



**AALBORG UNIVERSITY**  
DENMARK

**Aalborg Universitet**

## **Validity of Inflammatory Bowel Disease Diagnoses in the Danish National Patient Registry**

*A Population-Based Study from the North Denmark Region*

Albaek Jacobsen, Henrik; Jess, Tine; Larsen, Lone

*Published in:*  
Clinical Epidemiology

*DOI (link to publication from Publisher):*  
[10.2147/CLEP.S378003](https://doi.org/10.2147/CLEP.S378003)

*Creative Commons License*  
CC BY-NC 4.0

*Publication date:*  
2022

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Albaek Jacobsen, H., Jess, T., & Larsen, L. (2022). Validity of Inflammatory Bowel Disease Diagnoses in the Danish National Patient Registry: A Population-Based Study from the North Denmark Region. *Clinical Epidemiology*, 14, 1099-1109. <https://doi.org/10.2147/CLEP.S378003>

### **General rights**


Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### **Take down policy**

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# Validity of Inflammatory Bowel Disease Diagnoses in the Danish National Patient Registry: A Population-Based Study from the North Denmark Region

Henrik Albaek Jacobsen <sup>1,2</sup>, Tine Jess<sup>1,2</sup>, Lone Larsen<sup>1,2</sup>

<sup>1</sup>Department of Gastroenterology and Hepatology, Aalborg University Hospital, Aalborg, Denmark; <sup>2</sup>Center for Molecular Prediction of Inflammatory Bowel Disease, PREDICT, Department of Clinical Medicine, Aalborg University, Copenhagen, Denmark

Correspondence: Henrik Albaek Jacobsen, National Center of Excellence PREDICT, Department of Gastroenterology and Hepatology, Aalborg University Hospital, Mølleparkvej 4, Aalborg, 9000, Tel +45 97656267, Email henrik.jacobsen@rn.dk

**Purpose:** The Danish National Patient Registry (DNPR) is recognized for providing high-quality data. However, only a few minor studies have validated inflammatory bowel disease (IBD) diagnoses in the DNPR, reporting various degrees of validity. To pave the way for large-scale studies of IBD in Denmark, we aimed to investigate the validity of IBD among >8000 patients registered in the DNPR between 2002 and 2020 in the North Denmark Region.

**Patients and Methods:** To evaluate the reliability of the diagnoses in the DNPR, we initially compared all patients registered with one IBD diagnosis during 2002–2020 to a list of already verified patients in the regional IBD database GASTROBIO. Medical records on all DNPR registered patients not on the list were manually reviewed by a gastroenterologist to verify or dismiss the IBD diagnosis. Positive predictive values (PPV) were calculated.

**Results:** Of 8040 patients with at least one IBD diagnosis in DNPR, 5263 were already confirmed cases, leaving 2777 for medical record evaluation, of whom 849 had IBD. In total, 6112 were correctly registered with IBD based on one diagnosis, and 1343 were incorrectly registered, resulting in a PPV of 0.82 (95% CI, 0.81–0.83). For patients registered with at least two diagnoses, the PPV was 0.95 (95% CI, 0.95–0.96), and with at least three diagnoses, the PPV was 0.98 (95% CI, 0.98–0.99). Results were similar for UC and CD separately. Of note, the completeness of valid cases went from 6112 to 4606 (75%; 95% CI, 74%–76%) when demanding at least two registered diagnoses and to 3320 (54%; 95% CI, 53%–56%) when demanding at least three registered diagnoses.

**Conclusion:** Reassuringly, the validity of IBD diagnoses in DNPR is high, especially for patients registered more than once. However, the reduced completeness when applying a true case definition of at least two registered diagnoses should be considered.

**Keywords:** inflammatory bowel disease, Danish National Patient Registry, validation, diagnosis codes, positive predictive value, completeness

## Introduction

The Danish National Patient Registry (DNPR) is recognized for enabling register-based studies of high quality,<sup>1</sup> including studies on inflammatory bowel disease (IBD).<sup>2,3</sup> IBD is a multifactorial chronic disease triggered by genetic, environmental, and microbiobal factors and covers the main subtypes ulcerative colitis (UC) and Crohn's disease (CD). UC and CD are seen with rising prevalences worldwide,<sup>4</sup> especially in westernized countries.<sup>5,6</sup> In Denmark alone, more than 50,000 individuals suffer from IBD<sup>7</sup> and their disease course is highly unpredictable.<sup>8,9</sup>

Despite the identification of several genetic markers and environmental risk factors, the etiology of IBD remains largely unknown, and the disease course is highly uncertain.<sup>10</sup> This complexity has inspired researchers to look towards the use of big data to predict disease occurrence and disease course. The potential of integrating a comprehensive amount of clinical data, registry data and omics data, such as genomics, transcriptomics, proteomics, and metabolomics, to

develop algorithms for risk prediction is appealing.<sup>11,12</sup> However, in this regard, the validity of diagnostic data is essential for the correct and complete identification of patients to study.

In Denmark, the potential for performing large-scale data integration is evident due to a free and easily accessible health care system, nationwide health registries and a national biobank. Still, only three minor studies have evaluated the validity of IBD registration in the national health registers. A study from 1996, based on an older version of the International Classification of Diseases (ICD), reported high completeness and validity of IBD diagnoses.<sup>13</sup> However, in two recent studies of smaller patient series, varying degrees of validity have been reported.<sup>14,15</sup>

The aim of the present study was to perform a large-scale validation of IBD diagnoses in the DNPR through review of medical records on all adult patients registered with a diagnosis of UC or CD from January 1<sup>st</sup>, 2002, until December 31<sup>st</sup>, 2020, in a geographically well-defined part of Denmark, the North Denmark Region, to lay the ground for combining nationwide data with large-scale omics data for development of risk prediction algorithms.

## Patients and Methods

### Study Population

The North Denmark Region has approximately 591,740 inhabitants in year 2022.<sup>16</sup> It was established in year 2007 and covers the former North Jutland County, parts of the former Viborg County, and a minor part of the former Aarhus County. These regional changes were taken into consideration during the conduct of the present study by including all patients, from regions which later became part of the North Denmark Region, diagnosed with IBD during the study period.

### Data Sources

The Danish Civil Registration system gives all citizens a unique Central Person Registry (CPR) number at birth or immigration.<sup>17</sup> All health care information is linked to the CPR number in public health care registries, including the DNPR, established in 1977. From the establishment of the DNPR, all hospital admissions were registered, and from 1995, information from outpatient clinics was added. Diagnoses are recorded according to the Danish version of ICD-10.<sup>18</sup> ICD-10 was used in Denmark from 1994 and onward (ICD-9 was never used).

The Danish health care system provides free health care to all citizens, and only a few private endoscopy and gastroenterology clinics exist. Hence, patients with suspected IBD are typically referred to the public health care system by their general practitioner. This is also the case in the North Denmark Region, where no private consultant gastroenterologists are working outside the public health care system. The region has endoscopy clinics, but patients diagnosed with IBD are referred to a hospital department of gastroenterology.

When in contact with the health care system, IBD patients can be registered in the DNPR with an IBD diagnostic code more than once for various reasons such as flare-ups, endoscopy, change of disease extent, and starting follow-up in the outpatient clinic, etc. Clinicians managing the patients are responsible for the coding. Coding is always done after hospital discharge. However, there are no strict requirements that coding should always be performed in the outpatient clinic, and as such, coding is usually done when a change in disease severity or extent has developed. In our study, we searched for all IBD diagnostic codes registered in the DNPR including both out-patient and discharge codes.

### Case Ascertainment

Using the ICD-10 codes DK500-K509 for CD and DK510-519 for UC, we identified all adult patients from the North Denmark Region registered with at least one IBD diagnosis from January 1<sup>st</sup>, 2002, until December 31<sup>st</sup>, 2020 in the Danish National Patient Registry.

In addition to the national registries, a regional database with medical record information on verified IBD patients was established in 1995 and was in 2014 named GASTROBIO. The database contains data on IBD patients diagnosed from 1978 and onwards.<sup>19-21</sup> Patients have been added to the database only by experienced gastroenterologists and only when fulfilling the Copenhagen Criteria.<sup>22</sup>

For the present study, we first compared the entire list of patients with at least one IBD diagnosis in the DNPR living in the region with the complete list of patients from our regional database to identify cases not already verified through chart review. For all additional cases, a medical gastroenterologist evaluated the medical records using the Copenhagen Criteria ([Supplementary material](#)) proposed by Binder et al.<sup>22</sup> In cases of uncertainty regarding the diagnosis, a senior gastroenterologist would reevaluate the medical records. Patients with now verified diagnoses were included in our database, whereas the patients assessed not to have IBD were categorized into the following groups according to the correct diagnosis: 1) diverticulitis, 2) drug-induced colitis, 3) false coding, 4) gastrointestinal infection (GI-infection), 5) irritable bowel syndrome (IBS), 6) ischemic colitis, 7) malignancy, 8) microscopic colitis, 9) obstipation, 10) terminal ileitis unspecified, 11) unspecified diagnosis 12) and other. The latter was used for a variety of incorrect diagnoses with infrequent occurrences. *Unspecified diagnosis* was used when no other correct diagnoses were registered in the medical records. *Terminal ileitis unspecified* was used for patients with suspected terminal ileitis where the etiology was never determined. *False coding* was used when there was no association between the medical records and the IBD diagnosis.

Patients were excluded in case of insufficient data in medical records, when living outside the region, when age was below 18 on December 31<sup>st</sup>, 2020, or if total colectomy was performed on UC patients before moving to the region. Patients with microscopic colitis registered with an IBD diagnosis before 2012 were excluded since no specific diagnosis for microscopic colitis existed. If registered after 2012 as CD or UC, they were classified as incorrectly registered.

Misclassified patients recorded with CD or UC in the DNPR who were registered in GASTROBIO with the opposite diagnosis or unclassified IBD were accepted as diagnosed correctly with IBD in this study.

## Statistical Analysis

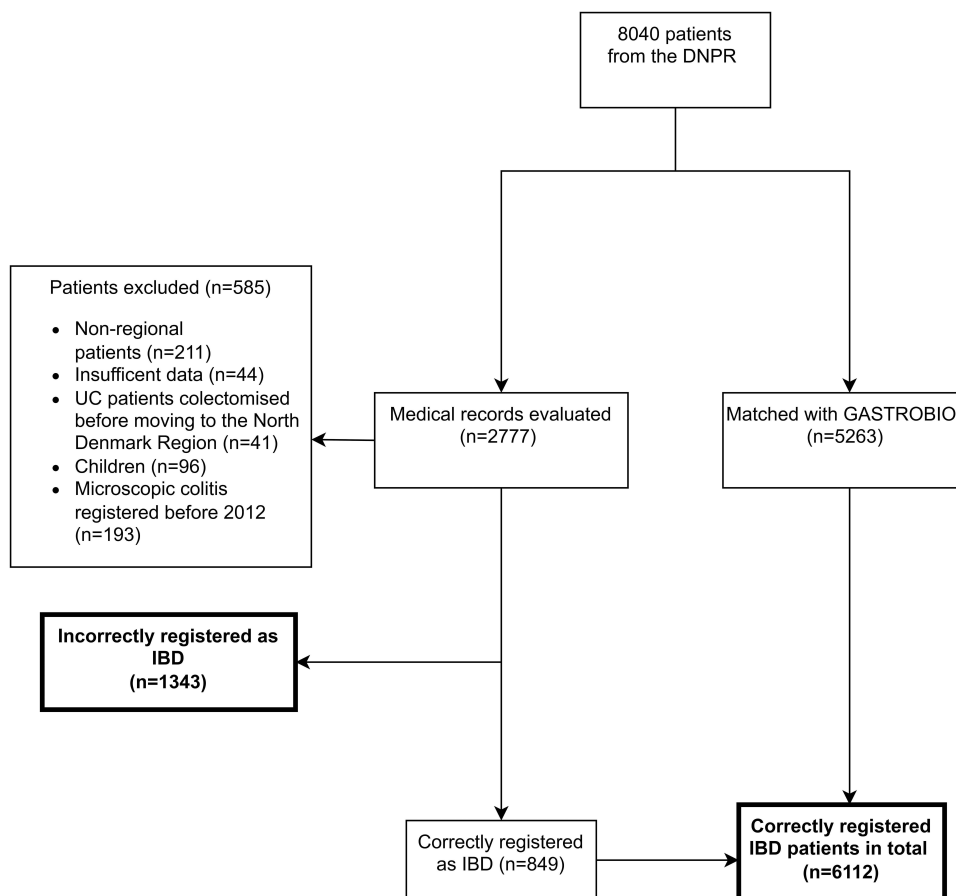
The positive predictive value (PPV) was calculated and used as a measure of validity. PPVs for IBD overall, UC, and CD were calculated for patients registered with at least one, two, and three diagnoses in the DNPR. When several diagnostic codes were registered the PPV was estimated based on the first registration within our search period. The PPV was defined as the proportion of valid cases in relation to the total number of registrations. Completeness refers to the proportion of true IBD cases correctly recorded in the registry. The completeness of the DNPR was calculated as the proportion of valid IBD cases captured when requiring at least two or three registrations compared to only one. A 95% Clopper-Pearson confidence interval was used for binominal data. The median age was calculated for the incorrect diagnoses, and age was defined as years between birth and registration date in the DNPR. Statistical analyses were performed using STATA (StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC).

## Results

A total of 8040 patients from Northern Denmark were recorded with at least one IBD diagnosis in the DNPR during years 2002–2020. Of these, 5263 had already been verified and included in our regional database GASTROBIO, leaving 2777 medical records for evaluation. Of these, 585 were excluded due to incomplete medical records or other reasons, 1343 were incorrectly registered with IBD, and 849 were correctly registered with IBD. In total, 6112 (76%) of the 8040 identified individuals were correctly registered with IBD based on one diagnosis, 3990 had UC, and 2122 had CD ([Figure 1](#)). Of the 2122 CD patients, 119 (5.6%) were registered as UC patients, and of the 3990 UC patients, 185 (4.6%) were registered as CD patients. Fifty-six percent of CD patients and 49% of UC patients were women.

## Positive Predictive Value

The positive predictive value (PPV) of a single IBD diagnosis in the DNPR was 0.82 (95% CI, 0.81–0.83), and PPVs for CD (0.82; 95% CI, 0.81–0.84) and UC (0.82; 95% CI, 0.81–0.83) were identical. When using at least two diagnoses, the PPV for IBD was 0.95 (95% CI, 0.95–0.96) and appeared to be slightly higher for CD (PPV, 0.96; 95% CI, 0.95–0.97) than UC (PPV, 0.95; 95% CI, 0.94–0.96). Based on at least three recordings with an IBD diagnosis in the registry, the PPV for IBD was 0.98 (95% CI, 0.98–0.99) and the same for CD (PPV, 0.98; 95% CI, 0.97–0.99) and UC (PPV, 0.98; 95% CI, 0.98–0.99) ([Table 1](#), [Figure 2](#)).



**Figure 1** Flowchart.

**Notes:** \*Name of regional database.

**Abbreviations:** DNPR, Danish National Patient Registry; UC, ulcerative colitis; IBD, inflammatory bowel disease.

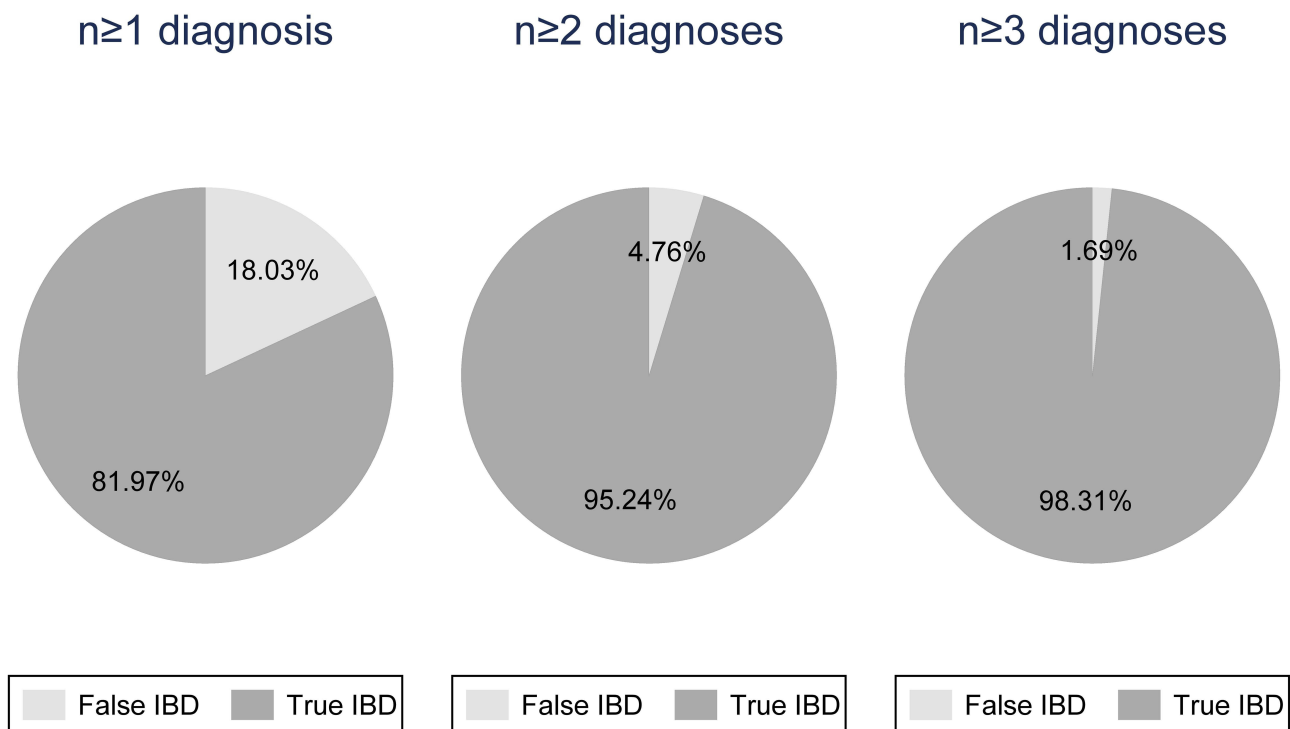
## Completeness

Notably, when requesting more than one IBD diagnosis in DNPR to confirm IBD diagnoses, several verified IBD cases were lost (Figure 3). The number of true IBD cases captured by the search in the DNPR decreased from 6112 to 4606

**Table 1** Positive Predictive Values (PPV) Based on Number of Registrations with a Diagnosis of Inflammatory Bowel Disease (IBD), Crohn's Disease (CD), or Ulcerative Colitis (UC) in the Danish National Patient Registry

No. of Registered Diagnoses	No. of Registered Cases	No. of Confirmed Cases	PPV (95% CI)
IBD (n≥1)	7455	6112	0.82 (0.81–0.83)
IBD (n≥2)	4836	4606	0.95 (0.95–0.96)
IBD (n≥3)	3377	3320	0.98 (0.98–0.99)
CD (n≥1)	2583	2122	0.82 (0.81–0.84)
CD (n≥2)	1816	1737	0.96 (0.95–0.97)
CD (n≥3)	1401	1377	0.98 (0.97–0.99)
UC (n≥1)	4872	3990	0.82 (0.81–0.83)
UC (n≥2)	3020	2869	0.95 (0.94–0.96)
UC (n≥3)	1976	1943	0.98 (0.98–0.99)

**Abbreviations:** No, number; IBD, inflammatory bowel disease; CD, Crohn's disease; UC, ulcerative colitis; PPV, positive predictive value.



**Figure 2** Validity of a diagnosis of inflammatory bowel disease in the Danish National Patient Registry based on at least one, two, or three registrations.  
**Abbreviation:** IBD, inflammatory bowel disease.

(75%; 95% CI, 74%-76%) when demanding at least two registered diagnoses, and it decreased further to 3320 (54%; 95% CI, 53%-56%) when demanding at least three diagnoses.

## Incorrect Diagnoses

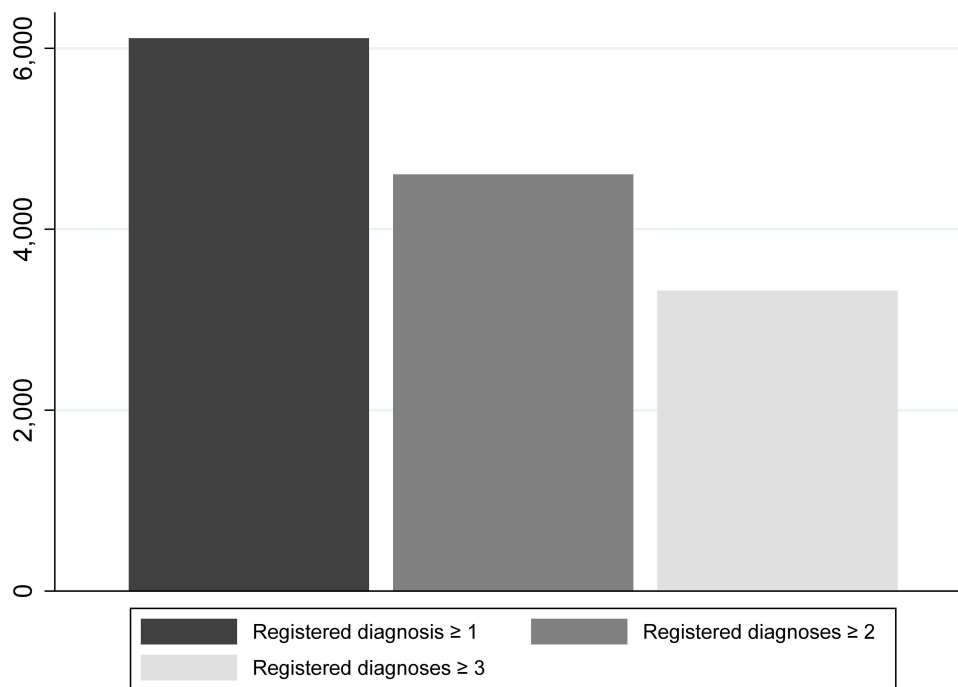
The distribution of diagnoses among the 1343 individuals incorrectly recorded with IBD is illustrated in Figure 4 and by IBD subtype in Table 2. The group unspecified diagnosis constituted 35.2% of all incorrect diagnoses. This group predominantly consisted of patients with various gastrointestinal symptoms and unspecific endoscopic findings, where a final diagnosis was never concluded in the medical records. The second most common group, *false coding*, which constituted 14.5%, included patients without any indication of IBD in the medical records. The third most common group was *IBS* which constituted 10%. This group included patients who were both difficult to distinguish from IBD and patients who were inaccurately diagnosed with IBD due to misconception. The rest of the groups each constituted 8% or less.

The distribution of incorrect diagnoses of UC and CD separately was proportionate to the overall number of correct UC and CD registrations for most diagnoses. However, the distribution of incorrect diagnoses differed somewhat between incorrect UC and CD cases, as the vast majority of *terminal ileitis unspecified* was registered as CD and *ischemic colitis* as UC, in accordance with the resemblance in disease presentation.

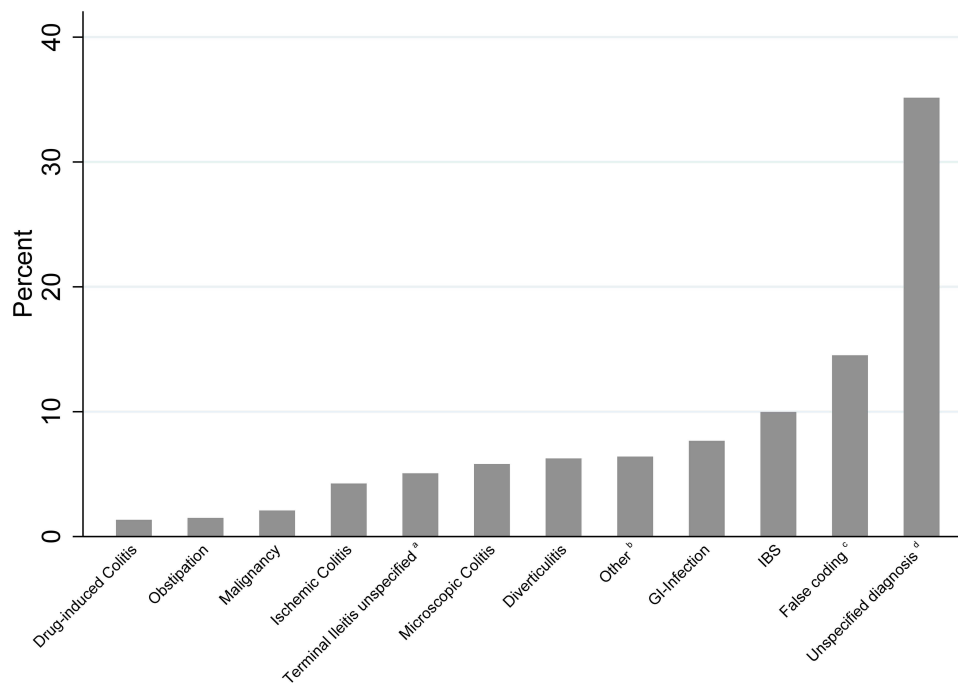
Figure 5 shows the percentual distribution and median age for each diagnostic group among the incorrectly registered. Only one group, *terminal ileitis unspecified*, had a median age below 40. Between the median age of 40 and 50, an increasing number of patients fell into the groups of *IBS*, *GI-infection*, *drug-induced colitis*, *obstipation*, and other. In individuals aged 60 to 70, *diverticulitis* and *microscopic colitis* were common causes of an incorrect IBD diagnosis, and above 70 years of age, *malignancy* and *ischemic colitis* were common causes.

## Discussion

This large-scale population-based validation study of IBD diagnoses in the DNPR based on manual scrutiny of thousands of medical records revealed a PPV of 82% for IBD based on one IBD diagnosis in the DNPR, increasing to 95% with two



**Figure 3** Loss of completeness when requiring more than one diagnosis for a true case definition. Columns represent true cases remaining when requiring two or three registrations in the Danish National Patient Registry to define a case of inflammatory bowel disease.



**Figure 4** Distribution (%) of incorrectly registered patients by subgroup.

**Notes:** a) Suspected terminal ileitis where the etiology was never concluded. b) Abscess (not IBD-related), adhesions, appendicitis/periappendicular abscess, bile acid malabsorption, celiac disease, diversion colitis, endometriosis, eosinophilic colitis, epiploic appendagitis, fistula (not IBD-related), Hirschsprung disease, ileus, radiation colitis, rectal prolaps, sarcoidosis, solitary rectal ulcer. c) No association between the medical records and the IBD diagnosis. d) No other correct diagnoses were registered in the medical records.

**Abbreviations:** IBS, irritable bowel syndrome; GI-infection, gastrointestinal infection.

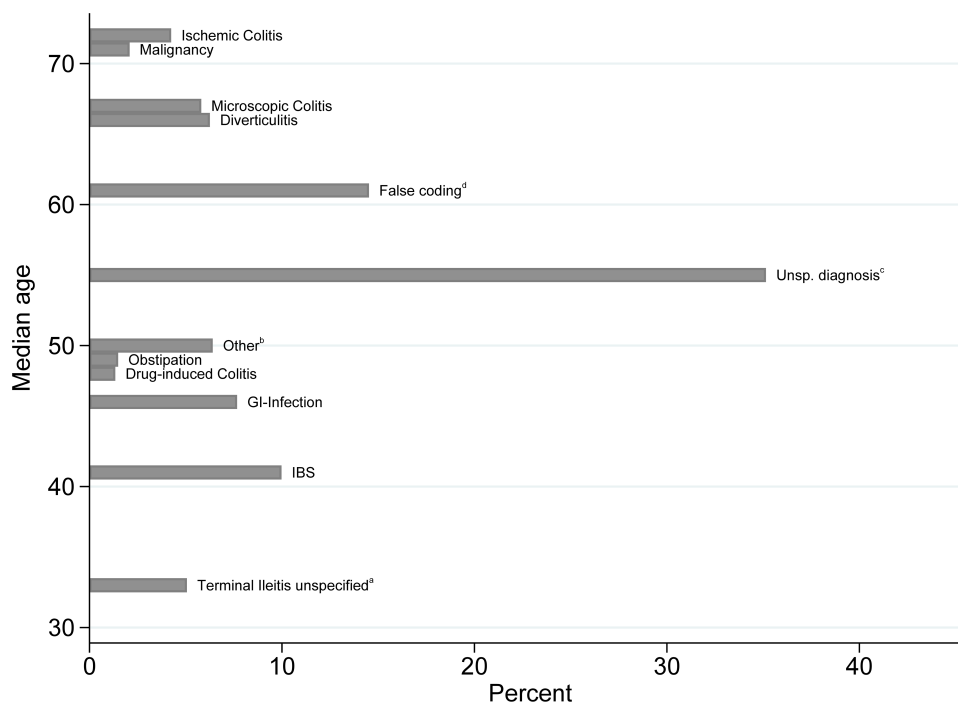
**Table 2** Distribution of Diagnoses by Incorrectly Recorded Inflammatory Bowel Disease Subtype, Crohn's Disease (CD) or Ulcerative Colitis (UC)

Diagnosis	Reg. as CD	Reg. as UC	Total
Diverticulitis	23	61	84
Drug-induced Colitis	5	13	18
False coding <sup>a</sup>	95	100	195
GI-Infection	36	67	103
IBS	45	89	134
Ischemic Colitis	7	50	57
Malignancy	16	12	28
Microscopic Colitis	14	64	78
Obstipation	13	7	20
Terminal Ileitis unspecified <sup>b</sup>	64	4	68
Unspecified diagnosis <sup>c</sup>	98	374	472
Other <sup>d</sup>	45	41	86

**Notes:** a) No association between the medical records and the IBD diagnosis. b) Suspected terminal ileitis where the etiology was never concluded. c) No other correct diagnoses were registered in the medical records. d) Abscess (not IBD-related), adhesions, appendicitis/periappendicular abscess, bile acid malabsorption, celiac disease, diversion colitis, endometriosis, eosinophilic colitis, epiploic appendagitis, fistula (not IBD-related), Hirschsprung disease, ileus, radiation colitis, rectal prolaps, sarcoidosis, solitary rectal ulcer.

**Abbreviations:** Reg, registered; IBS, irritable bowel syndrome; GI-infection, gastrointestinal infection.

diagnoses, and 98% with three diagnoses. Thus, epidemiological studies based on at least two IBD diagnoses in the Danish National Registries are highly valid. However, the study also demonstrated exclusion of true IBD cases when restricting to two or more diagnoses, which needs to be considered.

**Figure 5** Distribution (%) of patients with incorrect diagnoses subgrouped by identified reason for misdiagnosis according to median age at time of misdiagnosis.

**Notes:** a) Suspected terminal ileitis where the etiology was never concluded. b) Abscess (not IBD-related), adhesions, appendicitis/periappendicular abscess, bile acid malabsorption, celiac disease, diversion colitis, endometriosis, eosinophilic colitis, epiploic appendagitis, fistula (not IBD related), Hirschsprung disease, ileus, radiation colitis, rectal prolaps, sarcoidosis, solitary rectal ulcer. c) No other correct diagnoses were registered in the medical records. d) No association between the medical records and the IBD diagnosis.

**Abbreviations:** IBS, irritable bowel syndrome; GI-infection, gastrointestinal infection; Unsp; unspecified.



This study has several strengths. We have had unique tools to assess validity of IBD diagnoses in DNPR due to systematic manual scrutiny of patient files in the region for decades with inclusion of verified IBD cases in an existing database. With examination of additional potential cases by a trained gastroenterologist supported by a senior gastroenterologist according to the Copenhagen Criteria,<sup>22</sup> we were able to sample a complete regional cohort of verified IBD patients to compare to registrations in the DNPR. The large number of cases in our study ensures a high degree of certainty in our PPV estimates, which exceeds any previous Danish study and ensures generalizability of our findings.

Some limitations should, nevertheless, be considered. We were not able to perform calculations of specificity and sensitivity as this requires the identification of both true- and false-negative cases. True negatives could have been estimated by calculating the background population during the search period in the North Denmark Region. However, defining the entire background population as true negative cases could be misleading and overestimate the specificity. Also, calculation of sensitivity requires identifying the false-negative cases, meaning patients with IBD that were not registered with an IBD diagnostic code in the DNPR. We could have searched for these patients in GASTROBIO. But since patients are registered on the year of diagnosis in GASTROBIO, and we did not exclusively search for incident cases in the DNPR, it was not possible to define matching cohorts in the two registries. Completeness also needs to be considered. The term refers to the proportion of true disease cases correctly captured by the registry. Determining the completeness of patients with at least one diagnosis in the DNPR was not possible due to the reasons stated above, as this would require capturing the false-negative cases. However, our experience is that all patients with IBD in the North Denmark Region are referred to the public health care system, and we know that almost no individuals have prevalent undiagnosed IBD in Denmark.<sup>23</sup> Thus, we estimate the completeness to be high when only requiring one diagnosis.

Our finding of PPVs increasing from 82% based on one IBD diagnosis in the DNPR, to 95% with two diagnoses, and 98% with three diagnoses is promising. Three previous studies based on ICD-8 or smaller and selected subsets of patients have validated registrations of IBD diagnostic codes in the DNPR. A study from 1996 by Fonager et al reported a PPV of 0.97 for CD and 0.90 for UC based on ICD-8 codes for IBD in the DNPR.<sup>13</sup> The higher estimates may reflect that the study population was based mainly on diagnoses from specialized departments in relation to hospitalization. In a more recent Danish study limited to 513 patients from Copenhagen City and County, the PPV based on at least one diagnosis in DNPR was 0.85 for UC and 0.91 for CD. The PPV based on at least two diagnoses was 0.89 for UC and 0.94 for CD.<sup>14</sup> Compared to our study, the PPVs were higher based on at least one diagnosis and somewhat lower for UC patients with at least two diagnoses. However, since the study was not based on an initial full sample of cases from the DNPR, all incorrectly registered patients or false-positive cases might not have been captured, hence, potentially leading to an overestimation of the PPV, especially among individuals registered only once.

In another recent Danish study based on the Diet, Cancer and Health cohort,<sup>15</sup> the validity of IBD diagnoses was assessed among individuals aged 50–64 years at enrollment. Among 339 individuals in this age group, the PPV was 0.58 based on at least one IBD diagnosis, 0.78 based on two diagnoses and 0.92 based on at least three diagnoses. These PPVs are noticeably lower compared to our estimates and are in accordance with the increasing number of incorrectly registered patients with increasing median age in our study.

Among individuals in our study with incorrect IBD diagnoses in the DNPR, many had well-known mimickers of IBD with similar symptoms and endoscopic findings, yet the majority were categorized as either *unspecified diagnosis*, *false coding* or *IBS* and reflected inaccurate use of the Disease Classification System.<sup>24</sup> Hence, inaccuracy and not differential diagnoses was the primary cause of false registration. This might be due to inadequate awareness of the diagnostic criteria for IBD among physicians from other specializations recording in the DNPR.

According to the European Crohn's and Colitis Organisation (ECCO) guidelines, no golden standard to diagnose IBD exists. The diagnosis is established by a non-strictly defined combination of clinical presentation, endoscopic appearance, radiology, histology, surgical findings and serology.<sup>9,25</sup> We used the Copenhagen Criteria from 1982 for case ascertainment.<sup>22</sup> Although more novel criteria have been developed, none of these has been proven superior to the Copenhagen Criteria, which are still used in the current literature.<sup>14,15</sup> Furthermore, applying the Copenhagen Criteria ensured consistency between patients verified through chart review and previously confirmed patients in our database.

Only a few other countries besides Denmark have nationwide registries and a health care system that facilitates the potential for large population-based register studies. A Swedish study from 2017 validated IBD in the Swedish National

Patient Registry using at least two registered diagnoses and reported a PPV of 0.93.<sup>26</sup> An older Swedish study from 1991 had shown a validity of only 0.74 when including patients with one diagnosis.<sup>27</sup> In France, population-based register studies of IBD have also been conducted using administrative hospital databases,<sup>28</sup> and although validity has not been the focus of these studies, the case definition may serve as inspiration for future studies. In a recent French study of anti-TNF treated IBD patients, the study population was defined by combining inpatient diagnostic codes from the French hospital discharge database with outpatient drug reimbursements from the French National Health Insurance anonymized claim database.<sup>29</sup> Also, in a study of antibiotic use and IBD development in the United Kingdom, a first-time diagnosis of IBD in the Clinical Practice Research Datalink (CPRD GOLD) and a subsequently recorded IBD-related prescription were used as a case definition.<sup>30</sup> Likewise, combining diagnostic codes from the DNPR, which holds both data from in- and outpatient management, with data from the Danish National Prescription Registry could be an approach to further optimize validity and completeness.<sup>31</sup>

Finally, a consequence of using more strict case definitions may be the exclusion of true IBD cases from analyses. The interval between the first and second recorded diagnosis in the DNPR varies depending on coding procedure and disease development. Therefore, the risk of missing true cases when requiring two diagnoses will increase towards the end of a study period. Furthermore, when performing incidence studies, the interval between registrations could lead to a false decrease in incidence towards the end of the study.<sup>7</sup>

In our study, a sizeable number of cases was lost when requiring more registrations for accurate case definition, with a 25% loss of true cases when using two instead of one diagnosis and a 46% loss when using three instead of one diagnosis to define an IBD case. Hence, the pros and cons in case definitions have to be taken into account in study designs based on DNPR and are for the first time clearly demonstrated in this paper.

## Conclusion

Our large-scale validation study of IBD diagnoses in the DNPR based on manual scrutiny of thousands of medical files, we observed reassuring PPVs for IBD overall, as well as for UC and CD, especially when using two or more recorded diagnoses. However, we also demonstrated reduced completeness of the cohort when increasing the number of registrations required for a true case definition. Considering both validity and completeness, we suggest using at least two registrations with IBD in future studies based on the DNPR, while acknowledging that especially recently diagnosed patients may then be missed. Future studies should examine the potential for developing algorithms for the definition of IBD based on a combination of DNPR diagnoses and medication data, to further enhance the quality of DNPR based studies.

## Abbreviations

DNPR, Danish National Patient Registry; IBD, inflammatory bowel disease; PPV, positive predictive value; CI, confidence interval; UC, ulcerative colitis; CD, Crohn's disease; ICD, international classification of diseases; CPR, central person registry; IBS, irritable bowel syndrome; GI-infection, gastrointestinal infection.

## Data Sharing Statement

Data Sharing is not applicable due to General Data Protection Regulation (GDPR).

## Ethical Considerations

The study was approved by The North Denmark Region (2020-043900) according to the Danish Health Care Act and by the Danish Data Protection Agency: F2022-041. This approval was obtained before initiating the study and, according to Danish health law, waves off the requirement for written informed consent in order to gather data from medical records. Public disclosure is permitted when data is reported in a non-personally identifiable manner.

## Funding

The present study is supported by a National Center of Excellence Grant (DNRF148) from the Danish National Research Foundation.

## Disclosure

The authors report no conflicts of interest in this work

## References

- Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National patient registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449–490. doi:10.2147/CLEP.S91125
- Jess T, Simonsen J, Jorgensen KT, Pedersen BV, Nielsen NM, Frisch M. Decreasing risk of colorectal cancer in patients with inflammatory bowel disease over 30 years. *Gastroenterology*. 2012;143(2):375–381.e1. doi:10.1053/j.gastro.2012.04.016
- Andersen NN, Pasternak B, Basit S, et al. Association between tumor necrosis factor- $\alpha$  antagonists and risk of cancer in patients with inflammatory bowel disease. *J Am Med Assoc*. 2014;311(23):2406–2413. doi:10.1001/jama.2014.5613
- Nørgård BM, Nielsen J, Fonager K, Kjeldsen J, Jacobsen BA, Qvist N. The incidence of ulcerative colitis (1995–2011) and Crohn's disease (1995–2012) - Based on nationwide Danish registry data. *J Crohn's Colitis*. 2014;8(10):1274–1280. doi:10.1016/j.crohns.2014.03.006
- Uhlig HH, Muise AM. Clinical Genomics in Inflammatory Bowel Disease. In: *Trends in Genetics*. 33. Elsevier Ltd; 2017:629–641.
- Kaplan GG, Jess T. The Changing Landscape of Inflammatory Bowel Disease: east Meets West. *Gastroenterology*. 2016;150(1):24–26. doi:10.1053/j.gastro.2015.11.029
- Lophaven SN, Lyng E, Burisch J. The incidence of inflammatory bowel disease in Denmark 1980–2013: a nationwide cohort study. *Aliment Pharmacol Ther*. 2017;45(7):961–972. doi:10.1111/apt.13971
- Magro F, Gionchetti P, Eliakim R, et al. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and ileo-anal pouch disorders. *J Crohn's Colitis*. 2017;11(6):649–670.
- Magro F, Gionchetti P, Eliakim R, et al. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. *J Crohn's Colitis*. 2017;11(6):649–670. doi:10.1093/ecco-jcc/jjx008
- Jess T, Riis L, Vind I, et al. Changes in clinical characteristics, course, and prognosis of inflammatory bowel disease during the last 5 decades: a population-based study from Copenhagen, Denmark. *Inflamm Bowel Dis*. 2007;13(4):481–489. doi:10.1002/ibd.20036
- Seyed Tabib NS, Madgwick M, Sudhakar P, Verstockt B, Korcsmaros T, Vermeire S. Big data in IBD: big progress for clinical practice. *Gut*. 2020;69(8):1520–1532. doi:10.1136/gutjnl-2019-320065
- Agrawal M, Allin KH, Petralia F, Colombel JF, Jess T. Multiomics to elucidate inflammatory bowel disease risk factors and pathways. *Nat Rev Gastroenterol Hepatol*. 2022;19(6):399–409. doi:10.1038/s41575-022-00593-y
- Fonager K, Sorensen HT, Rasmussen SN, Moller-Petersen J, Vyberg M. Assessment of the diagnoses of Crohn's disease and ulcerative colitis in a Danish hospital information system. *Scand J Gastroenterol*. 1996;31(2):154–159. doi:10.3109/00365529609031980
- Lo B, Vind I, Vester-Andersen MK, Burisch J. Validation of ulcerative colitis and Crohn's disease and their phenotypes in the Danish National Patient Registry using a population-based cohort. *Scand J Gastroenterol*. 2020;55(10):1171–1175. doi:10.1080/00365521.2020.1807598.
- Rye C, Rubin KH, Moller FT, Julsgaard M, Jess T, Andersen V. Positive predictive value of diagnostic codes for inflammatory bowel disease in the Danish National Patient Registry among individuals 50+ years, using patient records as reference standard. *Clin Epidemiol*. 2021;13:335–344. doi:10.2147/CLEP.S298770
- STATISTICS DENMARK. BEV107: summary vital statistics by municipality, new increases/stock and sex [Internet]; 2022 Available from: <https://www.statbank.dk/BEV107>. Accessed September 29, 2022.
- Pedersen CB. The Danish civil registration system. *Scand J Public Health*. 2011;39(7):22–25. doi:10.1177/1403494810387965
- World Health Organization. International Statistical Classification of Diseases and Related Health Problems (ICD) [Internet]. ICD-10. Available from: <https://www.who.int/standards/classifications/classification-of-diseases>. Accessed September 29, 2022.
- Larsen L, Drewes AM, Fallingborg J, Jacobsen BA, Jess T. Touch screens as a tool in patient care in the IBD outpatient clinic. *Scand J Gastroenterol*. 2016;51(9):1106–1110. doi:10.1080/00365521.2016.1174879
- Larsen L, Drewes AM, Broberg MCH, et al. Changing Infliximab Prescription Patterns in Inflammatory Bowel Disease: a Population-Based Cohort Study, 1999–2014. *Inflamm Bowel Dis*. 2018;24(2):433–439. doi:10.1093/ibd/izx038
- Jacobsen BA, Fallingborg J, Rasmussen HH, et al. Increase in incidence and prevalence of inflammatory bowel disease in northern Denmark: a population-based study, 1978–2002. *Eur J Gastroenterol Hepatol*. 2006;18(6):601–606. doi:10.1097/00042737-200606000-00005
- Binder V, Both H, Hansen PK, Hendriksen C, Kreiner S, Torp-Pedersen K. Incidence and Prevalence of Ulcerative Colitis and Crohn's Disease in the County of Copenhagen, 1962 to 1978. *Gastroenterology*. 1982;83(3):563–568. doi:10.1016/S0016-5085(82)80190-X
- Jess T, Vestergaard MV, Iversen AT, Allin KH. Undiagnosed inflammatory bowel disease among individuals undergoing colorectal cancer screening: a nationwide Danish cohort study 2014–2018. *Gut*. 2022. doi:10.1136/gutjnl-2022-327296
- Gece KB, Vermeire S, Vermeire S. Differential diagnosis of inflammatory bowel disease: imitations and complications. *lancet Gastroenterol Hepatol*. 2018;3(9):644–653. doi:10.1016/S2468-1253(18)30159-6
- Gomollón F, Dignass A, Annesse V, et al. 3rd European Evidence-based Consensus on the Diagnosis and Management of Crohn's Disease 2016: part 1: diagnosis and Medical Management. *J Crohn's Colitis*. 2017;11(1):3–25. doi:10.1093/ecco-jcc/jjw168
- 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(2):216–221. doi:10.1080/00365521.2016.1246605
- Ekbom A, Helmick C, Zack M, Adami HO. The epidemiology of inflammatory bowel disease: a large, population-based study in Sweden. *Gastroenterology*. 1991;100(2):350–358. doi:10.1016/0016-5085(91)90202-V
- Chantry AA, Deneux-Tharoux C, Cans C, Ego A, Quantin C, Bouvier-Colle MH. Hospital discharge data can be used for monitoring procedures and intensive care related to severe maternal morbidity. *J Clin Epidemiol*. 2011;64(9):1014–1022. doi:10.1016/j.jclinepi.2010.11.015
- Kirchgesner J, Lemaitre M, Rudnichi A, et al. Therapeutic management of inflammatory bowel disease in real-life practice in the current era of anti-TNF agents: analysis of the French administrative health databases 2009–2014. *Aliment Pharmacol Ther*. 2017;45(1):37–49. doi:10.1111/apt.13835

30. Troelsen FS, Jick S. Antibiotic Use in Childhood and Adolescence and Risk of Inflammatory Bowel Disease: a Case-Control Study in the UK Clinical Practice Research Datalink. *Inflamm Bowel Dis*. 2020;26(3):440–447. doi:10.1093/ibd/izz137
31. Wallach Kildemoes H, Toft Sørensen H, Hallas J. The Danish national prescription registry. *Scand J Public Health*. 2011;39(7):38–41. doi:10.1177/1403494810394717

### Clinical Epidemiology

Dovepress

### Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>