Building a Robust Investigator-Initiated Platform: The I-CARE Experience

Julien Kirchgesner[1,*], Laurent Beaugerie[1], Cédric Baumann[2,3], Filip Baert[4], Jean-François Rahier[5], Daniel Bergemalm[6], Fiorella Cañete[7,8], Bénédicte Caron[9], Marta Cossignani[10], Sandra Dias[11], Klaudia Farkas[12,13], Parker Girod[14], Idan Goren[15,16], Tom Holvoet[17], Susanne Krasz[18], Aditi Kumar[19], Peter Rimmer[20], Juliette Sheridan[12,13], Jessica Thut[21], Mathieu Uzzan[22], Nikos Viazis[23], Petra Weimers[24], Edyta Zagórowicz[25], Anne Buisson[26], Luisa Avedano[27], Helene Rousseau[2,3], Laurent Peyrin-Biroulet[9], Shaji Sebastian[28] and I-CARE Collaborator Group

Investigator-initiated studies that include information collected by patients are rising, but limited data is available on patient and investigator experience in this setting. The I-CARE cohort included patients with inflammatory bowel disease (IBD) monthly collecting clinical information in 15 countries for up to 6 years. We describe patients and investigators' involvement in I-CARE and identify predictors of early withdrawal due to patient non-engagement. Patients' characteristics according to the number of electronic Patient-reported outcomes (ePRO) completed during follow-up were assessed. Predictors of early withdrawal due to patient non-engagement were identified using logistic regression. The coding of outcomes reported by patients and corrections by investigators on patients' ePROs were assessed. Among 12,846 patients included by 502 investigators, 79.3% and 77.3% filled more than one ePRO and at least one ePRO within 6 months before the study end date, respectively. All ePROs were completed in 72.8% and 56.4% of patients during year 1 and 3, respectively. Male gender, younger age (< 20), being unemployed or a student, and no previous history of abdominal surgery were associated with early withdrawal. Investigators corrected 52.5% of cancer or dysplasia reported by patients compared to 10% of serious infections. Investigators added or removed a treatment sequence in 19.6% of the 6708 patients treated with biologics. These results highlight the implication of patients in research and the importance of data validation by investigators alongside the challenge and potential of collecting medical data from patients. These findings can inform similar future initiatives in other diseases. (EudraCT, Number: 2014-004728-23; ClinicalTrials.gov, Number: NCT02377258).

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ While patient-reported data in investigator-initiated studies enhances real-world evidence, it faces challenges, notably attrition, and validation. Although administrative health databases offer large samples, they lack clinical detail. Primary data collection is more accurate, but it is costly and requires innovative solutions to balance feasibility and data quality.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This study aimed to assess the involvement of patients and investigators in the I-CARE cohort, which included 12,846 patients with inflammatory bowel disease (IBD). Clinical information was collected monthly from these patients in 15 countries for up to 6 years.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ Overall, 79.3% of patients completed at least one patient-reported outcomes (ePRO), with 77.3% remaining active near the end of the study. Early withdrawal was predicted by younger age, male gender, unemployment/student status, and no prior surgery. Investigators corrected 52.5% of patient-reported cancers/dysplasia, emphasizing the need for oversight. High engagement rates suggest that this model is viable for long-term observational research.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ This approach shows that patient-driven data collection, when combined with investigator validation, can produce high-quality real-world data on a large scale. Identifying attrition predictors enables targeted retention strategies, while hybrid data collection (from patients and clinicians) improves accuracy. These insights could inform more patient-centered research in pharmacoepidemiology.

Received September 4, 2025; accepted January 28, 2026. doi:10.1002/cpt.70231

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Team FORECAST, Assistance Publique-Hôpitaux de Paris, Hôpital Saint-Antoine, Department of Gastroenterology, F-75012, Paris, France; [2]Methodology, Data Management and Statistics Unit, University Hospital of Nancy, 54000, Nancy, France; [3]UMR 1319 INSPIIRE, University of Lorraine, 54000, Nancy, France; [4]Department of Gastroenterology, AZ Delta, Roeselare, Belgium; [5]Department of Gastroenterology and Hepatology, CHU UCL Namur, Université Catholique de Louvain, Yvoir, Belgium; [6]Faculty of Medicine and Health, Department of Gastroenterology, Örebro University, Örebro, Sweden; [7]Gastroentrology Department, Hospital Germans Trias i Pujol, Badalona, Catalonia, Spain; [8]CIBEREHD, Madrid, Spain; [9]INFINY Institute, Department of Gastroenterology, CHRU Nancy, INSERM NGERE, Université de Lorraine, F-54500, Vandœuvre-lès-Nancy, France; [10]Department of Gastroenterology, Ospedale dei Castelli Ariccia, Rome, Italy; [11]GEDII, IBD Portuguese Group, Portugal; [12]Center for Gastroenterology, Department of Internal Medicine, University of Szeged, Szeged, Hungary; [13]HCEMM-USZ Translational Colorectal Research Group, Szeged, Hungary; [14]Department of Gastroenterology, St Vincent's University Hospital & School of Medicine, University College Dublin, INITIATIVE, Investigator Network for IBD Research in Ireland; [15]Division of Gastroenterology, Rabin Medical Center, Petah Tikva, Israel, Affiliated With Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [16]Division of Gastroenterology, State University of New York Upstate Medical University, Syracuse, New York, USA; [17]Department of Gastroenterology, VITAZ, Moerlandstraat 1, Sint Niklaas, Belgium; [18]Universitätsklinikum Dresden, Medizinische Klinik 1, Abteilung Gastroenterologie, Fetscherstr. 74, 01307, Dresden, Germany; [19]Department of Gastroenterology, The Royal Wolverhampton NHS Trust, UK; [20]Department of Gastroenterology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK; [21]Hull University Teaching Hospitals NHS Trust, Hull Royal Infirmary, Hull, UK; [22]Paris Est Créteil University UPEC, Assistance Publique-Hôpitaux de Paris (AP-HP), Henri Mondor Hospital, Gastroenterology Department, Fédération Hospitalo-Universitaire TRUE InnovaTive theRapy for immUne disordErs, Créteil, France; [23]Gastroenterology Department, Evangelismos General Hospital, Athens, Greece; [24]Department of Gastroenterology, North Zealand University Hospital, Frederikssund, Denmark; [25]Department of Oncological Gastroenterology, The Maria Sklodowska-Curie National Research Institute of Oncology, Department of Gastroenterology, Hepatology and Clinical Oncology, Centre of Postgraduate Medical Education, Warsaw, Poland; [26]Association François-Aupetit, AFA Crohn RCH, Paris, France; [27]European Federation of Crohn's & Ulcerative Colitis Associations (EFCCA), Brussels, Belgium; [28]IBD Unit, Hull University Teaching Hospitals, Hull, UK. *Correspondence: Julien Kirchgesner (julien.kirchgesner@aphp.fr)

Contributors of the I-CARE Collaborator group are listed in the Appendix.

Inflammatory bowel diseases (IBD) including Crohn's disease (CD) and ulcerative colitis (UC) are lifelong chronic immune mediated inflammatory diseases characterized by relapsing and remitting episodes that may lead to irreversible intestinal lesions and severe disability over the life course.[1] IBD affect more than 2.5 and 1 million residents in Europe and the United States, respectively.[2,3] Biologics including drugs targeting tumor necrosis factor have revolutionized the medical treatment of IBD since 30 years, but the benefit–risk balance of biologics remained poorly studied.[4] Notably, randomized clinical trials (RCTs) cannot assess the risk of severe adverse events with a low incidence such as cancer, due to an inadequate sample size, in addition to the selected nature of patients included.

Real-world data (RWD) is usually defined as data routinely collected outside a clinical trial and can be collected through multiple methods.[5] RWD can be primarily collected for research (primary data), but 90% of RWD used for research is first collected for another purpose than research (secondary data),[6] notably administrative health databases. Administrative health databases potentially include a large sample size in a population-based setting.[7] On the other hand, treatment exposure is based on delivery, diagnoses are based on billing codes, and clinical parameters such as clinical symptoms are not collected. In the setting of IBD, disease activity based on clinical, imaging, and endoscopic parameters are not collected, and surrogate markers based on IBD related hospitalizations and surgeries are considered in studies based on administrative health databases. While the use of such surrogate markers has led to successful replication of RCTs,[8,9] it raises the question of residual confounding related to an inaccurate assessment of longitudinal IBD disease activity.

The low use of primary research data highlights the challenges to collect data for research, which is due to much higher costs and lack of dedicated time for data collection compared to secondary data. Usually, research assistants or healthcare practitioners are required to collect data, which has a major impact on cost and limits the feasibility in a setting of limited time to collect data for research. The implication of patients in research has evolved in various ways throughout the research process, from the involvement in the study design and data collection to the dissemination of the findings.[10,11] Data collection by patients has the potential to increase the feasibility of investigator-initiated cohorts by reducing the burden of data collection for healthcare practitioners, while increasing the relevance of data collected, notably for patient-reported outcomes. A major challenge in prospective cohorts is sustaining patient engagement over time. Identifying predictors of early withdrawal is critical to design interventions that improve representation and retention, particularly for high-risk groups.

The I-CARE (Ibd CAncer and seRious infections in Europe) study is a European prospective cohort study assessing the safety of biologics in patients with IBD, which included more than 10,000 patients in 15 countries.[12] This study aimed to report the experience of I-CARE regarding the patient and investigators involvement and data collection, while identifying predictors of early withdrawal due to patient non-engagement.

MATERIALS AND METHODS

The I-CARE study design and the full protocol are provided in the **Supplementary Material**.

Patients' implication

SANOIA, a service provider specializing in Electronic Patient-Reported Outcomes (ePRO), developed a dedicated web-based tool for gathering information directly from patients. This tool facilitated the regular collection of patient data by sending timely electronic reminders to ensure completion of the ePRO. Patients engaged in monthly reporting, including notably details on (i) current IBD-related treatment; (ii) clinical disease activity, quality of life, and disability based on validated scores; (iii) the occurrence of imaging or gastrointestinal endoscopy, including date and type of procedures; (iv) hospitalizations, surgeries, infections, cancer, and dysplasia diagnoses. Cancers of interest were not limited to gastrointestinal sites. Patients could report cancers across all anatomical locations, including but not limited to anus, colon, rectum, small bowel, and stomach, as well as non-gastrointestinal sites such as breast, lung, and prostate. Dysplasia reporting was focused on colon, esophagus, and uterus (Cervical

Intraepithelial Neoplasia grade 3 [CIN3] only). All questionnaires were in their local language. Patients were further asked to provide the hospitalization report in case of hospitalization or the pathology report in case of cancer or dysplasia. An alert system was established to notify study coordinators and project managers of incomplete data or reports concerning hospitalizations, cancer, or dysplasia. The alerts were exclusively for research purposes, ensuring data completeness by prompting coordinators to request supporting documents when patients reported outcomes.

In patients not filling the ePRO in due time after having received an electronic notification within 15 days, two text message reminders were sent over the following month. In case of no ePRO filled during a period of 6 months despite electronic reminders and investigator solicitation, follow-up was censored and patients could not fill further ePROs.

Investigator implication

Each investigator identified eligible patients and obtained written informed consent. At the inclusion visit, investigators collected the baseline demographic and disease characteristics data in an electronic case report form (eCRF). During follow-up, investigators were annually asked to validate an annual summary and, if needed, rectify any erroneous information that patients had entered prospectively on a monthly basis regarding treatment exposure. Additionally, endoscopic and imaging disease activity were assessed using a predetermined and simplified scoring system for imaging or gastrointestinal endoscopy declared by patients. As the patient's referring physician, investigators were already aware of any clinical events prior to patient reporting. Investigators could declare events not reported or modify events reported by the patients, including imaging and endoscopy occurrence, hospitalizations, surgeries, dysplasia, and cancer. If the patient did not respond to the ePRO after the two text message reminders, the investigator was notified and encouraged to contact the patient to motivate continued participation.

Outcomes coding methodology

Anonymized hospitalization and pathology reports were provided by patients or investigators in case of hospitalization, dysplasia, or cancer reported. After the upload of the reports by the study coordinators or project managers on the web platform, each report was validated and the related information was collected on the eCRF by a gastroenterologist. A scientific coding committee was set up including at least one gastroenterologist from each country. Each member of the committee attended online training and guidelines were available to ensure that the coding process was reproducible across countries.

For hospitalizations, discharge diagnoses were coded using the International Classification of Diseases, 10th edition (ICD-10). The health condition for which the patient was hospitalized was defined as the primary diagnosis and up to 9 associated diagnoses could be coded. Additionally, surgical procedures were collected, and the Charlson comorbidity index was calculated. For cancers and dysplasia, the site and histological type were coded using the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3).

Statistical analysis

Patients were excluded from this study if they withdrew secondarily consent to the use of their data, if data on demographics, disease characteristics, or treatment exposure at baseline were missing, or in case of a duplicate entry.

Patients' attrition

We assessed the number of ePRO filled for each patient during follow-up. The collection of ePROs over time was defined by the proportion of ePROs collected according to the expected follow-up (3 or 6 years according to the countries, or follow-up and censoring due to death).

Three groups of patients were considered: patients who filled no ePRO, patients who filled only one ePRO, and patients who filled more than one ePRO. Demographics and clinical characteristics were assessed in each group. Additionally, we assessed the predictors of no effective participation related to early withdrawal, defined as either none or only one ePRO filled during follow-up. Binomial logistic regression analysis was used to assess the relationship of demographical and clinical variables to the risk of no effective participation. Variable selection for the multivariable logistic regression model was guided by a directed acyclic graph (DAG) approach to identify potential confounding pathways (**Figure S1**, developed with DAGitty[13]). The DAG identified age as a key confounder for associations between working status, disease duration, previous surgery, and early withdrawal. Disease duration was identified as a confounder for the previous surgery–early withdrawal relationship. Based on this causal framework, the multivariable model included age, sex, working status, IBD disease duration, previous surgery, and personal history of cancer to estimate direct effects while controlling for confounding. All variables were entered simultaneously without stepwise selection.

Investigators

We first assessed the number of patients included per investigator and per country. The investigator profile based on the number of patients included was defined by quartiles on the basis of the overall cohort of investigators.

Among patients with at least two ePROs collected, the proportion of data corrected by the investigator in the annual summary was assessed. The number of patients for whom the investigators: (i) added an event or biologic exposure that was not reported by the patient; (ii) removed an event or biologic exposure that was reported by the patient (i.e., an event or biologic exposure that was incorrectly reported); (iii) corrected information related to an event that was reported on the correct date, was assessed. Event considered were cancer or dysplasia, serious infections defined as infections requiring hospitalization, and gastrointestinal endoscopy or imaging. Considering separately each ePRO, we finally calculated the sensitivity and specificity of patient-reported data for cancer or dysplasia, and serious infections. Calculations were separately performed for each outcome.

Event coding process

Among patients with at least two ePROs collected, the number of events reported by patients was assessed, including the number of reports retrieved and missing. The mean number of hospital discharge diagnosis codes was assessed.

Analyses were performed using SAS (V.9.4) statistical software (SAS Institute).

RESULTS

Patients

A total of 13,262 patients gave their written consent, among whom 139 subsequently refused the use of their data, 265 were duplicates of patients already included, and 12 had insufficient data collected at baseline, leaving 12,846 patients included in the analysis (**Figure 1**). More than 75% ($n = 10{,}192$, 79.3%) of patients filled more than one ePRO, while 425 (3.3%) and 2229 (17.4%) filled only one or no ePRO, respectively. The majority of patients were included in France ($n = 3697$, 28.8%) and the United Kingdom ($n = 2841$, 22.1%).

Baseline characteristics according to each group are provided in **Table 1**. After pooling patients with only one and no ePRO, male sex, younger age, and being unemployed or student were associated with early withdrawal in univariate analysis. (**Table 2**). IBD disease duration and previous history of abdominal surgery was

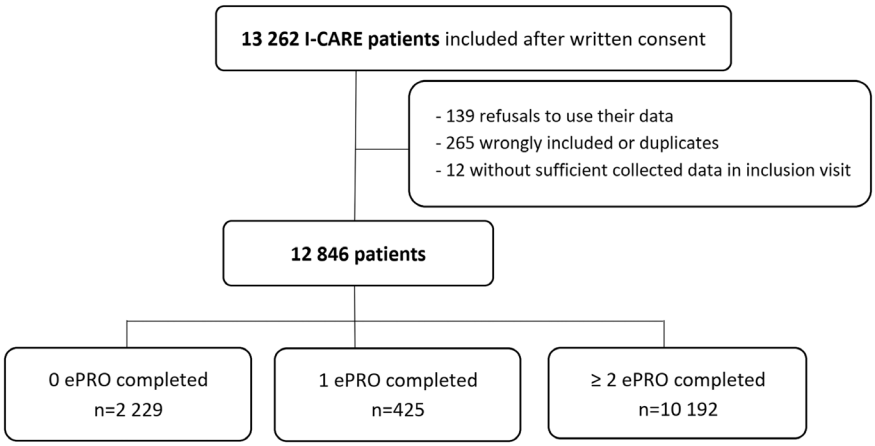

Figure 1 Flowchart.

associated with a lower risk of early withdrawal. In multivariable analysis, male sex, younger age, working status, and no previous history of abdominal surgery were associated with early withdrawal. (**Table 2**).

Among patients who filled more than one ePRO, 7877 patients (77.3%) filled an ePRO in the last 6 months before the end of the study (either 3 years or up to 6 years of follow-up according to the end of study date). Follow-up extension to 6 years was proposed to 3340 patients (32.8% of patients with assessable data), among whom 85.8% (2867) agreed to extend their follow-up and 83.4% (2787) filled at least one ePRO during the follow-up extension. (**Figure S2**) The cumulative risk of study withdrawal from cohort entry to year 3 is provided in **Figure 2** and from year 3 to year 6 in patients with follow-up extension in **Figure S3**. Withdrawal rate per month was stable over time during the first year (**Figure S4**). **Figure 3** provides the proportion of collected ePROs over time. During the first year, 72.8% of patients filled all ePROs, while 56.4% filled all ePROs during year 3.

Investigators

A total of 502 investigators participated in the study. The median number of investigators in each country was 24 (1st interquartile–3rd interquartile, 11–41), with the highest number of investigators in France ($n = 115$, 22.9%), followed by the United Kingdom ($n = 86$, 17.1%).

The median number of patients included per investigator was 22 (IQR, 14–31). There were no statistically significant differences regarding the number of patients included per investigator according to each country. The distribution of inclusion per investigator according to each country is provided in **Figure S5**.

The investigators validated 29,676 (90.3%) of 32,863 annual summaries from 374,549 ePROs of 10,192 patients. Overall, 13.9% of patients ($n = 1,422$) reported a cancer or dysplasia (4.8%, $n = 486$), or serious infections (10.0%, $n = 1,015$) during follow-up. Of the 486 patients who reported having cancer or dysplasia, 52.5% ($n = 252$) had their data modified by the investigators. The investigators made 284 modifications across these 252 patients. Of these modifications, 23.6% involved adding a cancer that had not been reported by the patients, 74.3% involved removing a cancer that had been erroneously reported by the patients, and 2.1% involved modifying information related to a cancer that had been reported on the correct date. Overall, 10% of patients ($n = 1,015$) reported serious infections; of these patients, 8.3% had their data modified by the investigators. Data modification accounted for 1.6% of cases where a serious infection was added that had not been reported by the patients, 6.6% of cases where a serious infection was removed that had been falsely reported by patients, and 0.6% of cases where the seriousness of infections reported at the correct date was modified. Gastrointestinal endoscopy or imaging was reported in 8021 patients (78.7%), of whom 33.4% ($n = 2,675$) had data modified. Data modification accounted for 9.4% of cases where a procedure was added that had not been reported by the patients, 20.4% of cases where a procedure was removed that had been falsely reported by patients, and 7% of cases where the procedure type was modified and reported at the correct date. Among the 6,708 patients treated with biologics, the investigators added or removed a treatment sequence declared by the patients in 19.6% of cases ($n = 1,316$). The modification was the addition or removal of a treatment sequence in 91.7% and 6.4% of patients, respectively. The drug type was modified in 1.4% of patients. Even after removing duplicate cancer reports, 40.9% of cancers reported by patients were still removed by the investigators. For cancers or dysplasia, the sensitivity was 87.1% and specificity was 99.8% before accounting for duplicate reports. After removing duplicates, which primarily affected specificity by reducing false positives from repeated entries, the revised sensitivity remained 87.1%, while specificity improved to 99.9%. For serious infections, the sensitivity and specificity were 98.9% and 99.9%, respectively.

Among 22,473 procedures (13,054 GI endoscopies and 9419 GI imaging) notified in 7864 patients, disease activity was filled in for 74.5% of procedures (78.5% for GI endoscopies and 69.4% for GI imaging).

Event coding process

Among the 1,152 cancers or dysplasia reported by patients or investigators, 385 (33.4%) were considered valid and coded.

Table 1 Baseline characteristics according to the number of ePROs completed

	Total		0 ePRO completed		1 ePRO completed		>1 ePRO completed	
			n = 2229		*n* = 425		*n* = 10,192	
	N = 12,846		(17.4%)		(3.3%)		(79.3%)	
	N/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD
Sex								
Female	6626	51.6	1049	47.1	200	47.1	5377	52.8
Male	6220	48.4	1180	52.9	225	52.9	4815	47.2
Age	39.3	13.2	37.9	13.7	35.1	12.5	39.8	13.1
Working status								
Having a job	9,515	74.1	1570	70.4	310	72.9	7635	74.9
Unemployed	1234	9.6	269	12.1	44	10.4	921	9
Retired	849	6.6	137	6.1	15	3.5	697	6.8
Student	1248	9.7	253	11.4	56	13.2	939	9.2
IBD sub type								
Crohn's Disease	7800	60.7	1360	61	279	65.6	6161	60.4
Indeterminate Colitis	231	1.8	41	1.8	7	1.6	183	1.8
Ulcerative Colitis	4815	37.5	828	37.1	139	32.7	3848	37.8
Disease duration (years)	10.4	8.8	9.8	8.7	9.1	7.5	10.6	8.9
Primary sclerosing cholangitis								
No	11,350	97.9	1728	96.9	369	95.6	9253	98.2
Yes	247	2.1	56	3.1	17	4.4	174	1.8
Missing	1,249		445		39		765	
Personal history of cancer								
No	12,490	97.2	2182	97.9	414	97.4	9894	97.1
Yes	356	2.8	47	2.1	11	2.6	298	2.9
Crohn's disease characteristics*								
Location								
L1	2662	37.6	396	36	97	38.6	2169	37.9
L2	1354	19.1	202	18.4	52	20.7	1100	19.2
L3	3065	43.3	502	45.6	102	40.6	2461	42.9
Missing	719	9.2	260	19.1	28	10	431	7.0
Isolated upper disease (L4)								
No	6467	91.3	992	90.8	224	89.6	5251	91.5
Yes	613	8.7	100	9.2	26	10.4	487	8.5
Missing	720	9.2	268	19.7	29	10.4	423	6.9
Behavior								
B1	3,663	51.9	561	51.5	129	51.6	2973	52
B2	1983	28.1	307	28.2	65	26	1611	28.2
B3	1412	20	222	20.4	56	22.4	1134	19.8
Missing	742	9.5	270	19.9	29	10.4	443	7.2
Any perianal disease								
No	5116	72.1	785	71.6	189	75.3	4142	72.1
Yes	1977	27.9	311	28.4	62	24.7	1604	27.9
Missing	707	9.1	264	19.4	28	10.0	415	6.7
Ulcerative colitis characteristics**								

(Continued)

Table 1 (Continued)

	Total		0 ePRO completed		1 ePRO completed		>1 ePRO completed	
			n = 2229		*n* = 425		*n* = 10,192	
	N = 12,846		(17.4%)		(3.3%)		(79.3%)	
	N/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD
UC involvement (E1, E2, E3)								
(E1) Proctitis	693	15	116	16.3	28	20.4	549	14.6
(E2) Left sided colitis	1815	39.4	260	36.5	55	40.1	1,500	39.9
(E3) Extensive UC	2102	45.6	336	47.2	54	39.4	1712	45.5
Missing	436	8.6	157	18.1	9	6.2	270	6.7
Previous surgery								
No	9511	74	1722	77.3	327	76.9	7462	73.2
Yes	3335	26	507	22.7	98	23.1	2730	26.8
Previous treatment								
No immunosuppressors	6210	53.4	930	52.5	192	49.7	5088	53.7
Conventional immunosuppressors with no biologic	2361	20.3	372	21	85	22	1904	20.1
Biologics	3055	26.3	470	26.5	109	28.2	2476	26.2
Missing	1220		457		39		724	
Treatment at cohort entry								
No immunosuppressors	2804	21.8	478	21.4	87	20.5	2239	22
Conventional immunosuppressors with no biologic	2365	18.4	389	17.5	70	16.5	1906	18.7
Biologics	7677	59.8	1,362	61.1	268	63.1	6047	59.3

*Among the 7800 CD patients included. **Among the 5046 UC or IC included.

Overall, 527 (46%) reports were duplicates. Among the 625 unique dysplasia or cancer reports, 256 (40.9%) were removed by the investigators in the annual summary. The reasons for removal were notably lesions that were considered benign based on pathology reports, as well as non-dysplastic polyps, such as post-inflammatory polyps.

Overall, 8141 hospitalizations have been reported by patients, among whom 85.3% ($n = 6{,}944$) have been linked to an uploaded hospitalization report. Among the 6944 uploaded hospitalization reports, the study coordinators or the members of the scientific committee excluded 3634 events, including outpatient visits without hospitalization or day hospitalization for outpatient perfusion. The median number of diagnoses coded was 1 (IQR, 1–2). Less than 2% of reports (1.9%) had more than three diagnoses coded, and 38.8% of hospitalizations reported had a CD or UC related diagnosis code as primary discharge diagnosis, defined as the condition leading to the hospitalization.

DISCUSSION

Among 12,846 patients included by 502 investigators, more than 75% of patients filled more than one ePRO including at least one within 6 months before study end date. While 72.8% of patients filled all ePROs during the first year, this rate dropped to 56.4% during year 3. Sex, age, working status, and no previous history of abdominal surgery were associated with early withdrawal. Investigators had to correct more than half of patients' reports on cancer or dysplasia, but only 10% on serious infections. These findings support the utility of the use of "patient as researcher" concept in large prospective observational studies.

Male gender, younger age, and being unemployed or a student were associated with a higher risk of early withdrawal due to patient non-engagement. These characteristics have been previously associated with loss to follow-up in RCTs investigating other diseases.[14] Previous history of abdominal surgery was the only IBD characteristic associated with a lower risk of early withdrawal, but its effect magnitude seems not clinically meaningful (21.5% of patients dropped out early in the no previous surgery group vs. 18.1% in the previous surgery group). It suggests that no characteristic specifically related to IBD may have significantly influenced early withdrawal. This finding is important for the assessment of the validity of prospective findings from the I-CARE cohort.

Many studies have assessed the characteristics of patients lost to follow-up in clinical practice,[15] but very few were performed in the setting of observational research. More than 75% of patients completed information within 6 months before the study end date. Simulation studies have assessed in cohort studies the bias associated with varying degrees of loss to follow-up across different missing data mechanisms.[16] Assuming that loss to follow-up is driven in the I-CARE cohort by either missing completely at random (MCAR) or missing at random (MAR) mechanisms, minimal bias related to loss of follow-up, if any, is expected in future studies

Table 2 Predictors of early withdrawal due to patient non-engagement

			Univariate analysis				Multivariate analysis			
	*N**	%	Odds ratio	IC* 95%		*p*	Odds ratio	IC* 95%		*p*
Age						<0.0001				<0.0001
<20	126	33.4	Ref				Ref			
[20–25]	424	28.1	0.78	0.61	0.99		0.72	0.56	0.92	
[25–30]	424	22.5	0.58	0.46	0.74		0.49	0.37	0.64	
[30–35]	368	20.5	0.51	0.4	0.66		0.43	0.32	0.56	
[35–40]	316	19.1	0.47	0.37	0.6		0.39	0.29	0.52	
[40–45]	255	17.1	0.41	0.32	0.53		0.34	0.26	0.46	
[45–50]	238	18.2	0.44	0.34	0.57		0.38	0.28	0.51	
[50–55]	180	16.7	0.4	0.31	0.52		0.34	0.25	0.47	
[55–60]	130	18.2	0.44	0.33	0.59		0.37	0.27	0.51	
≥60	193	18.7	0.46	0.35	0.6		0.4	0.28	0.56	
Sex						<0.0001				<0.0001
Male	1,405	22.6	Ref				Ref			
Female	1,249	18.8	0.8	0.73	0.87		0.79	0.72	0.86	
Working status						<0.0001				<0.0001
Having a job	1880	19.8	Ref				Ref			
Unemployed	313	25.4	1.38	1.2	1.58		1.42	1.24	1.64	
Retired	152	17.9	0.89	0.74	1.06		0.99	0.77	1.27	
Student	309	24.8	1.34	1.16	1.53		0.76	0.64	0.91	
Disease duration (years)						<0.0001				0.0845
<5	952	22.2	Ref				Ref			
[5–10]	732	22.6	1.02	0.92	1.14		1.1	0.99	1.23	
[10–15]	407	18.9	0.82	0.72	0.93		0.97	0.85	1.11	
[15–20]	272	19.8	0.86	0.74	1		1.09	0.93	1.28	
≥20	291	16.2	0.68	0.59	0.79		0.91	0.78	1.08	
Previous surgery						<0.0001				0.0011
No	2049	21.5	Ref				Ref			
Yes	605	18.1	0.81	0.73	0.89		0.84	0.76	0.93	
Personal history of cancer						0.0398				0.3727
No	2,596	20.8	Ref				Ref			
Yes	58	16.3	0.74	0.56	0.99		0.88	0.65	1.17	

*Combined number of patients with 0 (N=2,229) or 1 (N=425) ePRO completed, N=2,654.

based on the I-CARE cohort. In I-CARE, we found no association between surrogate markers of IBD severity and early withdrawal. Since IBD severity is a primary driver of the outcomes of interest in I-CARE, this suggests that loss to follow-up is unlikely to depend on the measured outcome. Instead, the observed predictors (e.g., age, gender, employment status) align with MCAR or MAR mechanisms, where missingness depends on measured variables. However, unobserved confounders cannot be entirely ruled out.

The high proportion of patients with follow-up data until the end of the I-CARE study supports the feasibility of observational studies involving patients in data collection. The I-CARE study design emphasizes the strong interaction between patients and investigators, a complementary interaction regarding the burden of data collection and quality but also original through the annual summary to validate ePROs.

The proportion of ePROs completed each year expectedly decreased over time, as 72.8% of patients completed all ePROs during year 1 vs. 56.4% during year 3. However, early withdrawal due to patient non-engagement was stable over time during the first year and throughout follow-up. The very high engagement rates in the first year support the concept of "patients as a researcher." It also highlights the challenges to involve patients at the inclusion of the study compared to a long-term basis. In the I-CARE cohort, patients were asked to fill an ePRO on a monthly basis, we can assume that the frequence of ePRO completion impacts the proportion of non-responders. Additionally, the software used may impact patients'

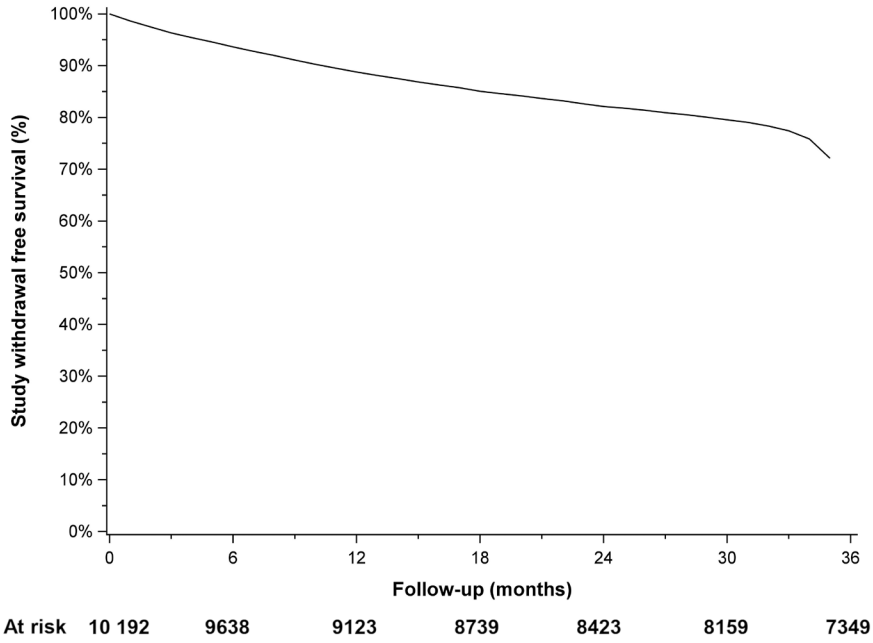

Figure 2 Cumulative risk of study withdrawal from cohort entry to year 3 among patients who filled more than one electronic patient-reported outcome (ePRO).

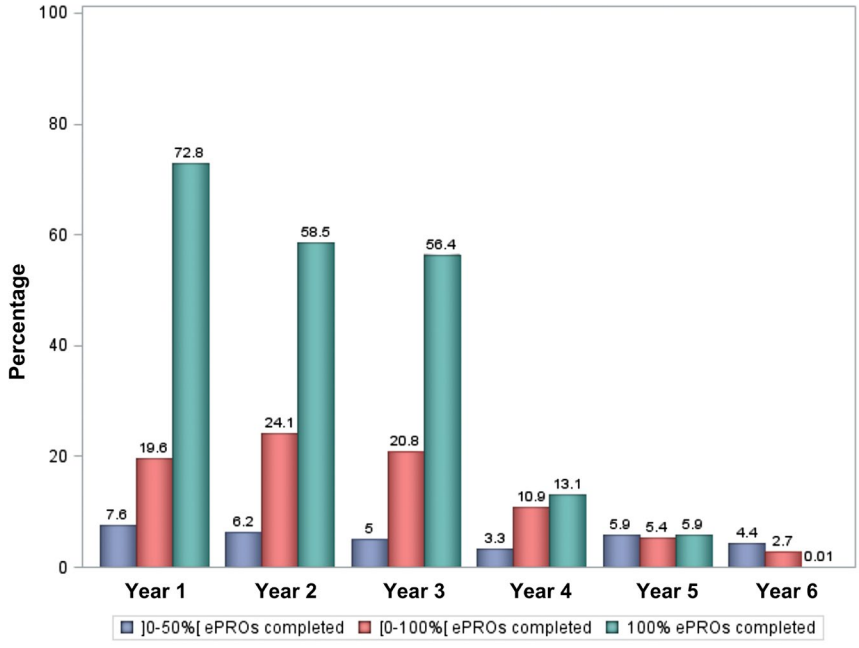

Figure 3 Proportion of electronic patient-reported outcomes (ePROs) completed each year among the 12,846 patients included.

data collection. This may limit the generalization of our findings to future studies using other software. Investigators corrected 52.5% of cancer or dysplasia reported by patients compared to 10% of serious infections and 19.6% of treatment sequences. This discrepancy may be partially explained by the way the question regarding cancer diagnosis was asked to the patient. The question "Have you been diagnosed with cancer?" was asked monthly in the ePRO, without excluding patients who had previously reported a cancer diagnosis. This led to duplicate reports when patients consistently answered "yes" in subsequent forms, as the system did not automatically account for prior responses. This design choice ensured comprehensive capture of new or evolving diagnoses but required investigator validation to remove redundancies. However, even after removing duplicate cancer reports, 40.9% of cancers reported by patients were still removed by the investigators. Our study highlights the novelty of assessing the accuracy of patient-reported data

in IBD, an area that has been poorly explored in other conditions so far.[17] While specificity was very high for both outcomes, sensitivity was numerically lower for cancers or dysplasia than for serious infections. Although these findings support the accuracy of patient-reported information in research studies, they also highlight that accuracy varies according to the outcome type, and that investigator validation remains essential.

More than 32,000 annual summaries were validated by 502 investigators. It underscores the burden for investigators related to primary research data collection, although this burden was reduced by patients directly reporting a substantial amount of information. Primary research data may be mandatory to accurately collect not only outcomes but also key parameters on disease characteristics that are known confounding factors or key subgroup parameters. Secondary research data such as electronic health records (EHRs) or administrative health databases are available for research at a far lower cost, but major variables may not be collected. Linking these data sources with primary research data would reduce redundancy by automating the collection of clinical events already documented in secondary research data, improve accuracy by cross-referencing patient-reported treatments with prescription records to clarify discrepancies between prescription and actual adherence, and enrich insights by combining clinical information with patient experiences. This approach is already being implemented in the ongoing I-CARE 2 study launched in 2024, which includes direct linkage with the French administrative healthcare database.[18]

Some limitations need to be discussed. The I-CARE cohort mainly included patients from France and United Kingdom and our findings may mainly reflect patient and investigator experience in these countries. Still, the I-CARE study included patients from 15 countries in Europe, and no statistically significant differences were observed in the mean number of patients included per investigator according to each country. Additionally, the findings reported may only be generalizable to the setting of IBD. However, IBD shares similar features with various immune mediated inflammatory diseases,[19] notably inflammatory rheumatic disorders, and our findings can inform future studies including patients with other IMIDs.

CONCLUSION

This study provides extensive data that highlights the value of using the "patient as a researcher" model including good engagement rates and details factors associated with loss of follow-up and validation of patient-reported outcomes by investigators. Overall, it supports the absence of differential follow-up according to IBD disease characteristics in the I-CARE cohort. This data increases the transparency and will facilitate the interpretation of future findings based on the I-CARE cohort. Such initiatives should be encouraged for future investigator-initiated studies including data directly collected by patients. IBD patients can be "their own researchers."

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENT

Open access publication funding provided by COUPERIN CY26.

FUNDING

The I-CARE cohort was supported by the following partners: Abbvie; Amgen; Ferring; Pfizer; Janssen; MSD; Takeda.

CONFLICT OF INTEREST

GETAID reports financial support was provided by Abbvie, Amgen, Ferring, Pfizer, Janssen, MSD, Takeda. Julien Kirchgesner reports a relationship with Abbvie, Alfasigma, Amgen, Celltrion, Lilly, Janssen, Takeda, Tillots, and Pfizer that includes: consulting or advisory and speaking and lecture fees. Laurent Beaugerie reports a relationship with Takeda, Janssen, Nordic Pharma, and Viatris that includes: consulting or advisory and speaking and lecture fees. Filip Baert reports a relationship with AbbVie, Amgen, Arena Pharmaceuticals, Celgene, Ferring, Fresenius Kabi, Galapagos, Janssen, Merck, Pfizer, and Takeda that includes: consulting or advisory, funding grants, and speaking and lecture fees. Jean-Francois Rahier reports a relationship with Abbvie, Takeda, Hospira, Mundipharma, MSD, Pfizer, GlaxoSK, Janssen, Takeda, Galapagos, Ferring, Falk, Biogen, Amgen, Celltrion, and Lilly that includes: consulting or advisory, funding grants, and speaking and lecture fees. Daniel Bergemalm reports a relationship with Abbvie, BMS, Ferring, Janssen, Pharmacosmos, Pfizer, Takeda and Sandoz that includes: consulting or advisory and speaking and lecture fees. Benedicte Caron reports a relationship with Abbvie, Amgen, Celltrion, Ferring, Galapagos, Janssen, Lilly, Nordic Pharma, Pfizer, and Takeda that includes: consulting or advisory and speaking and lecture fees. Idan Goren reports a relationship with Gilead sciences and Boehringer Ingelheim that includes: funding grants. Tom Holvoet reports a relationship with Galpagos, Abbvie, Takeda, Ferring, and Dr Falk that includes: consulting or advisory, funding grants, and speaking and lecture fees. Aditi Kumar reports a relationship with Abbvie Dr Falk, and Ferring that includes: Peter Rimmer reports a relationship with Abbvie, Ferring, Janssen, Bristol Myers Squibb, and F. Hoffman La Roche that includes: funding grants and speaking and lecture fees. Mathieu Uzzan reports a relationship with Abbvie, Janssen, Pfizer, Takeda, Celltrion, Amgen, Ferring that includes: consulting or advisory and speaking and lecture fees. Nikos Viazis reports a relationship with Abbvie, Ferring, Janssen, MSD, Pfizer, Sandoz, Takeda, Biocon, Faran, Vianex that includes: consulting or advisory and speaking and lecture fees. Petra Weimers reports a relationship with Novo Nordisk that includes: employment and equity or stocks. Edyta Zagorowicz reports a relationship with Abbvie, Lilly, Janssen-Cilag, AbbVie, Takeda Pharma, BMS that includes: consulting or advisory and speaking and lecture fees. Laurent Peyrin-Biroulet reports a relationship with Abbvie, Abivax, Adacyte, Alimentiv, Amgen, Applied Molecular Transport, Arena, Banook, Biogen, BMS, Celltrion, Connect Biopharm, Cytoki Pharma, Enthera, Ferring, Fresenius Kabi, Galapagos, Genentech, Gilead, Gossamer Bio, GSK that includes: consulting or advisory and speaking and lecture fees. Laurent Peyrin-Biroulet reports a relationship with IAC Image Analysis, Index Pharmaceuticals, Inotrem, Janssen, Lilly, Medac, Mopac, Morphic, MSD, Nordic Pharma, Novartis, Oncodesign Precision Medicine, ONO Pharma, OSE Immunotherapeuthics, Pandion Therapeuthics, Par Immune, Pfizer, Prometheus that includes: consulting or advisory and speaking and lecture fees. Laurent Peyrin-Biroulet reports a relationship with Protagonist, Roche, Samsung, Sandoz, Sanofi, Satisfay, Takeda, Telavant, Theravance, Thermo Fischer, Tigenix, Tillots, Viatris, Vectivbio, Ventyx, Ysopia that includes: consulting or advisory and speaking and lecture fees. Shaji Sebastian reports a relationship with Celltrion, Lilly, Takeda, AbbVie, Merck, Ferring, Pharmacocosmos, Warner Chilcott, Janssen, Falk Pharma, Biohit, TriGenix, Celgene, Olympus, and Tillots Pharma that includes: consulting or advisory, funding grants, and speaking and lecture fees. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AUTHOR CONTRIBUTIONS

J.K. wrote the article. J.K., H.R., and S.S. designed and performed the research. L.B., C.B., F.B., J.F.R., D.B., F.C., B.C., M.C., S.D., K.F., I.G., T.H., S.K., A.K., P.G., P.R., J.S., J.T., M.U., N.V., P.W., E.Z., A.B.,

L.A., H.R., and L.P.B. analyzed the data. All authors approved the final manuscript.

ETHICS STATEMENT

Master ICARE Protocol received an approval from Saint Louis Ethic Committee on 15/05/2015 and CNIL approval on 27/10/2015. Each country participating in I-CARE had the responsibility to submit the protocol and the amendments in agreement with their own local regulation. I-CARE Eudract Number: 2014-004728-23. N° INDS: TP305083. CEREES approval on 20 Feb 2019.

1. Cosnes, J., Gower-Rousseau, C., Seksik, P. & Cortot, A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* **140**, 1785–1794 (2011).
2. Kaplan, G.G. The global burden of IBD: from 2015 to 2025. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 720–727 (2015).
3. Kirchgesner, J. *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.* **45**, 37–49 (2017).
4. Beaugerie, L. & Kirchgesner, J. Balancing benefit vs risk of immunosuppressive therapy for individual patients with inflammatory bowel diseases. *Clin. Gastroenterol. Hepatol.* **17**, 370–379 (2018).
5. FDA Real-World Data. Assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products - guidance for industry. *FDA* <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory> (2021). Accessed 1 January 2026.
6. Franklin, J.M. & Schneeweiss, S. When and how can real world data analyses substitute for randomized controlled trials? *Clin. Pharmacol. Ther.* **102**, 924–933 (2017).
7. Kirchgesner, J., Desai, R.J., Schneeweiss, M.C., Beaugerie, L., Schneeweiss, S. & Kim, S.C. Decreased risk of treatment failure with vedolizumab and thiopurines combined compared with vedolizumab monotherapy in Crohn's disease. *Gut* **71**, 1781–1789 (2022).
8. Kirchgesner, J., Desai, R.J., Beaugerie, L., Kim, S.C. & Schneeweiss, S. Calibrating real-world evidence studies against randomized trials: treatment effectiveness of infliximab in Crohn's disease. *Clin. Pharmacol. Ther.* **111**, 179–186 (2022).
9. Kirchgesner, J., Desai, R.J., Schneeweiss, M.C., Beaugerie, L., Kim, S.C. & Schneeweiss, S. Emulation of a randomized controlled trial in ulcerative colitis with US and French claims data: infliximab with thiopurines compared to infliximab monotherapy. *Pharmacoepidemiol. Drug Saf.* **31**, 167–175 (2022).
10. Vergnas, O.L. Patients' participation in Health Research: a classification of cooperation schemes. *Journal of Participatory Medicine* **9**, e8933 (2017).
11. Honap, S., Buisson, A., Danese, S., Beaugerie, L. & Peyrin-Biroulet, L. Patient and public involvement in research: lessons for inflammatory bowel disease. *Journal of Crohn's and Colitis* **17**, 1882–1891 (2023).
12. Peyrin-Biroulet, L. *et al*. I-CARE, a European prospective cohort study assessing safety and effectiveness of biologics in inflammatory bowel disease. *Clin. Gastroenterol. Hepatol.* **21**, 771–788.e10 (2023).
13. Textor, J., van der, Z.B., Gilthorpe, M.S., Liskiewicz, M. & Ellison, G.T. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int. J. Epidemiol.* **45**, 1887–1894 (2016).
14. Madden, K. *et al*. Predicting and preventing loss to follow-up of adult trauma patients in randomized controlled trials. *J. Bone Joint Surg. Am.* **99**, 1086–1092 (2017).
15. Tong, C.Y.M., Koh, R.Y.V. & Lee, E.S. A scoping review on the factors associated with the lost to follow-up (LTFU) amongst patients with chronic disease in ambulatory care of high-income countries (HIC). *BMC Health Serv. Res.* **23**, 1–21 (2023).
16. Kristman, V., Manno, M. & Côté, P. Loss to follow-up in cohort studies: how much is too much? *Eur. J. Epidemiol.* **19**, 751–760 (2004).
17. Watanabe, T., Ichinose, Y., Toida, T. & Higashi, T. Validity of patient-reported information: agreement rate between patient reports and registry data. *BMC Health Serv. Res.* **25**, 182 (2025).
18. *Groupe d'Etude Therapeutique des Affections Inflammatoires Digestives Ibd CAncer and seRious Infections in France (I-CARE 2)* <https://clinicaltrials.gov/study/NCT06089590> (2023). Accessed 1 January 2024.
19. Schett, G., McInnes, I.B. & Neurath, M.F. Reframing immune-mediated inflammatory diseases through signature cytokine hubs. *N. Engl. J. Med.* **385**, 628–639 (2021).