BMJ Open Gastroenterology

# Clinical codes combined with procedure codes increase diagnostic accuracy of Crohn's disease in a US military health record

Manish Singla [1,2] Susan Hutfless [3] Elie Al Kazzi,[4] Benjamin Rodriguez,[5] John Betteridge,[6] Steven R Brant [3,7,8]

**To cite:** Singla M, Hutfless S, Al Kazzi E, *et al*. Clinical codes combined with procedure codes increase diagnostic accuracy of Crohn's disease in a US military health record. *BMJ Open Gastro* 2020;**7**:e000378. doi:10.1136/bmjgast-2020-000378

Received 10 January 2020
Revised 6 February 2020
Accepted 7 February 2020



For numbered affiliations see end of article.

**Correspondence to**
Dr Manish Singla;
manishsingla@gmail.com

## ABSTRACT

**Background and aims** Previous examinations of International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) codes to predict accuracy of diagnosis in inflammatory bowel disease have had limited chart review to confirm diagnosis. We aimed to evaluate using the ICD-9-CM for identifying Crohn's disease (CD) in a large electronic health record (EHR) database.

**Methods** This is a retrospective case-control study with a 3:1 allocation of EHRs of active duty service members diagnosed with CD from 1996 to 2012. Subjects were selected by having two ICD-9-CM codes for CD and none for ulcerative colitis during the study period. Gastroenterologists reviewed each chart and confirmed the diagnosis of CD by analysing medication history and clinical, endoscopic, histological, and radiographic exams.

**Results** 300 cases of CD were selected; 14 cases were discarded due to lack of data, limiting analysis to 284 subjects. Two diagnostic codes for CD had sensitivity and specificity of 1.0 and 0.53 respectively, for confirmed CD. If two or more encounters listing CD were with a gastroenterologist, the sensitivity and specificity was 0.71 and 0.87 respectively. If two encounters included a colonoscopy was performed at the same time as a CD code, sensitivity and specificity was 0.49 and 0.88 respectively.

**Conclusions** The relatively poor specificity of ICD-9-CM codes in making the diagnosis of CD should be taken into consideration when interpreting results and when conducting research using such codes. Limiting these codes to patients given this diagnosis by a gastroenterologist, or to those who had a colonoscopy at the time of a diagnosis, increases the specificity, although at cost of sensitivity, especially for colonoscopy.

## Summary box

### What is already known about this subject?

- Using clinical billing codes can allow big data analysis of healthcare outcomes in patients with Crohn's disease (CD).

### What are the new findings?

- Using only clinical billing codes had a poor specificity and positive predictive value (PPV) in predicting patients with CD.
  Requiring a gastroenterology encounter or adding a code for colonoscopy greatly increased specificity and PPV.

### How might it impact on clinical practice in the foreseeable future?

- Future studies identifying patients with CD using billing codes should include gastroenterology encounters or procedure codes to increase specificity and PPV.

## INTRODUCTION

Crohn's disease (CD) is a chronic idiopathic inflammatory disease of transmural inflammation of the gastrointestinal tract, primarily the ileum or colon. The disease is diagnosed based on biopsies indicative of chronic inflammation by endoscopy or surgery without a history of chronic infectious diseases (ie, tuberculosis) or other factors (eg, ovarian abscesses or diverticulitis) that may cause a similar appearance of chronic gut inflammation.[1]

Clinically coded data, used primarily for billing or encounter tracking, can be used to identify and study large cohorts of patients with CD in an efficient and cost-effective manner. However, clinically coded data and electronic health records (EHRs) are not designed for research purposes. The codes can reflect 'working diagnoses', and are often incomplete descriptions of the severity or complications of disease. Although the EHR provides more details, the notes and uploaded documents do not always capture the longitudinal phenotype and disease activity of patients that may be collected in a recruitment-based prospective study or randomised trial. The volume of patients that can be studied using clinically coded data

can add substantially to the knowledge base. Identifying a validated case definition for codes using the EHR associated with a particular cohort can add substantially to the value of the cohort.

Previous studies have examined the accuracy of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) and similar codes based on the reference standard for diagnosis, documentation of inflammatory bowel disease (IBD) in the medical record.[2] Previous studies of accuracy of diagnostic codes in the USA found that 67.5% of patients with CD were correctly classified based on at least one ICD-9-CM 555 encounter[3] and 88% with two encounters.[4] Some cohorts have not performed their own validation studies; rather, they have relied on a case definition of two encounters based on prior evidence.[5–10] One study showed a positive predictive value (PPV) of 91% when a CD code was present without any UC codes, although this appears to be an outlier.[11] The studies have used various methods to confirm CD from a mention of CD in medical record notes to review endoscopic or radiological images or reports, operative notes, and pathology reports.

The goal of our study was to assess the diagnostic accuracy of several ICD-9-CM definitions in the active duty US military population. The US military provides a unique opportunity for research on IBD and other significant chronic conditions because IBD and related conditions (including chronic diarrhoea and chronic abdominal pain) preclude entry in the US military. Overwhelmingly, first diagnoses entered will be those from initial disease presentations. It is a diverse population but with homogeneous and universal access to medical evaluation and treatment. At a minimum, we required at least two ICD-9-CM 555 encounters.[9] In addition, we aimed to examine other definitions (to include timing of diagnosis, procedure codes, and provider specialty) to maximise sensitivity, specificity, and the PPV of CD. The expansive military EHR including clinical notes, endoscopy reports, operative reports, images, and laboratory and pathology results was used to confirm CD diagnoses.

## METHODS

We conducted a retrospective case control study with a 3:1 allocation. Eligible patients included those with active military service between 1 January 1996 and 1 December 2012 with at least three serum samples available in the Armed Forces Repository of Specimen Samples required for a related IBD study. Individuals with at least two outpatient ICD-9-CM codes of 555.x (n=300), no codes of 556.x (ulcerative colitis (UC)) and 100 individuals with similar age, sex, race, and service, but no codes of 555 or 556, were selected for chart review. Electronic versions of clinical notes, pharmacy data, endoscopy reports, radiology reports, and laboratory values were reviewed from the Department of Defense EHR, the Armed Forces Health Longitudinal Technology Application (AHLTA), by medical doctors with subspecialty fellowship training in gastroenterology and clinical practices focused in IBD. All ICD-9-CM and Current Procedural Terminology (CPT) codes and the associated clinically coded information (ie, provider specialty and location of encounter) for all reviewed individuals were available.

Data extracted from the EHR included age, gender, Montreal classification (disease location, disease behaviour, and duration of disease), and histories of smoking, intestinal surgery (to include indication and location), medications, colonoscopies, radiological studies, and diagnoses of CD, UC, irritable bowel syndrome (IBS) and infection. Records were reviewed by four IBD specialists. A chart review confirmed case of CD was defined by clinical symptoms consistent and specific to CD accompanied either by mucosal ulceration on endoscopy or a surgical specimen with pathology confirming chronic histological inflammation.

All cases were reviewed by at least two specialists, with the ruling of the second specialist maintained.

These definitions of interest included different numbers of encounters for 555.x in combination with site of service (gastroenterology (Medical Expense and Performance Reporting System codes AAF for inpatient, BAG for outpatient) or general surgery (ABA)), hospitalisation for CD, and colonoscopy (CPT 45355, 45378, 45379, 45380, 45381, 45382, 45383, 45388, 45384, 45385, 45386, 45387, 45389, 45391, 45392, 45390, 45393, 45398, 45399). A 2×2 table was created for each potential case definition classifying each individual as a true negative, true positive, false negative and false positive based on the definition and chart review determination. Using this table, sensitivity, specificity, PPV and diagnostic accuracy (defined by true positives plus true negatives over the total denominator) were calculated. Exact binomial confidence limits were calculated.[4]

## RESULTS

Our analysis included 284 patients and 100 controls; no medical encounters were available in our EHR for 16 patients. Of the 284 evaluated patients, 196 had a confirmed diagnosis of CD (69%). Twenty cases had no mention of CD in their medical record nor any gastrointestinal or immunological condition (7%). Nine patients had mention of CD in their records but lacked endoscopy or pathology information to make a definitive diagnosis (3%). Multiple patients (6.0%) had other chronic IBDs including indeterminate colitis (n=4), radiographic ileitis without endoscopic inflammation (n=4), lymphocytic colitis (n=5), UC (n=3), and possible UC (n=1). Other intestinal inflammatory conditions were observed in 2.4% of subjects including eosinophilic gastrointestinal disease (n=3), Behcet's disease (n=1), acute colitis followed by normal endoscopic findings (n=2), and jejunal enteritis seen on radiographic imaging without endoscopic or pathological confirmation (n=1). In 3.5% of subjects, chart review showed complications or features found in CD but had no evidence to confirm the finding was

Table 1 Diagnostic accuracy characteristics of case definitions based on 284 chart reviewed cases and 100 controls

| Definition tested | True positive | False positive | False negative | True negative | Diagnostic accuracy 95% CI | Sensitivity | Specificity | Positive predictive value |
|---|---|---|---|---|---|---|---|---|
| ≥2 555.x codes | 196 | 88 | 0 | 100 | 77<br>73 to 81 | 100 (by definition) | 53<br>46 to 60 | 69<br>63 to 74 |
| ≥3 555.x codes | 189 | 66 | 7 | 122 | 81<br>77 to 85 | 96<br>93 to 99 | 65<br>58 to 72 | 74<br>68 to 79 |
| ≥2 555.x codes and ≥1 CD hospitalisation | 83 | 25 | 113 | 163 | 64<br>59 to 69 | 42<br>35 to 50 | 87<br>81 to 91 | 77<br>68 to 84 |
| ≥2 555.x codes and ≥2 CD hospitalisations | 39 | 7 | 157 | 181 | 57<br>52 to 62 | 20<br>15 to 26 | 96<br>92 to 98 | 85<br>71 to 94 |
| ≥2 555.x codes with ≥1 recorded by a gastroenterologist | 148 | 36 | 48 | 152 | 78<br>74 to 82 | 76<br>69 to 82 | 81<br>74 to 86 | 80<br>74 to 86 |
| ≥2 555.x codes with ≥2 recorded by a gastroenterologist | 140 | 25 | 56 | 163 | 79<br>74 to 83 | 71<br>65 to 88 | 87<br>81 to 91 | 85<br>78 to 90 |
| ≥3 555.x codes with ≥2 recorded by a gastroenterologist | 135 | 25 | 61 | 163 | 78<br>73 to 82 | 69<br>62 to 75 | 87<br>81 to 91 | 84<br>78 to 90 |
| ≥2 555.x codes with ≥1 recorded by a gastroenterologist or general surgeon | 149 | 36 | 47 | 152 | 78<br>74 to 83 | 76<br>69 to 82 | 81<br>75 to 86 | 81<br>74 to 86 |
| 2+555.x codes with ≥2 recorded by a gastroenterologist or general surgeon | 141 | 25 | 55 | 163 | 79<br>75 to 83 | 72<br>66 to 78 | 87<br>82 to 92 | 85<br>80 to 90 |
| 2+555.x codes with ≥1 colonoscopy at same time as a 555.x code | 96 | 22 | 100 | 166 | 68<br>63 to 73 | 49<br>42 to 56 | 88<br>83 to 93 | 81<br>73 to 88 |

True positive: met inclusion criteria and chart confirmed a case.
False positive: met inclusion criteria but not chart confirmed a case.
Diagnostic accuracy, Sensitivity, Specificity and Positive predictive value are in *percent*.
CD, Crohn's disease.

due to CD (ie, intra-abdominal abscess (n=1), cryptitis (n=1), mucosal thickening on CT (n=5), and recurrent anal fissures or perianal fistula without mucosal disease (n=3)). Other gastrointestinal diagnoses included most commonly IBS (n=16), small bowel obstruction (n=1), haemorrhoids (n=1), gastro-oesophageal reflux disorder (n=1), dyspepsia (n=1), chronic abdominal pain (n=1), carcinoid tumour (n=1), appendicitis (n=1) and traveller's diarrhoea (n=1). One patient had hidradenitis suppurativa, found more frequently among patients with CD (see online supplementary table). None of the 100 control patients had evidence for a diagnosis of IBD following similar examination of their medical records.

Having two diagnostic codes for CD and no codes for UC had sensitivity, specificity, and PPV (with 95% CIs) of 1.0 (by definition as only those with at least two codes were examined so no CI calculated), 0.53 (95% CI 0.46 to 0.60), and 0.69 (95% CI 0.63 to 0.74), respectively (see table 1). When two or more encounters listing CD were with a gastroenterologist,the sensitivity, specificity, and PPV was 0.71 (95% CI 0.65 to 0.88), 0.87 (95% CI 0.81 to 0.91), and 0.85 (95% CI 0.78 to 0.90), respectively. Sensitivity, specificity and PPV were nearly identical if two encounters were with a gastroenterologist or a general surgeon (table 1). If a colonoscopy was performed at the same time as a CD code, the sensitivity, specificity, and PPV was 0.49 (95% CI 0.42 to 0.56), 0.88 (95% CI 0.83 to 0.93), and 0.81 (95% CI 0.73 to 0.88), respectively.

## DISCUSSION

Retrospective review of charts to identify patients with CD can be difficult due to the varying presentations of CD; the absence of common, objective clinical tests to confirm diagnoses with high negative predictive values complicates the nature of large database studies to identify patients with CD. ICD9 (and now, ICD-10-CM) codes are frequently used as substitutes for chart review, especially in large database studies where chart reviews are impractical. The poor specificity and PPV we observed (0.69) of even two isolated ICD9 codes in making the diagnosis of CD should be taken into consideration when interpreting results of large population studies.

After starting with a preselected population, requiring at least two CD ICD9 codes be given by gastroenterologists, or requiring a colonoscopy at the time of a diagnostic code, substantially increased the specificity and PPV although at a cost of sensitivity, especially for a colonoscopy requirement. This has some implications for future 'big data' research, and suggests that we should continue to interpret database studies extracted from

EHRs with caution, particularly without a validation cohort.

Compare our results to these other studies: a study examining medical charts from Massachusetts General Hospital and Brigham and Women's Hospital of 600 patients with at least one ICD-9-CM code for CD confirmed CD in 67.5% of patients. They found evidence to support a diagnosis of UC instead of CD in 11.0% of the remaining 32.5% of patients.[3] These authors included as positives patients with EHRs that included multiple references to having CD without an endoscopic confirmation. In our study, we often found intestinal conditions or non-specific radiographs suggestive of CD (ie, thickening on CT) but endoscopic or pathology evidence was non-specific or supported a related diagnosis (ie, eosinophilic gastrointestinal disease). Additionally, our study had relatively few patients with UC; this was not surprising given we excluded patients with any ICD-9-CM codes for UC for increased CD specificity. A study of the Manitoba Health database used administrative case definitions and found a 91.3% specificity comparing to a self-report questionnaire of patients and a 93.7% specificity compared with a chart review gold standard.[12] A study of the General Practice Research Database to validate the diagnosis of CD using OXMIS codes and surveying general practitioners to confirm these diagnoses categorised 86% of 49 patients identified by EHR as having CD.[13] A study of the Kaiser Permanente membership randomly selected 2325 patients with at least two outpatient or inpatient ICD-9-CM codes for CD (ie, 555.x), and confirmed CD in 88% of patients with chart review.[14] These authors included those with radiological evidence of CD without confirmation with endoscopy. Another study identified patients with IBD using an endoscopy database, and found that an ICD-9-CM diagnostic code for IBD in addition to two medical contacts in the Alberta's Ambulatory Care Classification System yielded 97.4% PPV for IBD.[15] This study began with patients who were undergoing endoscopy with an ICD-9-CM code for IBD, so presumably the patients were starting with endoscopic confirmation. The study that correlates with our findings the best is a study that analysed algorithms to predict diagnosis of CD from discharge and billing data in two large cohorts of Ontario patients which required five physician contacts in 4 years listing IBD in discharge coding to achieve 81.4% PPV for predicting IBD.[16]

Our study has many strengths. The military health system is a single payer system, so all pathology specimen data for patients during their active duty time were available for analysis. In addition, all endoscopies and biopsies done while the patient was active were available. Rather than having medical billers analyse charts, all 400 charts were analysed by gastroenterologists with specialty training and interest in IBD, likely increasing the reliability of confirmation. We only confirmed patients who had endoscopic/surgical and pathological evidence of CD; this improved the reliability of our findings, but had a negative effect on our sensitivity. We only included those patients with available data on active duty both before and after diagnosis of CD, which may limit generalisability to other EHR systems. A drawback of previous studies is that many cases had long-standing IBD with the diagnosis occurring years before their entry into an evaluated database or health system. As noted, a history of IBD (or chronic intestinal maladies such as chronic diarrhoea) is disqualifying for enlistment and commissioning in the US Armed Forces. This study represents the first evaluation of CD in subjects who have all had their first CD diagnosis in the same EHR. One may have expected our study to find a higher sensitivity than reported by others, since physicians often bill patients from prior evaluations or from the notes of their previous physicians without supporting documentation.

The study has some limitations. In addition, use of codes and EHR databases for research can be affected by misclassification, given that ICD-9-CM codes (and most EHRs) do not have 'rule out' or 'presumed diagnosis' codes. This can affect the use of 'big data' to assess healthcare outcomes in patients identified with CD based on ICD-9-CM codes. In contrast to other studies, if information was available from radiology reports but no endoscopy and pathology information was available, the case was not considered a confirmed diagnosis. We also had to exclude 16 patients due to a lack of reviewable encounters despite billing codes for CD. This may be due to patients being evaluated at clinics billing TRICARE without AHLTA access or during a period when AHLTA was unavailable.

In summary, our study shows the poor specificity and PPV of two ICD9 billing codes for CD, and their significant increase when multiple appropriate ICD9 codes made during a specialist encounter or a colonoscopy procedure code are added to the case definition. To some extent, this should not be surprising as medical providers often give a billing code based on the 'working' or 'historical' diagnosis as opposed to the confirmed diagnosis. We urge our fellow researchers to include validation of billing codes when reporting results from EHR or other database-based research.

**Author affiliations**
[1]Gastroenterology Service, Department of Internal Medicine, Walter Reed National Military Medical Center, Bethesda, Maryland, USA
[2]Uniformed Services University, Bethesda, MD, United States
[3]Harvey M. and Lyn P. Meyerhoff Inflammatory Bowel Diseases Center, Gastroenterology Division, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[4]Department of Medicine, MedStar Union Memorial Hospital, Baltimore, Maryland, USA
[5]Gastroenterology Service, Department of Internal Medicine, US Naval Hospital Jacksonville, Jacksonville, Florida, USA
[6]Regional GI, Lancaster, Pennsylvania, USA
[7]Division of Gastroenterology and Hepatology, Crohns and Colitis Center of New Jersey, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA
[8]Department of Genetics, Rutgers University, Piscataway, New Jersey, USA

**Contributors** MS: chart review, analysis, drafting of the manuscript, and guarantor of the article. SH: study design, analysis, manuscript revision, critical review of

the manuscript. EAK: analysis, critical review of the manuscript. BR: chart review, critical review of the manuscript. JB: study design, manuscript revision. SRB: study design, interpretation of results and analysis, manuscript revision, critical review of the manuscript.

**Funding** SH and SRB were supported in part by Congressionally Directed Medical Research Program Grant PR110833.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study was approved by the Institutional Review Boards of all involved institutions and complied with the highest standards of ethical research.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information. The data are collated from the military's Electronic Health Record.



**ORCID iDs**
Manish Singla http://orcid.org/0000-0002-0445-7134
Susan Hutfless http://orcid.org/0000-0002-6311-2611
Steven R Brant http://orcid.org/0000-0003-3019-4240

## REFERENCES

1 Silverberg MS, Satsangi J, Ahmad T, *et al*. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;19:5A–36.
2 Jakobsson GL, Sternegård E, Olén O, *et al*. Validating inflammatory bowel disease (IBD) in the Swedish national patient register and the Swedish Quality Register for IBD (SWIBREG). *Scand J Gastroenterol* 2017;52:216–21.
3 Ananthakrishnan AN, Cai T, Savova G, *et al*. Improving case definition of Crohn's disease and ulcerative colitis in electronic medical records using natural language processing: a novel informatics approach. *Inflamm Bowel Dis* 2013;19:1411–20.
4 Zhu W, Zeng N, Wang N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG proceedings: health care and life sciences*, Baltimore, Maryland, 2010:1–9.
5 Buckley JP, Kappelman MD, Allen JK, *et al*. The burden of comedication among patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2013;19:2725–36.
6 Singh S, Heien HC, Sangaralingham LR, *et al*. Comparative Effectiveness and Safety of Anti-Tumor Necrosis Factor Agents in Biologic-Naive Patients With Crohn's Disease. *Clin Gastroenterol Hepatol* 2016;14:1120–9.
7 Arora G, Singh G, Vadhavkar S, *et al*. Incidence and risk of intestinal and extra-intestinal complications in Medicaid patients with inflammatory bowel disease: a 5-year population-based study. *Dig Dis Sci* 2010;55:1689–95.
8 Akhuemonkhan E, Parian A, Miller K, *et al*. Prevalence and screening for anaemia in mild to moderate Crohn's disease and ulcerative colitis in the United States, 2010-2014. *BMJ Open Gastroenterol* 2017;4:e000155.
9 Betteridge JD, Armbruster SP, Maydonovitch C, *et al*. Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. *Inflamm Bowel Dis* 2013;19:1421–7.
10 Hou JK, Tan M, Stidham RW, *et al*. Accuracy of diagnostic codes for identifying patients with ulcerative colitis and Crohn's disease in the Veterans Affairs health care system. *Dig Dis Sci* 2014;59:2406–10.
11 Thirumurthi S, Chowdhury R, Richardson P, *et al*. Validation of ICD-9-CM diagnostic codes for inflammatory bowel disease among veterans. *Dig Dis Sci* 2010;55:2592–8.
12 Bernstein CN, Blanchard JF, Rawsthorne P, *et al*. Epidemiology of Crohn's disease and ulcerative colitis in a central Canadian province: a population-based study. *Am J Epidemiol* 1999;149:916–24.
13 Lewis JD, Brensinger C, Bilker WB, *et al*. Validity and completeness of the general practice research database for studies of inflammatory bowel disease. *Pharmacoepidemiol Drug Saf* 2002;11:211–8.
14 Liu L, Allison JE, Herrinton LJ. Validity of computerized diagnoses, procedures, and drugs for inflammatory bowel disease in a northern California managed care organization. *Pharmacoepidemiol Drug Saf* 2009;18:1086–93.
15 Rezaie A, Quan H, Fedorak RN, *et al*. Development and validation of an administrative case definition for inflammatory bowel diseases. *Can J Gastroenterol* 2012;26:711–7.
16 Benchimol EI, Guttmann A, Mack DR, *et al*. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. *J Clin Epidemiol* 2014;67:887–96.