, we expect that the impact on our findings is small.

Conclusion

We rigorously validated 5 colonoscopy data elements that are routinely used in health administrative data studies of colonoscopy, for quality-assurance purposes and to guide health policy. Validation enhances the credibility and transparency of studies using health administrative data, which is important for uptake of findings[36] by scientific and policy consumers and for application of algorithms by other scientists.[28] We expect that our findings will be of interest to scientists and policymakers interested in using health administrative data to study and improve the quality of colonoscopy.

References

1. Rabeneck L, Paszat LF, Hilsden RJ, et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. *Gastroenterology* 2008;135:1899-906.
2. Hilsden RJ, Dube C, Heitman SJ, et al. The association of colonoscopy quality indicators with the detection of screen-relevant lesions, adverse events, and postcolonoscopy cancers in an asymptomatic Canadian colorectal cancer screening population. *Gastrointest Endosc* 2015;82:887-94.
3. Bowles CJ, Leicester R, Romaya C, et al. A prospective study of colonoscopy practice in the UK today: Are we adequately prepared for national colorectal cancer screening tomorrow? *Gut* 2004;53:277-83.
4. Alharbi O, Rabeneck L, Sutradhar R, et al. A population-based analysis of outpatient colonoscopy in adults assisted by an anesthesiologist. *Anesthesiology* 2009;111:734-40.
5. Dominitz JA, Baldwin LM, Green P, et al. Regional variation in anesthesia assistance during outpatient colonoscopy is not associated with differences in polyp detection or complication rates. *Gastroenterology* 2013;144:298-306.
6. Khiani VS, Soulos P, Gancayco J, et al. Anesthesiologist involvement in screening colonoscopy: temporal trends and cost implications in the Medicare population. *Clin Gastroenterol Hepatol* 2012;10:58-64 e1.
7. Bielawska B, Hookey LC, Sutradhar R, et al. Anesthesia assistance in outpatient colonoscopy and risk of aspiration pneumonia, bowel perforation, and splenic injury. *Gastroenterology* 2018;154:77-85.e3.
8. Wernli KJ, Brenner AT, Rutter CM, et al. Risks associated with anesthesia services during colonoscopy. *Gastroenterology* 2016;150:888-94; quiz e18.
9. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014;370:1298-306.
10. Rees CJ, Thomas Gibson S, Rutter MD, et al. UK key performance indicators and quality assurance standards for colonoscopy. *Gut* 2016;16:1923-9.
11. Tinmouth J, Kennedy E, Baron D, et al. Colonoscopy quality assurance in Ontario: systematic review and clinical practice guideline. *Can J Gastroenterol Hepatol* 2014;28:251-74.
12. Kaminski MF, Anderson J, Valori R, et al. Leadership training to improve adenoma detection rate in screening colonoscopy: a randomised trial. *Gut* 2016;65: 616-24.
13. Coe SG, Crook JE, Diehl NN, et al. An endoscopic quality improvement program improves detection of colorectal adenomas. *Am J Gastroenterol* 2013;108: 219-26.
14. Quality assurance. Edmonton: Forzani & MacPhail Colon Cancer Screening Centre, Alberta Health Services; 2016. Available: www.albertahealthservices.ca/info/Page8336.aspx (accessed 2016 Sept. 21).
15. Quality Management Partnership. Colonoscopy. Toronto: Cancer Care Ontario. Available: https://www.qmpontario.ca/colonoscopy/ (accessed 2016 Mar. 10).
16. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12:e1001885.
17. Baxter NN, Goldwasser MA, Paszat LF, et al. Association of colonoscopy and death from colorectal cancer. *Ann Intern Med* 2009;150:1-8.
18. Bressler B, Paszat LF, Chen Z, et al. Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis. *Gastroenterology* 2007;132:96-102.
19. Baxter NN, Sutradhar R, Forbes SS, et al. Analysis of administrative data finds endoscopist quality measures associated with postcolonoscopy colorectal cancer. *Gastroenterology* 2011;140:65-72.
20. Rothbard A. Quality issues in the use of administrative data records. In: Fantuzzo J, Culhane DP, editors. *Actionable intelligence: using integrated data systems to achieve a more effective, efficient, and ethical government*. New York: Palgrave Macmillan US; 2015:77-103.
21. Logan JR, Lieberman DA. The use of databases and registries to enhance colonoscopy quality. *Gastrointest Endosc Clin N Am* 2010;20:717-34.
22. De Coster C, Li B, Quan H. Comparison and validity of procedures coded with ICD-9-CM and ICD-10-CA/CCI. *Med Care* 2008;46:627-34.
23. Groenen MJ, van Buuren HR, van Berge Henegouwen GP, et al. Validation study of automatically generated codes in colonoscopy using the endoscopic report system Endobase. *Scand J Gastroenterol* 2010;45:1121-6.
24. Lee JK, Jensen CD, Lee A, et al. Development and validation of an algorithm for classifying colonoscopy indication. *Gastrointest Endosc* 2015;81:575-82.e4.
25. Fisher DA, Grubber JM, Castor JM, et al. Ascertainment of colonoscopy indication using administrative data. *Dig Dis Sci* 2010;55:1721-5.
26. Quality Management Partnership. *Building on strong foundations: inaugural report on quality in colonoscopy, mammography and pathology*. 2015. Available: https://www.qmpontario.ca/common/pages/UserFile.aspx?fileId=350473 (accessed 2016 Mar. 10).
27. Steele LS, Glazier RH, Lin E, et al. Using administrative data to measure ambulatory mental health service provision in primary care. *Med Care* 2004;42: 960-5.
28. Benchimol EI, Manuel DG, To T, et al. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol* 2011;64:821-9.
29. Rosner B. *Fundamentals of biostatistics*. 6th ed. Boston: Duxbury Press; 2006.
30. Casella G, Berger RL. *Statistical inference*. Belmont (CA): Duxbury Press; 1990.
31. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5.
32. Li X, Hilsden R, Hossain S, et al. Validation of administrative data sources for endoscopy utilization in colorectal cancer diagnosis. *BMC Health Serv Res* 2012;12:358.
33. Ko CW, Dominitz JA, Green P, et al. Accuracy of Medicare claims for identifying findings and procedures performed during colonoscopy. *Gastrointest Endosc* 2011;73:447-53.e1.
34. Wyse JM, Joseph L, Barkun AN, et al. Accuracy of administrative claims data for polypectomy. *CMAJ* 2011;183:E743-7.
35. Murthy SK, Benchimol EI, Tinmouth J, et al. Temporal trends in postcolonoscopy colorectal cancer rates in screen-eligible persons: a population-based study. *Gastrointest Endosc* 2018;87:1334.e4.
36. Morrato EH, Elias M, Gericke CA. Using population-based routine data for evidence-based health policy decisions: lessons from three examples of setting and evaluating national health policy in Australia, the UK and the USA. *J Public Health (Oxf)* 2007;29:463-71.