Identifying Patients With Inflammatory Bowel Diseases in an Administrative Health Claims Database: Do Algorithms Generate Similar Findings?

Yizhou Ye, PhD[1] iD, Sudhakar Manne, MS[1], and Dimitri Bennett, MD[1,2] iD

Abstract

Application of selective algorithms to administrative health claims databases allows detection of specific patients and disease or treatment outcomes. This study identified and applied different algorithms to a single data set to compare the numbers of patients with different inflammatory bowel disease classifications identified by each algorithm. A literature review was performed to identify algorithms developed to define inflammatory bowel disease patients, including ulcerative colitis, Crohn's disease, and inflammatory bowel disease unspecified in routinely collected administrative claims databases. Based on the study population, validation methods, and results, selected algorithms were applied to the Optum Clinformatics® Data Mart database from June 2000 to March 2017. The patient cohorts identified by each algorithm were compared. Three different algorithms were identified from literature review and selected for comparison (A, B, and C). Each identified different numbers of patients with any form of inflammatory bowel disease (323 833; 246 953, and 171 537 patients, respectively). The proportions of patients with ulcerative colitis, Crohn's disease, and inflammatory bowel disease unspecified were 32.0% to 47.5%, 38.6% to 43.8%, and 8.7% to 26.6% of the total population with inflammatory bowel disease, respectively, depending on the algorithm applied. Only 5.1% of patients with inflammatory bowel disease unspecified were identified by all 3 algorithms. Algorithm C identified the smallest cohort for each disease category except inflammatory bowel disease unspecified. This study is the first to compare numbers of inflammatory bowel disease patients identified by different algorithms from a single database. The differences between results highlight the need for validation of algorithms to accurately identify inflammatory bowel disease patients.

Keywords

administrative claims database, Crohn's disease, inflammatory bowel diseases, observational study, retrospective studies, ulcerative colitis

---

What do we already know about this topic?

Using validated case definitions in secondary databases is critical for ensuring comparability between retrospective observational studies. To our knowledge, no study has directly compared the differences in IBD patients identified using different algorithms in the same claims database.

How does your research contribute to the field?

By investigating difference between numbers of identified IBD patients with different algorithms in the same database, this study further quantitatively demonstrates how much different disease definitions can affect patient cohort definition, incidence and prevalence calculation, disease burden estimation, and health outcomes assessment in descriptive studies.

What are your research's implications toward theory, practice, or policy?

Our research implicates the importance of maintaining good practice in using real-world data to generate real-world evidence and we strongly urge fellow real-world data scientists to consider the strengths and limitations of each patient definition used thoroughly.

Introduction

Real-world data (RWD) is a vital source of data that complements those gathered from clinical trials, and is increasingly used to inform health care decision-making processes in health technology assessment, regulatory, clinical practice, and post-marketing lifecycle management settings.[1] Large, longitudinal health care data sets can be derived from a wide range of real-world data sources, such as routine health care electronic medical records (EMRs), administrative health claims databases, registries, surveys, and personal health care devices. Electronic medical records and administrative health claims databases are the most common real-world data sources used to generate RWE for understanding disease burden, studying the use of medications, assessing effectiveness and safety of medication in real-world practice, and identifying unmet medical needs.[2,3] Although administrative health claims databases are not designed for research, the information within them can yield valuable insights for specific research objectives, providing that robust methodology is used to detect patients and monitor their disease outcomes and treatment use.

Specific diseases and medical conditions are classified within administrative health claims databases using diagnostic codes, such as the International Classification of Diseases (ICD) codes.[4] These unique alphanumeric codes can be used alone or combined with claims information to create an algorithm for detecting patients of interest, or selecting disease or treatment outcomes.

Algorithms for a certain disease state or treatment outcome require validation in the context that they are going to be applied. Validation is performed by verifying patients identified using an algorithm against patients identified from medical records to evaluate the capacity of the algorithm for accurate detection. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are commonly used quantitative measurements of this capacity; a high PPV is considered to reflect an efficient and robust tool.[5] Retrospective studies of administrative health claims data should be conducted, when possible, using validated algorithms to ensure quality and comparability with other research. However, validated algorithms may not be available from the literature and studies may make use of different algorithms. Therefore, researchers should establish the best methodology for performing these studies to ensure that robust and relevant research is conducted.

Several observational studies have been performed using routinely collected health data sources such as claims databases in patients with inflammatory bowel disease (IBD), which encompasses disorders characterized by chronic gastrointestinal inflammation, including ulcerative colitis (UC) and Crohn's disease (CD).[6] Various algorithms for the identification of IBD patients from claims-based databases have been employed in these studies[7,8]; however, no study has tested these different algorithms in the same data set to establish their accuracy for detecting patients with UC or CD.[7,9-11] This study identified algorithms through a comprehensive literature review for application to the same database and aimed quantitatively to assess and compare the numbers of IBD patients identified, as well as the different classifications of IBD used.

Methods

Literature Review to Identify Algorithms

A comprehensive literature review was performed in the PubMed database to identify all algorithms in peer-reviewed journal articles used to detect IBD patients. We included studies that met the following criteria: (1) published before September 30, 2017; (2) retrospective; (3) identified IBD patients using claims databases; and (4) published in English. After screening of titles and abstracts, the remaining studies underwent full-text review for final selection of relevant algorithms. Publications from the bibliographies of selected studies were also examined against the above inclusion criteria. Information extracted from identified studies included the operational algorithm and administrative codes used for IBD identification, publication year, study setting, and population. Algorithms to be tested in this study were selected based on their quality, novelty, validation results, and citation numbers.

Data Source

Patients with UC, CD, or inflammatory bowel disease unspecified (IBDU; also known as indeterminate colitis) were identified from the Clinformatics® Data Mart (CDM; Optum, Eden Prairie, MN, USA) database, using the selected algorithms.[12] CDM is a large administrative health claims database that is geographically diverse, and includes patients from all 50 states in the United States. It comprises administrative health claims for patients with a health insurance plan from 1 large, national provider of health care insurance including Medicare Advantage coverage. The database was developed for billing and reimbursement purposes for all

[1]Takeda Pharmaceutical Company Limited, Cambridge, MA, USA
[2]University of Pennsylvania, Philadelphia, USA

Corresponding Author:
Yizhou Ye, Epidemiology, Takeda Pharmaceutical Company Limited, 40 Lansdowne Street, Cambridge, MA 02139, USA.
Email: james.ye@takeda.com

Table 1. Details of Algorithms Identified and Selected for Comparison.

	Algorithm A (Herrinton et al[7])	Algorithm B (McAuliffe et al[10])	Algorithm C (Rezaie et al[11])
Data source and population	• Health Maintenance Organization Research Network Centers for Education and Research in Therapeutics (HMORN CERT) core data set • Privately insured population in the United States	• HealthCore Integrated Research Database • Privately insured population in the United States	• Alberta Health administrative databases • Residents with universal coverage in the Alberta Health Care Insurance Plan in Alberta, Canada
IBD case definition and classification	*UC* Having ≥2 UC diagnoses within any 30-month period of the available data, any setting, any location *CD* Having ≥2 CD diagnoses within any 30-month period of the available data, any setting, any location *IBDU* (any of the two) • Having ≥2 diagnoses of UC AND ≥2 diagnoses of CD • Having ≥2 pharmacy dispensing for 5-aminosalicylates with no recorded IBD diagnosis	*UC* Having ≥2 diagnoses of UC at least 30 days apart at any time, any setting, any location *CD* Having ≥2 diagnoses of CD at least 30 days apart at any time, any setting, any location *IBDU*[a] Having a diagnosis of both UC and CD, any setting, any location	Having ≥2 hospitalizations, or ≥4 physician office visits, or ≥2 ambulatory care visits with a diagnosis of UC or CD in any 2-year period, any location ***IBD score calculation*** IBD score is the sum of all scores: Each hospitalization for UC and CD is given a score of +1 and −1, respectively Each ambulatory care visit for UC and CD is given a score of +1 and −1, respectively Each physician office visit for UC and CD is given a score of +0.4 and −0.4, respectively ***IBD classification*** UC: IBD score > 2 CD: IBD score < 2 IBDU: −2 ≤ IBD score ≤ 2
Disease index date	The first of either the qualifying date of IBD diagnosis or prescription dispensing	The later qualifying date of IBD diagnosis	The earliest among all eligible dates of service in claims

Note. IBD = inflammatory bowel disease; UC = ulcerative colitis; CD = Crohn's disease; IBDU = inflammatory bowel disease unspecified.
[a]Patients with diagnosis of both UC and CD were excluded from the original study to reduce disease misclassification. This category has been created for comparison purposes.

facility and physician claims in all settings including outpatient visits, hospitalizations, ambulatory care visits, emergency department visits, and so on. The CDM database encompasses 17 to 19 million lives annually and a total of approximately 57 million unique lives over the 10-year period from 2007 to 2017. The demographics of the population is similar to the US general population regarding age and sex: 22% enrollees are 0 to 17 years of age, 41% are 18 to 44 years old, 23% are 45 to 64 years old, and 14% are more than 65 years old; 49% are male and 51% are female.[13] The database contains statistically deidentified data in compliance with the Health Insurance Portability and Accountability Act (HIPAA), including demographic, diagnostic, and procedural information in the medical claims, and outpatient dispensed medications in the pharmacy claims. An important advantage of claims databases is complete data from a large number of participants, which provides the capability to study medication utilization longitudinally.[14] Eligibility files, medical, and pharmacy claims data from June 1, 2000 to March 31, 2017 were used for this study.

Application of Algorithms

Each selected algorithm was applied to the CDM database to identify cohorts of patients with any form of IBD, UC,

CD, or IBDU. Patient cohorts were examined to evaluate the numbers of patients identified by each algorithm for each disease classification, and whether patients were identified by more than 1 algorithm or exclusively by a single algorithm. A set of ICD-9 (555.x, 556.x) and ICD-10 (K50.x, K51.x) codes in any position were used to aid identification of IBD patients in the database (Table 1). Venn diagrams were created for a visual representation of the size of the cohorts identified by each algorithm, including patients identified by single or multiple algorithms. Cohen Kappa statistic has been used to measure agreement between algorithms selected.[15]

Results

Literature Review

Three appropriate algorithms were selected from the literature review (Table 1). Algorithm A, developed by Herrinton et al, was based on medical claims with diagnosis codes for CD and UC, as well as pharmacy claims for IBD-specific medications to aid the identification of patients with IBDU.[7] Algorithm B was selected from a paper by McAuliffe et al[10] and was based on patient medical claims with diagnosis codes of UC or CD to identify IBD patients.
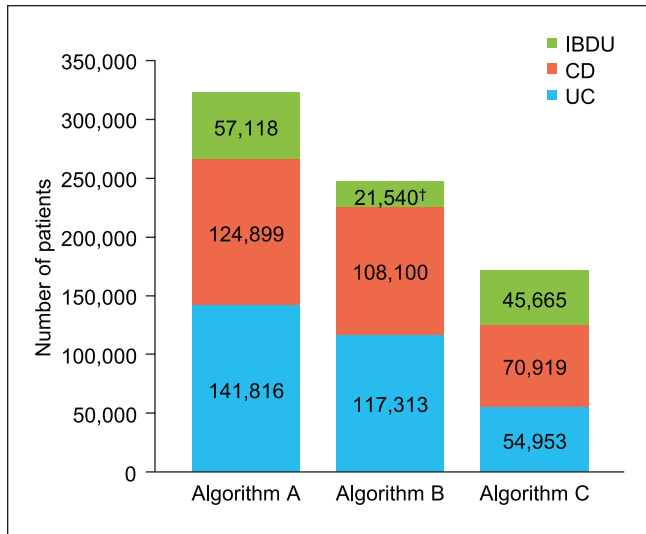
Figure 1. Comparison of numbers of IBD patients identified by each algorithm.
Note. IBDU = inflammatory bowel disease unspecified; CD = Crohn's disease; UC = ulcerative colitis.
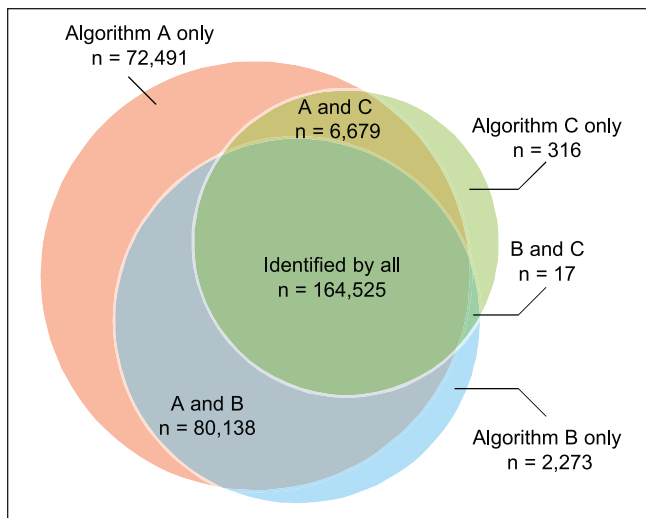

Figure 2. Venn diagram showing the numbers and overlap of IBD patients identified by the three algorithms.

Algorithm C was taken from a study by Rezaie et al,[11] who developed and validated more than 150 case definitions with an aim to provide a validated administrative case definition for IBD. This algorithm was based on health care utilization captured in medical claims to identify IBD patients. Algorithms A and C have both been validated by chart review (although this was in a different population from the CDM database) and have PPVs of 81% and 97%, respectively.[7,11] Algorithm B was initially developed using algorithm A and then adapted for use in another privately insured US administrative health claims database. It has not been validated using any data source.
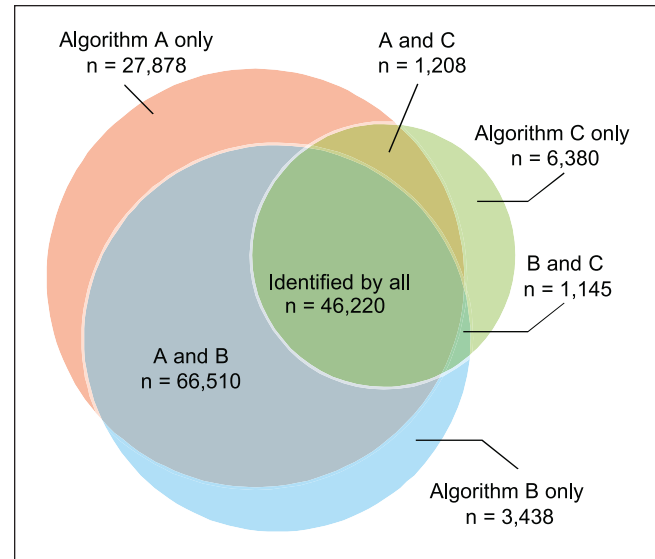

Figure 3. Venn diagram showing the numbers and overlap of UC patients identified by the three algorithms.

Patient Identification Using Individual Algorithms

The largest number of patients was identified using algorithm A, which identified 323 833 patients with any form of IBD. Of these, 141 816 patients had UC (43.8%), 124 899 had CD (38.6%), and 57 118 patients had IBDU (17.6%; Figure 1). Application of algorithm B captured 246 953 patients with any form of IBD. This total comprised 117 313 patients with UC (47.5%), 108 100 with CD (43.8%), and 21 540 patients with IBDU (8.7%; Figure 1). It should be noted that algorithm B originally excluded patients with diagnoses of both UC and CD, but these patients were retained in our study to create an IBDU group for comparison. The smallest number of patients was identified using algorithm C, which retrieved 171 537 patients with any form of IBD, including 54 953 patients with UC (32.0%), 70 919 with CD (41.3%), and 45 665 patients with IBDU (26.6%; Figure 1).

The proportion of patients with UC was 32.0% to 47.5% of the overall IBD population across the algorithms, depending on the algorithm applied, whereas the corresponding proportions of patients with CD or IBDU were 38.6% to 43.8% and 8.7% to 26.6% of the total population with IBD, respectively.

Patient Identification: Overlap Across the Algorithms

A total of 326 439 patients with any form of IBD were identified from the database using the combination of all 3 algorithms, with algorithms A, B, and C identifying 99.2%, 75.7%, and 52.5% of these patients, respectively (Figure 2). Approximately half of the identified patients were captured by all 3 algorithms (164 525 patients, 50.4%). In all, 50.8% of the patients with any form of IBD identified by
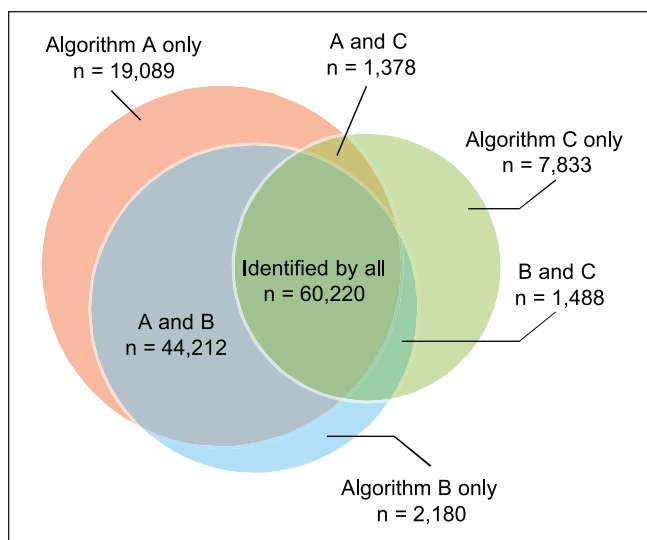
Figure 4. Venn diagram showing the numbers and overlap of CD patients identified by the three algorithms.
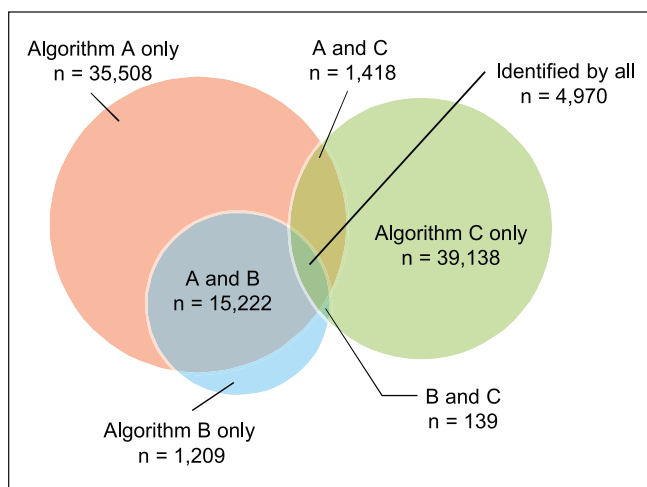

Figure 5. Venn diagram showing the numbers and overlap of IBDU patients identified by the three algorithms.

algorithm A were also captured by the other 2 algorithms, compared with 66.6% of patients detected by algorithm B that were also identified by algorithms A and C, and 95.9% of patients by algorithm C which were also detected by algorithms A and B. Algorithm A exclusively identified 72 491 patients (22.4% of all patients identified by algorithm A for this category) not captured by any other algorithm, compared with 2273 patients (0.9%) by algorithm B and 316 patients (0.2%) by algorithm C.

In total, 152 779 patients with UC were identified by any of the algorithms, with algorithms A, B, and C identifying 92.8%, 76.8%, and 36.0% of these patients, respectively (Figure 3). A cohort of 46 220 patients (30.3%) was captured by all 3 algorithms. Algorithm A exclusively

identified 27 878 patients (8.6% of all patients identified by algorithm A for this category) not captured by any other algorithm, compared with 3438 patients (1.4%) by algorithm B and 6380 patients (3.7%) by algorithm C.

In total, 136 400 patients with CD were identified by any of the algorithms, with 91.6% of these being identified by algorithm A and 79.2% and 52.0% detected by algorithms B and C, respectively (Figure 4). Of this total, 60 220 patients (44.1%) were identified by all 3 algorithms; algorithm A identified the lowest proportion relative to its total for this category (48.2%), followed by algorithm B (55.7%) and algorithm C (84.9%). Algorithm A exclusively identified 19 089 patients (15.3% of all patients identified by algorithm A for this category), compared with 2180 patients (2.0%) being exclusively identified by algorithm B and 7833 patients (11.0%) by algorithm C.

For IBDU, 97 604 patients were identified by any of the algorithms, with the largest proportion of these being identified by algorithm A (58.5%), followed by 46.8% for algorithm C and 22.1% for algorithm B (Figure 5). The total proportion of patients captured by all 3 algorithms was much lower than that for the other categories (5.1%, 4970 patients). Of this shared cohort, only 8.7% contributed to the total identified by algorithm A for this category, versus 10.9% by algorithm C and 23.1% by algorithm B. Algorithm A exclusively identified 35 508 patients (62.2% of all patients identified by algorithm A for this category) not retrieved by any other algorithm, compared with 1209 patients (5.6%) by algorithm B and 39 138 patients (85.7%) by algorithm C.

Kappa measures have been explored to study measures of agreement between the 3 algorithms. Results showed a simple Kappa estimated value of 0.949 (95% confidence interval [CI], 0.947-0.950), indicating a strong agreement between algorithms A and B. A Kappa value of 0.494 (95% CI, 0.490-0.497) was observed for algorithms A and C, indicating a moderate agreement. Similarly, algorithms B and C yielded a similar Kappa value of 0.5184 (95% CI, 0.515-0.525), suggesting a moderate agreement as well.

Discussion

This study quantitatively investigated different claims-based algorithms to detect IBD patients in the same US-based claims database, to evaluate the differences in the numbers of patients identified with any form of IBD, UC, CD, or IBDU. There are several other algorithms that have been validated and applied to other health administrative databases to identify IBD patients.[9,16] For algorithm selection, we considered differences and similarities between various validation populations and our study population to ensure that selected algorithms are representative of all validation studies. Algorithms validated in population largely different from privately insured US population were therefore ruled out for consideration. Algorithm A was validated in a regional comprehensive

health records database containing both administrative claims and electronic health records. Algorithm B was used in another US national administrative health claims database which is the most similar to our study population. Last but not the least, algorithm C was validated in a single-payer universal health system in 1 Canadian province. Algorithm A by Herrinton et al demonstrated the most lenient criteria and identified the largest patient cohorts for each disease category, whereas algorithm C by Rezaie et al had high sensitivity and specificity and proved to be the most rigorous algorithm, consistently identifying the smallest patient cohort in each disease category except IBDU. The largest overlap between captured patient cohorts was seen with algorithms A and B, whereas algorithm C had the smallest overlap. This demonstrates that even validated algorithms can result in the selection of different patient cohorts from the same data set.

Algorithm A was the most lenient, with the least restrictions among the 3 algorithms included. Unsurprisingly, this algorithm picked up the most patients. The time between diagnoses is not specified in the criteria of this algorithm, and this may have resulted in the inclusion of patients that had been misclassified as having UC or CD.

Algorithm B by McAuliffe et al, adapted from algorithm A, is refined by the requirement of at least 2 diagnoses of UC or CD separated by at least 30 days. The 2 outpatient diagnoses more than 30 days apart is a commonly used method for identifying chronic conditions in claims databases due to the potential for clustering of care immediately following the first diagnosis.[17,18] As anticipated, there was a large degree of overlap between the patient cohorts captured by these 2 algorithms. The timeframe specified between diagnoses as part of the algorithm B criteria may have helped exclude patients who had only a "rule-out" diagnosis recorded in their claims; it may also have increased the accuracy of detection by filtering out patients receiving multiple diagnoses as part of their initial presentation, thus reducing the number of potentially misclassified patients. This timeframe of separation between diagnoses is commonly used in algorithms to identify patients with other diseases in administrative health claims databases, and these algorithms have been shown to have greater accuracy for identifying patients than those with unspecified timeframes.[19,20]

Algorithm C identified the smallest IBD cohorts and can be considered the most rigorous and conservative algorithm. It differs from the other algorithms because it includes more stringent criteria, based on health care utilization, including numbers of hospitalizations, physician office visits, and ambulatory care visits, and requires the generation of a patient "IBD score" to distinguish between UC and CD. This algorithm not only addresses the clustering of care but also evaluates a patient's claims comprehensively to address for potential diagnosis switch from CD to UC. It has higher specificity and PPV compared with algorithm A (about 80%)[7] and can be assumed to identify IBD patients with more certainty than the other 2 algorithms, resulting in a lower number of patients inappropriately identified as having IBD. This high specificity is achieved by sacrificing sensitivity (although specificity was still more than 80%) and generalizability. The patient cohorts identified are different to the patients captured by the other 2 algorithms, exemplified by the reduced overlap between the patient cohorts identified by this algorithm and the other 2 algorithms. The stringent inclusion criteria may also select for patients with more severe and active IBD who require regular health care visits. As a result, algorithm C may be useful for performing studies in large data sets in which patients with confirmed disease and potentially more severe disease are of interest.

The algorithms differed greatly in the criteria used to define IBDU, and there were large differences in the patient cohorts identified for this disease category. Of the 97 604 patients identified by any of the algorithms, only 4970 patients (5.1%) were identified by all 3 algorithms. This inconsistency was also illustrated in the original Herrinton et al study, where the application of the single algorithm to a range of US health plan data sets retrieved different percentages of patients with IBDU.[7] The results of this study highlight how each data source and the behavior of each algorithm can differ, and the necessity of selecting an appropriate algorithm for use in different data sources. Using pharmacy information alone to identify IBD patients, without access to medical records detailing a specific diagnosis or IBD classification, may increase the number of patients recorded as having IBDU. Clinical notes are not available in health claims data, which can lead to further difficulty in identifying appropriate patients with certainty.[12]

Variation in real-world settings may also contribute to discrepancies in data. Inconsistencies in clinical practice regarding IBD classification and treatment may lead to irregular coding across different points of health care contact and can result in 1 patient being registered with multiple diagnoses.[12] Variation in characteristics across the patient population within the database, such as age or disease severity, may also affect the sensitivity of the algorithms. For example, older patients may have multiple comorbidities and require increased health care interventions for reasons other than IBD, which may lead to the appropriate IBD code potentially being omitted on patient documentation as a result. The variation in the ability of the algorithms to identify patients with IBDU indicates that caution should be used when applying these algorithms, and, without further research on this topic, only algorithms for confirmed UC and CD or any form of IBD should be used to ensure greater certainty of detection.[12]

No studies have previously compared different claims-based algorithms for identifying IBD patients in the same database. The findings of this study help establish standardized methodology for conducting analyses of claims-based data on IBD, as well as helping interpret the validity and results for observational studies in IBD. In addition, this study supports and highlights the importance of

defining disease in determining the outcome for prevalence estimation and also shows that modification in case definitions can result in considerable variance in prevalence estimates. In the area of observational research, we know that prevalence estimates are a cornerstone of health epidemiology and can provide critical RWE that impacts numerous aspects of the health care decision-making process.[21,22] The number of patients with IBDU found by each algorithm may relate to newly developed or incident cases of IBD, which may explain the lack of clarity around official diagnoses. Equally, patients with confirmed diagnoses of CD or UC may feature in the number of IBD patients.

This study has several strengths. We used a large US longitudinal population–based data source that contained many IBD patients and provided sufficiently sized patient cohorts for comparison of algorithms. Importantly, the population in this database was diverse, spanning all 50 US states. In addition, this data source can be considered typical of a large proprietary claims database, in which such research would be performed, and is representative of a privately insured employed US population.

Conversely, we also note several limitations to our study. The database used was claims based and therefore this study is subject to the same limitations as other studies using similar data sources, such as potential misclassification due to inaccurate and inconsistent coding practice.[23] In addition, algorithm A was validated in Kaiser Permanent database which contains both administrative health claims and electronic health records,[7] whereas algorithm C was validated in a single-payer universal health system in 1 Canadian province.[11] As we know that claims data were collected for billing and reimbursement purposes, algorithms might perform differently in populations with different practice, billing, and coding patterns. However, it is not uncommon to conduct high-quality RWE studies using algorithms validated in different but similar populations.[10,24-28] Another limitation of this study is the generalizability outside of the insured population, as elderly and low-income populations are underrepresented, and its results may not be entirely applicable to countries other than the United States. Also, we did not include all published IBD algorithms in our study, but we wanted to open a dialog for further similar studies where all IBD algorithms can be compared and included. Our initial study is a first step in that direction. Finally, without confirmation from EMR review, we were unable to compare the validity of the 3 algorithms and identify one that best suits this database. However, we have demonstrated that algorithms can be reliable tools for performing real-world observational studies and provide valuable insights for future research. Procedure codes and pharmacy claims data are not often included in these algorithms. We recommend future studies to explore the inclusion of this valuable information in claims databases. Finally, the difference between primary and secondary diagnoses is worthy of further investigation in claims database methodology research.

This study did not aim to determine the superiority of the selected algorithms, but to demonstrate the capability of these algorithms in real-world observational studies by comparing the numbers of IBD patients identified from a single database, and emphasizes the importance of appropriate methods for identifying patients from a database. Although the algorithms tested are robust tools for performing observational studies in IBD, the choice of algorithm depends on the desired population of interest or the type of outcomes to be analyzed. The algorithms should be validated in identical or similar populations to those in which research will be performed to ensure that observational studies are conducted as robustly as possible, and with limited bias.

Author Contributions

Y.Y., S.M., and D.B. contributed to conception and design. Y.Y. and S.M. conducted the analysis. All authors made substantial contributions to the interpretation of data and drafting of the article. All authors reviewed and revised the article critically for crucial intellectual content and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: All authors are employees of Takeda Pharmaceutical Company, Ltd. This study did not relate to any investigational products or have any patient contact that necessitated ethical approval at the national or institutional level.

ORCID iDs

Yizhou Ye [iD] https://orcid.org/0000-0002-9090-9431
Dimitri Bennett [iD] https://orcid.org/0000-0002-8387-9342

References

1. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence—what is it and what can it tell us. *N Engl J Med*. 2016;375(23):2293-2297.
2. Yuan H, Ali MS, Brouwer ES, et al. Real-world evidence: what it is and what it can tell us according to the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG). *Clin Pharmacol Ther*. 2018;104(2):239-241. doi:10.1002/cpt.1086.

3. Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials. *Clin Pharmacol Ther*. 2017;102(6):924-933. doi:10.1002/cpt.857.

4. World Health Organization. Classification of Diseases (ICD). http://www.who.int/classifications/icd/en/. Accessed May 23, 2018.

5. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol*. 2008;56(1):45-50.

6. Cho JH, Abraham C. Inflammatory bowel disease genetics: Nod2. *Annu Rev Med*. 2007;58:401-416.

7. Herrinton LJ, Liu L, Lafata JE, et al. Estimation of the period prevalence of inflammatory bowel disease among nine health plans using computerized diagnoses and outpatient pharmacy dispensings. *Inflamm Bowel Dis*. 2007;13(4):451-461. doi:10.1002/ibd.20021.

8. Loftus EV Jr. The burden of inflammatory bowel disease in the United States: a moving target. *Clin Gastroenterol Hepatol*. 2007;5(12):1383-1384. doi:10.1016/j.cgh.2007.10.016.

9. 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(8):887-896. doi:10.1016/j.jclinepi.2014.02.019.

10. McAuliffe ME, Lanes S, Leach T, et al. Occurrence of adverse events among patients with inflammatory bowel disease in the HealthCore Integrated Research Database. *Curr Med Res Opin*. 2015;31(9):1655-1664. doi:10.1185/03007995.2015.1065242.

11. Rezaie A, Quan H, Fedorak RN, Panaccione R, Hilsden RJ. Development and validation of an administrative case definition for inflammatory bowel diseases. *Can J Gastroenterol*. 2012;26(10):711-717. doi:10.1155/2012/278495.

12. Gunaseelan V, Kenney B, Lee JS, Hu HM. Invited commentary: databases for surgical health services research: Clinformatics Data Mart. *Surgery*. 2018;165:669-671.

13. US Census Bureau. Age and sex composition: 2010. https://www.census.gov/content/dam/Census/library/publications/2011/dec/c2010br-03.pdf. Published 2011. Accessed September 24, 2019.

14. Diehr P, Yanez D, Ash A, Hornbrook M, Lin DY. Methods for analyzing health care utilization and costs. *Annu Rev Public Health*. 1999;20:125-144. doi:10.1146/annurev.publhealth.20.1.125.

15. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257-268.

16. Thirumurthi S, Chowdhury R, Richardson P, Abraham NS. Validation of ICD-9-CM diagnostic codes for inflammatory bowel disease among veterans. *Dig Dis Sci*. 2010;55(9):2592-2598. doi:10.1007/s10620-009-1074-z.

17. Sumner W, Stwalley DL, Asaro PV, Hagen MD, Olsen MA. Adding flexible temporal constraints to identify chronic comorbid conditions in ambulatory claims data. *AMIA Annu Symp Proc*. 2014;2014:1088-1097.

18. Nickel KB, Wallace AE, Warren DK, Ball KE, Mines D, Fraser VJ, Olsen MA. Modification of claims-based measures improves identification of comorbidities in non-elderly women undergoing mastectomy for breast cancer: a retrospective cohort study. *BMC Health Serv Res*. 2016;16(a):388. doi:10.1186/s12913-016-1636-7.

19. Whyte JE-NNM, Teitelbaum A, Gomez Rey G, Kallich JD. An evaluation of algorithms for identifying metastatic breast, lung, or colorectal cancer in administrative claims data. *Medical Care*. 2015;53(7):e49-e57.

20. Capkun G, Lahoz R, Verdun E, et al. Expanding the use of administrative claims databases in conducting clinical real-world evidence studies in multiple sclerosis. *Curr Med Res Opin*. 2015;31(5):1029-1039. doi:10.1185/03007995.2015.1014029.

21. Luther J, Dave M. Rising inflammatory bowel disease prevalence highlights the need for effective, cost-effective therapies. *Inflamm Bowel Dis*. 2019;izz203. doi:10.1093/ibd/izz203.

22. Ye Y, Manne S, Treem WR, Bennett D. Prevalence of inflammatory bowel disease in pediatric and adult populations: recent estimates from large national databases in the United States, 2007-2016. *Inflamm Bowel Dis*. 2019. doi:10.1093/ibd/izz182.

23. Ladha KS, Eikermann M. Codifying healthcare—big data and the issue of misclassification. *BMC Anesthesiol*. 2015;15:179. doi:10.1186/s12871-015-0165-y.

24. Kappelman MD, Rifas-Shiman SL, Kleinman K, et al. The prevalence and geographic distribution of Crohn's disease and ulcerative colitis in the United States. *Clin Gastroenterol Hepatol*. 2007;5(12):1424-1429. doi:10.1016/j.cgh.2007.07.012.

25. Kappelman MD, Moore KR, Allen JK, Cook SF. Recent trends in the prevalence of Crohn's disease and ulcerative colitis in a commercially insured US population. *Dig Dis Sci*. 2013;58(2):519-525. doi:10.1007/s10620-012-2371-5.

26. Chen C, Hartzema AG, Xiao H, et al. Real-world pattern of biologic use in patients with inflammatory bowel disease: treatment persistence, switching, and importance of concurrent immunosuppressive therapy. *Inflamm Bowel Dis*. 2019. doi:10.1093/ibd/izz001.

27. Hutfless S, Fireman B, Kane S, Herrinton LJ. Screening differences and risk of cervical cancer in inflammatory bowel disease. *Aliment Pharmacol Ther*. 2008;28(5):598-605. doi:10.1111/j.1365-2036.2008.03766.x.

28. Caldera F, Saha S, Wald A, et al. Comparing guideline-based care quality for inflammatory bowel disease and rheumatoid arthritis patients within a medical home. *Expert Rev Gastroenterol Hepatol*. 2016;10(6):759-766. doi:10.1586/17474124.2016.1169920.