Published in final edited form as:

Gastrointest Endosc. 2015 March ; 81(3): 575–582.e4. doi:10.1016/j.gie.2014.07.031.

Development and validation of an algorithm for classifying colonoscopy indication

Jeffrey K. Lee, MD, MAS^{*}, Christopher D. Jensen, PhD, MPH[†], Alexander Lee, MD^{*}, Chyke A. Doubeni, MD, MPH[‡], Ann G. Zauber, PhD[§], Theodore R. Levin, MD[†], Wei K. Zhao, MPH[†], and Douglas A. Corley, MD, PhD[†]

^{*}Department of Medicine, Division of Gastroenterology, University of California San Francisco, San Francisco, California

[†]Division of Research, Kaiser Permanente Northern California, Oakland, California

[‡]Department of Family Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

[§]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York

Abstract

Background—Accurate determination of colonoscopy indication is required for managing clinical programs and performing research; however, existing algorithms using available electronic databases (e.g., diagnostic and procedure codes) have yielded limited accuracy.

Objective—To develop and validate an algorithm for classifying colonoscopy indication that uses comprehensive electronic medical data sources.

Design—We developed an algorithm for classifying colonoscopy indication using commonly available electronic diagnostic, pathology, cancer, and laboratory databases and validated its performance characteristics in comparison to a comprehensive review of patient medical records.

© 2014 American Society for Gastrointestinal Endoscopy. Published by Mosby, Inc. All rights reserved.

Address correspondence to: Jeffrey Lee, MD, MAS, University of California, San Francisco, Division of Gastroenterology, 513 Parnassus Ave. Room S-357, San Francisco, CA, Tel 510-891-5918, Fax 510-891-3802, jeff.lee@ucsf.edu.

Conflict of Interest Summary:

All authors have no conflict of interest in this manuscript. All authors have not published or submitted any related papers from the same study.

Author Contributions:

JKL had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: J. Lee, Jensen, Corley, Levin, Doubeni, Zauber

Acquisition of data: J. Lee, A. Lee, Zhao, Jensen

Analysis and interpretation of data: J. Lee, Jensen, Corley, Levin, Doubeni, Zauber

Drafting the manuscript: J. Lee, Jensen, Corley

Critical revision of the manuscript for important intellectual content: J. Lee, Jensen, Corley, Levin, Doubeni, Zauber, A. Lee.

Statistical Analysis: J. Lee, Zhao, Jensen, Corley.

Obtained funding: J. Lee, Doubeni, Zauber, Levin, Corley.

Study Supervision: Corley.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

We also evaluated the influence of each data source on the algorithm's performance characteristics.

Setting—Kaiser Permanente Northern California healthcare system.

Patients—300 patients who underwent a colonoscopy between 2007 and 2010.

Main Outcome Measurement—Algorithm's sensitivity, specificity, and positive predictive value (PPV) for classifying screening, surveillance, and diagnostic colonoscopies. The reference standard was the indication assigned after comprehensive medical record review.

Results—For screening indications, the algorithm's sensitivity was 88.5% (95% confidence interval (CI), 80.4%–91.7%), specificity was 91.7% (95% CI, 87.0%–95.1%), and PPV was 83.3% (95% CI, 74.7%–90.0%). For surveillance indications, the algorithm's sensitivity was 93.4% (95% CI, 86.2%–97.5%), specificity was 92.8% (95% CI, 88.4%–95.9%), and PPV was 85.0% (95% CI, 76.5%–91.4%). The algorithm's sensitivity, specificity, and PPV for diagnostic indications were 81.4% (95% CI, 73.0%–88.1%), 96.8% (95% CI, 93.2%–98.8%), and 93.9% (95% CI, 87.2%–97.7%).

Limitations—Validation was confined to a single healthcare system.

Conclusion—An algorithm that uses commonly available modern electronic medical data sources yielded a high sensitivity, specificity, and PPV for classifying screening, surveillance, and diagnostic colonoscopy indications. This algorithm had greater accuracy than the indication listed on the colonoscopy report.

INTRODUCTION

Colonoscopy is a widely used procedure for colorectal cancer (CRC) screening, surveillance, and diagnostic workup, and is one of the most commonly performed medical procedures in the United States.¹ Observational studies have shown that colonoscopy reduces the incidence and mortality of CRC.^{2–4} In 2006, the American Society for Gastrointestinal Endoscopy and the American College of Gastroenterology task force published a list of quality indicators for colonoscopy, including physician adenoma detection rate for screening examinations, adherence to recommended post-polypectomy surveillance intervals, cecal intubation rates, and withdrawal times; and reducing examination-related perforation rates.⁵ Of the 14 recommended colonoscopy quality indicators, several (e.g., adenoma detection rate, adherence to surveillance intervals, and perforation rate) require knowledge of examination indication. However, identifying examination indication from endoscopy reports or progress notes within electronic medical records can be difficult due to the text-based nature of reports in many settings, uncertainty regarding the accuracy of codes from only procedure reports, and the high cost, potential for reviewer bias, and fatigue associated with manual chart review.

An accurate and accessible method for classifying colonoscopy indication is crucial for measuring colonoscopy quality indicators, performing colonoscopy-related research, and monitoring CRC screening rates. To date, five studies have tested algorithms using electronically-available administrative, diagnostic, and procedure codes to classify colonoscopy indication; the reported accuracy of these algorithms varied widely, with

sensitivity for screening indications ranging from 29% to 84% and specificity ranging from 58% to 93%.^{6–10} Also, these algorithms were limited in their ability to differentiate surveillance examinations from screening or diagnostic examinations due to their inability to link administrative, procedure, and diagnostic codes with pathology and cancer registry data.

The availability of electronically accessible databases makes it possible to integrate diagnostic and procedure codes with pathology and laboratory data. Thus, the aim of our study was to use these resources to develop and validate a comprehensive algorithm for classifying colonoscopy indication and to evaluate the influence of each data source on the algorithm's overall performance.

METHODS

Study setting

This cross-sectional study was conducted among members of Kaiser Permanente Northern California (KPNC), an integrated health care delivery organization with over 3.3 million members across 21 medical centers and hospitals in urban, suburban, and semirural regions within a large geographic area.¹¹ This study was conducted as part of the National Cancer Institute funded consortium, Population-Based Research Optimizing Screening through Personalized Regimens. The overall aim of this consortium is to conduct multi-site, coordinated, transdisciplinary research to evaluate and improve cancer screening processes. This research was approved by the KPNC Institutional Review Board.

Evaluation of electronic data sources

KPNC uses the EPIC platform (Epic Systems Corp, Verona, Wisconsin) for health care diagnoses, procedure codes and laboratory test results; the EPIC platform will soon cover approximately half of the population in the United States.¹² Pathology data use the Cerner CoPathPlus (Cerner Corporation, Kansas City, Kansas) platform and Systematized Nomenclature of Medicine coding; this system is used by over 250 medical centers in the United States.¹³ Additional data sources included an electronic consult/referral database for gastroenterology specialty and procedure referrals and the KPNC cancer registry. Prior validation studies were conducted to evaluate whether the electronic data elements were capturing the desired procedures and laboratory data, compared with the results from manual medical record review (details not shown). These evaluations confirmed a very high level of agreement and/or sensitivity for capture of colonoscopy examination performance compared with manual procedure log books (99%); and assignment of adenoma and CRC status from pathology and cancer registry databases compared with text-based reports, respectively (100%).¹¹

Development of an algorithm to classify colonoscopy indication

Algorithm development—An expert panel that included gastroenterologists, internists, a family medicine physician, and data analysts was convened to identify diagnoses, pathology findings, laboratory values, and procedures associated with screening, diagnostic, and surveillance indications for colonoscopy. Electronic definitions of these factors were developed using International Classification of Disease 9th edition (ICD-9) procedure and

diagnostic codes, Current Procedural Terminology (CPT) codes, KPNC laboratory codes, Systematized Nomenclature of Medicine codes, and International Classification of Disease (Oncology) codes from the KPNC cancer registry (Supplementary Tables 1–4). For the electronic gastroenterology referral system unique to KPNC, referral reasons were condensed into 17 standardized indication categories (e.g., abdominal pain, occult blood in stool, diarrhea), which corresponded to screening, surveillance, or diagnostic indications.

An algorithm was designed to classify examination indication hierarchically: diagnostic indications took precedence over surveillance indications, which took precedence over screening indications (Figure 1). The algorithm was refined using a development set of 150 colonoscopies. The algorithm's performance was compared with findings from the manual review of these medical records to identify correctable errors in logic and to modify inclusion of relevant data codes for assignment of examination indication categories. The algorithm excluded diagnoses and pathology codes assigned on the day of the index colonoscopy as their inclusion can negatively impact the sensitivity and specificity for classifying screening examinations.⁸

Colonoscopies were classified as diagnostic if the medical databases included any of the following in the 6 months before the examination (Figure 1): abdominal pain, gastrointestinal bleeding (e.g., melena, bright red blood per rectum, hematochezia), diarrhea, change in bowel habits, unexplained weight loss, diverticulitis, iron-deficiency anemia, abnormal imaging findings (i.e., barium enema, computed tomography [CT] or magnetic resonance imaging [MRI] scan of the abdomen/pelvis), or a prior CRC diagnosis. In addition, examinations were classified as diagnostic if any of the above data sources included: an adenoma diagnosis or a positive fecal-based screening test (i.e., fecal occult blood test or fecal immunochemical test) within 1 year before the examination; either a flexible sigmoidoscopy or a colonoscopy accompanied by a polyp diagnosis within 1 year before the examination; or a history in the last 10 years of an inflammatory bowel disease diagnosis (i.e., Crohn's disease or ulcerative colitis).

Colonoscopies were classified as surveillance examinations if the patient had a diagnosis of an adenoma (on pathology) >1 year or a diagnosis of CRC >6 months before the current examination. The earlier interval was allowed given that a surveillance examination may be performed approximately a year after a CRC diagnosis. Examinations were also classified as surveillance if the patient had an ICD-9 diagnosis code for a history of polyp, adenoma, CRC, or history of CRC but without any pathology record in our database. The latter was intended to capture diagnoses that occurred before membership within KPNC.

Colonoscopies in patients with a family history of CRC, defined as an ICD-9 diagnosis code or gastroenterology referral code for family history, were classified as high-risk screening examinations in the situation where four indications (diagnostic, surveillance, high-risk screening, and average-risk screening) were considered; otherwise, these examinations were included as screening examinations.

Last, the algorithm classified all other examinations as screening; thus, patients receiving a screening examination were asymptomatic, average-risk, and without any prior adenomas or CRC.

Reference method for ascertaining examination indication—We used a test sample of 100 additional colonoscopies for reviewer training and to evaluate reproducibility of manual indication assignment between reviewers. Two reviewers independently using the reference method described below had a summary 93% concordance across the three primary examination indications, which corresponded to a near-perfect agreement (kappa statistic: 0.90).¹⁴ Given the high level of concordance, the subsequent additional 300 examinations used for algorithm validation were evaluated by a single reviewer; these 300 examinations served as the reference method (i.e., criterion standard) for assignment of colonoscopy indication.

The reviewers followed a standardized protocol for examining each data source and assigning examination indication, and were blinded to the indication assigned by the algorithm. Data sources included colonoscopy procedure reports, pre-operative/procedure history and physical notes, clinic notes, pathology reports, laboratory reports, radiology reports, consultation reports, physician orders, and cancer registry data. Relevant clinical details and dates were entered into a database and colonoscopy examinations were assigned a diagnostic, surveillance, or screening indication by the reviewer based on the total patient information available at the time the examination was ordered (not the date it was performed) and consistent with the criteria outlined in Figure 1. Examinations determined to have a screening indication were further classified as high-risk or average-risk screening based on family history of CRC (defined as any first-degree relative with a history of CRC).

Algorithm validation: selection of colonoscopies—A total of 300 colonoscopies, representing all three indications (100 each of screening, surveillance, and diagnostic examinations) were randomly selected from all examinations performed between 2007 and 2010 (total colonoscopies performed during this period was 238,102) and used for our validation analysis. The sample size was chosen a priori and was comparable to prior validation studies.^{6–8} Colonoscopy examinations were identified using CPT codes and ICD-9 procedure codes.

Statistical analysis

Descriptive statistics were used to characterize the 300 patients who underwent colonoscopy examinations in the validation sample. The algorithm's sensitivities, specificities, positive predictive values (PPV), and negative predictive values (NPV), along with their 95% confidence intervals [CI], were calculated for each indication using the reference method. Sensitivity was defined as the proportion of colonoscopy examinations with a specified indication (per the reference method) that were correctly identified by the algorithm. Specificity was defined as the proportion of examinations without a specified indication that were correctly identified by the algorithm as without that specified indication. Thus, a high specificity for screening would indicate that colonoscopies performed for surveillance or diagnostic purposes were uncommonly misclassified as having been performed for

screening. PPV was defined as the proportion of all examinations classified by the algorithm with a specified indication that had the examination indication of interest (per the reference method). We performed sensitivity analyses by removing data sources (i.e., gastroenterology referral codes, pathology codes, cancer registry codes, and laboratory codes) individually and collectively from the algorithm to evaluate their impact on the algorithm's performance. In comparison to the reference method, we also calculated the percent of examinations whose indications were misclassified when indications were obtained solely from colonoscopy procedure reports versus indications ascertained from the algorithm. STATA 12.0 (Stata Corporation, College Station, Tex) was used for all statistical analyses.

RESULTS

Patient characteristics

The descriptive characteristics of the validation study sample are shown in Table 1. Mean age ($\pm$ SD) was 61.9 ± 10.6 years (range: 19–86), 50.5% of patients were female, 62.3% were white, and 54.7% had a Charlson comorbidity score of 0. The cecal intubation rate was 97.7% and 92.0% had an adequate-to-excellent bowel preparation. Of the 300 colonoscopies, 113 (37.7%) were determined by the reference method to be diagnostic, 91 (30.3%) were surveillance, and 96 (32.0%) were screening, including 23 (7.7%) high-risk screening and 73 (24.3%) average-risk screening.

Algorithm performance

The performance of the electronic algorithm did not vary widely across the three examination indications in comparison to the reference method (Table 2). Sensitivity, specificity, and PPV were high for screening (sensitivity: 88.5%; 95% CI, 80.4%–94.1%; specificity: 91.7%; 95% CI, 87.0%–95.1%; and PPV: 83.3%; 95% CI, 74.7%–90.0%), surveillance (sensitivity: 93.4%; 95% CI, 86.2%–97.5%; specificity: 92.8%, 95% CI: 88.4%–95.9%; and PPV: 85.0%; 95% CI, 76.5%–91.4%), and diagnostic (sensitivity: 81.4%; 95% CI, 73.0%–88.1%; specificity: 96.8%; 95% CI, 93.2%–98.8%; and PPV: 93.9%; 95% CI, 87.2%–97.7%) indications.

When high-risk screening (i.e., patients with a family history of CRC) was added as a fourth examination indication (Table 2), the algorithm's sensitivity and specificity for average-risk screening indications remained high, but PPV declined to 70.3% (95% CI, 59.8%–79.5%). The algorithm's sensitivity for detecting patients with a high-risk screening indication was low at 34.8% (95% CI, 16.4%–57.3%) and PPV was moderate-to-high at 72.7% (95% CI, 39.0–94.0).

Sensitivity analysis: elimination of specific data sources

Given that certain healthcare settings may not have access to a full array of electronic databases (e.g., pathology data, cancer registry data), we evaluated the impact of excluding specific data sources on the algorithm's performance characteristics (Table 3, Supplementary Table 5). For each indication, the algorithm's performance was mainly affected by the exclusion of pathology codes. For screening indications, this resulted in a decrease in the PPV to 75.9% (95% CI, 66.9%–83.5%) and specificity to 86.8% (95% CI,

81.3%–91.1%). For surveillance indications, sensitivity and PPV decreased to 84.6% (95% CI, 75.5%–91.3%) and 81.9% (95% CI, 72.6%–89.1%), respectively. For diagnostic indications, this resulted in a decrease in sensitivity to 77.9% (95% CI, 69.1%–85.1).

Removing all data sources except diagnostic ICD-9 codes and procedure codes reduced PPV to 72.1% (95% CI, 63.3–79.9) for screening indications; reduced sensitivity to 82.4% (95% CI, 73.0–89.6) for surveillance indications; and reduced sensitivity to 75.2% (95% CI, 66.2–82.9) for diagnostic indications, primarily due to the impact of excluding pathology codes.

Examination indication misclassification: colonoscopy reports versus algorithm

To assess the utility of using colonoscopy reports alone versus the algorithm for determining examination indication, we compared the agreement between the reference method with both the examination indication noted on colonoscopy reports and the algorithm's ascertained indication. The indication on the colonoscopy report had good overall performance characteristics compared with the reference standard: sensitivity 90.6% (95% CI, 83.0%–95.6%), specificity: 90.2 (95% CI, 85.3%–93.9%), and PPV: 81.3% (95% CI, 72.6%–88.2%) for screening indications; sensitivity: 76.9 (95% CI, 66.9%–85.1%), specificity: 98.1% (95% CI, 95.2%–99.5%), and PPV: 94.6 (95% CI, 86.7%–98.5%) for surveillance indications; and sensitivity: 74.3% (95% CI, 65.3%–82.1%), specificity: 100% (95% CI, 98.1%–100%), and PPV: 100% (95% CI, 95.7%–100%) for diagnostic indications.

In comparison to the reference method, the algorithm misclassified 12.7% (concordance: 87.3%, kappa statistic: 0.81) and 17.0% (concordance: 83.0%, kappa statistic: 0.76) of examinations across the three and four examination indications, respectively. By comparison, use of colonoscopy report indication alone to ascertain examination indication misclassified 19.7% (concordance, 80.3%; kappa statistic, 0.72) and 20.7% (concordance 79.3%; kappa statistic, 0.72) of examinations across the 3 and 4 examination indications, respectively. Furthermore, 11.7% (35 examinations) of the colonoscopy reports were missing a precise examination indication.

DISCUSSION

We developed an algorithm using a comprehensive combination of administrative, diagnostic, and procedure codes linked with laboratory, pathology, and cancer registry data to classify colonoscopy indications. Our results demonstrate that the algorithm had a simultaneously high sensitivity, specificity, and PPV for classifying screening (89%, 92%, and 83%) surveillance (93%, 93%, and 85%), and diagnostic (81%, 97%, and 94%) indications for colonoscopies. Although our study was not designed to compare the accuracy of colonoscopy reports alone with the algorithm in determining examination indications, the percent of examinations misclassified (across three indications) using the algorithm was 12.7% versus 19.7% when colonoscopy reports alone were used. These findings have important implications for effectively and efficiently monitoring colonoscopy quality, outcomes, and utilization patterns. The accurate identification of examination indication using electronic data sources makes it possible to monitor trends in CRC screening, evaluate the effectiveness of CRC screening improvement strategies, calculate adenoma detection

rates from screening colonoscopies, and track the appropriate use of surveillance intervals in patients with prior adenomas or CRC, all without the use of costly manual record review.

To ascertain the potential applicability of the algorithm for use in other healthcare systems that may not have electronic access to the same full array of databases, we removed individual data sources including the gastroenterology referral, pathology, cancer registry, and laboratory codes from the algorithm. Removing the pathology data source appeared to have the most impact on the algorithm's sensitivity and PPV for screening and surveillance examinations; PPV for screening examinations was reduced to 75.9% and for surveillance examinations, sensitivity and PPV were reduced to 84.6% and 81.9%, respectively. Removal of either the gastroenterology referral or cancer registry data had minimal impact on the algorithm's overall performance. These findings suggest the importance of accessing pathology data when ascertaining surveillance and screening indications.

To date, few algorithms have been developed for classifying colonoscopy indication using electronically-available data.⁶⁻¹⁰ In a Veterans Health administration (VA) study, El-Serag et al. developed an algorithm using administrative, diagnostic, and procedure codes from the 4 years before the index colonoscopy, which showed a sensitivity of 70% and specificity of 72% for screening examinations.⁷ Using the same VA database, Fisher et al. developed 3 modified algorithms based on El-Serag's algorithm⁷ incorporating diagnostic codes most related to lower gastrointestinal conditions along with removing diagnostic codes recorded on the day of index colonoscopy.⁸ The best-performing of these algorithms demonstrated a sensitivity of 83% and a specificity of 76% for screening examinations. In a study featuring a community-based population from Kaiser Permanente Southern California, Haque et al. developed an algorithm based on administrative, diagnostic, and procedures codes and reported a sensitivity and specificity of 84% and 76%, respectively, for screening examinations.⁶ Limitations of these studies⁶⁻⁸ include the lack of a simultaneously high sensitivity and specificity for screening examinations; evaluating colonoscopy history and diagnostic codes only in the 1 to 4 years before the index colonoscopy, which can lead to higher examination misclassification rates (e.g., a patient with a code for personal history of colorectal adenomas at 5 years before the index colonoscopy would be misclassified as screening instead of surveillance using the algorithms above); inability to incorporate pathology, cancer registry, and laboratory data sources into the algorithm; only attempting to distinguish screening versus non-screening examinations; and no reporting of the PPV for screening examinations.⁶⁻⁸

Model-based algorithms using diagnostic and procedure codes have also been developed to improve the accuracy of classifying colonoscopy indications.^{9, 10} In a study of the Medicare population, Ko et al. developed and validated three model-based algorithms to classify colonoscopy indications based on ICD-9 diagnostic codes and procedure codes within 12 months before the index colonoscopy; these demonstrated high specificity for screening and surveillance indications (>95%), but only moderate sensitivity (55%-58%).⁹ More recently, Sewitch et al. developed 2 model-based algorithms using administrative health data from Canada, which demonstrated a sensitivity for screening examinations between 84.5% and 85.1% and a specificity between 62.2% and 63.3%.¹⁰ Both studies^{9, 10} relied on the indication from colonoscopy procedure reports as the reference standard rather than a

comprehensive medical chart review; findings from the current study suggest use of this data source may misclassify a moderate proportion of subjects.

Our study has several strengths. First, our reference method for ascertaining examination indication featured a comprehensive manual review of patient medical records, including prior progress, referral, and endoscopy procedure notes; the index colonoscopy report; and pathology, laboratory, and radiology reports. This enabled the reviewer, who was blinded to the algorithm's results, to consider the complex and detailed clinical information behind each colonoscopy when ascertaining examination indication. Second, there was a high degree of agreement between 2 reviewers who independently used the reference method to ascertain examination indication. Third, including pathology and cancer registry data allowed for more accurate classification of surveillance indication. Fourth, we included administrative, diagnostic, and procedures codes dating back to 1995 for the algorithm instead of using codes only in the 1–4 years before the index colonoscopy. This important time element, along with incorporating pathology and cancer registry data sources in our algorithm, may explain some of the improved sensitivity and specificity estimates for screening and surveillance indications compared to prior published algorithms.^{6–10} Last, the 300 colonoscopy examinations in our validation sample were randomly selected from all examinations performed during the study interval.

There are also several limitations to the study that should be considered. First, colonoscopies were classified by the algorithm as surveillance if codes for polyps or adenomas were detected, even if no adenoma was detected in the pathology database. This was permitted because patients may have had a history of adenomas before joining the health plan; however, this could lead to misclassification of some patients. Second, the algorithm had limited ability to detect patients with a family history of CRC given that these data are incomplete for many patients. However, prior large colonoscopy studies have included patients with a family history of CRC in their large screening cohorts.^{15, 16} In addition, a recent population-based survey study showed that the reported family history of CRC in the general population was highly inaccurate.¹⁷ Third, the validation was performed in a single healthcare setting with a sample of 300 subjects. Finally, it is theoretically possible that the algorithm could perform differently with a different case mix of patients; thus, the findings should be replicated in other studies, including those in other healthcare settings.

In summary, proper classification of colonoscopy indication is crucial in observational research studying the effectiveness of CRC screening and surveillance, and for evaluating quality improvement benchmarks for colonoscopy. In our study, an algorithm using the commonly available EPIC and Cerner electronic data sources for endoscopic procedures, diagnoses, laboratory and pathology findings, combined with data from a referral database and a cancer registry, demonstrated a high sensitivity, specificity, and PPV for differentiating screening, surveillance, and diagnostic colonoscopies. Selectively excluding individual data sources other than the pathology database (including exclusion of the cancer registry and electronic referral system), had little impact on the algorithm's performance, suggesting that the algorithm may have utility in settings with less comprehensive access to data sources. Further validation studies are needed to test this algorithm in other healthcare

settings to determine its generalizability, and against other methods for classifying examination indication.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Primary Funding Source: National Cancer Institute (U54 CA163262), National Institute of Diabetes and Digestive and Kidney Diseases (T32 DK007007)

Funding Source:

Grants from the National Institute of Diabetes and Digestive and Kidney Diseases (T32DK007007, JKL), and the National Cancer Institute (U54 CA163262, DAC, TRL, AGZ, CAD)

Acronyms

PPV	positive predictive value
CI	confidence interval
CRC	colorectal cancer
KPNC	Kaiser Permanente Northern California
PROSPR	Population-Based Research Optimizing Screening through Personalized Regimens
NCI	National Cancer Institute
ICD-9	International Classification of Disease 9 th edition
SNOMED	Systematized Nomenclature of Medicine
CPT	Current Procedural Terminology
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
FIT	fecal immunochemical test
VA	Veterans Health Administration

References

1. Seeff LC, Richards TB, Shapiro JA, et al. How many endoscopies are performed for colorectal cancer screening? Results from CDC's survey of endoscopic capacity. *Gastroenterology*. 2004; 127:1670–1677. [PubMed: 15578503]
2. Zauber AG, Winawer SJ, O'Brien, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012; 366:687–696. [PubMed: 22356322]
3. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 1993; 329:1977–81. [PubMed: 8247072]
4. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal cancer incidence and mortality after lower endoscopy. *N Engl J Med*. 2013; 369:1095–105. [PubMed: 24047059]

5. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. *Gastrontest Endosc.* 2006; 63:S16–28.
6. Haque R, Chiu V, Mehta KR, et al. An automated data algorithm to distinguish screening and diagnostic colorectal cancer endoscopy exams. *J Natl Cancer Inst Monogr.* 2005; 35:116–118. [PubMed: 16287897]
7. El-Serag HB, Petersen L, Hampel H, et al. The use of screening colonoscopy for patients cared for by the Department of Veterans Affairs. *Arch Intern Med.* 2006; 166:2202–2208. [PubMed: 17101937]
8. Fisher DA, Grubber JM, Castor JM, et al. Ascertainment of colonoscopy indication using administrative data. *Dig Dis Sci.* 2010; 55:1721–1725. [PubMed: 20393875]
9. Ko, CW.; Dominitz, JA.; Neradilek, M.; Polissar, N.; Green, P.; Kreuter, W.; Baldwin, LM. *Med Care.* 2012. epub ahead of print
10. Sewitch MJ, Jiang M, Joseph L, et al. Developing model-based algorithms to identify screening colonoscopies using administrative health databases. *BMC Medical Informatics and Decision Making.* 2013; 13:45. [PubMed: 23574795]
11. Corley DA, Jensen CD, Marks AR, et al. Variation of adenoma prevalence by age, sex, race, and colon location in a large population: implications for screening and quality programs. *Clin Gastroenterol Hepatol.* 2013; 11(2):172–180. [PubMed: 22985608]
12. Moukheiber, Z. [Accessed April 25, 2014] An interview with the most powerful woman in health care. Available at: <http://www.forbes.com/sites/zinamoukheiber/2013/05/15/a-chat-with-epic-systems-ceo-judy-faulkner/>
13. [Accessed April 25, 2014] CERNER CoPATHplus solution. Available at: http://www.cerner.com/public/Cerner_3.asp?id=30881
14. Byrt T. How good is that agreement? *Epidemiology.* 1996; 7:561. [PubMed: 8862998]
15. Lieberman DA, Holub JL, Moravec MD, et al. Prevalence of colon polyps detected by colonoscopy screening in asymptomatic black and white patients. *JAMA.* 2008; 300:1417–22. [PubMed: 18812532]
16. Lieberman DA, Weiss DG, Bond JH, et al. Veteran Affairs Cooperative Study Group 380. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. *N Engl J Med.* 2000; 343:162–8. [PubMed: 10900274]
17. Mai PL, Garceau AO, Graubard BI, et al. Confirmation of family cancer history reported in a population-based survey. *J Natl Cancer Inst.* 2011; 103:788–97. [PubMed: 21562245]

Take home message

- An algorithm using commonly available electronic diagnostic, pathology cancer, and laboratory databases demonstrated high accuracy in classifying screening, surveillance, and diagnostic indications for colonoscopies.
- This algorithm, which uses commonly available modern electronic medical data sources, can efficiently and accurately ascertain colonoscopy indications without the use of costly manual chart review.

Figure 1. Algorithm for classifying colonoscopy indication

CT: computed tomography

MRI: magnetic resonance imaging

CRC: colorectal cancer

ICD-9: International Classification of Disease 9th edition

Table 1

Patient characteristics of the validation study sample.

Characteristics	n (%)
Total study sample	300 (100.0)
Age, year	
<50	19 (6.3)
50–59	113 (37.7)
60–69	93 (31.0)
70–79	60 (20.0)
80	15 (5.0)
Female	151 (50.5)
Race/ethnicity	
White	187 (62.3)
African American	16 (5.3)
Hispanic	34 (11.3)
Asian	37 (12.3)
Other	19 (6.3)
Missing	7 (2.3)
Charlson co-morbidity score	
0	164 (54.7)
1	57 (19.0)
2	47 (15.7)
3	12 (4.0)
4	8 (2.7)
5	12 (4.0)
Colonoscopy indication*	
Diagnostic	113 (37.7)
Surveillance	91 (30.3)
Screening	96 (32.0)
High-risk screening	23 (7.7)
Average-risk screening	73 (24.3)
Cecal intubation	293 (97.7)
Bowel preparation	
Excellent/good	200 (66.7)
Fair/adequate	76 (25.3)
Poor	6 (2.0)
Not reported	18 (6.0)

* By comprehensive medical chart review.

Table 2

Algorithm performance for classifying diagnostic, surveillance, and screening indications.

Indication (n) [*]	Sensitivity (%) (95% confidence interval)	Specificity (%) (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)
Diagnostic (113)	81.4 (73.0–88.1)	96.8 (93.2–98.8)	93.9 (87.2–97.7)	89.6 (84.6–93.5)
Surveillance (91)	93.4 (86.2–97.5)	92.8 (88.4–95.9)	85.0 (76.5–91.4)	97.0 (93.6–98.9)
Screening [†] (96)	88.5 (80.4–94.1)	91.7 (87.0–95.1)	83.3 (74.7–90.0)	94.4 (90.3–97.2)
High-risk screening [‡] (23)	34.8 (16.4–57.3)	98.9 (96.9–99.8)	72.7 (39.0–94.0)	94.8 (91.6–97.1)
Average-risk screening (73)	87.7 (77.9–94.2)	88.1 (83.2–92.0)	70.3 (59.8–79.5)	95.7 (92.0–98.0)

^{*} Number of examinations in each indication category, as determined by comprehensive medical record review

[†] Includes average-risk and high-risk screening subjects

[‡] Defined as subjects with a family history of colorectal cancer

Table 3

Sensitivity analyses: algorithm performance for classifying examination indication after exclusion of data sources

Indication (n) [*]	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)
<i>Without use of gastroenterology referral codes</i>				
Diagnostic (113)	78.8 (70.1–85.9)	96.8 (93.2–98.8)	93.7 (86.8–97.7)	88.3 (83.1–92.4)
Surveillance (91)	91.2 (83.4–96.3)	94.7 (90.8–97.3)	88.3 (80.0–94.0)	96.1 (92.5–98.3)
Screening [†] (96)	91.7 (84.2–96.3)	88.7 (83.6–92.7)	79.3 (70.6–86.4)	95.8 (91.8–98.2)
<i>Without use of pathology codes</i>				
Diagnostic (113)	77.9 (69.1–85.1)	96.8 (93.2–98.8)	93.6 (86.6–97.6)	87.9 (82.6–92.0)
Surveillance (91)	84.6 (75.5–91.3)	91.9 (87.3–95.2)	81.9 (72.6–89.1)	93.2 (88.9–96.2)
Screening [†] (96)	88.5 (80.4–94.1)	86.8 (81.3–91.1)	75.9 (66.9–83.5)	94.2 (89.8–97.0)
<i>Without use of laboratory codes</i>				
Diagnostic (113)	77.9 (69.1–85.1)	96.8 (93.2–98.8)	93.6 (86.6–97.6)	87.9 (82.6–92.0)
Surveillance (91)	93.4 (86.2–97.5)	90.9 (86.2–94.4)	81.7 (73.0–88.6)	96.9 (93.5–98.9)
Screening [†] (96)	86.5 (78.0–92.6)	90.7 (85.8–94.3)	81.4 (72.5–88.4)	93.4 (89.0–96.5)
<i>Without use of cancer registry codes</i>				
Diagnostic (113)	80.5 (72.0–87.4)	96.8 (93.2–98.8)	93.8 (87.0–97.7)	89.2 (84.1–93.1)
Surveillance (91)	93.4 (86.2–97.5)	92.8 (88.4–95.9)	85.0 (76.5–91.4)	97.0 (93.6–98.9)
Screening [†] (96)	88.5 (80.4–94.1)	91.2 (86.4–94.7)	82.5 (73.8–89.3)	94.4 (90.2–97.2)
<i>Without use of gastroenterology referral, pathology, laboratory, and cancer registry codes</i>				
Diagnostic (113)	75.2 (66.2–82.9)	96.8 (93.2–98.8)	93.4 (86.2–97.5)	86.6 (81.2–90.9)
Surveillance (91)	82.4 (73.0–89.6)	94.3 (90.2–97.0)	86.2 (77.2–92.7)	92.5 (88.1–95.7)
Screening [†] (96)	91.7 (84.2–96.3)	83.3 (77.5–88.2)	72.1 (63.3–79.9)	95.5 (91.3–98.0)

^{*} Number of examinations in each indication category, as determined by comprehensive medical record review

[†] Includes patients with a family history of colorectal cancer