United European Gastroenterology Journal

WILEY

UNITED EUROPEAN GASTROENTEROLOGY ueg journal

ORIGINAL ARTICLE OPEN ACCESS

Assessment of PredictSURE IBD Assay in a Multinational Cohort of Patients With Inflammatory Bowel Disease

Dahham Alsoud[1] | Nurulamin M. Noor[2,3] | Lea Ann Chen[4] | Vivian Abadom[5] | Simon H. C. Anderson[6] | Lediona Ardolli[7] | Jordan Axelrad[8] | Peter Bossuyt[9] | Kenneth Croitoru[10] | Oriana M. Damas[11] | Lily Deng[12] | Parakkal Deepak[13] | Juan De La Revilla Negro[3] | Shanika de Silva[14] | Marc Ferrante[1,15] | Karen Hills[16] | Peter M. Irving[6] | James O. Lindsay[17] | Dana J. Lukin[18] | Paul A. Lyons[2,19] | Eoin F. McKinney[2,19] | Maria Oliva-Hemker[5] | Caterina Oneto[20] | Roohi Patel[21] | Miles Parkes[2,3] | Lieven Pouillon[9] | João Sabino[1,15] | Lawrence J. Saubermann[22] | Jenny S. Sauk[23] | Sarah Sheibani[24] | Kenneth G. C. Smith[25,26] | Keith S. Sultan[27] | Tony C. Tham[28] | Sare Verstockt[1] | Raluca Vrabie[29] | Melissa Weidner[30] | Huimin Yu[31] | Bram Verstockt[1,15] | James C. Lee[32,33,34] | Séverine Vermeire[1,15]

[1]Department of Chronic Diseases and Metabolism, Translational Research in Gastrointestinal Disorders, KU Leuven, Leuven, Belgium | [2]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK | [3]Department of Gastroenterology, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, UK | [4]Division of Gastroenterology, Department of Medicine, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA | [5]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA | [6]Department of Gastroenterology, Guy's & St Thomas' Hospital, London, UK | [7]PredictImmune Ltd, Cambridge, UK | [8]Division of Gastroenterology, Department of Medicine, New York University Grossman School of Medicine, New York, New York, USA | [9]Department of Gastroenterology, Imelda GI Clinical Research Centre, Imelda General Hospital, Bonheiden, Belgium | [10]Division of Gastroenterology, University of Toronto, Mount Sinai Hospital, Toronto, Canada | [11]Division of Digestive Health and Liver Diseases, University of Miami Miller School of Medicine, Miami, Florida, USA | [12]Department of Medicine, Temple University Hospital, Philadelphia, Pennsylvania, USA | [13]Division of Gastroenterology and Inflammatory Bowel Diseases Center, Washington University in Saint Louis School of Medicine, St. Louis, Missouri, USA | [14]Department of Gastroenterology, Dudley Group NHS Foundation Trust, Dudley, UK | [15]Department of Gastroenterology and Hepatology, University Hospitals Leuven, Leuven, Belgium | [16]Juneberry Ltd, Cambridge, UK | [17]Department of Gastroenterology, Barts Health NHS Trust, London, UK | [18]Department of Gastroenterology and Hepatology, New York Presbyterian Hospital-Weill Cornell Medicine, New York, New York, USA | [19]Cambridge Institute for Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, Cambridge, UK | [20]Vanguard Gastroenterology Clinic, New York, New York, USA | [21]Division of Gastroenterology and Hepatology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA | [22]Division of Gastroenterology and Hepatology, University of Rochester Medical Center, Rochester, New York, USA | [23]Vatche and Tamar Manoukian Division of Digestive Diseases, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California, USA | [24]Division of Gastrointestinal and Liver Diseases, University of Southern California Keck School of Medicine, Los Angeles, California, USA | [25]Walter and Eliza Hall Institute of Medical Research, Parkville, Australia | [26]Department of Medical Biology, The University of Melbourne, Parkville, Australia | [27]Department of Gastroenterology, North Shore University Hospital-Long Island Jewish Medical Center, Northwell Health, Manhasset, New York, USA | [28]Division of Gastroenterology, Ulster Hospital, Dundonald, UK | [29]Division of Gastroenterology, Northwell Long Island Jewish-Forest Hills Hospital, Queens, New York, USA | [30]Division of Pediatric Gastroenterology, Department of Pediatrics, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA | [31]Division of Gastroenterology and Liver Diseases, Department of Medicine, The George Washington University School of Medicine and Health Sciences, Washington, DC, USA | [32]Genetic Mechanisms of Disease Laboratory, The Francis Crick Institute, London, UK | [33]Department of Gastroenterology, Royal Free Hospital, London, UK | [34]Division of Medicine, Institute for Liver and Digestive Health, University College London, London, UK

Correspondence: Séverine Vermeire (severine.vermeire@uzleuven.be)

Received: 2 February 2025 | **Revised:** 15 March 2025 | **Accepted:** 31 March 2025

Funding: The PRECIOUS study was funded by the Crohn's and Colitis Foundation.

Abbreviations: CD, Crohn's disease; CRP, C-reactive protein; FC, fecal calprotectin; IBD, inflammatory bowel disease; IQR, interquartile range; PWP, Prentice, Williams, and Peterson; UC, ulcerative colitis.

Dahham Alsoud and Nurulamin M. Noor share joint first authorship.

Bram Verstockt, James C. Lee, and Séverine Vermeire share joint last authorship.

Keywords: biomarker | blood tests | corticosteroids | crohn's disease | inflammatory bowel diseases | personalised management | prognosis | treatment | ulcerative colitis

ABSTRACT

Background and Aims: PredictSURE IBD is a prognostic blood test that classifies newly diagnosed, treatment-naïve Inflammatory Bowel Disease (IBD) patients into 'IBDhi' (high-risk) or 'IBDlo' (low-risk) groups (risk of future aggressive disease). We evaluated this assay in a multinational cohort and explored the effect of concomitant corticosteroids on its discrimination.

Methods: One hundred thirty-six (71 Ulcerative colitis [UC], 65 Crohn's Disease [CD]) and 41 (15 UC, 26 CD) patients with active IBD were 'unexposed' and 'exposed', respectively, to corticosteroids at baseline blood sampling. The number of treatment escalations, time to first escalation, and need for repeated escalations were compared between the biomarker subgroups. Another 20 patients (13 UC, 7 CD) were longitudinally sampled over 6 weeks after commencing corticosteroids.

Results: In corticosteroids-naïve UC and CD patients, all bowel surgeries ($n = 6$) and multiple therapy escalations ($n = 10$) occurred in IBDhi patients. IBDhi UC patients required significantly more treatment escalations, had a shorter time to first escalation, and a greater need for multiple escalations than IBDlo patients. No statistically significant differences were observed among CD patients. In corticosteroid-exposed patients, 66.6% of 'misclassifications' were IBDlo patients who required escalations. Among corticosteroid-treated patients with longitudinal sampling, 81.3% of those classified as IBDhi before steroids switched to IBDlo during therapy.

Conclusions: No significant differences in treatment escalations were observed between biomarker-defined subgroups in CD. However, IBDhi UC patients required significantly earlier and more frequent therapy escalations, highlighting the need to further investigate PredictSURE IBD in UC. Notably, the discrimination ability of the biomarker was unreliable in patients receiving corticosteroid therapy.

1 | Introduction

The management of inflammatory bowel diseases (IBD), of which the two most common types are Crohn's disease (CD) and ulcerative colitis (UC), has rapidly advanced in the last 20 years, partly driven by the accompanying increase in therapeutic options [1, 2]. Nevertheless, the possibility of a personalised therapy plan for every patient has remained aspirational due to the unpredictable and variable nature of IBD [3, 4]. Currently, the choice of treatment strategy is mainly based on clinical features at diagnosis including: endoscopic severity, presence of perianal disease or other disease complications, inflammatory burden, co-morbidities and disease impact on quality of life [5]. Although widely used in clinical practice, there is limited evidence to support the utility of clinical features as predictors of disease prognosis. Empirical and/or reactive treatment escalations (in response to worsening disease) inevitably delay the timely introduction of effective therapies in patients destined for severe and refractory forms of IBD [6–9]. Conversely, milder forms of IBD, which can occur in 20%–30% of patients, may not require exposure to immunosuppressants, and their associated costs and side effects [10, 11].

To address this gap in IBD management, attempts have been made to identify molecular biomarkers that could predict disease course and facilitate risk stratification [12]. Following prior discovery in other immune-mediated diseases [13], a CD8+ T-cell gene expression signature was identified that was associated with subsequent disease course in IBD patients [14], with later mechanistic work revealing a likely role for T-cell exhaustion [15]. This CD8+ T-cell signature was subsequently translated into a 17-gene whole blood qPCR-based test (PredictSURE IBD, PredictImmune, United Kingdom) and validated in cohorts of predominantly newly diagnosed CD and UC patients, all of whom were not receiving concomitant corticosteroids, immunomodulators, or biologics at the time of testing [16]. This classifier stratifies patients into two distinct groups that correspond with their likelihood of experiencing an aggressive disease course: 'IBDhi' patients who are more likely to have a shorter time to first relapse and a higher number of treatment escalations over time, and 'IBDlo' patients who are more likely to have a longer time to relapse, fewer flares, and a lower need for treatment escalations.

Recently, the PROFILE (PRedicting Outcomes For Crohn's disease using a moLecular biomarker) trial investigated the clinical utility of PredictSURE IBD in adults with newly-diagnosed active CD [17]. The trial randomly assigned patients to either a 'top-down' or an 'accelerated step-up' treatment strategy with stratification of treatment based on biomarker subgroup (IBDhi or IBDlo). The 'top-down' approach of using infliximab combination therapy immediately resulted in significantly higher rates of remission, improved quality of life, and fewer flares requiring treatment escalation compared with the 'accelerated step-up' approach. However, no significant differences were observed between the biomarker subgroups, highlighting a lack of clinical utility for the biomarker in this group of patients. A 'top-down' approach was both more efficacious and safer for patients. The results from PROFILE have been adopted in international guidelines where a 'top-down' strategy of early effective therapy has been proposed as the standard of care for patients who would have met the inclusion criteria for the trial, that is, being symptomatic, with raised inflammatory markers and endoscopic evidence of disease activity [18].

It is important to note that patients with asymptomatic mild CD were not enrolled in PROFILE. Moreover, PROFILE only assessed PredictSURE IBD in CD and not in UC. Additionally,

Summary

- Summarise the established knowledge on this subject
  - A CD8+ T-cell gene expression signature has been identified as being associated with subsequent disease progression in IBD patients. This signature was translated into a 17-gene whole blood qPCR-based test (PredictSURE IBD; PredictImmune, United Kingdom).
  - The prognostic value of PredictSURE IBD has been validated in prospective UK cohorts of newly diagnosed CD and UC patients, none of whom were receiving corticosteroids, immunomodulators, or biologics at the time of testing.
  - The PROFILE trial did not demonstrate clinical utility for PredictSURE IBD in CD patients who met the trial's specific eligibility criteria and followed its predefined follow-up and treatment escalation protocol.
- What are the significant and/or new findings of this study?
  - This is the first multinational prospective study to evaluate a prognostic biomarker in IBD. PredictSURE IBD correctly identified all IBD patients who later required bowel surgery or multiple therapy escalations as high-risk at baseline.
  - PredictSURE IBD showed significant prognostic value in UC, where predicted high-risk UC patients had a shorter time to the first escalation and a greater need for multiple escalations compared with low-risk UC patients.
  - The discriminatory ability of PredictSURE IBD was unreliable in patients who were receiving corticosteroid therapy at the time of testing.

while the test was developed and evaluated in newly-recruited prospective cohorts, the predictive performance of PredictSURE IBD has not been tested outside of the United Kingdom. Given the growing understanding of biological diversity between populations [19, 20], and of the differences between clinical practice in distinct healthcare settings [21], an important unanswered question is whether the assay would be predictive in other populations.

Furthermore, the development and validation of this biomarker was performed in treatment-naïve cohorts, which was reflected in the manufacturer's instructions for PredictSURE IBD to avoid use in patients concurrently taking corticosteroids. However, in many countries, newly-diagnosed IBD patients have already commenced corticosteroids before being seen in specialist centres [9].

In this study, we hypothesised that PredictSURE IBD could identify IBD patients at a higher risk of therapy escalation. To test this hypothesis, we assessed the test's ability to classify newly diagnosed IBD patients into high- and low-risk prognostic groups by comparing predicted risk to the real-life need for therapy escalation. Additionally, we posited that performing a test based on gene expression in patients receiving corticosteroid therapy could affect its discriminative ability. Therefore, we aimed to investigate whether concomitant corticosteroid use would influence the test's ability to discriminate escalation risk or the consistency of the test results over time.

2 | Materials and Methods

2.1 | Study Design and Patient Selection

This was a multicentre prospective observational cohort study, in which four distinct cohorts were recruited. Cohorts 1, 2 and 3 consisted of patients with newly diagnosed CD or UC who were recruited largely from specialised outpatient IBD clinics:

Cohort 1 recruited adult patients (aged ≥ 18 years) from two sites in Belgium during routine clinical practice between December 2016 and July 2020. Follow-up in this cohort concluded in March 2022.

Cohort 2 recruited adult patients from five sites in the United Kingdom during routine clinical practice between May 2015 and February 2021. Follow-up in this cohort concluded in December 2021.

Cohort 3 recruited patients between 16 and 80 years of age from 12 sites in North America between March 2020 and July 2022 as part of a multicentre observational study (PRECIOUS, Predicting Crohn's & Colitis Outcomes in the United States). Follow-up in this cohort was pre-specified to be 12 months, and the study was terminated early by the sponsor (PredictImmune Ltd.) for commercial reasons.

In all three cohorts, patients were not taking any immunomodulators, biologics, or small molecules at inclusion, and needed to have active disease confirmed by one or more objective measures (presence of ulcers on endoscopy, C-reactive protein (CRP) ≥ 5 mg/L, or faecal calprotectin (FC) ≥ 250 μg/g). Cohort 1 included patients regardless of their exposure to corticosteroids at the time of testing, whereas cohorts 2 and 3 included only patients who were not receiving concomitant corticosteroids. Patients were prospectively followed up and therapy escalations were made at the discretion of the treating physicians. Escalations were defined as the initiation of immunomodulators (thiopurines or methotrexate), biologics or small molecules, or bowel surgery due to non-responsiveness to previous therapies.

In cohort 4, adult patients with active CD or UC who were scheduled to start oral or intravenous corticosteroids between March 2017 and February 2018 were recruited from outpatient clinics or gastroenterology specialist wards at Cambridge University Hospitals, Cambridge, United Kingdom, and followed for 6 weeks.

Inpatients were started on intravenous methylprednisolone at a dose of 40 mg to be taken twice daily. In all reported cases, patients continued intravenous methylprednisolone for at least 5 days. Decisions about escalation or de-escalation were typically made between days three and five. For patients who were then de-escalated following a good response to intravenous steroids, they were switched to oral prednisolone 40 mg daily,

and this was weaned down by 5 mg every week over the next 2 months in the outpatient/community setting. Outpatients were started on oral prednisolone at a dose of 40 mg daily, and this was weaned down by 5 mg every week over the next 2 months in the outpatient/community setting.

2.2 | Ethical Approval

For patients recruited from Cambridge University Hospitals, ethical approval was obtained from the Cambridgeshire Regional Ethics Committee (REC08/H0306/21), and written informed consent was obtained from each patient. At other centres in the United Kingdom, the study was classified as a service evaluation project, so formal IRB approval was not required. However, verbal or written informed consent was obtained from each patient in accordance with Good Clinical Practice guidelines. In Belgium, all included patients provided written consent to participate in IBD repositories approved by the IRBs of Imelda General Hospital and University Hospitals Leuven (B322201213950/S53684). This consent already covered the use of data for subsequent research, including the current study, and thus, no additional IRB approval was necessary. For sites participating in the PRECIOUS study, PredictImmune Ltd. coordinated a centralised IRB process through WCG IRB, and written informed consent was obtained from each patient.

2.3 | Samples and Biomarker Testing

PAXgene whole-blood RNA tubes were drawn at inclusion in cohorts 1, 2 and 3. In cohort 4, a PAXgene whole-blood RNA sample was taken prior to the initial dose of corticosteroids, and repeat samples were taken at week 1 and week 6 in outpatients, or on day 3, day 5, and at week 6 in inpatients. Samples were then sent for PredictSURE IBD assessment (PredictImmune, UK) to determine whether patients were IBDhi or IBDlo [16].

2.4 | Clinical Data Variables

Patient demographic and disease characteristics at the time of enrolment were collected and included age, haemoglobin, CRP, serum albumin, disease location and behaviour in CD and disease extent in UC following the Montreal classification.

2.5 | Endpoints

In this study, our objective was to assess the ability of PredictSURE IBD to distinguish between IBD prognosis subtypes based on the number of needed treatment escalations and the time to the first escalation in patients who were not exposed to corticosteroids at the time of testing in cohorts 1, 2 and 3. Additionally, we evaluated the effect of exposure to steroids on the accuracy of the test's classifications and their longitudinal stability in patients who were exposed to corticosteroids in cohorts 1 and 4, respectively.

2.6 | Statistical Analysis and Reporting

Statistical analyses and data visualisations were performed using the R programming language (v. 4.4.1; R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics were reported as counts and percentages for binary and categorical variables, and as medians and interquartile ranges (IQR) (quartile 1–quartile 3) for continuous variables, using the 'gtsummary' package (v. 2.0.3).

Since the follow-up period in cohorts 1 and 2 was longer than the 18 months of the original non-randomised validation study [16], while follow-up in cohort 3 was limited to 12 months, we evaluated the prognostic performance of the biomarker at 12 months, 18 months, and the entire follow-up period for patients unexposed to steroids in cohorts 1, 2 and 3, and at 18 months and the entire follow-up period for those exposed to steroids in cohort 1. To evaluate the discrimination ability of PredictSURE IBD in these patients, we compared the number of treatment escalations, time to first escalation, and time to first and recurrent escalations between the IBDhi and IBDlo predicted groups. Consistent with the original validation approach [16], our alternative hypothesis was that IBDhi patients would experience a higher number of escalations than those in the IBDlo group. Based on this directional hypothesis and prior knowledge of the expected effect, we used a one-sided Mann-Whitney *U* test to compare the number of treatment escalations between groups [22]. Since our statistical comparisons were predefined and hypothesis-driven rather than exploratory, we did not apply multiple testing corrections.

Time to first escalation was plotted using Kaplan-Meier curves, and the statistical significance of any differences between prognosis groups was assessed by log-rank tests. We also fitted multivariable Cox regression models to adjust for known and measured covariates. Since traditional Cox regression considers only the first escalation, we also employed multivariable Prentice, Williams, and Peterson (PWP) gap time models to analyse recurrent escalations. Among the methods available for modelling recurrent events, we selected the PWP gap time model because it allows for different baseline hazards for each escalation event within the same patient [23]. In addition to adjusting for covariates, the PWP models stratified escalations by their sequence to account for disease progression over follow-up and the impact of prior escalations. The models were also clustered by individual patients to account for within-patient correlations and the heterogeneity in disease progression among IBD patients. Given the smaller sample size of steroid-exposed patients in cohort 1 and the lower number of escalations, we fitted only univariable models to avoid the uncertainty in estimated coefficients associated with a low 'events per variable' ratio [24].

In time-to-event analyses, censoring was applied at the end of the defined follow-up period (12 months, 18 months, or the entire follow-up), or at the time of total colectomy in UC patients, whichever occurred first. If a patient refused a physician's recommendation to escalate therapy despite objectively documented active disease, an escalation was counted, and follow-up was terminated. Combining immunomodulators with biologics to prevent drug antibodies was considered a single escalation. In

addition, dose escalations of the same therapy and therapy switches due to intolerance or immunogenicity were not regarded as treatment escalations. Time-to-event analyses were implemented using the 'survival' package (v. 3.6-4) and visualised using the 'ggsurvfit' package (v. 1.1.0).

The available covariates included in the regression models were: age (continuous, years), haemoglobin (continuous, g/L), CRP (continuous, mg/L), albumin (continuous, g/L), disease location (binary, extensive colitis [UC] or ileocolonic [CD] vs. limited types), and perianal disease for CD. Current smoking status was excluded because there were only two cases in the UC cohort unexposed to steroids, both within the same prognostic group. This is known to produce infinite coefficient estimates and invalid p values [25]. To handle missing data in our datasets, we applied multiple imputation using the Multivariate Imputation by Chained Equations approach, implemented using the 'mice' package (v. 3.16.0) [26]. We applied different imputation methods tailored to the type of covariate, under the assumption that missingness was at random: predictive mean matching for continuous variables and random forest for binary variables. We specified ($m = 5$) to generate five complete datasets, and used the default number of iterations (maxit = 5) to ensure convergence of the imputation models. We fitted multivariable Cox and PWP models in the main analysis using the imputed datasets and additionally provided models based on complete cases as a sensitivity analysis.

To explore potential sex-based differences in biomarker performance, we compared the number of treatment escalations and time to first escalation within the first 12 months separately for males and females in the steroids-unexposed cohort.

Lastly, we estimated additional discrimination metrics across all cohorts, jointly based on exposure to steroids status and separately by recruitment location, including sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios, using the 'epiR' package (v. 2.0.76).

3 | Results

3.1 | Included Cohorts and Characteristics

Cohorts 1, 2 and 3 included 136 (71 UC and 65 CD) and 41 (15 UC and 26 CD) patients who were unexposed and exposed, respectively, to corticosteroids at the time of testing (Tables 1 and Table 2, respectively). One hundred twenty-four (70.1%) patients had ulcers on baseline endoscopy at inclusion, and whole-blood samples were drawn from all 177 patients for baseline assessment by PredictSURE IBD. Within each cohort, patients who were predicted to be IBDhi or IBDlo had comparable age, laboratory markers and disease distribution.

Cohort 4 included 20 patients (13 UC and 7 CD), all of whom were not on corticosteroids at inclusion. Median (IQR) baseline CRP was 47.5 (25.5–102.5) mg/L, with 16 (80%) patients having CRP $\geq$ 5 mg/L. Baseline and follow-up whole-blood samples were taken at the scheduled timepoints, although samples from week 6 were missing in 3 inpatients (hence, $n = 67$ for total samples). No significant differences were observed in patient demographics or disease characteristics between IBDhi and IBDlo patients (Table 3).

3.2 | Performed Treatment Escalations

Among patients unexposed to steroids, 57 (41.9%) required a total of 70 escalations, of which 64 (91.4%) were medical treatment escalations (Table 4). Within this group, a smaller percentage of UC patients required escalations compared with CD patients (26.8% vs. 58.5%). In patients exposed to steroids in cohort 1, 10 (66.7%) UC patients and 18 (69.2%) CD patients required a total of 36 escalations, most of which, 33 (91.7%), were medical.

3.3 | PredictSURE IBD Performance in Patients Unexposed to Steroids

During the first 12 months of follow-up, IBDhi UC patients needed significantly more treatment escalations than IBDlo UC patients (19 vs. 5 in total; $p = 0.011$, Figure 1). In addition, 4 patients needed more than one escalation, and one patient required colectomy due to severe refractory disease, all of whom were IBDhi patients. Of the 19 patients requiring escalations, 14 were correctly classified as IBDhi by PredictSURE IBD, resulting in a sensitivity of 73.7% (Table 5). However, sensitivity varied across geographic sub-cohorts, ranging from a minimum of 62.5% at sites in Belgium to a maximum of 83.3% at sites in the United Kingdom.

In CD patients, no statistically significant difference in the number of required escalations was observed between IBDhi and IBDlo patients (29 vs. 17 in total; $p = 0.16$, Figure 1). However, 5 patients required bowel surgeries (four intestinal resections and one peri-anal procedure), and 6 patients needed multiple escalations, all of whom were predicted as IBDhi. The test's sensitivity to correctly identify high-risk patients was 55.3%, with the highest sensitivity of 80% observed in the United Kingdom sub-cohort (Table 5).

IBDhi UC patients showed a significantly shorter time to first escalation compared with IBDlo UC patients ($p = 0.031$, Figure 2). After adjusting for known and measured covariates using the imputed datasets, IBDhi UC patients also demonstrated significantly higher hazards for both first (Table 6) and recurrent (Table 7) escalations. This significant association persisted in the complete case analyses (Table S1 and Table S2). Notably, a few covariates, such as albumin, age and haemoglobin, showed sporadic significant associations with escalation risk. However, IBDhi status was the only variable that consistently showed a significantly higher hazard across all models (Cox and PWP) and both imputed and complete case datasets.

In contrast, we observed no statistically significant difference in the time to first escalation between risk groups in CD patients ($p > 0.9$, Figure 2). This lack of association between PredictSURE IBD classification and escalation risk was consistent for both first and recurrent escalations across imputed (Tables 6 and Table 7) and complete case data (Table S1 and Table S2).

TABLE 1 | Baseline characteristics of patients unexposed to steroids at the time of testing in cohorts 1, 2 and 3.

	Ulcerative colitis			Crohn's disease		
Characteristic	**IBDhi** ***N* = 37**	**IBDlo** ***N* = 34**	***p* value**[a]	**IBDhi** ***N* = 35**	**IBDlo** ***N* = 30**	***p* value**[a]
Recruitment location, *n* (%)			0.31			0.29
Belgium	15 (40.5)	8 (23.5)		9 (25.7)	13 (43.3)	
United Kingdom	10 (27.0)	12 (35.3)		11 (31.4)	6 (20.0)	
North America	12 (32.4)	14 (41.2)		15 (42.9)	11 (36.7)	
Endoscopic baseline disease activity, *n* (%)	25 (67.6)	20 (58.8)	0.44	25 (67.6)	20 (58.8)	0.46
Age (years), median (IQR)	36.0 (23.8–46.0)	33.5 (23.9–41.6)	0.43	27.1 (22.8–32.2)	30.1 (23.3–46.4)	0.28
(missing)	0	2		1	0	
Female, *n* (%)	12 (44.4)	12 (50.0)	0.69	13 (48.1)	12 (48.0)	> 0.99
(missing)	10	10		8	5	
Current smoker, *n* (%)	2 (8.3)	0 (0.0)	0.50	4 (20.0)	6 (31.6)	0.48
(missing)	13	16		15	11	
Haemoglobin (g/L), median (IQR)	13.2 (12.0–14.9)	14.2 (13.0–14.9)	0.17	13.3 (11.4–14.4)	13.0 (11.8–14.3)	0.93
(missing)	8	9		4	6	
C-reactive protein (mg/L), median (IQR)	4.0 (2.0–5.6)	4.0 (1.0–7.6)	0.85	12.0 (2.8–20.0)	6.4 (1.1–19.0)	0.44
(missing)	7	6		2	0	
Serum albumin (g/L), median (IQR)	42.2 (38.0–47.0)	43.0 (41.0–46.8)	0.70	40.4 (34.5–44.3)	44.6 (41.0–47.0)	0.031
(missing)	12	9		7	8	
Disease distribution, *n* (%)			0.78			0.12
Proctitis	7 (20.0)	6 (19.4)				
Left-sided colitis	13 (37.1)	14 (45.2)				
Extensive colitis	15 (42.9)	11 (35.5)				
(missing)	2	3		2	0	
Ileal				10 (30.3)	16 (53.3)	
Colonic				10 (30.3)	4 (13.3)	
Ileocolonic				13 (39.4)	10 (33.3)	
Perianal disease				6 (17.6)	6 (20.0)	0.81
(missing)				1	0	

[a]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test.

While both the number of escalations and time to first escalation were significantly different between IBDhi and IBDlo patients in the full steroids-unexposed cohort, the number of escalations remained significant only in males ($p = 0.034$), whereas no significant differences were observed in females ($p = 0.47$). On the other hand, differences in time to first escalation between prognosis groups were not significant in either males or females ($p = 0.06$ and $p > 0.9$, respectively).

The biomarker's prognostic performance for predicting escalations within the first 12 months was largely maintained at 18 months in cohorts 1 and 2, with a significant difference in the number of escalations between risk groups for UC ($p = 0.028$) but not for CD ($p = 0.22$). However, while the biomarker continued to predict escalation numbers, its performance for identifying differences in the time to first escalation between risk groups in UC decreased, shifting from significant at 12 months to non-significant at 18 months ($p = 0.07$). In CD, the difference in time to first escalation between risk groups remained non-significant ($p = 0.52$).

When considering the entire follow-up period in cohorts 1 and 2 (UC: IBDhi, $n = 25$, median [IQR] follow-up of 34.7 [21.3–56.7] months; IBDlo, $n = 20$, median [IQR] follow-up of 27.4 [18.9–35.4] – CD: IBDhi, $n = 20$, median [IQR] follow-up of 19.2 [13.4–33.1] months; IBDlo, $n = 19$, median [IQR]

TABLE 2 | Baseline characteristics of patients exposed to steroids at the time of testing in cohort 1.

Characteristic	Ulcerative colitis			Crohn's disease		
	IBDhi $N = 7$	IBDlo $N = 8$	p value[a]	IBDhi $N = 14$	IBDlo $N = 12$	p value[b]
Endoscopic baseline disease activity, *n* (%)	7 (100.0)	8 (100.0)	> 0.99	14 (100.0)	12 (100.0)	> 0.99
Age (years), median (IQR)	25.8 (18.6–41.2)	27.5 (23.4–30.3)	0.87	26.7 (21.5–37.6)	28.1 (23.3–38.7)	0.66
Female, *n* (%)	6 (85.7)	2 (25.0)	0.041	10 (71.4)	4 (33.3)	0.052
Current smoker, *n* (%)	0 (0.0)	1 (12.5)	> 0.99	5 (35.7)	5 (41.7)	> 0.99
Haemoglobin (g/L), median (IQR)	10.8 (9.7–12.2)	13.2 (12.4–13.8)	0.032	12.9 (12.6–13.9)	13.6 (12.1–14.2)	0.76
(missing)				1	0	
C-reactive protein (mg/L), median (IQR)	1.6 (0.6–17.4)	4.5 (1.8–9.1)	0.61	14.8 (4.9–46.7)	15.2 (1.1–40.3)	0.63
Serum albumin (g/L), median (IQR)	40.6 (36.5–44.8)	39.2 (37.2–47.8)	> 0.99	42.2 (41.0–45.2)	44.1 (38.9–48.5)	0.62
(missing)	0	1		3	4	
Disease distribution, *n* (%)			> 0.99			0.30
Proctitis	0 (0.0)	0 (0.0)				
Left-sided colitis	2 (28.6)	2 (25.0)				
Extensive colitis	5 (71.4)	6 (75.0)				
Ileal				3 (21.4)	6 (50.0)	
Colonic				4 (28.6)	1 (8.3)	
Ileocolonic				7 (50.0)	5 (41.7)	
Perianal disease				1 (7.1)	2 (16.7)	0.58

[a]Fisher's exact test; Wilcoxon rank sum exact test; Wilcoxon rank sum test.
[b]Fisher's exact test; Wilcoxon rank sum test.

follow-up of 37.4 [22.1–46.7]), a significant difference in the total number of escalations between risk groups remained in UC but not in CD ($p = 0.025$ and $p = 0.22$, respectively). Consistent with observations at 18 months, the time to first escalation remained comparable between prognostic groups in both UC and CD ($p = 0.12$ and $p = 0.4$, respectively) over the entire follow-up period. This suggests that while the biomarker retains a stable ability to predict the total number of escalations, particularly in UC patients, its ability to predict the timing of the first escalation may diminish over extended follow-up periods.

3.4 | PredictSURE IBD Performance in Patients Exposed to Steroids

Among steroid-exposed patients in cohort 1, there was no statistically significant difference in the need for treatment escalation between the IBDhi and IBDlo groups during the first 18 months of follow-up in UC (6 vs. 8; $p = 0.33$) or CD (13 vs. 9; $p = 0.74$). Similarly, no difference was observed in the time to first escalation between risk groups among steroid-exposed patients ($p > 0.9$ for UC and $p = 0.5$ for CD, Figure 3) or in the hazard for escalation risk (Table 8). This lack of prognostic value for the biomarker persisted when considering the entire follow-up period, with no significant differences observed between prognostic groups in the number of escalations ($p = 0.48$ for UC, $p = 0.83$ for CD) or in the time to first escalation ($p = 0.9$ for UC, $p = 0.3$ for CD). Notably, 70% (14/20) of steroid-exposed IBDlo patients required treatment escalation, while only 33.3% (7/21) of exposed IBDhi patients maintained quiescent disease.

Of the 20 longitudinally sampled patients in cohort 4, 16 (80%) and 4 (20%) were classified as IBDhi and IBDlo, respectively, prior to initiation of corticosteroids. Compared to baseline, median (IQR) CRP dropped significantly at the first follow-up sampling timepoint (follow-up, 16.5 [4–34.8] mg/L; baseline, 47.5 [25.5–102.5] mg/L; $p < 0.001$). Over the course of follow-up, only 3 patients remained IBDhi throughout longitudinal sampling, whereas the remaining 13 switched to IBDlo in at least one subsequent sample. In contrast, the 4 patients who were classified as IBDlo at baseline did not experience any change in their classification.

These findings underscore the recommendation to use PredictSURE IBD in patients before commencing steroids, since steroid use at the time of sampling may result in erroneous classification of individuals as IBDlo, when they may have been IBDhi before steroids.

TABLE 3 | Baseline characteristics of longitudinally sampled patients in cohort 4.

Characteristic	Ulcerative colitis			Crohn's disease		
	IBDhi $N = 11$	IBDlo $N = 2$	p value[a]	IBDhi $N = 5$	IBDlo $N = 2$	p value[b]
Age (years), median (IQR)	38.0 (26.0–56.0)	39.5 (24.0–55.0)	0.92	26.0 (21.0–27.0)	38.0 (24.0–52.0)	0.57
Female, *n* (%)	5 (45.5)	0 (0.0)	0.49	1 (20.0)	0 (0.0)	> 0.99
Current smoker, *n* (%)	1 (9.1)	0 (0.0)	> 0.99	0 (0.0)	1 (50.0)	0.29
Haemoglobin (g/L), median (IQR)	11.9 (8.5–13.3)	11.3 (9.0–13.6)	0.92	12.8 (11.3–13.6)	12.8 (12.4–13.1)	> 0.99
C-reactive protein (mg/L), median (IQR)	36.0 (6.0–140.0)	39.0 (4.0–74.0)	0.62	97.0 (49.0–119.0)	48.5 (45.0–52.0)	0.57
Serum albumin (g/L), median (IQR)	27.0 (24.0–34.0)	33.0 (27.0–39.0)	0.49	33.0 (28.0–33.0)	32.0 (32.0–32.0)	0.84
Disease distribution, *n* (%)			> 0.99			0.14
Proctitis	0 (0.0)	0 (0.0)				
Left-sided colitis	3 (27.3)	1 (50.0)				
Extensive colitis	8 (72.7)	1 (50.0)				
Ileal				1 (20.0)	2 (100.0)	
Colonic				4 (80.0)	0 (0.0)	
Ileocolonic				0 (0.0)	0 (0.0)	
Perianal disease				0 (0.0)	1 (50.0)	0.29

[a]Wilcoxon rank sum test; Fisher's exact test.
[b]Wilcoxon rank sum exact test; Fisher's exact test; Wilcoxon rank sum test.

TABLE 4 | Treatment escalations among included patients were based on disease type and steroid exposure status.

	Total patients (*n*)	Patients with escalation (*n*, %)	Total escalations (*n*)	Sequence of escalation		Type of escalation	
				First (*n*, %)	Recurrent (*n*, %)	Medical (*n*, %)	Surgical (*n*, %)
UC patients unexposed to steroids in cohorts 1, 2, and 3	71	19 (26.8)	24	19 (79.2)	5 (20.8)	23 (95.8)	1 (4.2)
CD patients unexposed to steroids in cohorts 1, 2, and 3	65	38 (58.5)	46	38 (82.6)	8 (17.4)	41 (89.1)	5 (10.9)
UC patients exposed to steroids in cohort 1	15	10 (66.7)	14	10 (71.4)	4 (28.6)	14 (100)	0 (0)
CD patients exposed to steroids in cohort 1	26	18 (69.2)	22	18 (81.8)	4 (18.2)	19 (86.4)	3 (13.6)

Abbreviations: CD, Crohn's disease; UC, ulcerative colitis.

4 | Discussion

In line with the PROFILE results, we did not observe significant differences in disease course between the predicted risk subgroups in CD patients. However, our study validated the discrimination ability of the transcriptomic signature measured by PredictSURE IBD in a new steroid-naïve multinational UC cohort. We found that IBDhi UC patients required significantly more treatment escalations, had a shorter time to first escalation, and exhibited higher hazards for first and recurrent escalations within the first 12 months compared with IBDlo patients. In addition, in both UC and CD, all bowel surgeries and multiple escalations within the year following inclusion were required in IBDhi patients. Therefore, our findings suggest that PredictSURE IBD may have clinical utility in UC patients, warranting further investigation.

While these findings highlight the potential short-term clinical utility of PredictSURE IBD in UC patients, our additional analyses suggest that its prognostic performance in predicting the

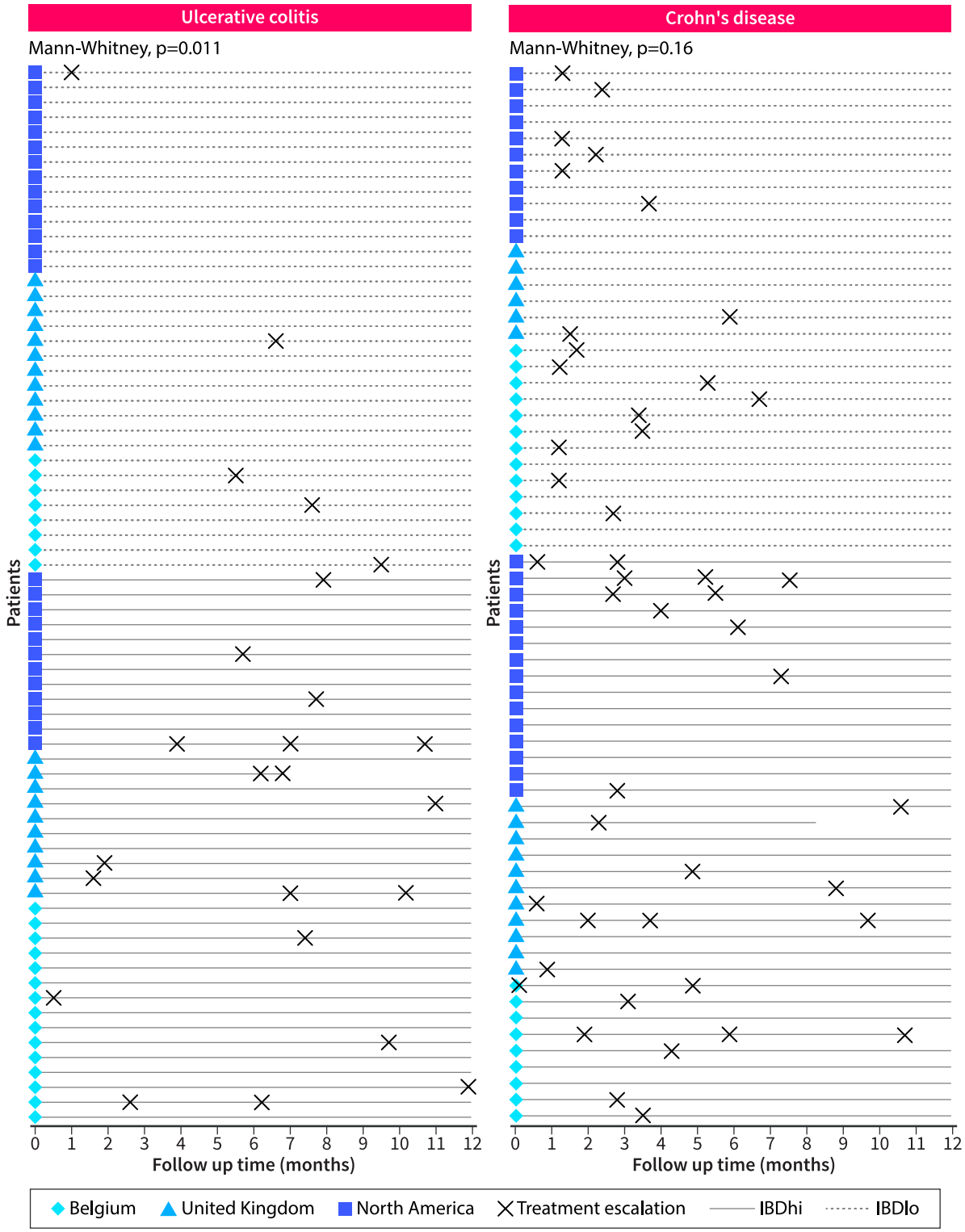

FIGURE 1 | Real-life treatment escalations during the first 12 months of follow-up in ulcerative colitis and Crohn's disease patients unexposed to steroids in cohorts 1, 2 and 3.

total number of escalations is largely maintained at 18 months and over longer follow-up periods. However, its ability to predict the timing of the first escalation diminishes over extended follow-up periods. This decline in the predictive performance for escalation timing may be influenced by the confounding effects of differences in clinical management, including variations in monitoring intensity, physician decisions on when to escalate therapy, and patient compliance with treatment recommendations. These factors may introduce inconsistencies in how soon patients receive their first escalation. Moreover, IBD is

TABLE 5 | Discrimination metrics of PredictSURE IBD for identifying high-risk patients likely to require treatment escalations.

	Patients with escalation (*n*, %)	Patients with correct IBDhi classification (*n*, %)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	PLR (95% CI)	NLR (95% CI)
Unexposed UC—All sites	19 (26.8)	14 (73.7)	73.7 (48.8–90.9)	55.8 (41.3–69.5)	37.8 (22.5–55.2)	85.3 (68.9–95)	1.67 (1.11–2.5)	0.47 (0.21–1.04)
Unexposed UC—Belgium	8 (34.8)	5 (62.5)	62.5 (24.5–91.5)	33.3 (11.8–61.6)	33.3 (11.8–61.6)	62.5 (24.5–91.5)	0.94 (0.49–1.79)	1.12 (0.36–3.54)
Unexposed UC—United Kingdom	6 (27.3)	5 (83.3)	83.3 (35.9–99.6)	68.8 (41.3–89)	50 (18.7–81.3)	91.7 (61.5–99.8)	2.67 (1.19–5.99)	0.24 (0.04–1.5)
Unexposed UC—North America	5 (19.2)	4 (80)	80 (28.4–99.5)	61.9 (38.4–81.9)	33.3 (9.9–65.1)	92.9 (66.1–99.8)	2.1 (1.04–4.23)	0.32 (0.05–1.93)
Unexposed CD—All sites	38 (58.5)	21 (55.3)	55.3 (38.3–71.4)	48.1 (28.7–68.1)	60 (42.1–76.1)	43.3 (25.5–62.6)	1.07 (0.67–1.69)	0.93 (0.55–1.57)
Unexposed CD—Belgium	15 (68.2)	6 (40)	40 (16.3–67.7)	57.1 (18.4–90.1)	66.7 (29.9–92.5)	30.8 (9.1–61.4)	0.93 (0.32–2.68)	1.05 (0.49–2.25)
Unexposed CD—United Kingdom	10 (58.8)	8 (80)	80 (44.4–97.5)	57.1 (18.4–90.1)	72.7 (39–94)	66.7 (22.3–95.7)	1.87 (0.75–4.64)	0.35 (0.09–1.41)
Unexposed CD—North America	13 (50)	7 (53.8)	53.8 (25.1–80.8)	38.5 (13.9–68.4)	46.7 (21.3–73.4)	45.5 (16.7–76.6)	0.87 (0.45–1.7)	1.2 (0.49–2.96)
Exposed UC	10 (66.7)	4 (40)	60 (26.2–87.8)	60 (14.7–94.7)	75 (34.9–96.8)	42.9 (9.9–81.6)	1.5 (0.46–4.91)	0.67 (0.23–1.89)
Exposed CD	18 (69.2)	10 (55.6)	44.4 (21.5–69.2)	50 (15.7–84.3)	66.7 (34.9–90.1)	28.6 (8.4–58.1)	0.89 (0.37–2.11)	1.11 (0.5–2.49)

Abbreviations: CD, Crohn"s disease; CI, confidence interval; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value; UC, ulcerative colitis.

a dynamic disease influenced by evolving immune responses, environmental triggers, and gut microbiota changes, which may gradually diminish the ability of a baseline immune signature to accurately predict the precise timing of disease progression. These results reinforce the need to develop a more dynamic prognostic approach—one that adapts to disease progression as well as patient- and care-specific factors over time—to better anticipate long-term disease trajectories and therapeutic needs in a timely and accurate manner.

Notably, the assay's discriminative performance in UC patients varied across recruitment locations. These variations may be attributed to differences in disease monitoring practices and thresholds for treatment escalation. Contributing factors include physicians' approaches to patient follow-up, patient compliance with monitoring protocols, and variations in local regulatory guidelines. Such variations underscore the importance of harmonised monitoring and therapeutic plans across sites to facilitate a fair evaluation of prognostic performance.

In addition, PredictSURE IBD showed poor discriminative ability in newly-diagnosed patients who had already been exposed to steroids at the time of testing of the assay. This was mainly driven by a higher proportion of apparent IBDlo patients who required treatment escalation(s). In the longitudinally sampled IBD cohort, both before and after steroid exposure, the majority of patients classified as IBDhi before corticosteroid treatment switched to an IBDlo classification in at least one sample during the 6-week follow-up period after commencing corticosteroids. This classification switch was accompanied by decreased inflammatory activity, as evidenced by the decline in CRP levels. It is known that corticosteroids exert their therapeutic (and adverse) effects through altering gene expression [27], which may explain the changes in PredictSURE IBD classification after initiating corticosteroids. Therefore, the interpretation of PredictSURE IBD results must be considered unreliable in patients on corticosteroid therapy.

In contrast to the initial, non-randomised validation study [16], the PROFILE trial showed no clinical utility of PredictSURE IBD in patients with newly-diagnosed active CD [17]. The lack of utility of the biomarker in PROFILE has been attributed to unexpectedly high treatment escalation rates in the IBDlo subgroup. This was likely driven by

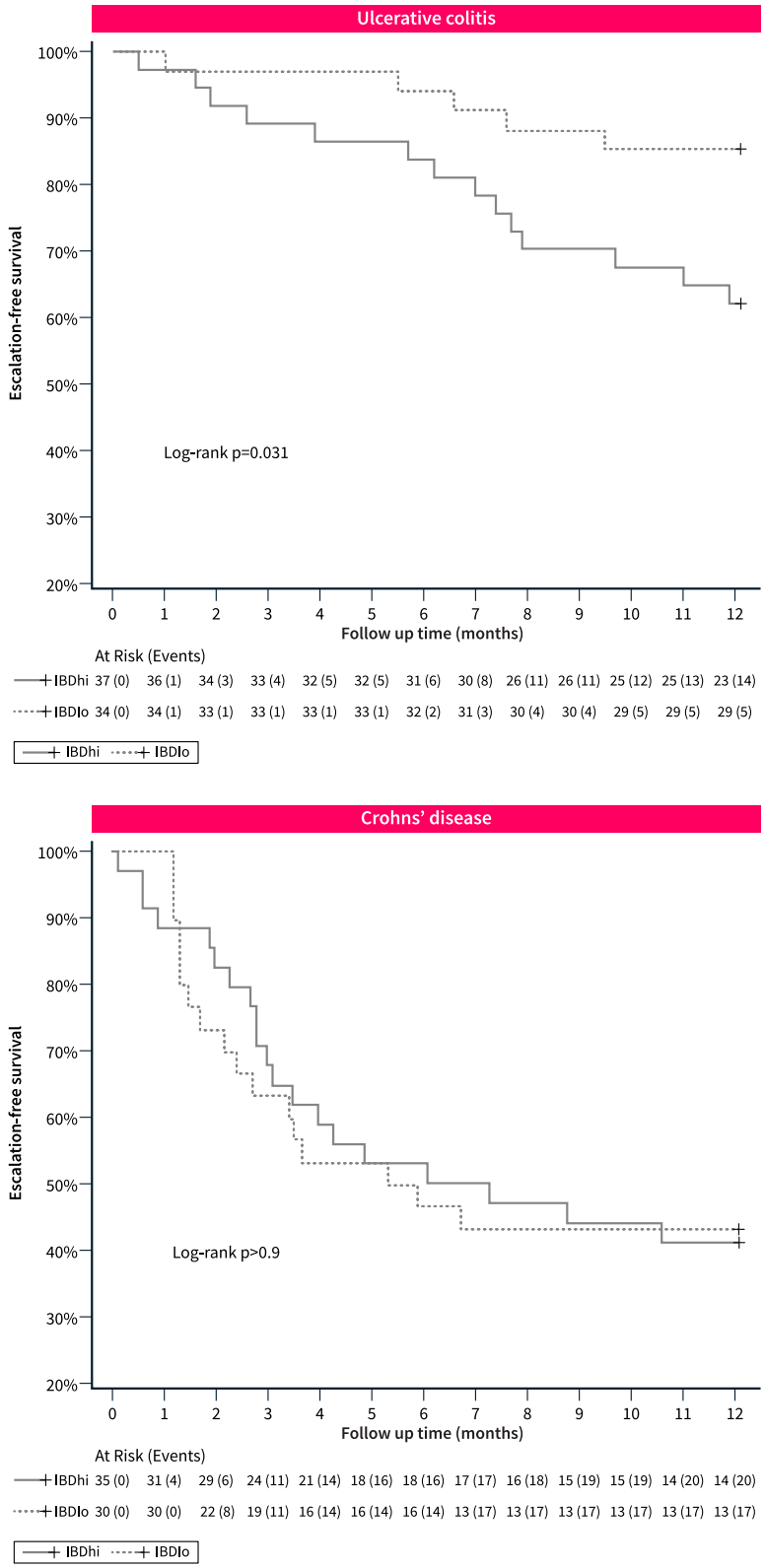

FIGURE 2 | Time to first treatment escalation during the first 12 months of follow-up in ulcerative colitis and Crohn's disease patients unexposed to steroids in cohorts 1, 2 and 3.

differing thresholds for treatment escalation in the interventional trial compared with routine clinical practice in previous validation cohorts. When combined with the early and more intensive monitoring of clinical responses in PROFILE, it is possible that treatment escalations reflected not only those with disease flares but also those responding slowly to initial

TABLE 6 | Hazard ratios for first escalation in ulcerative colitis and Crohn's patients unexposed to steroids in cohorts 1, 2 and 3 were estimated using multivariable Cox regression models in imputed datasets.

	Ulcerative colitis			Crohn's disease		
Characteristic	***N***	**HR (95% CI)**	***p* value**	***N***	**HR (95% CI)**	***p* value**
IBDhi prognosis	71	4.29 (1.38–13.3)	0.012	65	0.97 (0.46–2.05)	0.94
Age (years)	71	0.96 (0.92–1.00)	0.032	65	0.99 (0.96–1.02)	0.48
Haemoglobin (g/L)	71	1.46 (1.04–2.06)	0.031	65	1.00 (0.81–1.25)	0.98
C-reactive protein (mg/L)	71	1.00 (0.98–1.02)	0.97	65	1.00 (0.99–1.01)	0.74
Serum albumin (g/L)	71	0.91 (0.83–1.00)	0.048	65	0.99 (0.93–1.06)	0.83
Extensive colitis (UC)/Ileocolonic (CD)	71	1.65 (0.84–3.24)	0.15	65	1.26 (0.73–2.16)	0.41
Perianal disease				65	1.35 (0.57–3.17)	0.49

Abbreviations: CD, Crohn's disease; CI, confidence interval; HR, hazard ratio; UC, ulcerative colitis.

TABLE 7 | Hazard ratios for first and recurrent escalations in ulcerative colitis and Crohn's patients unexposed to steroids in cohorts 1, 2 and 3 were estimated using multivariable Prentice, Williams, and Peterson gap time models in imputed datasets.

	Ulcerative colitis		Crohn's disease	
Characteristic	**HR (95% CI)**	***p* value**	**HR (95% CI)**	***p* value**
IBDhi prognosis	4.06 (1.41–11.7)	0.005	1.32 (0.67–2.61)	0.42
Age (years)	0.97 (0.94–1.00)	0.30	0.99 (0.96–1.01)	0.39
Haemoglobin (g/L)	1.22 (0.95–1.58)	0.079	0.98 (0.81–1.20)	0.87
C-reactive protein (mg/L)	1.00 (0.98–1.01)	0.34	1.00 (0.99–1.01)	0.80
Serum albumin (g/L)	0.92 (0.85–0.99)	0.021	0.99 (0.93–1.05)	0.65
Extensive colitis (UC)/Ileocolonic (CD)	1.42 (0.77–2.63)	0.26	1.11 (0.67–1.85)	0.69
Perianal disease			1.52 (0.71–3.22)	0.27

Abbreviations: CD, Crohn's disease; CI, confidence interval; HR, hazard ratio; UC, ulcerative colitis.

therapy—making comparison with initial reports of PredictSURE IBD utility in CD unreliable. Similarly, one possible explanation for the discrepancy between UC and CD findings in this study is that treatment escalation decisions varied between countries. Additionally, for CD in particular, the standard of care has considerably evolved over time, with a growing trend toward earlier initiation of effective or advanced therapies.

Our study has several limitations. First, our evaluation in the full steroids-unexposed cohort was limited to 12 months due to the early termination of the PRECIOUS study. Nonetheless, as PredictSURE IBD was originally developed and validated for an 18-month period, our study focused on assessing its performance within this intended scope. Additionally, sample sizes were relatively small across all cohorts. Furthermore, in a minority of patients ($n = 10$), treating physicians may not have been blinded to the biomarker results during follow-up. However, treatment escalations were performed according to routine clinical practice. Moreover, we did not implement a protocol to ensure that therapies were pharmacokinetically optimised before determining therapy failure and requiring a therapy switch, which may have overestimated the need for escalations in our cohort. Additionally, the value of fecal calprotectin in predicting disease course compared to the biomarker could not be assessed as baseline fecal calprotectin data were missing in more than 80% of patients. Similarly, BMI data were not collected, preventing us from assessing its potential influence on disease progression and the biomarker's prognostic performance. Lastly, the small sample size and missing sex data prevented us from drawing firm conclusions about whether true sex-based differences in biomarker performance exist.

The strength of our study lies in being the first multinational performance assessment of a blood-based transcriptomic biomarker for predicting the course of IBD, with a broad diversity of population in which a biomarker was prospectively evaluated. In addition, the findings of our study help to shed light on the impact of common treatments such as corticosteroids on biomarkers.

In conclusion, our study found no significant differences in treatment escalations between biomarker risk groups in CD patients, similar to the PROFILE trial. However, IBDhi UC patients required significantly earlier and more frequent treatment escalations, suggesting potential utility for PredictSURE IBD in UC that may warrant further study. Additionally, the

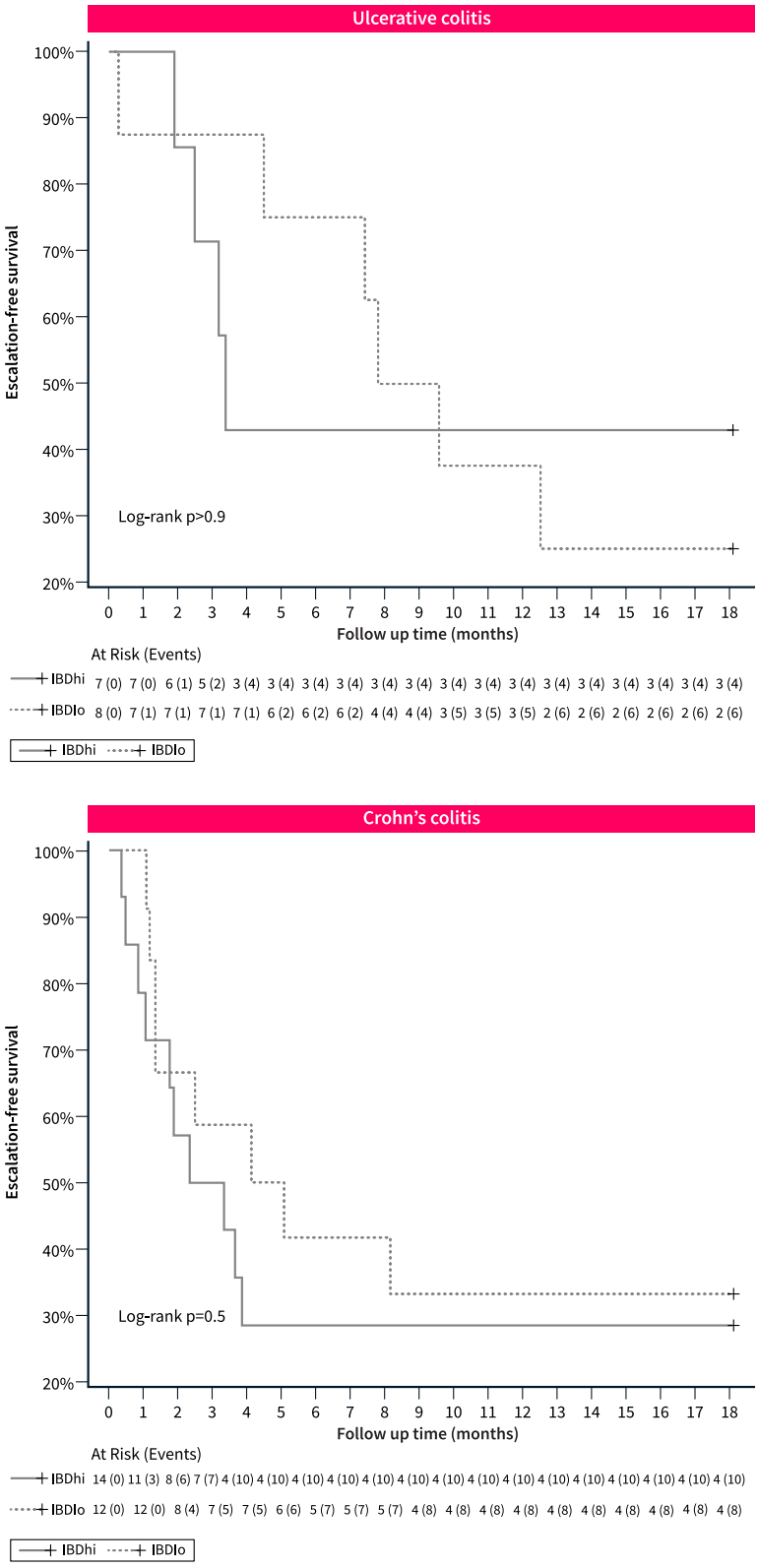

FIGURE 3 | Time to first treatment escalation during the first 18 months of follow-up in ulcerative colitis and Crohn's disease patients exposed to steroids in cohort 1.

TABLE 8 | Hazard ratios for first and recurrent escalations in ulcerative colitis and Crohn's patients exposed to steroids in cohort 1 were estimated using univariable Cox and Prentice, Williams, and Peterson gap time models.

	Ulcerative colitis		Crohn's disease	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value
Cox regression model				
IBDhi prognosis	0.94 (0.26–3.37)	0.92	1.35 (0.53–3.45)	0.53
Prentice, Williams, and Peterson gap time model				
IBDhi prognosis	1.22 (0.41–3.65)	0.72	1.53 (0.68–3.47)	0.31

Abbreviations: CI, confidence interval; HR, hazard ratio.

discrimination ability of the biomarker was found to be unreliable in patients on steroid therapy, which is likely due to steroid-induced gene expression effects.

Author Contributions

Dahham Alsoud: visualization, writing – original draft, formal analysis, conceptualization, methodology. **Nurulamin M. Noor:** visualization, writing – original draft, formal analysis, conceptualization, methodology, resources, investigation. **Lea Ann Chen:** investigation, resources, writing – review and editing, project administration. **Vivian Abadom:** investigation, writing – review and editing, resources. **Simon H. C. Anderson:** investigation, writing – review and editing, resources. **Lediona Ardolli:** writing – review and editing, resources, data curation. **Jordan Axelrad:** writing – review and editing, resources, investigation. **Peter Bossuyt:** investigation, writing – review and editing, resources. **Kenneth Croitoru:** investigation, writing – review and editing, resources. **Oriana M. Damas:** investigation, writing – review and editing, resources. **Lily Deng:** investigation, writing – review and editing, resources. **Parakkal Deepak:** investigation, writing – review and editing, resources. **Juan De La Revilla Negro:** investigation, writing – review and editing, resources. **Shanika de Silva:** investigation, writing – review and editing, resources. **Marc Ferrante:** investigation, writing – review and editing, resources. **Karen Hills:** writing – review and editing, resources, data curation. **Peter M. Irving:** investigation, writing – review and editing, resources. **James O. Lindsay:** investigation, writing – review and editing, resources. **Dana J. Lukin:** investigation, writing – review and editing, resources. **Paul A. Lyons:** writing – review and editing, conceptualization, methodology. **Eoin F. Mckinney:** writing – review and editing, methodology, conceptualization. **Maria Oliva-Hemker:** investigation, writing – review and editing, resources. **Caterina Oneto:** investigation, writing – review and editing, resources. **Roohi Patel:** investigation, writing – review and editing, resources. **Miles Parkes:** investigation, writing – review and editing, resources. **Lieven Pouillon:** investigation, writing – review and editing, resources. **Joao Sabino:** investigation, writing – review and editing, resources. **Lawrence J. Saubermann:** investigation, writing – review and editing, resources. **Jenny S. Sauk:** investigation, writing – review and editing, resources. **Sarah Sheibani:** investigation, writing – review and editing, resources. **Kenneth G. C. Smith:** conceptualization, methodology, writing – review and editing. **Keith S. Sultan:** investigation, writing – review and editing, resources. **Tony C. Tham:** investigation, writing – review and editing, resources. **Sare Verstockt:** conceptualization, methodology, writing – review and editing. **Raluca Vrabie:** investigation, writing – review and editing, resources. **Melissa Weidner:** investigation, writing – review and editing, resources. **Huimin Yu:** investigation, writing – review and editing, resources. **Bram Verstockt:** conceptualization, methodology, investigation, writing – review and editing, resources, supervision. **James C. Lee:** supervision, project administration, conceptualization, methodology, investigation, writing – review and editing, resources, **Severine Vermeire:** supervision, resources, writing – review and editing, methodology, conceptualization, investigation.

Acknowledgements

N.M.N. and M.P. are supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312). B.V. is supported by the Clinical Research Fund (KOOR) at the University Hospitals Leuven and the Research Council at the KU Leuven. M.F. and J.S. are Senior Clinical Researchers and S.V. is a postdoctoral fellow of the Fonds Wetenschappelijk Onderzoek (FWO), Flanders, Belgium. J.C.L. is a Lister Institute Prize Fellow and is supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC2219), the UK Medical Research Council (CC2219), and the Wellcome Trust (CC2219). J.A. is supported by grants from the Crohn's and Colitis Foundation and the NIH NIDDK Diseases K23DK124570.

S.V. and B.V. have received a research grant from the Leona M. and Harry B. Helmsley Charitable Trust, and D.A. is funded by this grant. D. A. received a research grant from the IBD Patient's Association Flanders (CCV VZW).

Conflicts of Interest

D.A. declares no conflicts of interest.

N.M.N. has received educational grants and/or speaker fees from AbbVie, B.M.S., Celltrion, Dr Falk Pharma, Ferring, Galapagos, Janssen, Lilly, Pfizer, Pharmacosmos, Takeda and Tillotts Pharma AG.

L.A.C. has received consulting fees from Pfizer, Bristol Myers Squibb, Eli Lilly, and PredictImmune.

V.A. declares no conflicts of interest.

S.H.C.A. declares no conflicts of interest.

L.A. declares no conflicts of interest.

J.A. has received research grants from BioFire Diagnostics and Genentech; consultancy fees, advisory board members, or honorarium from Abbvie, Adiso, BioFire Diagnostics, bioMérieux, Bristol-Myers Squibb, Celltrion, Fresnius, Ferring, Johnson and Johnson, Pfizer, and Vedanta.

P.B. has received financial support for research from AbbVie, Amgen, Celltrion, Mylan, Pfizer, and Takeda; lecture fees from AbbVie, Celltrion, Janssen, Lilly, and Takeda; and advisory board fees from AbbVie, Arena Pharmaceuticals, BMS, Celltrion, Dr Falk, Galapagos, Janssen, Lilly, Pentax, PSI-CRO, Roche, Takeda, and Tetrameros.

K.C. declares that it has received financial support for research from Abbvie and Janssen; speaking fees from Ferring and Janssen and advisory board fees from Abbvie, Takeda, Celltrion, Sandoz, Pfizer; and consulting fees from Goodcap and Microbiome Insights.

O.M.D. has received consultancy and/or advisory board fees from Abbvie and Janssen, and research funding from Pfizer.

L.D. declares no conflicts of interest.

P.D. has received research support under a sponsored research agreement unrelated to the data in the paper and/or consulting from AbbVie, Arena Pharmaceuticals, Boehringer Ingelheim, Bristol Myers Squibb, Johnson and Johnson, Pfizer, Prometheus Biosciences, Takeda Pharmaceuticals, Roche Genentech, Fresenius Kabi, Teva Pharmaceuticals,

Landos Pharmaceuticals, Sandoz, AGMB, ExeGi, Tr1X, Direct Biologics, Astra Zeneca, Alimentiv, Merck, LEK consulting, Celltrion and CorEvitas LLC.

J.D.R.N. has received an educational grant from Galapagos.

S.d.S. has received speaker fees/and or consulting fees- Abbvie, Tillots, Janssen, Dr Falk, Takeda.

M.F. has received research grants from AbbVie, Biogen, EG, Janssen, Pfizer, Takeda and Viatris; consultancy fees from: AbbVie, AgomAb Therapeutics, Boehringer Ingelheim, Celgene, Celltrion, Eli Lilly, Janssen-Cilag, MRM Health, MSD, Pfizer, Takeda and ThermoFisher; and speakers' fees from AbbVie, Biogen, Boehringer Ingelheim, Falk, Ferring, Janssen-Cilag, MSD, Pfizer, Takeda, Truvion Healthcare and Viatris.

K.H. declares no conflicts of interest.

P.M.I. has served as a consultant for AbbVie and Bristol Myers Squibb; received lecture fees from AbbVie, Bristol Myers Squibb, Celgene, Celltrion, Falk Pharma, Galapagos, Janssen, Lilly, Pfizer, and Takeda; served on Data Safety Monitoring Boards or advisory boards for AbbVie, Arena, Boehringer Ingelheim, Bristol Myers Squibb, Celgene, Celltrion, Genentech, Gilead, Hospira, Janssen, Lilly, MSD, Pfizer, Prometheus, Roche, Sandoz, Samsung Bioepis, and Takeda; and financial support for research from Celltrion, Galapagos, Pfizer, and Takeda.

J.O.L. declares consultant and advisory board participant for AbbVie, Bristol Myers Squibb, Celgene, Celltrion, Eli Lilly, Engytix Ferring Pharmaceuticals, Galapagos, Gilead, GSK, Janssen, MSD, Napp, Pfizer, Shire, Takeda, and Vifor Pharma; has received speaker fees and sponsorship for academic meetings from AbbVie, Ferring Pharmaceuticals, Janssen, MSD, Napp, Norgine, Pfizer, Shire, Tillotts Pharma, and Takeda; and investigator-led research grants from AbbVie, Gilead, Pfizer, Shire and Takeda.

D.J.L. declares consultancy fees from Abbvie, Altrubio, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly, Johnson & Johnson, Palatin Technologies, Pfizer, Prime Therapeutics, Prometheus Laboratories, PSI, Takeda, and Vedanta; has received speaker fees from Abbvie and Johnson & Johnson; and has received grants from Abbvie, Boehringer Ingelheim, Crohn's & Colitis Foundation, Johnson & Johnson, and Takeda.

P.A.L. was a co-founder of PredictImmune Ltd. and received consultancy fees from the company.

E.F.M. declares that he was the Chief Scientific Officer of PredictImmune Ltd 2017–2023.

M.O.H. has received research grants from Takeda Pharmaceuticals and Pfizer and has served on advisory boards for Takeda and Janssen Research& Development.

C.O. declares speaker bureau and/or advisory board participants for Salix, Abbvie, BMS, Pfizer, Ferring, Takeda, and Phathom.

R.P. declares no conflicts of interest.

M.P. has received research funding from Astra Zeneca, Gilead, Lilly, and Pfizer and speaker fees from Takeda and Janssen.

L.P. has received advisory board fees from Celltrion, Galápagos, Janssen-Cilag, Sandoz and Takeda; consultancy fees from Ipsos NV and Ismar Healthcare funded by Viatris; presentation fees from AbbVie, Celltrion, Ferring and Galápagos; and personal fees (congress support) from AbbVie, Ferring, Galápagos, Norgine and Takeda.

J.S. has received speaker's fees from Pfizer, Abbvie, Ferring, Falk, Takeda, Janssen, Fresenius, and Galapagos; consultancy fees from Pfizer, Janssen, Ferring, Fresenius, Abbvie, Galapagos, Celltrion, Pharmacosmos, and Pharmanovia; and research support from Galapagos and Viatris.

L.J.S. declares no conflicts of interest.

J.S.S. has received speaker's fees from Abbvie and Prometheus; consultancy fees or advisory board fees from CorEvitas, Roivant and Janssen; and research support from Crohn's and Colitis Foundation.

S.S. declares no conflicts of interest.

K.G.C.S. was a co-founder of PredictImmune Ltd. and has received consultancy fees from GSK.

K.S.S. declares no conflicts of interest.

T.C.T. has received honoraria for advisory boards, speaking and chairing from Amgen, Bristol Myers Squibb, Lilley.

S.V., R.V., M.W., and H.Y., declares no conflicts of interest.

B.V. has received research support from AbbVie, Biora Therapeutics, Landos, Pfizer, Sossei Heptares and Takeda; speaker's fees from Abbvie, Biogen, Bristol Myers Squibb, Celltrion, Chiesi, Falk, Ferring, Galapagos, Janssen, Lily, MSD, Pfizer, R-Biopharm, Sandoz, Takeda, Tillots Pharma, Truvion and Viatris; consultancy fees from Abbvie, Alfasigma, Alimentiv, Applied Strategic, Astrazeneca, Atheneum, BenevolentAI, Biora Therapeutics, Boxer Capital, Bristol Myers Squibb, Galapagos, Guidepont, Landos, Lily, Merck, Mylan, Nxera, Inotrem, Ipsos, Janssen, Pfizer, Progenity, Sandoz, Sanofi, Santa Ana Bio, Sapphire Therapeutics, Sosei Heptares, Takeda, Tillots Pharma and Viatris; stock options from Vagustim.

J.C.L. has received consultancy and/or advisory board fees from PredictImmune, C4X Discovery, Abbvie and Janssen.

S.V. has received grants from AbbVie, J&J, Pfizer, Galapagos and Takeda; consulting and/or speaking fees from AbbVie, Abivax, AbolerIS Pharma, AgomAb, Alimentiv, Arena Pharmaceuticals, AstraZeneca, Avaxia, BMS, Boehringer Ingelheim, Celgene, CVasThera, Dr Falk Pharma, Ferring, Galapagos, Genentech-Roche, Gilead, GSK, Hospira, Imidomics, Janssen, J&J, Lilly, Materia Prima, MiroBio, Morphic, MrMHealth, Mundipharma, MSD, Pfizer, Prodigest, Progenity, Prometheus, Robarts Clinical Trials, Second Genome, Shire, Surrozen, Takeda, Theravance, Tillots Pharma AG and Zealand Pharma.

Data Availability Statement

Pseudonymised data underlying this article will be shared upon reasonable request to the corresponding author.

References

1. M. Zhao, M. Sall Jensen, T. Knudsen, et al., "Trends in the Use of Biologicals and Their Treatment Outcomes Among Patients With Inflammatory Bowel Diseases – a Danish Nationwide Cohort Study," *Alimentary Pharmacology & Therapeutics* 55, no. 5 (2022): 541–557, https://doi.org/10.1111/apt.16723.

2. T. Kobayashi and T. Hibi, "Improving IBD Outcomes in the Era of Many Treatment Options," *Nature Reviews Gastroenterology & Hepatology* 20, no. 2 (2023): 79–80, https://doi.org/10.1038/s41575-022-00738-z.

3. I. L. Monstad, I. C. Solberg, M. Cvancarova, et al., "Outcome of Ulcerative Colitis 20 Years After Diagnosis in a Prospective Population-Based Inception Cohort From South-Eastern Norway, the IBSEN Study," *Journal of Crohn's and Colitis* 15, no. 6 (2021): 969–979, https://doi.org/10.1093/ecco-jcc/jjaa232.

4. J. Burisch, S. Lophaven, E. Langholz, and P. Munkholm, "The Clinical Course of Crohn's Disease in a Danish Population-Based Inception Cohort With More Than 50 Years of Follow-Up, 1962-2017," *Alimentary Pharmacology & Therapeutics* 55, no. 1 (2022): 73–82, https://doi.org/10.1111/apt.16615.

5. C. A. Siegel, C. B. Whitman, B. M. R. Spiegel, et al., "Development of an Index to Define Overall Disease Severity in IBD," *Gut* 67, no. 2 (2018): 244–254, https://doi.org/10.1136/gutjnl-2016-312648.

6. S. Hamdeh, M. Aziz, O. Altayar, M. Olyaee, M. H. Murad, and S. B. Hanauer, "Early vs Late Use of Anti-TNFa Therapy in Adult Patients With Crohn Disease: A Systematic Review and Meta-Analysis,"

Inflammatory Bowel Diseases 26, no. 12 (2020): 1808–1818, https://doi.org/10.1093/ibd/izaa031.

7. R. C. Ungaro, S. Aggarwal, O. Topaloglu, W. J. Lee, R. Clark, and J. F. Colombel, "Systematic Review and Meta-Analysis: Efficacy and Safety of Early Biologic Treatment in Adult and Paediatric Patients With Crohn's Disease," *Alimentary Pharmacology & Therapeutics* 51, no. 9 (2020): 831–842, https://doi.org/10.1111/apt.15685.

8. N. Cleveland, N. Candela, J. Carter, et al., "Early Versus Delayed Initiation of Vedolizumab In Ulcerative Colitis: Treatment Response In The Real World (RALEE)," *Gastroenterology* 162, no. 3 (2022): S105–S106, https://doi.org/10.1093/ibd/izac015.171.

9. C. A. Siegel, F. Yang, S. Eslava, and Z. Cai, "Treatment Pathways Leading to Biologic Therapies for Ulcerative Colitis and Crohn's Disease in the United States," *Clinical and Translational Gastroenterology* 11, no. 2 (2020): e00128, https://doi.org/10.14309/ctg.0000000000000128.

10. D. Wintjens, F. Bergey, E. Saccenti, et al., "Disease Activity Patterns of Crohn's Disease in the First Ten Years After Diagnosis in the Population-Based IBD South Limburg Cohort," *Journal of Crohn's and Colitis* 15, no. 3 (2021): 391–400, https://doi.org/10.1093/ecco-jcc/jjaa173.

11. M. Aloi, M. Bramuzzo, L. Norsa, et al., "Disease Activity Patterns in the First 5 Years After Diagnosis in Children With Ulcerative Colitis: A Population-Based Study," *Journal of Crohn's and Colitis* 15, no. 3 (2021): 367–374, https://doi.org/10.1093/ecco-jcc/jjaa203.

12. B. Verstockt, N. M. Noor, U. M. Marigorta, et al., "Results of the Seventh Scientific Workshop of ECCO: Precision Medicine in IBD-Disease Outcome and Response to Therapy," *Journal of Crohn's and Colitis* 15, no. 9 (2021): 1431–1442, https://doi.org/10.1093/ecco-jcc/jjab050.

13. E. F. McKinney, P. A. Lyons, E. J. Carr, et al., "A CD8+ T Cell Transcription Signature Predicts Prognosis in Autoimmune Disease," *Nature Medicine* 16, no. 5 (2010): 586–591, https://doi.org/10.1038/nm.2130.

14. J. C. Lee, P. A. Lyons, E. F. McKinney, et al., "Gene Expression Profiling of CD8+ T Cells Predicts Prognosis in Patients With Crohn Disease and Ulcerative Colitis," *Journal of Clinical Investigation* 121, no. 10 (2011): 4170–4179, https://doi.org/10.1172/jci59255.

15. E. F. McKinney, J. C. Lee, D. R. Jayne, P. A. Lyons, and K. G. Smith, "T-Cell Exhaustion, Co-Stimulation and Clinical Outcome in Autoimmunity and Infection," *Nature* 523, no. 7562 (2015): 612–616, https://doi.org/10.1038/nature14468.

16. D. Biasci, J. C. Lee, N. M. Noor, et al., "A Blood-Based Prognostic Biomarker in IBD," *Gut* 68, no. 8 (2019): 1386–1395, https://doi.org/10.1136/gutjnl-2019-318343.

17. N. M. Noor, J. C. Lee, S. Bond, et al., "A Biomarker-Stratified Comparison of Top-Down Versus Accelerated Step-Up Treatment Strategies for Patients With Newly Diagnosed Crohn's Disease (PROFILE): A Multicentre, Open-Label Randomised Controlled Trial," *Lancet Gastroenterology & Hepatology* 9, no. 5 (2024): 415–427, https://doi.org/10.1016/s2468-1253(24)00034-7.

18. H. Gordon, S. Minozzi, U. Kopylov, et al., "ECCO Guidelines on Therapeutics in Crohn's Disease: Medical Treatment," *Journal of Crohn's and Colitis* 18, no. 10 (2024): 1531–1555, https://doi.org/10.1093/ecco-jcc/jjae091.

19. W. W. Liao, M. Asri, J. Ebler, et al., "A Draft Human Pangenome Reference," *Nature* 617, no. 7960 (2023): 312–324, https://doi.org/10.1038/s41586-023-05896-x.

20. K. Gettler, R. Levantovsky, A. Moscati, et al., "Common and Rare Variant Prediction and Penetrance of IBD in a Large, Multi-Ethnic, Health System-Based Biobank Cohort," *Gastroenterology* 160, no. 5 (2021): 1546–1557, https://doi.org/10.1053/j.gastro.2020.12.034.

21. J. Burisch, H. Vardi, D. Schwartz, et al., "Health-Care Costs of Inflammatory Bowel Disease in a Pan-European, Community-Based, Inception Cohort During 5 Years of Follow-Up: A Population-Based Study," *Lancet Gastroenterology & Hepatology* 5, no. 5 (2020): 454–464, https://doi.org/10.1016/s2468-1253(20)30012-1.

22. J. Ludbrook, "Should We Use One-Sided or Two-Sided Values in Tests of Significance?," *Clinical and Experimental Pharmacology and Physiology* 40, no. 6 (2013): 357–361, https://doi.org/10.1111/1440-1681.12086.

23. W. Yang, C. Jepson, D. Xie, et al., "Statistical Methods for Recurrent Event Analysis in Cohort Studies of CKD," *Clinical Journal of the American Society of Nephrology* 12, no. 12 (2017): 2066–2073, https://doi.org/10.2215/cjn.12841216.

24. P. Peduzzi, J. Concato, E. Kemper, T. R. Holford, and A. R. Feinstein, "A Simulation Study of the Number of Events Per Variable in Logistic Regression Analysis," *Journal of Clinical Epidemiology* 49, no. 12 (1996): 1373–1379, https://doi.org/10.1016/s0895-4356(96)00236-3.

25. M. C. Bryson and M. E. Johnson, "The Incidence of Monotone Likelihood in the Cox Model," *Technometrics* 23, no. 4 (1981): 381–383, https://doi.org/10.2307/1268228.

26. S. van Buuren and K. Groothuis-Oudshoorn, "Mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software* 45, no. 3 (2011): 1–67, https://doi.org/10.18637/jss.v045.i03.

27. O. J. Schoneveld, I. C. Gaemers, and W. H. Lamers, "Mechanisms of Glucocorticoid Signalling," *Biochimica et Biophysica Acta* 1680, no. 2 (2004): 114–128, https://doi.org/10.1016/j.bbaexp.2004.09.004.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.