

Cohort profile: ESPRESSO (Epidemiology Strengthened by histoPathology Reports in Sweden)

Jonas F Ludvigsson¹⁻⁴
Mariam Lashkariani¹

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; ²Department of Pediatrics, Örebro University Hospital, Örebro, Sweden; ³Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Clinical Sciences Building 2, City Hospital, Nottingham, UK; ⁴Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA

Abstract: The ESPRESSO study constitutes a novel approach to examine the etiology and prognosis of gastrointestinal disease in which histopathology plays a prominent role. Between 2015 and 2017, all pathology departments (n=28) in Sweden were contacted and asked to procure histopathology record data from the gastrointestinal tract (pharynx to anus), liver, gallbladder, and pancreas. For each individual, local histopathology IT personnel retrieved data on personal identity number, date of histopathology, topography (where the biopsy is taken), morphology (biopsy appearance), and where available free text. In total, between 1965 and 2017, histopathology record data were available in 2.1 million unique individuals, but the number of data entries was 6.1 million because more than one biopsy was performed in many of the study participants. Index individuals with histopathology data were matched with up to five controls from the general population. We also identified all first-degree relatives (parents, children, full siblings), and the index individual's first spouse. The total study population consisted of 13.0 million individuals. Data from all the study participants have been linked to Swedish National Healthcare Registers allowing research not only on such aspects as fetal and perinatal conditions and the risk of future gastrointestinal disease but also on the risk of comorbidity and complications (including cancer and death). Furthermore, the ESPRESSO database allows researchers and practitioners to identify diagnoses and disease phenotypes not currently indexed in national registers (including disease precursors). The ESPRESSO database increases the sensitivity and specificity of already-recorded diseases in the national health registers. This paper is an overview of the ESPRESSO database.

Keywords: cohort, gallbladder, gut, liver, pancreas, population-based, colon, pathology, endoscopy, celiac, inflammatory bowel disease, colitis, esophagus

Introduction to the ESPRESSO cohort

Sweden has a population of about 10 million inhabitants,¹ with a total life expectancy of 82.4 years (males 80.7 years, females 84.0 years).² Healthcare delivery is funded by a decentralized, taxpayer-funded system. In 2014, healthcare expenditures comprised 11.9% of the country's gross domestic product.³ Healthcare use is monitored by "Statistics Sweden" (Swedish: Statistiska Centralbyrån, SCB), and healthcare registers have been the source of invaluable population-based registry linkage research.⁴ Despite the importance of histopathology data in routine healthcare, none of the Swedish government-administered national healthcare registers (with the exception of limited cancer data in the "Swedish Cancer Register")⁵ contains any histopathology information.

Correspondence: Jonas F Ludvigsson
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
Tel +46 8 5248 0000
Fax +46 8 31 4975
Email jonasludvigsson@yahoo.com

From 2004 to 2007, the principal investigator (JFL) of the ESPRESSO database published several scientific papers on coeliac disease^{6–8} using relevant International Classification of Disease (ICD) codes in the Swedish Inpatient Register.⁹ However, it soon became evident that using inpatient data for the identification of coeliac disease jeopardized overestimating the risk of complications given that coeliac patients admitted to hospital represent individuals with a more severe disease (selection bias)¹⁰ than the average patient.^{11,12}

We therefore hypothesized that small intestinal biopsy record data might help to identify individuals with coeliac disease, and furthermore, are more representative of all patients with coeliac disease. Between 2006 and 2008, all Swedish pathology departments (n=28) were contacted and consented to export data for a histopathology-based coeliac database. An important characteristic of this coeliac database is that it contains information on inflammation and “normal mucosa”^{13,14} – phenotypes that lack ICD codes and therefore cannot be identified through national healthcare registers – but the database also enabled us to identify patients with coeliac disease who healed over time (mucosal healing)^{15,16} (first biopsy with villous atrophy, but the second biopsy showed no persistent villous atrophy).

It has become clear over the years that research on other gastrointestinal diseases (eg, microscopic colitis and fatty liver disease) may also benefit from histopathology data. Accordingly, between October 12, 2015, and April 10, 2017, we collected all gastrointestinal histopathology report data from the 28 Swedish pathology departments.

Data collection

In Sweden, gastrointestinal biopsies and surgical specimens are classified according to the SnoMed classification system. SnoMed (or SnoMed CT) is a multilingual clinical healthcare terminology (<http://www.ihtsdo.org/snomed-ct/>) that was jointly developed by the National Health Service (NHS) in England and the College of American Pathologists.

To minimize heterogeneity due to different reporting from individual pathology departments and to ensure that standard information was delivered from each centre, JFL in 2015 contacted the two companies that deliver the technical systems needed for SnoMed recording and storage in Sweden, namely “CompuGroup Medical” (system platform: “Analytix”) and “Tieto-Enator” (system platform: “Sympathy”). Together with these companies, JFL constructed a dedicated structured query language (SQL) question, which was then delivered as part of an update of the Analytix and

Sympathy systems. This SQL question enabled local users to perform the search automatically, with a predefined search algorithm (roughly “the gastrointestinal tract including the liver, gallbladder, and pancreas” up until the date of search).

The algorithm was distributed in 2015, but because of local requirements (including a specific request for locally tailored research applications), the search was delayed and performed in most pathology departments in 2016 (up until April 10, 2017). All Swedish pathology departments contributed data (percentages of all biopsies from Sweden are reported within brackets): Borås (contributed 1.95% of all gastrointestinal histopathology reports in Sweden), Eskilstuna (1.56%), Falun (2.96%), Gävle (2.14%), Halmstad (2.03%), Jönköping (3.28%), Kalmar (2.30%), Karlskrona (1.42%), Karlstad (3.31%), Medilab (Stockholm) (11.43%), St Göran Hospital (1.51%), Göteborg (6.61%), Lund-Malmö-Helsingborg-Kristianstad (14.14%), Skövde (3.27%), Stockholm (12.89%), Sunderbyn (1.98%), Sundsvall (1.65%), Trollhättan (2.69%), Umeå (4.57%), Uppsala (4.14%), Västerås (1.96%), Växjö (2.07%), Örebro (6.37%), and Linköping-Norrköping (3.77%).

Local IT technicians performed searches and saved data on 1) date of biopsy, 2) personal identity number (PIN),⁴ 3) morphology, 4) topography, and 5) free text of the histopathology report (available from ten departments: Eskilstuna, Falun, Jönköping, Karlskrona, Medilab [Stockholm], St Göran [Stockholm], Skövde, Sunderbyn, Västerås, and Örebro). Computerized search results were delivered to the researchers. ML then cleaned the data and merged all the files. It should be noted that the histological classification of the ESPRESSO database is based on information recorded in histopathology reports. We did not request any tissue collection or DNA and were unable to collect histopathology image data.

Eligibility criteria

We requisitioned all electronic histopathology reports in Sweden with a topography code from T56 to T69 corresponding to the liver, gallbladder, pancreas, and the bowel (from the pharynx to the anus; Figure 1). A small number of histopathology reports occurred before 1965 (n=13) and were omitted from the cohort for fear of misclassification. In total, we identified 2,109,579 unique individuals with a histopathology report between 1965 and April 2017. The number of data sets contained more data entries (6.1 million) than the number of individuals because many of the individuals (53.8%) had been biopsied more than once.

Figure 2 shows the number of histopathology reports collected in the ESPRESSO database per year. The number of histopathology reports increased gradually until 2015. The lower number of reports in 2016 and 2017 (for which we have data up until April) reflects that the first centers sent their data to us (the researchers) in the late 2015 but not thereafter.

Each (index) individual with a histopathology report was matched with up to five controls for age, sex, calendar year (of biopsy), and county of residence from the Total Population Register.¹⁷ We also identified all first-degree relatives of each index individual and his or her controls (parents, children, full siblings) as well as first spouse. First-degree relatives and spouses may or may not have had a biopsy.

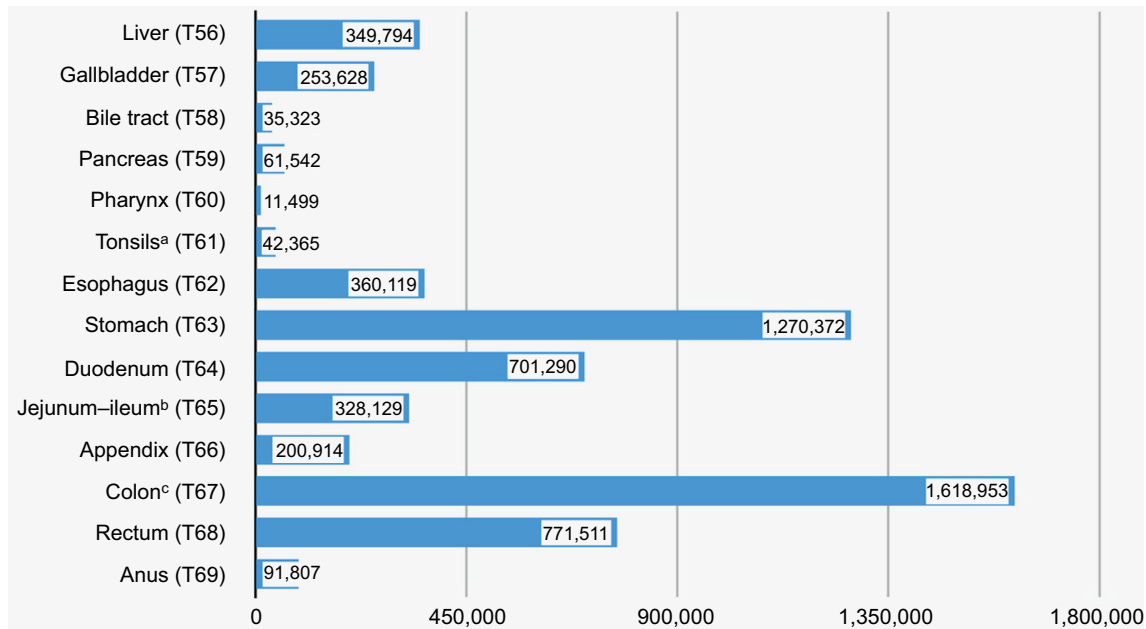


Figure 1 Number of gastrointestinal histopathology reports from 1965 to April 2017 in Sweden.

Notes: The number of unique individuals is smaller than the number of entries. ^aIncludes adenoid tissue. ^bT652 (for ileum) allows for differentiation between the jejunum and ileum. ^cT67 can be further divided into caecum, pars ascendens, right flexure, transverse, left flexure, descendens, and the sigmoid.

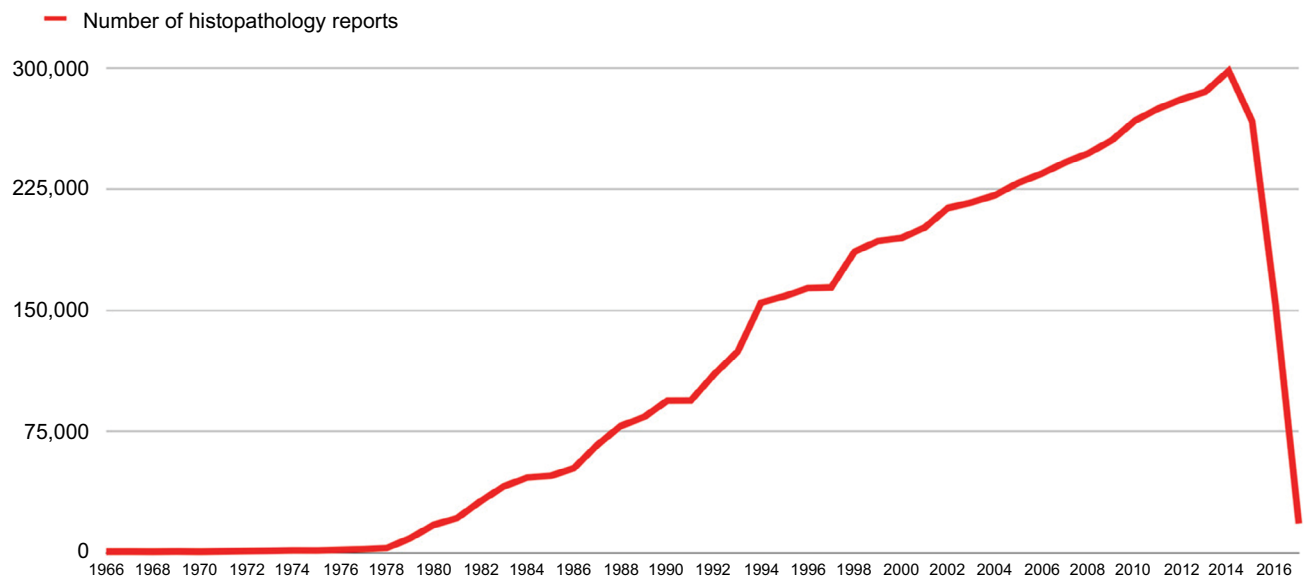


Figure 2 Annual number of gastrointestinal histopathology reports in Sweden from 1965 to 2017.

Notes: The ESPRESSO database contains information on 661 histopathology reports from 1965. The annual number of histopathology reports first exceeded 20,000 in 1981 and exceeded 100,000 in 1992.

The total study population contained 12,983,573 individuals. The study population exceeded the total Swedish population (roughly 10 million people) because >10 million unique individuals had lived in Sweden during the years when histopathology data were collected (Figure 2).

Through linkage to the Total Population Register with the date of death and emigration, it is possible to follow individuals over time. The median follow-up from the first biopsy date until 2017 was 12 years (we had data on the date of death and emigration until December 31, 2017). Males constituted 45.8% of the individuals having a histopathology record and the median age at first biopsy was 58 years.

Given the nature of the study and the almost 100% follow-up data of individuals (tracked through their PIN), we expect virtually no loss to follow-up. Concerning the histopathological follow-up, most patients had one histopathology report (median =1). Participants were never contacted directly nor did they receive any questionnaires or undergo any clinical

assessments as part of the ESPRESSO database. Instead, all data originated from the histopathology reports and from linked national Swedish registers (Figure 3).

Linkage to background data

For each study participant (n=13.0 million), Statistics Sweden delivered information on vitality status (dates of birth and death) and immigration/emigration. Statistics Sweden also supplied data on sex, age, county of residence, civil status, income, education, number of children, occupation, socioeconomic status, and nationality. Most of the background data originated from the LISA (longitudinal integrated database for health insurance and labor market studies) database.

Linkage to other national registers

Health data (serving as exposures, outcome measures, and covariates) were obtained from the Swedish national registers

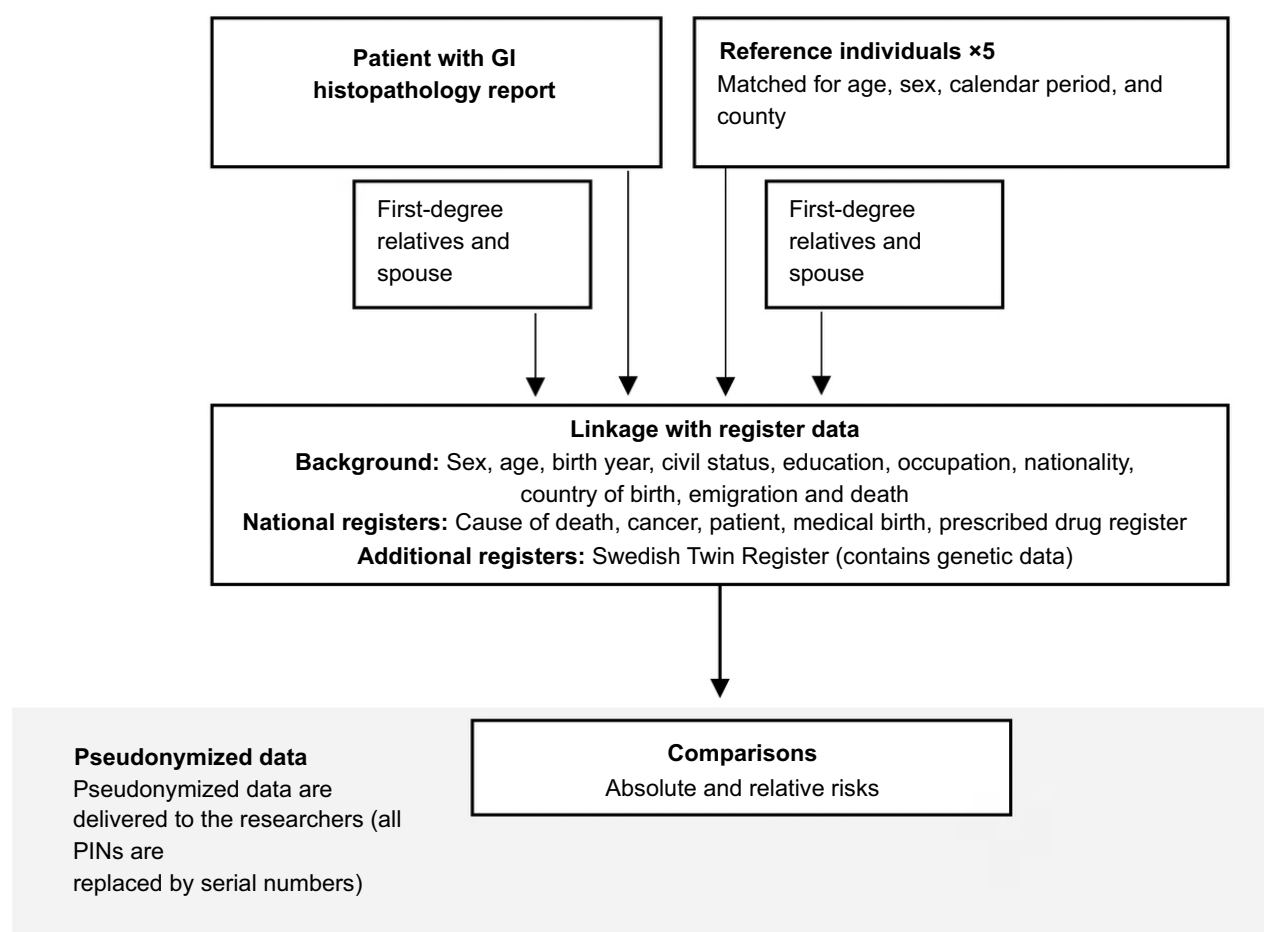


Figure 3 Main linkages in the ESPRESSO database.

Note: Pseudonymized data, and the explanation is given in the "Ethics" section.

Abbreviations: PIN, personal identity number; GI, gastrointestinal.

maintained by the National Board of Health and Welfare (Swedish: Socialstyrelsen). Approximately 97% of the Swedish healthcare system is under public authority, permitting easy and inexpensive access to comprehensive registry data. For each individual, we obtained data from the “Causes of Death Register” (covering >99% of all deaths),¹⁸ “the Swedish Cancer Register” (began in 1958, where >96% of all malignancies are recorded), the “Patient Register” (began in 1964, with hospital-based outpatient data since 2001; the positive predictive value (PPV) of the Patient Register is usually ~90%),⁹ the “Medical Birth Register” (antenatal and perinatal data on >98% of all births in Sweden since 1973 and includes data on early pregnancy smoking since 1982 and early pregnancy body mass index since 1992), and the “Swedish Prescribed Drug Register” (established in July 2005).¹⁹ Additional linkage, including the Swedish Twin Register, is expected in the future.²⁰ Figure 3 shows the main linkages of the ESPRESSO study.

Validation of histopathology codes

In 2009, we validated small intestinal histopathology data for coeliac disease against patient chart data.²¹ This chart validation found that 95% (108/114) of the individuals with small intestinal villous atrophy had a clinical diagnosis of coeliac disease (PPV, 95%). This figure is higher than when receiving a physician-assigned diagnosis of coeliac disease in Sweden (86%).²² Patient chart reviews demonstrated that 79% of the individuals with villous atrophy had gastrointestinal symptoms before biopsy and that the 88% with villous atrophy had positive coeliac disease serology before the first biopsy, further confirming that individuals with villous atrophy were likely to suffer from coeliac disease. The coeliac disease validation also contained individuals with small intestinal inflammation but no villous atrophy (n=39). We reviewed symptoms and signs of these individuals.²¹ In a related validation, we were able to examine patient charts of 112 individuals with normal mucosa but positive coeliac antibodies.²³

In a third validation, we examined patient charts from individuals with a histopathology report indicating microscopic colitis.²⁴ In total, 200/211 patients with a histopathology diagnosis of microscopic colitis were confirmed as also having a clinical diagnosis of microscopic colitis after chart review, yielding a PPV of 95% (95% CI =91%–97%). The most common symptoms in patients with microscopic colitis were diarrhea (seen in 96%), weight loss (24%), and abdominal pain (13%).

Relationship to IBD

Histopathology use can serve as a means to identify patients with inflammatory bowel disease (IBD) (Figure 4). Up until 2001, only inpatient diagnoses of IBD were registered in the Swedish Patient Register;⁹ even after that, only patients admitted to hospitals are recorded. Gastrointestinal histopathology data of Swedish IBD patients^{1,25,26} have revealed that such histopathology data can be used to ascertain the actual date of IBD diagnosis. Figure 4 illustrates the different incidence rates for IBD in Sweden using different definitions, where we planned to use “≥2 relevant ICD codes for IBD or one relevant SnoMed code and one relevant ICD code for IBD” (green line) as opposed to “only ≥2 relevant ICD codes for IBD” since the latter may underestimate incidence before the introduction of outpatient data in 2001. Serial biopsy data can also be used to evaluate mucosal healing in IBD. Additional data on validation against IBD can be found in the Supplementary Material.

Linkage to procedure codes in the Patient Register also allows the ESPRESSO database to add information on endoscopy (Esophagogastroduodenoscopy: 2861, 2880, 2881, 4480, 4483, 4486, 4487, 4488, 4489, 4490, 4686, 4687, 9003, 9004, 9021, UJC, UJD, UJF02, UJF05. Colonoscopy and sigmoidoscopy: 9011, 9012*, 9023, 4685*, 4688, 4689, 4674, 4684, UJF32, UJF35, UJF42*, UJF45* (sigmoidoscopy codes are marked with *). Endoscopic retrograde cholangiopancreatography: 9014, 5388, 5394, UJK02, UJK05).

Morphology data

The Swedish version of the SnoMed system allows users to assign morphology (M codes) and diagnostic codes (D codes). A diagnostic code is regarded as a kind of morphology code. A selection of diagnostic codes relevant to the ESPRESSO database is listed in Table 1.

M-codes range from normal (macroscopically: M00100, microscopically: M00110, chromosomally normal: M00150) to minimal abnormality (M01110) and to more severe abnormalities. A few M-codes are organ specific (eg, ectopic pancreas, M26020 but also ectopic bowel and stomach tissue), ulcers (M38000 and then specifically peptic ulcer M38090), chronic persisting hepatitis (M43001), chronic autoimmune hepatitis (M43005), and reactive gastritis (M69400). M-codes may also signal the existence of specific pathogens (eg *Helicobacter* [ME1370]) and *Giardia lamblia* (ME4416), but can also be used for cancer staging.

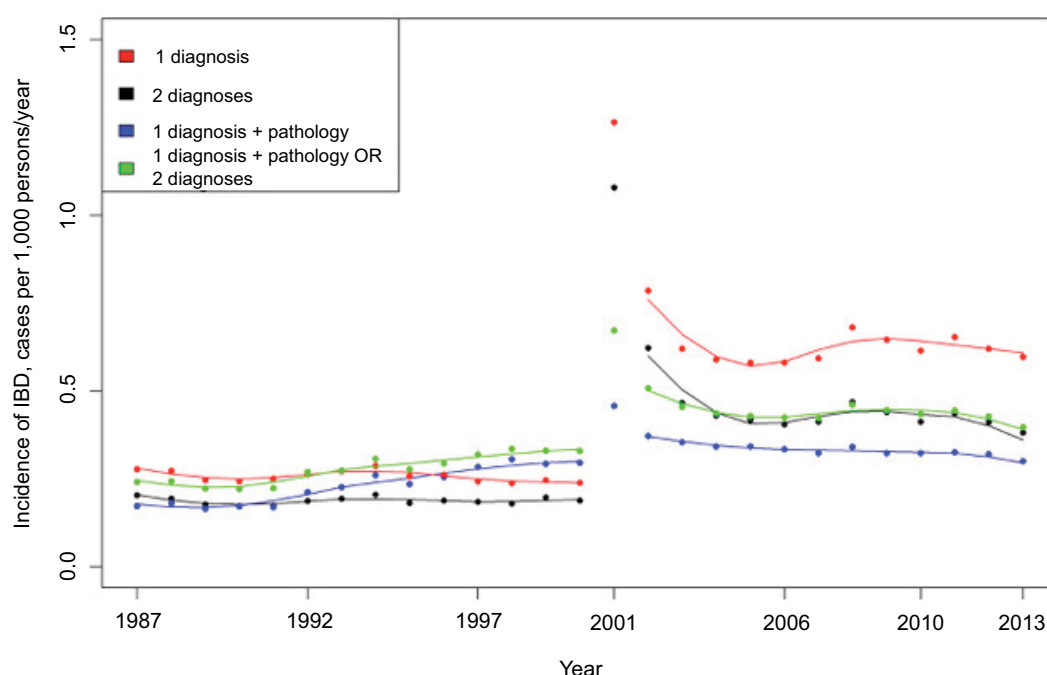


Figure 4 Using pathology data to improve incidence data for IBD in adults in Sweden.

Notes: Incidence of IBD based on data from the Swedish Patient Register (one or two diagnoses) and pathology (ESPRESSO data; green and blue lines). Using histopathology data increases incidence rates up until 2001 (when outpatient data were added to the Patient Register), and thereby ascertains a number of patients who without access to histopathology data had been diagnosed in 2001/2002 or later. Diagnosis refers to ICD diagnosis for IBD.

Abbreviation: IBD, inflammatory bowel disease.

Table 1 Diagnostic SnoMed codes

SnoMed code	Clinical condition
D0520	Viral hepatitis ^a
D100	Metabolic disease ^b
D6214	Indeterminate colitis (IBD-U)
D6216	Crohn's disease
D6218	Celiac disease
D6255	Ulcerative colitis
D8770/E5510	Alcohol-related disease

Notes: ^aViral hepatitis has subheadings for chronic active hepatitis, chronic persistent hepatitis, hepatitis B virus, and hepatitis C virus. ^bMetabolic disease: subheadings for inborn errors of metabolism and storage disease. D101-I07 also specifies nutritional disease and alpha-1-antitrypsin deficiency.

In a clinical setting, most M-codes must be combined with a topography code to make sense that a patient has a morphology code suggesting that unspecific inflammation is difficult to interpret without knowing if the inflammation occurs in the kidney, liver, or skin. M40-M43, M44000, and M470-471 are all different forms of inflammation, sometimes with additional information (when seen in the colorectal part: collagenous colitis M40600 and lymphocytic inflammation M4717 represent microscopic colitis). End-stage morphology, such as fibrosis (M49000-04) and cirrhosis (M495, with subheadings for primary biliary cirrhosis), fat infiltration

(M50080 and M55200), necrosis (M54 and necrotic fat tissue M54110), infarction (M547), and atrophy (M58) can also be registered.

The SnoMed system makes it possible for pathologists to classify degree of atypia (M697), squamous cell and glandular cell atypia (with subheadings for the degree of atypia), hyperplasia (M720), and then specific M-codes for adenomatous hyperplasia (M7242), metaplasia (M73), and dysplasia (M74), polyps (M76800), and in combination with other codes (fibrous dysplasia, M74910). Furthermore, there are M-codes for benign tumors (M800), suspected cancer (the code M801 also includes in situ cancer), and metastases (M80106). Other M80-M99 codes represent different forms of cancer (eg, hepatocellular cancer, M817, with several subheadings; hepatoblastoma, M897; carcinoids, M82401; and familial polyposis coli, M82200). Most of the M80-99 codes refer to cancers outside the gastrointestinal tract.

The ESPRESSO database complements the national registers in a number of ways (Table 2). The database can be used to identify diseases that lack a specific ICD code (diseases that are recognized, such as microscopic colitis, and also conditions that are more diffuse, eg, nonspecific inflammation in the gut). The database can also help

increase sensitivity and specificity in identifying disease (eg, fatty liver disease), identify disease precursors (early stages of dysplasia), and grade the severity of disease and disease extent (using topography codes to define the location of disease, eg, transverse colon, T674). Multiple longitudinal biopsies allow researchers to follow the mucosal response to treatment over time (eg, repeated biopsies in IBD patients who have started on biological treatment).

Free text

Certain histopathology reports contain free text. This text can be scanned manually or by computerized algorithms to identify diagnoses or to characterize patients and their lesions. Table 3 exemplifies reports of free text.

Ethics

Preceding data collection, the general outline of the ESPRESSO study was approved by the Stockholm Ethics

Table 2 Comparison between Swedish hospital-based data and the ESPRESSO histopathology data

Identification through hospital-based data	Identification through histopathology registers
Hospital admission may vary with age (especially in children and old people) and disease activity	When a diagnosis is conditional on histopathology, then, histopathology data represent average individuals with GI disease, irrespective of their age, sex, and disease activity, thereby minimizing selection bias
Very low sensitivity for conditions that lack ICD codes (eg, microscopic colitis) or when such codes were introduced late into the ICD system (eg, eosinophilic esophagitis)	Applying histopathology data is the only way to identify several diseases (eg, microscopic colitis and serrated adenomas) and has been the gold standard for coeliac disease in children ²⁷ and remains so in adults
The date of hospital admission may not correlate with the date of diagnosis. The first inpatient diagnosis of GI disease is sometimes registered only years later	The date of first histopathology usually correlates with the actual diagnostic date. This information is of particular importance when examining complications in, and risk factors for, diseases (eg, IBD; Figure 4)

Notes: ²⁷In 2012, Sweden adopted a non-biopsy option for selected children with high suspicion of coeliac disease who fulfilled certain criteria.²⁷ Biopsy is still recommended for all adults diagnosed with coeliac disease.²⁸

Abbreviation: GI, gastrointestinal.

Table 3 Examples of free text in the ESPRESSO database

<p>M-code: M72040: hyperplastic polyp</p> <p>Biopsy no. 3 shows a hyperplastic polyp. Biopsy no 7 demonstrates active ulcerative colitis with severely aberrant crypt pattern, intense inflammation and crypt abscesses, and partly the surface epithelium cannot be seen. The findings are consistent with an ulcerative proctitis. Biopsy I shows a normal ileum, biopsies II–VI. Normal colon. III. Hyperplastic polyp in the colon. VII. Active ulcerative colitis</p>
<p>M-code: M49590: primary biliary cirrhosis</p> <p>The biopsy shows microscopic liver tissue with a preserved basic structure. Hepatocytes are characterized by mild anisokaryosis. The sinusoids are partly dilated. There are only a small number of inflammatory cells in the lobuli. Ordinary Kupffer cell activity. Iron staining is negative. In the portal areas, there is a small increased amount of connective tissue. There are areas with a somewhat reduced number of biliary ducts. Preserved bile ducts show ordinary epithelium. There is a mild-to-moderate increase in the number of inflammatory cells in the portal areas. No granulomas are observed. The border between lobuli and the portal zones is somewhat irregular. The biopsy is consistent with primary biliary cirrhosis (stages I–II)</p>
<p>M-code: M47170: microscopic colitis (lymphocytic colitis)</p> <p>In the biopsies from the colonic mucosa is seen a chronic inflammation with an increased number of lymphocytes. I see no signs of collagenous colitis; nor is there any atypia or malignancy. Instead, the biopsy is consistent with microscopic colitis</p>
<p>M-code: M00110: microscopically normal mucosa in the duodenum</p> <p>In both specimens from the duodenal mucosa, the architecture is normal with long and slender villi. The villous epithelium is normal, and there is no increase in the number of intraepithelial lymphocytes. There is nothing pathological in the deeper parts of the mucosa. Histologically, this is a normal duodenal mucosa, strongly arguing against coeliac disease</p>
<p>M-code: M74009: severe dysplasia in the (pyloric) antrum</p> <p>In some areas, there is a substantially normal stomach mucosa of antrum type, but in most parts of the mucosa there is a strong atypia. The foveole are irregular, and the cylinder epithelium is severely atypical. I cannot confirm that there is invasive growth. In some areas, the atypical mucosa has eroded. In all biopsies, there is hence severe dysplasia. Invasiveness cannot be confirmed, but this is clearly a precancerous lesion in the antral mucosa</p>
<p>M-code: D6214: indeterminate colitis (IBD-U)</p> <p>The biopsies consist of separate colorectal biopsies. Microscopically, there is colorectal mucosa with a rich number of inflammatory cells in the lamina propria, among which granulocytes can be observed. There are an increased number of intraepithelial lymphocytes and granulocytes. In addition, the gland architecture is aberrant. No ulcerations. No granulomas or crypt abscesses. No atypia or malignancy. The specimen is suggestive of IBD, but there are no convincing signs of activity. In conclusion, the findings are consistent with IBD</p>

Review Board (No. 2014/1287-31/4) on August 27, 2014. Ethics amendments have since allowed us to add data from the Prescribed Drug Register, data on health economics, and data on country of birth and nationality and to validate a number of specific diagnoses (No. 2017/1497-32; July 19, 2017). Linkage to the Swedish Twin Register²⁰ was approved on October 30, 2017 (No. 2017/2087). A more detailed background of the study was submitted to the Stockholm Ethics Review Board (No. 2018/972-32) and approved on May 14, 2018.

The large-scale register-based nature of the ESPRESSO database determines that participants will not be directly contacted by the researchers. Therefore, the ethics review board has waived informed consent.²⁹ To protect the integrity of the data, linked data are “pseudonymized before delivery”. This pseudonymization procedure implies that the PINs have been replaced by serial numbers by the government agencies delivering data to the researchers. However, we (researchers) have requested that a key between the PIN and the serial number should be saved at Statistics Sweden until March 2021. Such a key allows for additional linkages (eg, if some data are found to be incorrect) or updates (after relevant ethics approval). Furthermore, requesting informed consent for 13 million people would make this research not only impossible to carry out (for the economic resources needed to contact everyone and because a number of participants have deceased) but also because selective opt out^{30–32} would damage the validity and statistical power of the study. Vulnerable participants (such as minorities and children) would be at greater risk of being excluded. Finally, the ESPRESSO study aims to examine mortality, and because the data are retrospective, it is not possible to go back in time to approve study participation. Through the ESPRESSO database, we can avoid selection bias and have access to general population-based data. It will also allow us to examine side effects of hazardous exposures that otherwise would have been unethical to investigate in a randomized clinical trial.

The main risk with register-based research is if someone, by mistake or purposely, reveals the identity of the study participants. That is possible either through stealing the key from the National Board of Health and Welfare or through “backtracking”. Assume, for instance, that Patient X is a woman, who was born in 1977, immigrated to Sweden in 1987, is of Jamaican heritage, and is now living in one of the Northern provinces of Sweden. These bits of information could be utilized to deduce the true identity of Patient X, a deduction which is illegal. To safeguard against this potentiality, the National Board of Health and Welfare and

Statistics Sweden have established a protocol in which most often only the year and month of birth (but not day) are supplied for individuals >2 years of age at first biopsy and where nationalities are aggregated into larger groups.

Ethics approvals in Sweden are not time limited, and hence no annual approval is needed.³³ The ESPRESSO database has been registered with the Karolinska Institutet personal data officer (Swedish: Personuppgiftsombud).

Potential research use of the ESPRESSO database

As noted by the ESPRESSO acronym, the strength of this study lies in its ability to combine histopathology data with the rich resources of the Swedish National Healthcare Registers. Although it is particularly relevant to examine gastrointestinal disease that is identified in the Patient Register (to determine the phenotype of patients or to increase the validity of a diagnosis), mucosal histopathology can serve other purposes. For instance, ESPRESSO helps in the identification of patients with irritable bowel syndrome (IBS) or IBD with a relevant ICD code in the 13 million person cohort but with a possibility to stratify for mucosal appearance (normal, inflammation, other appearance, etc).

We illustrate how gastrointestinal histopathology data can improve epidemiological research in the following sections.

Infections

Through linkage to the Patient Register, it is possible to examine gastrointestinal diseases and sepsis.³⁴ Infections may potentially cause gastrointestinal disease (eg, multiple reports indicate an association between rotavirus and coeliac disease),^{35,36} but gastrointestinal disease per se may also predispose to infections.³⁷ Other times, infections may mimic gastrointestinal disease (intestinal tuberculosis can be misjudged as Crohn’s disease, tuberculosis as coeliac disease, etc). It is therefore important to combine histopathology data and data from the Patient Register and the Cancer Register (as well as other registers).

Cancer

Several gastrointestinal diseases are characterized by inflammation, and inflammatory activity has been linked to a number of cancers.^{16,38–47} Earlier studies suggested that Barrett’s esophagus may be a risk factor for ear–nose–throat cancer. Although coeliac disease has been linked to a lower risk of lung cancer⁴⁸ (possibly because of lower rates of smoking among patients),⁴⁹ these same patients are probably at increased risk of both liver cancer and IBD.^{50,51} IBD of

itself has been associated with thymic cancer.⁵² Inflammation has also been linked to malignant melanoma⁵³ and to nonmelanoma skin cancer.^{54,55} The linkage to the Prescribed Drug Register allows researchers (and practitioners) to add data on medication and thereby to disentangle the effects of inflammation and treatment (sometimes for inflammation).

Psychiatric disease

In recent years, interest in the brain–gut axis has increased dramatically. In addition, research has increasingly focused on the potential implications of antibiotic use, the microflora, and incident gastrointestinal disorder in psychiatric disease.⁵⁶ Our research group has previously demonstrated an association between coeliac disease and depression⁵⁷ (and also other psychiatric disorders^{58,59} and even suicide).¹⁴ Crohn's disease has been linked to depression,⁶⁰ and some data suggest that the presence and extent of depression influences the risk of complications in IBD.^{60,61} Compellingly, antidepressive treatment can sometimes ameliorate the symptoms and situation of patients with gastrointestinal pain.

Through linkage to normal mucosa and serology data, we have detected a potential association between early (potential) gastrointestinal disease and autism.⁶² Some patients with gastrointestinal disease describe a phenomenon known as “brain fog”,⁶³ whereas others may experience ataxia.⁶⁴ Depression is also more common in patients with colorectal cancer;⁶⁵ however, the role of inflammation for this risk increase is unknown. Attention-deficit hyperactivity disorder is another neuropsychiatric disorder that has been linked to gastrointestinal disease, including liver disease.⁶⁶

Neurology

Neurological conditions may be congenital or acquired. Many children with early-life neurological disease suffer from concomitant gastrointestinal tract conditions. Here, ESPRESSO offers the advantage to differentiate between children with biopsy-verified abnormalities and children with a normal mucosa. Dementia has been linked to IBD,⁶⁷ and although a recent Swedish study found no association between IBD and Parkinson disease,⁶⁸ Danish researchers just reported a positive association.⁶⁹ Other disorders that have been linked to gastrointestinal disease include epilepsy,⁷⁰ migraine,^{71,72} and neuropathy.⁷³ Via the ESPRESSO database, it is possible to examine nonspecific gastrointestinal inflammation and its association with neurological disease.

Cardiovascular disease

The ESPRESSO database contains hundreds of thousands of individuals with chronic inflammation in the gut. Thus,

researchers are able to examine the association with chronic inflammation, specific gastrointestinal disease (other than coeliac diseases),^{74–79} and cardiovascular disease. Inflammation may not only be a risk factor for ischemic heart disease and stroke but also for thromboembolism.⁸⁰

Respiratory disease

Respiratory disease is a major cause of death worldwide^{81,82} and has been linked to a number of gastrointestinal disorders.^{83–85} Some data indicate that patients with gastrointestinal disease do not respond to certain vaccinations; if this assertion proves to be true, such patients may be at increased risk of severe infections (respiratory infections, including influenza^{37,86} and respiratory syncytial virus in small children⁸⁷ and also pneumonia).⁸⁴ Certain inflammatory gastrointestinal conditions, such as eosinophilic esophagitis (characterized by a typical histopathology in the esophagus) may be linked to allergic disorders, including asthma.

Musculoskeletal disease and connective tissue disorders

Some connective tissue disorders have autoimmune traits and may be linked to gastrointestinal autoimmune disease.⁸⁸ Because of malabsorption and calcium deficiency, bone mineral density is often decreased in patients with gastrointestinal disease. Although the ESPRESSO database does not contain data on bone mineral density, it offers an opportunity to calculate precise risk estimates for fractures (especially for osteoporotic fractures).⁸⁹ gastrointestinal diseases, with chronic inflammation and poor nutrient uptake, may predispose to osteoporosis and fractures.^{90–93}

Renal disease

A PubMed search in May 2018 revealed 120 publications on gastrointestinal disorders and glomerulonephritis. We have previously shown that coeliac disease is associated with end-stage renal disease⁹⁴ as well as with milder renal disease.⁹⁵ The Patient Register contains data on procedures, implicating that data on dialysis and renal transplantation can be ascertained as per the linked ESPRESSO database.

Pregnancy and perinatal health

By linkage to the Medical Birth Register and the Multigeneration Register (the latter is part of the Total Population Register),¹⁷ it is possible to examine pregnancy outcome in offspring to parents (both mothers and fathers) with gastrointestinal disease.^{96–101} Many women with chronic gastrointestinal diseases are encouraged to give birth through cesarean section, which may have serious consequences for offspring health. Of note,

the ESPRESSO database includes information on procedures. Two recent papers failed to detect any substantial impact on pregnancy outcome from endoscopy or liver biopsy during pregnancy.^{102,103} Yet, pregnancy outcome may not only be seen as a consequence of parental gastrointestinal disease,^{101,104} ie, it could also be a risk factor for future gastrointestinal disease in the offspring.^{105–107} Finally, certain gastrointestinal diseases have their onset during early life (eg, necrotising enterocolitis and pyloric stenosis) and in this context the ESPRESSO database may help to understand such diseases.

Other diseases

Endocrine, nutritional, and metabolic disorders may serve as confounders and others as differential diagnoses when using histopathology data to identify gastrointestinal disease. Thyroid disease, type 1 diabetes (T1D), and Addison's disease have all been linked to gastrointestinal disease.^{108–110} Our research group and others have used histopathology data (villous atrophy, ie, coeliac disease) to examine whether forms of autoimmune gastrointestinal disease influence the outcome of T1D.^{111–113}

Using the Patient Register, we can also identify alcohol abuse. This possibility is important for several reasons. Alcohol may be a risk factor for gastrointestinal disease but may also influence liver pathology. Linkage to the Patient Register will also allow researchers to calculate the Charlson comorbidity index or other comorbidity indices when needed for studies in the ESPRESSO database.

The ESPRESSO database also contains limited data on skin, eye, and ear–nose–throat disease.

Use of medications

Linkage to the Prescribed Drug Register unlocks opportunities for pharmaco-epidemiological research in gastrointestinal disease. Certain drugs aim to change the microenvironment in the gastrointestinal tract (eg, proton pump inhibitors);¹¹⁴ others, such as antibiotics, will influence the microflora, potentially impacting the risk of gastrointestinal disease. Medication data can also be used to identify individuals with milder disease in which hospital admission is rarely needed (eg, hypertension or hyperlipidaemia). NSAIDs have been linked to gastric ulcers and gastrointestinal bleeding.¹¹⁵ Oral contraceptives have been linked to IBD;¹¹⁶ tetracycline use may induce fatty liver development; and some data suggest that aspirin could have an antitumor effect. In all, linkage to Prescribed Drug data enhances the quality and usefulness of histopathology data.

Discussion

In this profile paper, we outline the background, structure, and potential use of the ESPRESSO cohort. The strengths of the

database include the large number of participants, the longitudinal records with histopathology, actual data on histopathology (including normal mucosa, inflammation, and cancer precursors), linkage to the Swedish National Registers, complementary medical data (medical diagnoses, medication use, etc), and data on first-degree relatives and first spouses, which could serve as important secondary controls to examine intrafamilial confounding when that is an issue. The ESPRESSO database also allows for comparisons between individuals with a certain disease, as diagnosed through histopathology (eg, microscopic colitis), and individuals with a normal mucosa on biopsy, hereby minimizing the influence of health-seeking pattern and reducing confounding due to underlying comorbidity.

The ESPRESSO cohort has some limitations. Figure 2 is likely to reflect both an increased use of histopathology and an increase in reporting of actual biopsies. The coverage of the database may be suboptimal, especially before the year 1990. We do not have access to biological data. This limitation prohibits DNA analyses, or any other laboratory analysis, if the ESPRESSO data are not linked to other data sources containing relevant biological data. That is the case with our planned linkage to the Swedish Twin Register.

We cannot rule out a certain amount of selection bias. Patient characteristics may differ from that of the general patient population, if histopathology report data are used to identify diseases that may otherwise not require biopsy, or where biopsy is typically carried out only for advanced disease. The ESPRESSO database is likely to have lower coverage for certain gastrointestinal diseases in patients where endoscopy¹⁰² and, eg, liver biopsy¹⁰³ are sometimes avoided due to other concurrent conditions (eg, pregnancy). The new recommendations allowing for a non-biopsy celiac diagnosis in certain children,²⁷ but not in adults,¹¹⁷ mean that not all children with incident celiac disease after 2012 may have been identified. It is also possible that in a population where cancer is more common (older people), endoscopies are more often carried out to make sure that a gastrointestinal symptom is not caused by malignancy, and this may lead to higher number of biopsies with both normal and abnormal findings among the elderly. In addition, reference individuals may differ with regard to, eg, age and sex, as they were matched on these factors with individuals having a histopathology report.

In conclusion, the ESPRESSO database offers a tantalizing opportunity to strengthen epidemiology through histopathology.

Abbreviations

ESPRESSO, Epidemiology Strengthened by histoPathology Reports in Sweden; GI, gastrointestinal; IBD, inflammatory

bowel disease; IBD-U, unclassified IBD; PIN, personal identity number; PPV, positive predictive value; SnoMed CT, Systematized Nomenclature of Medicine-Clinical Terms.

Disclosure

Dr JFL coordinates a study on behalf of the Swedish IBD quality register (SWIBREG). This study has received funding from Janssen corporation. The authors report no other conflicts of interest in this work.

References

- Jakobsson GL, Sterngård 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.
- Eurostat [homepage on the Internet]. Mortality and life expectancy statistics: EU; 2017. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Mortality_and_life_expectancy_statistics. Accessed December 21, 2018.
- CIA [homepage on the Internet]. The World Fact Book; 2014. Available from: <https://www.cia.gov/library/publications/resources/the-world-factbook/fields/2225.html#sw>. Accessed July 24, 2018.
- Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekblom A. The Swedish personal identity number: possibilities and pitfalls in health-care and medical research. *Eur J Epidemiol*. 2009;24(11):659–667.
- Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish cancer register: a sample survey for year 1998. *Acta Oncol*. 2009;48(1):27–33.
- Ludvigsson JF, Welander A, Lassila R, Ekblom A, Montgomery SM. Risk of thromboembolism in 14,000 individuals with coeliac disease. *Br J Haematol*. 2007;139(1):121–127.
- Ludvigsson JF, Montgomery SM, Ekblom A. Risk of pancreatitis in 14,000 individuals with celiac disease. *Clin Gastroenterol Hepatol*. 2007;5(11):1347–1353.
- Ludvigsson JF, Wahlström J, Grunewald J, Ekblom A, Montgomery SM. Coeliac disease and risk of tuberculosis: a population based cohort study. *Thorax*. 2007;62(1):23–28.
- Ludvigsson JF, Andersson E, Ekblom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11(1):450.
- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Philadelphia: Lippincott William & Wilkins; 2008.
- Ludvigsson JF, Sanders DS, Maeurer M, Jonsson J, Grunewald J, Wahlström J. Risk of tuberculosis in a large sample of patients with coeliac disease—a nationwide cohort study. *Aliment Pharmacol Ther*. 2011;33(6):689–696.
- Sadr-Azodi O, Sanders DS, Murray JA, Ludvigsson JF. Patients with celiac disease have an increased risk for pancreatitis. *Clin Gastroenterol Hepatol*. 2012;10(10):1136–1142.
- Ludvigsson JF, Montgomery SM, Ekblom A, Brandt L, Granath F. Small-intestinal histopathology and mortality risk in celiac disease. *JAMA*. 2009;302(11):1171–1178.
- Ludvigsson JF, Sellgren C, Runeson B, Långström N, Lichtenstein P. Increased suicide risk in coeliac disease—a Swedish nationwide cohort study. *Dig Liver Dis*. 2011;43(8):616–622.
- Lebwohl B, Granath F, Ekblom A, et al. Mucosal healing and mortality in coeliac disease. *Aliment Pharmacol Ther*. 2013;37(3):332–339.
- Lebwohl B, Granath F, Ekblom A, et al. Mucosal healing and risk for lymphoproliferative malignancy in celiac disease: a population-based cohort study. *Ann Intern Med*. 2013;159(3):169–175.
- Ludvigsson JF, Almqvist C, Bonamy AK, et al. Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol*. 2016;31(2):125–136.
- Brooke HL, Talbäck M, Hörnblad J, et al. The Swedish cause of death register. *Eur J Epidemiol*. 2017;32(9):765–773.
- Wettermark B, Hammar N, Foré CM, et al. The new Swedish Prescribed Drug Register—opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf*. 2007;16(7):726–735.
- Magnusson PK, Almqvist C, Rahman I, et al. The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res Hum Genet*. 2013;16(1):317–329.
- Ludvigsson JF, Brandt L, Montgomery SM, Granath F, Ekblom A. Validation study of villous atrophy and small intestinal inflammation in Swedish biopsy registers. *BMC Gastroenterol*. 2009;9(1):19.
- Smedby KE, Akerman M, Hildebrand H, Glimelius B, Ekblom A, Askling J. Malignant lymphomas in coeliac disease: evidence of increased risks for lymphoma types other than enteropathy-type T cell lymphoma. *Gut*. 2005;54(1):54–59.
- Ludvigsson JF, Brandt L, Montgomery SM. Symptoms and signs in individuals with serology positive for celiac disease but normal mucosa. *BMC Gastroenterol*. 2009;9(1):57.
- Svensson M, Bergman D, Olen O. Validating microscopic colitis (MC) in Swedish pathology registers. *Scand J Gastroenterol*. Epub 2019 Jan 2.
- Olén O, Askling J, Sachs MC, et al. Childhood onset inflammatory bowel disease and risk of cancer: a Swedish nationwide cohort study 1964–2014. *BMJ*. 2017;358:j3951.
- Everhov ÅH, Halfvarson J, Myrelid P, et al. Incidence and treatment of patients diagnosed with inflammatory bowel diseases at 60 years or older in Sweden. *Gastroenterology*. 2018;154(3):518–528.
- Husby S, Koletzko S, Korponay-Szabó IR, et al. European society for pediatric gastroenterology, hepatology, and nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr*. 2012;54(1):136–160.
- Ludvigsson JF, Bai JC, Biagi F, et al. Diagnosis and management of adult coeliac disease: guidelines from the British Society of Gastroenterology. *Gut*. 2014;63(8):1210–1228.
- WHO. World Health Organization; 2015. Available from: http://www.who.int/rpc/research_ethics/Process_seeking_IF_printing.pdf. Accessed December 21, 2018.
- Trevena L, Irwig L, Barratt A. Impact of privacy legislation on the number and characteristics of people who are recruited for research: a randomised controlled trial. *J Med Ethics*. 2006;32(8):473–477.
- Woolf SH, Rothenich SF, Johnson RE, Marsland DW. Selection bias from requiring patients to give consent to examine data for health services research. *Arch Fam Med*. 2000;9(10):1111–1118.
- Junghans C, Feder G, Hemingway H, Timmis A, Jones M. Recruiting patients to medical research: double blind randomised trial of “opt-in” versus “opt-out” strategies. *BMJ*. 2005;331(7522):940.
- Ludvigsson JF, Håberg SE, Knudsen GP, et al. Ethical aspects of registry-based research in the Nordic countries. *Clin Epidemiol*. 2015;7:491–508.
- Ludvigsson JF, Olén O, Bell M, Ekblom A, Montgomery SM. Coeliac disease and risk of sepsis. *Gut*. 2008;57(8):1074–1080.
- Rostami-Nejad M, Rostami K, Sanaei M, et al. Rotavirus and coeliac autoimmunity among adults with non-specific gastrointestinal symptoms. *Saudi Med J*. 2010;31(8):891–894.
- Stene LC, Honeyman MC, Hoffenberg EJ, et al. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am J Gastroenterol*. 2006;101(10):2333–2340.
- Mårild K, Fredlund H, Ludvigsson JF. Increased risk of hospital admission for influenza in patients with celiac disease: a nationwide cohort study in Sweden. *Am J Gastroenterol*. 2010;105(11):2465–2473.
- Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420(6917):860–867.
- Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. 2008;454(7203):436–444.
- Ullman TA, Itzkowitz SH. Intestinal inflammation and cancer. *Gastroenterology*. 2011;140(6):1807–1816.

41. Mellander MR, Ekbom A, Hultcrantz R, Löfberg R, Öst Å, Björk J. Microscopic colitis: a descriptive clinical cohort study of 795 patients with collagenous and lymphocytic colitis. *Scand J Gastroenterol*. 2016;51(5):556–562.
42. Wang SS, Vajdic CM, Linet MS, et al. Associations of non-Hodgkin Lymphoma (NHL) risk with autoimmune conditions according to putative NHL loci. *Am J Epidemiol*. 2015;181(6):406–421.
43. Kotlyar DS, Lewis JD, Beaugerie L, et al. Risk of lymphoma in patients with inflammatory bowel disease treated with azathioprine and 6-mercaptopurine: a meta-analysis. *Clin Gastroenterol Hepatol*. 2015;13(5):847–858.
44. Ludvigsson JF, Lebowl B, Rubio-Tapia A, et al. Does celiac disease influence survival in lymphoproliferative malignancy? *Eur J Epidemiol*. 2013;28(6):475–483.
45. Ludvigsson JF, Lebowl B, Rubio-Tapia A, Murray JA, Green PH, Ekbom A. Risk of lymphoproliferative malignancy in celiac patients with a family history of lymphoproliferative malignancy. *J Gastroenterol*. 2013;48(12):1324–1331.
46. Elfström P, Granath F, Ekström Smedby K, et al. Risk of lymphoproliferative malignancy in relation to small intestinal histopathology among patients with celiac disease. *J Natl Cancer Inst*. 2011;103(5):436–444.
47. Landgren O, Pfeiffer RM, Kristinsson SY, Björkholm M. Survival patterns in patients with Hodgkin's lymphoma with a pre-existing hospital discharge diagnosis of autoimmune disease. *J Clin Oncol*. 2010;28(34):5081–5087.
48. Ludvigsson JF, West J, Hubbard R, Card T. Neutral risk of lung cancer in adults with celiac disease--nationwide cohort study. *Lung Cancer*. 2012;78(3):179–184.
49. Snook JA, Dwyer L, Lee-Elliott C, Khan S, Wheeler DW, Nicholas DS. Adult coeliac disease and cigarette smoking. *Gut*. 1996;39(1):60–62.
50. Bitton A, Vutocivi M, Sewitch M, Suissa S, Brassard P. Mortality trends in Crohn's disease and ulcerative colitis: a population-based study in Québec, Canada. *Inflamm Bowel Dis*. 2016;22(2):416–423.
51. Piccirillo JF. Importance of comorbidity in head and neck cancer. *Laryngoscope*. 2000;110(4):593–602.
52. Peters HC, Liu X, Iqbal A, Cunningham LA, Tan SA. Colorectal cancer metastasis to the thymus gland: rare presentation of colorectal cancer as anterior mediastinal mass. *Case Rep Surg*. 2017;2017:6581965.
53. Raaschou P, Simard JF, Holmqvist M, Askling J; ARTIS Study Group. Rheumatoid arthritis, anti-tumour necrosis factor therapy, and risk of malignant melanoma: nationwide population based prospective cohort study from Sweden. *BMJ*. 2013;346:f1939.
54. Long MD, Martin CF, Pipkin CA, Herfarth HH, Sandler RS, Kappelman MD. Risk of melanoma and nonmelanoma skin cancer among patients with inflammatory bowel disease. *Gastroenterology*. 2012;143(2):390–399.
55. Singh H, Nugent Z, Demers AA, et al. Increased risk of nonmelanoma skin cancers among individuals with inflammatory bowel disease. *Gastroenterology*. 2011;141(5):1612–1620.
56. Dinan TG, Cryan JF. Gut-brain axis in 2016: brain-gut-microbiota axis - mood, metabolism and behaviour. *Nat Rev Gastroenterol Hepatol*. 2017;14(2):69–70.
57. Ludvigsson JF, Reutfors J, Ösby U, Ekbom A, Montgomery SM. Coeliac disease and risk of mood disorders – a general population-based cohort study. *J Affect Disord*. 2007;99(1–3):117–126.
58. Butwicki A, Frisén L, Almqvist C, Zethelius B, Lichtenstein P. Risks of psychiatric disorders and suicide attempts in children and adolescents with type 1 diabetes: a population-based cohort study. *Diabetes Care*. 2015;38(3):453–459.
59. Mårild K, Størdal K, Bulik CM, et al. Celiac disease and anorexia nervosa: a nationwide study. *Pediatrics*. 2017;139(5):e20164367.
60. Gaines LS, Slaughter JC, Horst SN, et al. Association between affective-cognitive symptoms of depression and exacerbation of Crohn's disease. *Am J Gastroenterol*. 2016;111(6):864–870.
61. Gracie DJ, Guthrie EA, Hamlin PJ, Ford AC. Bi-directionality of brain-gut interactions in patients with inflammatory bowel disease. *Gastroenterology*. 2018;154(6):1635–1646.
62. Ludvigsson JF, Reichenberg A, Hultman CM, Murray JA. A nationwide study of the association between celiac disease and the risk of autistic spectrum disorders. *JAMA Psychiatry*. 2013;70(11):1224–1230.
63. Lebowl B, Ludvigsson JF. Editorial: “brain fog” and coeliac disease – evidence for its existence. *Aliment Pharmacol Ther*. 2014;40(5):565.
64. Hadjivassiliou M, Sanders DD, Aeschlimann DP. Gluten-related disorders: gluten ataxia. *Dig Dis*. 2015;33(2):264–268.
65. Braamse AMJ, van Turenhout ST, Terhaar Sive Droste JS, et al. Factors associated with anxiety and depressive symptoms in colorectal cancer survivors. *Eur J Gastroenterol Hepatol*. 2016;28(7):831–835.
66. Fehér E, Pongor Éva, Altdorfer K, Kóbori L, Lengyel G. Neuroimmunomodulation in human autoimmune liver disease. *Cell Tissue Res*. 2013;354(2):543–550.
67. Papatheanasiou A, Nikakis P, Bonakis A, et al. Rapidly progressive dementia as presenting feature in inflammatory bowel disease. *Alzheimer Dis Assoc Disord*. 2014;28(3):294–295.
68. Weimers P, Halfvarson J, Sachs MC, et al. Inflammatory bowel disease and Parkinson's disease: a Nationwide Swedish Cohort Study. *Inflamm Bowel Dis*. 2019 Jan 1;25(1):111–123.
69. Villumsen M, Aznar S, Pakkenberg B, Jess T, Brudek T. Inflammatory bowel disease increases the risk of Parkinson's disease: a Danish nationwide cohort study 1977–2014. *Gut*. 2019 Jan;68(1):18–24.
70. Ludvigsson JF, Zingone F, Tomson T, Ekbom A, Ciacci C. Increased risk of epilepsy in biopsy-verified celiac disease: a population-based cohort study. *Neurology*. 2012;78(18):1401–1407.
71. Bushara KO. Neurologic presentation of celiac disease. *Gastroenterology*. 2005;128(4):S92–S97.
72. Hadjivassiliou M, Grunewald RA, Lawden M, Davies-Jones GA, Powell T, Smith CM. Headache and CNS white matter abnormalities associated with gluten sensitivity. *Neurology*. 2001;56(3):385–388.
73. Ludvigsson JF, West J, Card T, Appelros P. Risk of stroke in 28,000 patients with celiac disease: a nationwide cohort study in Sweden. *J Stroke Cerebrovasc Dis*. 2012;21(8):860–867.
74. Lebowl B, Emilsson L, Fröbert O, Einstein AJ, Green PH, Ludvigsson JF. Mucosal healing and the risk of ischemic heart disease or atrial fibrillation in patients with celiac disease; a population-based study. *PLoS One*. 2015;10(1):e0117529.
75. Ludvigsson JF, James S, Askling J, Stenestrand U, Ingelsson E. Nationwide cohort study of risk of ischemic heart disease in patients with celiac disease. *Circulation*. 2011;123(5):483–490.
76. Lebowl B, Cao Y, Zong G, et al. Long term gluten consumption in adults without celiac disease and risk of coronary heart disease: prospective cohort study. *BMJ*. 2017;357:j1892.
77. Emilsson L, Carlsson R, James S, Hambraeus K, Ludvigsson JF. Follow-up of ischaemic heart disease in patients with coeliac disease. *Eur J Prev Cardiol*. 2013;22(1):83–90.
78. Emilsson L, James S, Ludvigsson JF. Ischaemic heart disease in first-degree relatives to coeliac patients. *Eur J Clin Invest*. 2014;44(4):359–364.
79. Emilsson L, Andersson B, Elfström P, Green PHR, Ludvigsson JF. Risk of idiopathic dilated cardiomyopathy in 29 000 patients with celiac disease. *J Am Heart Assoc*. 2012;1(3):e001594.
80. Vazquez-Garza E, Jerjes-Sanchez C, Navarrete A, Joya-Harrison J, Rodriguez D. Venous thromboembolism: thrombosis, inflammation, and immunothrombosis for clinicians. *J Thromb Thrombolysis*. 2017;44(3):377–385.
81. Bernstein CN, Nugent Z, Targownik LE, Singh H, Lix LM. Predictors and risks for death in a population-based study of persons with IBD in Manitoba. *Gut*. 2015;64(9):1403–1411.
82. Jess T, Frisch M, Simonsen J. Trends in overall and cause-specific mortality among patients with inflammatory bowel disease from 1982 to 2010. *Clin Gastroenterol Hepatol*. 2013;11(1):43–48.
83. Grainge MJ, West J, Card TR, Holmes GK. Causes of death in people with celiac disease spanning the pre- and post-serology era: a population-based cohort study from Derby, UK. *Am J Gastroenterol*. 2011;106(5):933–9.

84. Zingone F, Abdul Sultan A, Crooks CJ, Tata LJ, Ciacci C, West J. The risk of community-acquired pneumonia among 9803 patients with coeliac disease compared to the general population: a cohort study. *Aliment Pharmacol Ther.* 2016;44(1):57–67.
85. Moayyedi P, Burch N, Akhtar-Danesh N, et al. Mortality rates in patients with Barrett's oesophagus. *Aliment Pharmacol Ther.* 2008;27(4):316–320.
86. Kårhus LL, Gunnes N, Størdal K, et al. Influenza and risk of later celiac disease: a cohort study of 2.6 million people. *Scand J Gastroenterol.* 2018;53(1):15–23.
87. Tjernberg AR, Ludvigsson JF. Children with celiac disease are more likely to have attended hospital for prior respiratory syncytial virus infection. *Dig Dis Sci.* 2014;59(7):1502–1508.
88. Laszkowska M, Roy A, Lebowitz B, Green PH, Sundelin HE, Ludvigsson JF. Nationwide population-based cohort study of celiac disease and risk of Ehlers-Danlos syndrome and joint hypermobility syndrome. *Dig Liver Dis.* 2016;48(9):1030–1034.
89. Ludvigsson JF, Michaëlsson K, Ekbom A, Montgomery SM. Coeliac disease and the risk of fractures – a general population-based cohort study. *Aliment Pharmacol Ther.* 2007;25(3):273–285.
90. Heikkilä K, Pearce J, Mäki M, Kaukinen K. Celiac disease and bone fractures: a systematic review and meta-analysis. *J Clin Endocrinol Metab.* 2015;100(1):25–34.
91. Lebowitz B, Michaëlsson K, Green PH, Ludvigsson JF. Persistent mucosal damage and risk of fracture in celiac disease. *J Clin Endocrinol Metab.* 2014;99(2):609–616.
92. Weiss RJ, Wick MC, Ackermann PW, Montgomery SM. Increased fracture risk in patients with rheumatic disorders and other inflammatory diseases – a case-control study with 53,108 patients with fracture. *J Rheumatol.* 2010;37(11):2247–2250.
93. Card T, West J, Hubbard R, Logan RFA. Hip fractures in patients with inflammatory bowel disease and their relationship to corticosteroid use: a population based cohort study. *Gut.* 2004;53(2):251–255.
94. Welander A, Prütz K-G, Foröd M, Ludvigsson JF. Increased risk of end-stage renal disease in individuals with coeliac disease. *Gut.* 2012;61(1):64–68.
95. Welander A, Sundelin B, Foröd M, Ludvigsson JF. Increased risk of IgA nephropathy among individuals with celiac disease. *J Clin Gastroenterol.* 2013;47(8):678–683.
96. Ludvigsson J, Montgomery S, Ekbom A. Celiac disease and risk of adverse fetal outcome: a population-based cohort study. *Gastroenterology.* 2005;129(2):454–463.
97. Ludvigsson JF, Montgomery SM, Ekbom A. Coeliac disease in the father and risk of adverse pregnancy outcome: a population-based cohort study. *Scand J Gastroenterol.* 2006;41(2):178–185.
98. Hagström H, Höijer J, Ludvigsson JF, et al. Adverse outcomes of pregnancy in women with non-alcoholic fatty liver disease. *Liver Int.* 2016;36(2):268–274.
99. Wikström Shemer E, Marschall HU, Ludvigsson JF, et al. Intrahepatic cholestasis of pregnancy and associated adverse pregnancy and fetal outcomes: a 12-year population-based cohort study. *BJOG.* 2013;120(6):717–723.
100. Ludvigsson JF, Bergquist A, Ajne G, Kane S, Ekbom A, Stephansson O. A population-based cohort study of pregnancy outcomes among women with primary sclerosing cholangitis. *Clin Gastroenterol Hepatol.* 2014;12(1):95–100.
101. Stokkeland K, Ludvigsson JF, Hultcrantz R, et al. Increased risk of preterm birth in women with autoimmune hepatitis – a nationwide cohort study. *Liver Int.* 2016;36(1):76–83.
102. Ludvigsson JF, Lebowitz B, Ekbom A, et al. Outcomes of pregnancies for women undergoing endoscopy while they were pregnant: a nationwide cohort study. *Gastroenterology.* 2017;152(3):554–563.
103. Ludvigsson JF, Marschall HU, Hagström H, Höijer J, Stephansson O. Pregnancy outcome in women undergoing liver biopsy during pregnancy: a nationwide population-based cohort study. *Hepatology.* 2018;68(2):625–633.
104. Stephansson O, Larsson H, Pedersen L, et al. Crohn's disease is a risk factor for preterm birth. *Clin Gastroenterol Hepatol.* 2010;8(6):509–515.
105. Kaijser M, Akre O, Cnattingius S, Ekbom A. Preterm birth, low birth weight, and risk for esophageal adenocarcinoma. *Gastroenterology.* 2005;128(3):607–609.
106. Forsell L, Cnattingius S, Bottai M, et al. Increased risk of Barrett's esophagus among individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* 2013;11(7):790–794.
107. Mårild K, Ludvigsson JF, Størdal K. Current evidence on whether perinatal risk factors influence coeliac disease is circumstantial. *Acta Paediatr.* 2016;105(4):366–375.
108. Elfström P, Montgomery SM, Kämpe O, Ekbom A, Ludvigsson JF. Risk of thyroid disease in individuals with celiac disease. *J Clin Endocrinol Metab.* 2008;93(10):3915–3921.
109. Elfström P, Montgomery SM, Kämpe O, Ekbom A, Ludvigsson JF. Risk of primary adrenal insufficiency in patients with celiac disease. *J Clin Endocrinol Metab.* 2007;92(9):3595–3598.
110. Elfström P, Sundström J, Ludvigsson JF. Systematic review with meta-analysis: associations between coeliac disease and type 1 diabetes. *Aliment Pharmacol Ther.* 2014;40(10):1123–1132.
111. Mollazadegan K, Kugelberg M, Montgomery SM, Sanders DS, Ludvigsson J, Ludvigsson JF. A population-based study of the risk of diabetic retinopathy in patients with type 1 diabetes and celiac disease. *Diabetes Care.* 2013;36(2):316–321.
112. Mollazadegan K, Sanders DS, Ludvigsson J, Ludvigsson JF. Long-term coeliac disease influences risk of death in patients with type 1 diabetes. *J Intern Med.* 2013;274(3):273–280.
113. Leeds JS, Hopper AD, Hadjivassiliou M, Tesfaye S, Sanders DS. High prevalence of microvascular complications in adults with type 1 diabetes and newly diagnosed celiac disease. *Diabetes Care.* 2011;34(10):2158–2163.
114. Lebowitz B, Spechler SJ, Wang TC, Green PH, Ludvigsson JF. Use of proton pump inhibitors and subsequent risk of celiac disease. *Dig Liver Dis.* 2014;46(1):36–40.
115. Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Ann Intern Med.* 1991;115(10):787–796.
116. Khalili H, Higuchi LM, Ananthakrishnan AN, et al. Oral contraceptives, reproductive factors and risk of inflammatory bowel disease. *Gut.* 2013;62(8):1153–1159.
117. Rubio-Tapia A, Hill ID, Kelly CP, Calderwood AH, Murray JA; American College of Gastroenterology. ACG clinical guidelines: diagnosis and management of celiac disease. *Am J Gastroenterol.* 2013;108(5):656–676.

Supplementary material

Additional data on linkage between the ESPRESSO (Epidemiology Strengthened by histoPathology Reports in Sweden) database and inflammatory bowel disease (IBD) in the Swedish Patient Register

The number of histopathology reports in IBD patients diagnosed since 2002 was 4.9 in Crohn's disease, 4.8 in ulcerative colitis (UC), and 2.7 in unclassified IBD (IBD-U). Patients with Crohn's disease with an ICD code indicative of ileal disease (L1) more often had a record of ileal biopsies in ESPRESSO than had Crohn's disease with other locations. Crohn's disease patients with a perianal modifier were more likely to have a biopsy from the anorectal area than patients

without this modifier. Looking specifically at the proportion of IBD patients with ≥ 1 histopathology report, normal mucosa was seen in only 8% of Crohn's disease patients, 6% of UC patients, and 12% of IBD-U patients, indicating that most IBD patients have abnormal mucosa.

Because histopathology data also include specimens from surgery, we examined all bowel surgeries in Swedish IBD patients. We found that the proportion of patients with a histopathology record after surgery ($n=27,933$) was 86% in Crohn's disease, 83% in UC, and 94% in IBD-U. The higher proportion among patients with IBD-U may be because this is a more recent diagnosis and surgeries for IBD-U are likely to have taken place when the coverage of the ESPRESSO database was higher.

Clinical Epidemiology

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,

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

Dovepress

systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and 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.