Digestive and Liver Disease xxx (xxxx) xxx

Contents lists available at ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Alimentary Tract

Vedolizumab in inflammatory bowel disease: Real-world outcomes and their prediction with machine learning—the IG-IBD LIVE study

Daniela Pugliese[a,b,c], Giuseppe Privitera[d], Nicola Cersullo[e], Harsh Bordekar[e], Federica Crispino[f], Nicolò Mezzina[g], Lucienne Pellegrini[h], Mariangela Allocca[i], Lucrezia Laterza[j], Anna Viola[k], Lorenzo Bertani[l], Pietro Soru[m], Barbara Scrivo[n], Brigida Barberio[o], Chiara Ricci[p], Paola Balestrieri[q], Marco Daperno[r], Dario Pluchino[s], Fernando Rizzello[t], Maria Lia Scribano[u], Renato Sablich[v], Luca Pastorelli[w], Francesco Manguso[x], Angela Variola[y], Antonio Di Sario[z], Laurino Grossi[aa], Davide Giuseppe Ribaldone[ab], Giuseppe Biscaglia[ac], Andrea Buda[ad], Giammarco Mocci[ae], Angelo Viscido[af], Maria Carla Di Paolo[ag], Sara Onali[ah], Stefano Rodino'[ai], Marina Coletta[aj], Mariabeatrice Principi[ak], Agnese Miranda[al], Arnaldo Amato[am], Cristina Bezzio[d,an], Carlo Petruzzellis[ao], Silvia Mazzuoli[ap], Stefano Festa[aq], Alessandro Sartini[ar], Davide Checchin[as], Libera Fanigliulo[at], Sara Gallina[au], Monica Cesarini[av], Giorgia Bodini[aw], Davide Stradella[ax], Rocco Spagnuolo[ay], Luisa Guidi[j], Edoardo Savarino[o], Maria Cappello[n], Flavio Caprioli[m], Francesco Costa[az], Walter Fries[k], Franco Scaldaferri[j], Gionata Fiorino[ba], Fabiana Castiglione[h], Alessandro Massari[g], Ambrogio Orlando[f], Alessandro Armuzzi[d,an,*]

[a] *Dipartimento di Medicina e Chirurgia Traslazionale, Università Cattolica del Sacro Cuore, L. Go A. Gemelli 8, 00168 Rome, Italy*
[b] *UOC Pronto Soccorso, Medicina d'Urgenza e Medicina Interna, Ospedale Isola Tiberina Gemelli Isola, 00186 Rome, Italy*
[c] *UOS Gastroenterologia, Ospedale Isola Tiberina Gemelli Isola, 00186 Rome, Italy*
[d] *Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy*
[e] *Institute of mechanics and adaptronics, TU Braunschweig*
[f] *IBD Unit, Villa Sofia–Cervello Hospital, Palermo, Italy*
[g] *Department of Biochemical and Clinical Science "L. Sacco" ASST Fatebenefratelli Sacco–University of Milan, Italy*
[h] *Gastroenterology, Federico II University Hospital, Napoli, Campania, Italy*
[i] *Gastroenterology and Endoscopy, IRCCS Ospedale San Raffaele, Milan, Italy*
[j] *CEMAD – IBD UNIT, Unità Operativa Complessa di Medicina Interna e Gastroenterologia, Dipartimento di Scienze Mediche e Chirurgiche, Fondazione Policlinico Universitario "A. Gemelli" IRCCS, Rome, Italy*
[k] *UOSD Malattie Intestinali Croniche, Dip. Di Medicina Clinica e Sperimentale, Policlinico Messina, Sicily, Italy*
[l] *Gastroenterology and Digestive Endoscopy Department of Medical Specialties Apuane Hospital, Tuscany North-West ASL, Massa, Italy*
[m] *Gastroenterology and Endoscopy Unit, La Fondazione IRCCS Ca' Granda Ospedale Maggiore di Milano Policlinico, Department of Pathophysiology and Transplantation, University of Milan, Milano, Lombardia, Italy*
[n] *Head IBD Clinic, Gastroenterology Section, Promise, University of Palermo, Sicily, Italy*
[o] *Gastroenterology Unit, Department of Surgery, Oncology and Gastroenterology, University of Padua, Padua, Italy*
[p] *Gastroenterology Unit, Spedali Civili Hospital, Department of Experimental and Clinical Sciences, University of Brescia, Brescia, Italy*
[q] *Unit of Digestive Disease of Campus Bio Medico University of Rome, Italy*
[r] *Gastroenterology Unit, Azienda Ospedaliera Ordine Mauriziano di Torino, Torino, Piemonte, Italy*
[s] *Gastroenterology Unit, A.O.U. Policlinico "Vittorio Emanuele" Catania Italy*
[t] *Policlinico Sant'Orsola Malpighi, Department of Internal Medicine and Gastroenterology, Bologna, Italy*
[u] *Villa Stuart Multi-Speciality Clinic, Rome, Italy*
[v] *Gastroenterology Unit, Santa Maria degli Angeli Hospital, Pordenone, Italy*
[w] *Gastroenterology and Hepatology Unit, ASST Santi Paolo e Carlo, 20142 Milan, Italy*
[x] *AO Cardarelli, Surgical Department, Naples, Italy*
[y] *IBD Unit, IRCCS Sacro Cuore Don Calabria, Negrar di Valpolicella, Verona, Italy*
[z] *Clinica di Gastroenterologia, Università Politecnica delle Marche, IBD-UNIT and Dipartimento Gastroenterologico e dei Trapianti, Polo Ospedaliero-Universitario "Umberto I-G.M. Lancisi- G. Salesi", Ancona, Italy*
[aa] *G D'Annunzio University-Digestive Physiopathology Ospedale Spirito Santo Pescara, Pescara, Italy*
[ab] *Department of Medical Sciences, University of Turin, Turin, Italy*

* Corresponding author at: IBD Center, IRCCS Humanitas Research Hospital, Via Alessandro Manzoni, 56, 20089 Rozzano, Milan, Italy.
E-mail address: alessandro.armuzzi@hunimed.eu (A. Armuzzi).

https://doi.org/10.1016/j.dld.2025.04.021

Please cite this article as: D. Pugliese, G. Privitera, N. Cersullo et al., Vedolizumab in inflammatory bowel disease: Real-world outcomes and their prediction with machine learning—the IG-IBD LIVE study, Digestive and Liver Disease, https://doi.org/10.1016/j.dld.2025.04.021

[ac] Division of Gastroenterology, IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Puglia, Italy
[ad] Department of Gastrointestinal Oncological Surgery, Gastroenterology and Endoscopy Unit, S. Maria del Prato Hospital, Feltre, Italy
[ae] SC Gastroenterologia Ospedale Brotzu, Cagliari, Italy
[af] Gastroenterology Unit, Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy
[ag] Department of Gastroenterology and Digestive Endoscopy, S. Giovanni Addolorata Hospital, Rome, Italy
[ah] Gastroenterology Unit, University Hospital AOU Cagliari, Department of Medical Science and Public Health, University of Cagliari, Cagliari, Italy
[ai] Division of Gastroenterology, 'Ciaccio-Pugliese' Hospital, Catanzaro, Italy
[aj] Gastroenterology and Endoscopy Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy
[ak] University of Bari, Gastroenterology, Bari, Italy
[al] Gastroenterology and Endoscopy Unit, University of Campania "L. Vanvitelli" Naples, Italy
[am] Ospedale Valduce, Gastroenterology, Como, Italy
[an] IBD Center, IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy
[ao] Department of Medicine, Gastroenterology and Endoscopy, Fondazione Poliambulanza, Brescia, Italy
[ap] Section of Gastroenterology & Artificial Nutrition, Hospital San Nicola Pellegrino, Bari, Italy
[aq] S. Filippo Neri Hospital, IBD Unit, Rome, Lazio, Italy,
[ar] Gastroenterology and Digestive Endoscopy Unit, Forlì-Cesena, AUSL della Romagna, Italy
[as] UO.C. Gastroenterologia, Ospedale HUB di Mestre, Venezia, Italy
[at] Gastroenterology Unit, Ospedale Santissima Annunziata, Taranto, Italy
[au] Division of Gastroenterology, 'Belcolle' Hospital, Viterbo, Italy
[av] Casa di Cura "Madonna della fiducia", Roma, Italy
[aw] Cattedra di Gastroenterologia, Dipartimento di Medicina Interna, Universita'di Genova, Genova, Italy
[ax] Gastroenterologia, A.O.U Maggiore della Caritá di Novara, Piemonte, University of Eastern Piedmont Amedeo Avogadro, Italy
[ay] Gastroenterology and Digestive Endoscopy Department, University of Catanzaro, Catanzaro, Italy
[az] AOUP, Gastroenterology, Pisa, Italy
[ba] IBD Unit, Gastroenterology and Digestive Endoscopy, San Camillo-Forlanini Hospital, I, Rome, Italy

ARTICLE INFO

Article history:
Received 8 March 2025
Accepted 10 April 2025
Available online xxx

Keywords:
Anti-integrin therapy
Machine learning (ML)
Shapley values (SHAP)

ABSTRACT

Background and aims: Real-world studies on vedolizumab in inflammatory bowel disease (IBD) are often limited by small sample size and short follow-up. In this study, we investigated the 2-year effectiveness and safety of vedolizumab in patients with IBD, and applied eXplainable Artificial Intelligence (XAI) to identify predictors of both.
Methods: The Long-term Italian Vedolizumab Effectiveness (LIVE) study is multicentric, ambispective, observational study enrolling 1111 IBD patients (563 Crohn's disease, CD, 542 ulcerative colitis, UC). Steroid-free clinical remission (SFCR) at 24 months was the primary endpoint. A XAI model (eXtreme Gradient Boosting, XGB) was applied to identify the main clinical predictors of SFCR and development of adverse events (AEs).
Results: Rates of SFCR at 24 months were 31.6 % and 39.7 % in CD and UC patients, and 0.14 AEs per patient-year was recorded. On XGB analysis, previous exposure to anti-TNFα and older age were the most important drivers for the prediction of SFCR; lower baseline CRP levels and fewer comorbidities were the most important features associated with no development of AEs.
Conclusions: Vedolizumab is effective and safe in IBD patients. XAI yielded promising results in identifying the most important predictors of SFCR and development of AEs.

1. Introduction

Inflammatory bowel disease (IBD), with its two main forms (Crohn's disease, CD, and ulcerative colitis, UC), is a chronic condition that primarily affects the gastrointestinal tract and can significantly reduce patients' quality of life [1,2]. For over a decade, anti-Tumor Necrosis Factor (TNF)α drugs had been the only treatment for patients who have failed or are intolerant to conventional therapies. Vedolizumab, a fully humanised monoclonal antibody that blocks the $\alpha4\beta7$ integrin to prevent lymphocyte homing into the gut mucosa [3], is the first non-anti-TNFα target therapy licensed for the treatment of IBD, based on the results of the GEMINI program [4–6].

Currently, treatment choice is mostly based on pharmacoeconomic considerations and physicians' confidence with a specific drug, whereas personalised medicine approaches are restricted to a few scenarios [7]. Hence, the quest for the identification of reliable predictors is of the utmost importance. Artificial intelligence (AI), specifically eXplainable AI (XAI), can be implemented in clinical research for the identification of the most accurate predictors and, potentially, to develop predictive algorithms [8–10]. Machine learning is a subfield of AI that can identify complex, non-linear relationships between inputs (*e.g.*, patients' baseline characteristics) and outputs (*e.g.*, clinically relevant outcomes), and it is used for the development of algorithms that allow computers to learn from and make predictions based on data [11,12].

With this study, we aim to assess the real-life effectiveness and safety of vedolizumab in a large cohort of patients with IBD, and to use a machine learning-based approach to identify the most relevant predictors for both outcomes.

2. Materials and methods

2.1. Study design and outcome measures

The Long-term Italian Vedolizumab Effectiveness (LIVE) study is an observational, ambispective study conducted at 47 Italian IBD centres affiliated with the Italian Group for the Study of Inflammatory Bowel Disease (IG-IBD), enrolling consecutive patients with IBD who had started vedolizumab between April 2016 and June 2017; all patients had a baseline endoscopic evaluation (no more than 3 months before the first vedolizumab infusion). All patients received standard vedolizumab induction therapy, followed by maintenance treatment with vedolizumab 300 mg *iv* every 8 weeks; for patients who did not achieve clinical remission or experienced symptomatic relapse, treatment was optimized to 300 mg

iv every 4 weeks, based on physician's clinical judgement. Data from vedolizumab initiation up to the time of inclusion in the study were retrospectively extracted from medical records; after enrolment (concluded in December 2018), patients were prospectively followed up until June 2019 or until drug discontinuation. Data collection methods are presented in Supplementary Materials. The primary outcome was steroid-free clinical remission (SFCR) at 24 months. Secondary outcomes included: 1) clinical response at 14 weeks and 6 months, 2) clinical remission at each time point, 1) SFCR at any other time point, 4) biochemical remission at each time point 5) endoscopic remission (for patients with an available follow-up endoscopy) during the first and second year of treatment, 6) occurrence of adverse events (AEs) and 7) treatment persistence. A XAI-based approach was adopted to identify predictors of SFCR at 12 and 24 months, and of occurrence of AEs. Additional information on study design and outcome measures can be found in Supplementary Materials.

2.2. Conventional statistics and artificial intelligence analysis

Details on statistical analyses are presented in Supplementary Materials. XAI was used to predict SFCR at 12 and 24 months and occurrence of AEs. Patients' age, sex, baseline CRP levels, previous anti-TNFα exposure, concomitant steroid therapy, number of comorbidities, disease type, baseline clinical and endoscopic activity, disease extent for UC, and location and behaviour for CD, were considered for the analysis. In order to maximize the performance of the model and, at the same time, provide explainability of the results, we adopted a three-step approach.

1) Synthetic Minority Over-sampling Technique (SMOTE) for data preparation. To account for imbalances in the dataset, SMOTE was used. In SMOTE [13], the minority class is oversampled by creating synthetic data instances, which not only improve the class imbalance, but also help to train a more robust model. To maintain the statistical coherence in the augmented dataset, several linear statistic metrics were employed, as described elsewhere [14].
2) eXtreme Gradient Boosting (XGBoost) to predict patients' outcomes. XGBoost is an machine learning technique that uses stochastic gradient boosting to train a sequence of simple decision tree models, which are later combined to make the final prediction with the help of ensemble learning [14].
3) Shapley values (SHAP) to provide visual interpretability of the model. SHAP [15] presents an explanation for a specific prediction by computing the contribution of each feature to the final prediction. In brief, SHAP provides a ranked list of the features contributing to the prediction, giving to each of them a specific 'weight' in the model, expressed as its SHAP value (*i.e.*, the higher the SHAP value, the more a specific variable is important in determining the final prediction of the model). Furthermore, it allows to visualize the directionality of the contribution of each feature: in the beeswarm plot, each dot represents a single prediction, the colour of the dots represents the value assumed by the feature (blue for lower values, red for higher ones), and a positive of negative SHAP value (represented on the x-axis) informs as to whether that specific observation is associated with a positive or negative outcome. A detailed description of the AI analysis is provided in Supplementary Materials.

2.3. Ethical considerations

The study protocol was approved on 4 June 2018 by the Ethics Committee of the coordinating centre (Fondazione Policlinico Universitario A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy). All patients included in the study provided written informed consent.

3. Results

3.1. Patients' population

Out of 1111 enrolled patients, 1105 were included in the final analysis (the remaining 6 were excluded, having a follow-up $<$ 14 weeks). Table 1 summarizes the characteristics of the cohort. The majority of patients (77.1 %) had been previously exposed to at least one anti-TNFα drug; no patient had previously received target therapies other than anti-TNFα, as they had not been commercialized in Italy when the study was conducted. The mean follow-up for the entire cohort was 78.5 $\pm$ 35.3 weeks (77.8 $\pm$ 34.9 weeks for CD, 79.3 $\pm$ 35.8 weeks for UC). With regard to treatment optimization, 398 (36.0%) patients required the additional vedolizumab infusion at week 10, and 357 (32.3 %) patients received dose escalation to every four weeks during follow-up.

3.2. Effectiveness

SFCR at 24 months was observed in 178 (31.6 %) patients with CD. Notably, a nearly significant increase in the rates of SFCR was observed from 14 weeks to 6 months (27.1 % vs. 35.7%, $p = 0.051$) (Fig. 1A). Similarly, the rates of clinical response and remission, and of biochemical remission increased from 14 weeks to 6 months (57.7 % vs 67.7 %, $p < 0.001$, 29.1 % vs. 35.7 %, $p = 0.019$, 16.2 % *vs.* 22.0 %, $p = 0.012$) (Fig. 1B–D). From 6 months onwards, a substantial stability was observed in the cohort (Fig. 1A–D). At least one follow-up endoscopy was available for 295 patients with CD (median time from vedolizumab initiation to the first endoscopic evaluation: 17.3 months, range 3–27.3). Endoscopic remission was recorded in 23/117 (19.7 %) and 45/178 (25.3 %) patients, in the first and subsequent year, respectively (Fig. 1E).

With regard to UC, SFCR at 24 months was recorded in 215 (39.7 %) patients (Fig. 2A). A significant increase in the rates of SFCR (28.0 % *vs.* 35.6%, $p = 0.008$), clinical remission (30.3 % *vs.* 38.7%, $p = 0.003$) and biochemical remission (22.9 % *vs.* 29.3%, $p = 0.016$) from 14 weeks to 6 months was recorded, with an overall stability during the subsequent observations (Fig. 2A–D). At least one follow-up endoscopy was available for 352 patients with UC (median time from vedolizumab initiation to the first endoscopic evaluation: 17.4 months, range 2.2–27.4). During the first year, endoscopic remission was recorded in 34/145 (23.4%) patients, while in the subsequent year in 81/207 (39.1 %) (Fig. 2E).

3.3. Safety

In the entire cohort, 318 AEs were experienced by 308 patients, with an incidence rate of 0.14 AE per patient-year. Three deaths were recorded during the follow-up: a 26.5-year-old man who died from metastatic colon cancer (duration of vedolizumab therapy 84.7 weeks), a 53 years-old man who dies from an advanced small bowel carcinoma (duration of therapy 63 weeks) and a 90-year-old man who died after a complicated colectomy (duration of therapy 21.1weeks). Table 2 reports the incidence of AEs: infections were the most frequent (120, 38.0 %), and47 cases of arthralgias were reported, but clinical signs of arthritis were evident only in 10 patients. New diagnosis or recurrence of cancer/dysplasia were reported in 26 (8.1 %) patients (a detailed list is presented in Supplementary Table 1); 3 patients developed intestinal cancer (in the colon or small bowel), with an incidence rate of 1.3 per 1000 patient-year. AEs caused 63 patients to withdraw from vedolizumab treatment, with malignancies and infections being the most common reasons (AEs leading to vedolizumab discontinuation are presented in Supplementary Table 2).

Table 1
Patients' baseline characteristics.

Patients, n	1105	CD (n = 563)	UC (n = 542)
Female, n (%)	467 (42.3)	251 (44.6)	216 (39.9)
Age, years, median (range)	46.7 (18.0–90.0)	45.7 (17.8–85.9)	48.1 (17.8–90.0)
Weight, kg, mean (SD)	67.2 (14.0)	64.6 (13.5)	69.7 (14.0)
Current smokers, n (%)	346 (31.3)	216 (38.4)	130 (24.0)
Disease duration, years, median (range)	10.0 (0.0–59.7)	10.9 (0.0–59.7)	8.8 (0.1–40.7)
Disease, n (%)			
CD	563 (51.0)		
UC	542 (49.0)		
CD location, n (%)			
L1		159 (28.3)	
L2		76 (13.5)	
L3		316 (56.1)	
L4		12 (2.1)	
CD behaviour, n (%)			
B1		169 (29.9)	
B2		272 (48.3)	
B3		123 (21.8)	
Perianal CD, n (%)		165 (29.3)	
CD clinical activity (HBI), n (%)			
Quiescent (<5)		47 (8.3)	
Mild (5–7)		135 (24.0)	
Moderate (8–16)		340 (60.4)	
Severe (>16)		41 (7.3)	
CD endoscopic activity (SES-CD), n (%)			
Quiescent (0–2)		20 (3.6)	
Mild (3–6)		78 (13.8)	
Moderate (7–15)		319 (56.7)	
Severe (>15)		146 (25.9)	
UC extent, n (%)			
E1			21 (3.9)
E2			213 (39.3)
E3			308 (56.8)
UC clinical activity (PMS), n (%)			
Quiescent (0–1)			18 (3.3)
Mild (2–4)			95 (17.5)
Moderate (5–7)			317 (58.5)
Severe (>7)			112 (20.7)
UC endoscopic activity (endoscopic MS), n (%)			
Quiescent (0)			5 (0.9)
Mild (1)			31 (5.7)
Moderate (2)			258 (47.6)
Severe (3)			248 (45.8)
Previous exposure to anti-TNFα, n (%)	852 (77.1)	462 (82.0)	390 (71.9)
Previous cancer diagnosis, n (%)	90 (8.1)	55 (9.8)	35 (6.5)
Previous CD-related surgery, n (%)		324 (57.5)	
Concomitant therapies, n (%)			
5-ASA	412 (37.3)	114 (20.2)	298 (55.0)
IMM	59 (5.3)	34 (6.0)	25 (4.6)
5-ASA+IMM	45 (4.1)	14 (2.5)	31 (5.7)
Baseline steroid therapy, n (%)	502 (45.4)	229 (40.7)	273 (50.4)
CRP, mg/L, mean (SD)	12.4 (17.9)	13.5 (17.6)	11.3 (18.2)

5-ASA, aminosalicylates; CD, Crohn's disease; CRP, C-reactive protein; HBI, Harvey-Bradshaw Index; IMM, Immunosuppressants; MS, Mayo score; PMS, partial Mayo score; SES-CD; Simple Endoscopic Score for Crohn's Disease; TNF, tumour necrosis factor; UC, ulcerative colitis.

3.4. Persistence

Overall, 488 (44.2 %) patients suspended vedolizumab during the observation: 261 (46.4 %) CD and 227 (41.9 %) UC patients. Median vedolizumab therapy duration was 98.7 (range 14.0–117.1) and 102.2 (14.0–117.4) weeks for patients with CD and UC, respectively. Vedolizumab ineffectiveness (on intestinal, perianal and/or extra-intestinal manifestations) was the cause of treatment discontinuation in 399 (81.8 %) patients: 79 (16.2 %) primary failures and 320 (65.6 %) secondary failures were recorded. More specifically, ineffectiveness on EIMs was experienced by 28 patients (5.7 %, 8 of whom also had active intestinal disease) and ineffectiveness on perianal disease was experienced by 15 patients (3.1 %, 2 of whom also had active intestinal disease). Reasons for vedolizumab discontinuation are presented extensively in Supplementary Table 3.

Kaplan-Meier survival analysis revealed no differences between patients with CD and UC (Fig. 3A). Upon stratification by previous anti-TNFα exposure, a significant difference emerged in favour of bionaïve (Fig. 3B). Notably, previous anti-TNFα exposure significantly impacted on vedolizumab persistence only in patients with CD (70.0 % and 66.5 % at 12 and 24 months in bionaïve, *vs.* 73.7 % at 12 51.9 % at 12 and 24 months in bioexposed, log-rank test $p = 0.02$); conversely, no such differences were observed in patients with UC (Fig. 3C, D).

3.5. Outcome prediction from the trained machine learning model

Data augmentation with SMOTE was used to create more balanced datasets for the machine learning model to work with: three cohorts were generated for the outcomes considered for this analysis (SFCR at 12 and 24 months, and developments of AEs). The cohorts' characteristics after data augmentation are presented in Supplementary Table 4.

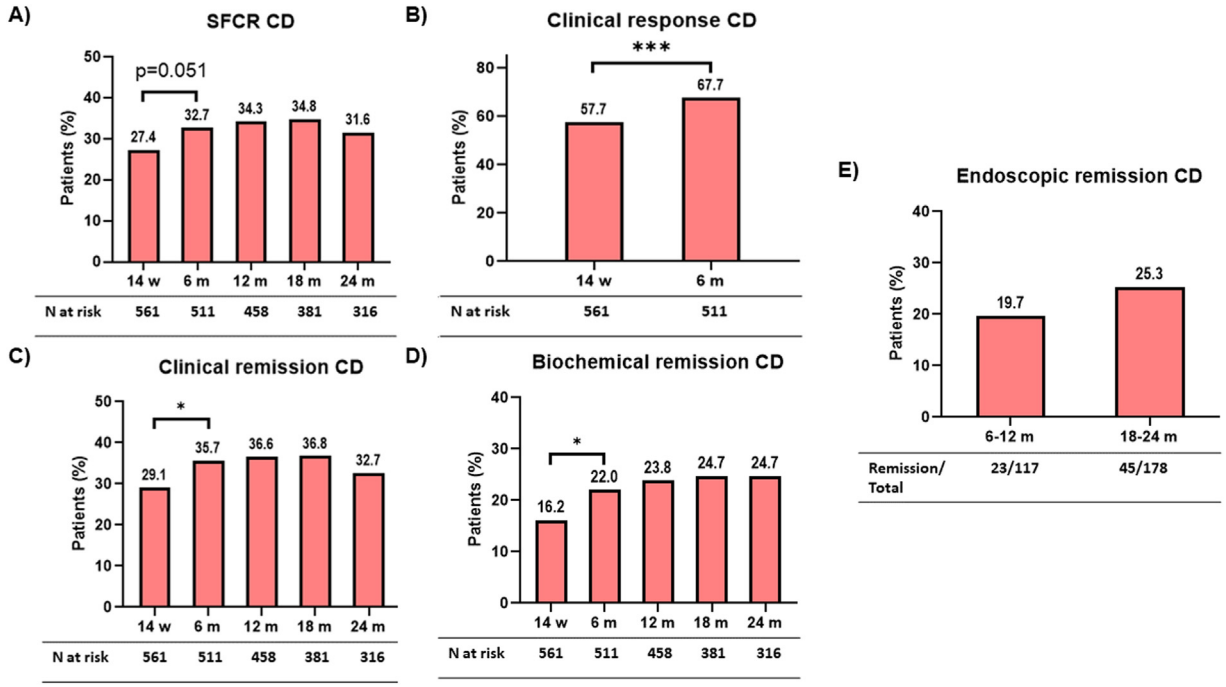

Fig. 1. Effectiveness of vedolizumab in patients with Crohn's disease. Effectiveness outcomes in CD by time of assessment: A) Clinical response (intention-to-treat), B) Clinical remission (intention-to-treat), C) Steroid-free clinical remission (intention-to-treat), D) Biochemical remission (intention-to-treat), E) Endoscopic remission (per-protocol). *, $p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

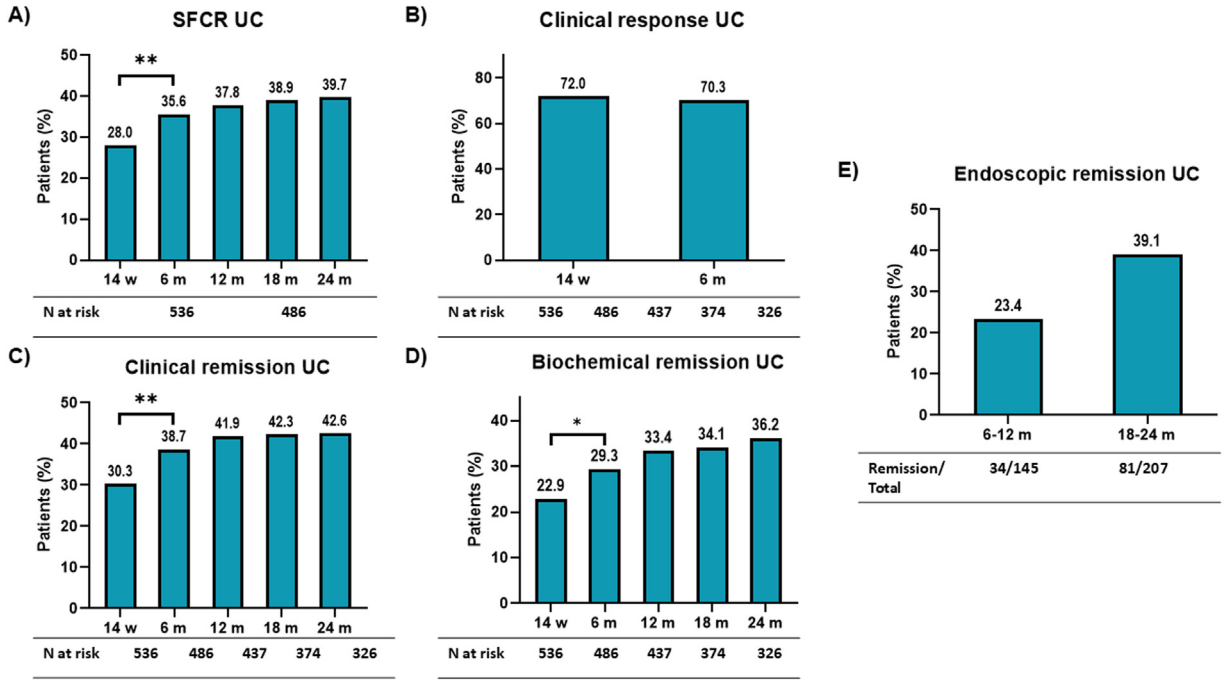

Fig. 2. Effectiveness of vedolizumab in patients with ulcerative colitis. Effectiveness outcomes in UC by time of assessment: A) Clinical response (intention-to-treat), B) Clinical remission (intention-to-treat), C) Steroid-free clinical remission (intention-to-treat), D) Biochemical remission (intention-to-treat), E) Endoscopic remission (per-protocol). *, $p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

The classification results and confusion matrixes of the XGBoost model are presented in Fig. 4 and in Supplementary Fig. 1. For each outcome, the model was trained on 80 % of the augmented dataset (*i.e.*, training dataset), then its performance was tested on a sample of 20% randomly selected patients (*i.e.*, testing dataset), as summarized in Fig. 4A: the F-1 score, which provides an overall estimate of the model's ability to identify true positive and true negative cases, is ≥ 0.75 for all the considered scenarios, suggesting an overall good performance of the model for each outcome. The models were successfully trained to achieve the AUC of 0.74 for SFCR at 24 months, of 0.84 for SFCR at 12 months and of 0.78 for the development of AEs (as outlined in Supplementary Figure 1).

Fig. 4B-G presents the SHAP analysis of the models for both outcomes. The global feature contribution plot (left panels) hierarchically lists the features included in the model and the beeswarm

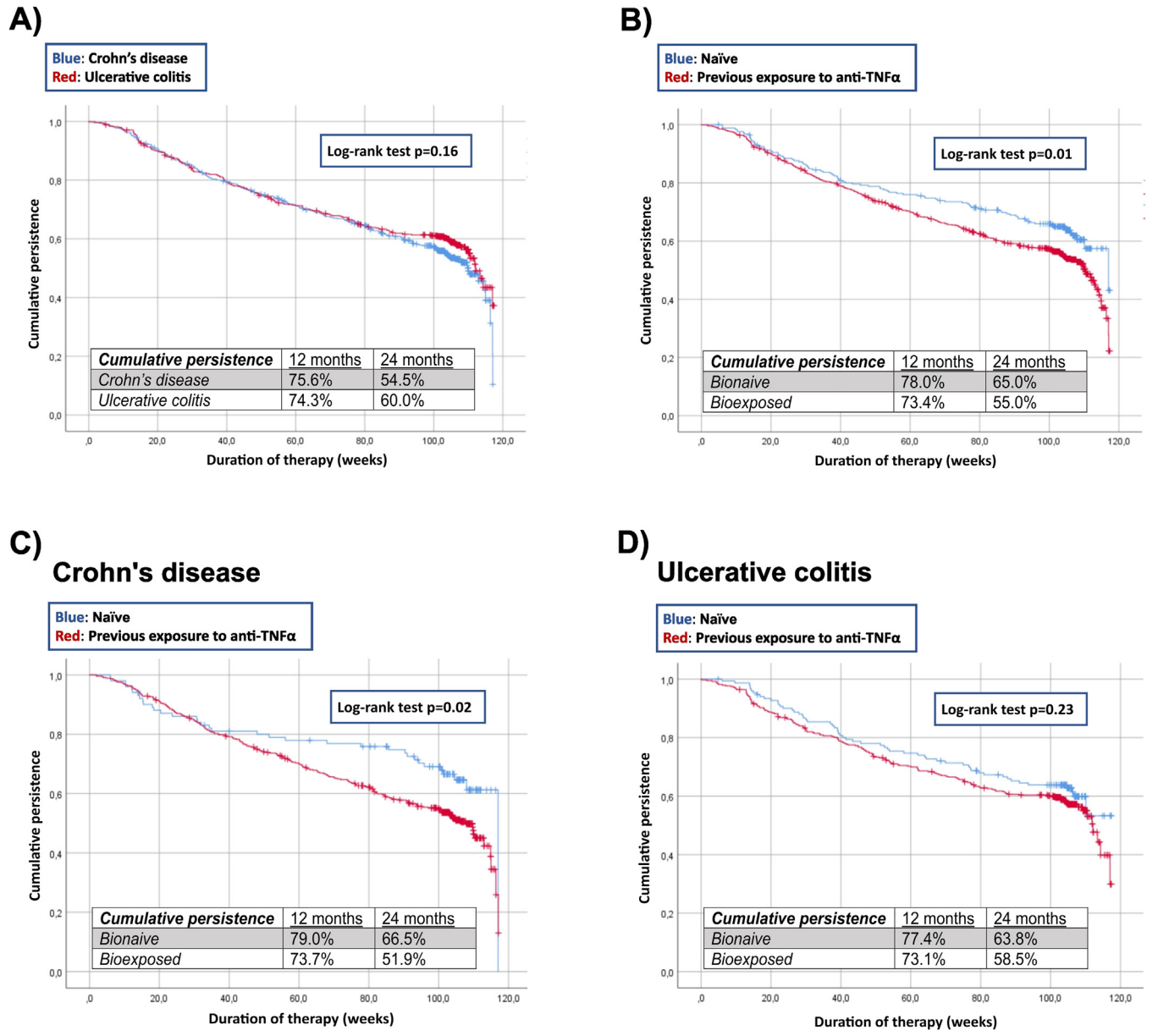

Fig. 3. Persistence on vedolizumab. Kaplan–Meier survival curves for persistency of vedolizumab therapy: A) Crohn's disease *vs* ulcerative colitis, B) Bionaïve *vs* bioexposed, C) Bionaïve *vs* bioexposed, in Crohn's disease, D) Bionaïve *vs* bioexposed, in ulcerative colitis.

plot (right panel) shows how the value assumed by each feature contributes to the final prediction. Previous exposure to anti-TNFα, older age and shorter disease duration were the most important drivers for the prediction of SFCR at 24 months (Fig. 4B and C); similarly, previous exposure to anti-TNFα, older age and female sex were the most important drivers for the prediction of SFCR at 12 months (Fig. 4D and E). In regard to safety, lower baseline CRP levels, fewer comorbidities and lower age were the most important features associated with no development of AEs (Fig. 4F and G).

4. Discussion

Our study confirms the real-life effectiveness and safety of vedolizumab in a large cohort of patients affected by IBD. About one-third of our patients achieved SFCR at 12 and 24 months, which is in line with the results from a recent meta-analysis of observational studies [16]. Of note, our cohort comprised the initial patients treated with vedolizumab: the majority had prior exposure to biologic therapies and a prolonged disease duration, both of which are typically linked to decreased treatment effectiveness in real-world clinical practice (at least for CD) [17–19]; furthermore, nearly one-fifth of our patients were elderly, a population previously noted—especially among those with ulcerative colitis[20]—to experience comparatively reduced efficacy with vedolizumab. We report an increase in the rates of SFCR, clinical and biochemical remission from 14 weeks to 6 months, suggesting that vedolizumab might have a slower onset of action in a subgroup of partial responders/non-responders; on the other hand, after 6 months, percentages for all explored outcomes remained stable, confirming the long-term durability of vedolizumab. Vedolizumab persistence was comparable between patients with CD and UC. Interestingly, previous exposure to anti-TNF-α agents was associated with reduced persistence only in CD patients, but not in those with UC. While the underlying reason remains unclear, the specific anti-TNF-α agent and its route of administration (infliximab *iv vs* adalimumab or golimumab *sc*) might have been relevant in that regard;

A)

Output		Precision	Recall	F1-score
SFCR at 24 months	Yes	0.76	0.97	0.86
	No	0.96	0.71	0.81
SFCR at 12 months	Yes	0.83	0.87	0.85
	No	0.85	0.81	0.83
Development of AEs	Yes	0.78	0.78	0.78
	No	0.78	0.78	0.78

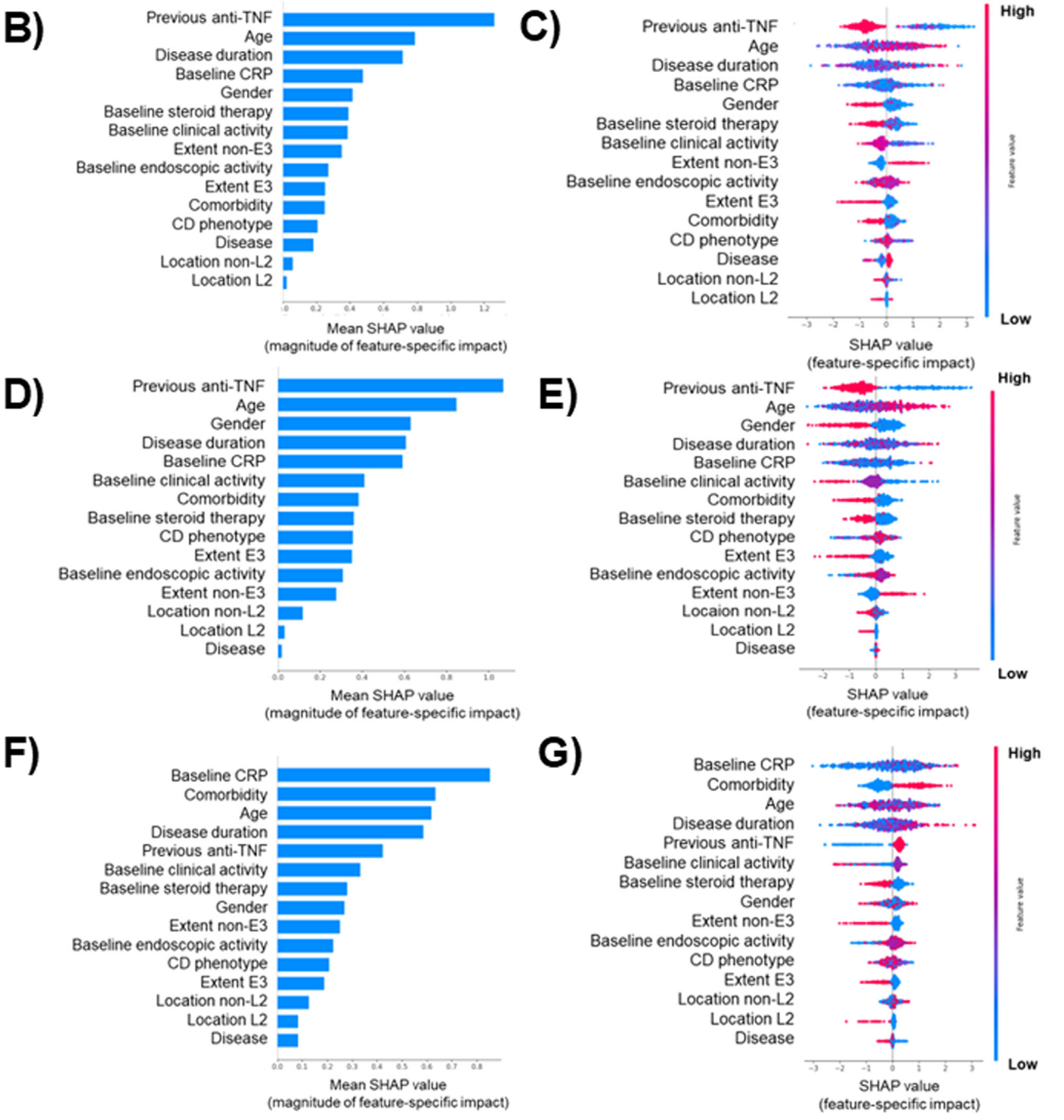

Fig. 4. Performance and results of XGBoost. A) Accuracy metrics for the performance of XGBoost for steroid-free clinical remission at 12 months and development of adverse events. B and C) Magnitude (left panel) and directionality (right panel) of feature-specific impact for prediction steroid-free clinical remission at 12 months. D and E) Magnitude (left panel) and directionality (right panel) of feature-specific impact for prediction development of adverse events.

Table 2
Adverse events reported during vedolizumab treatment.

Adverse event	Occurrence, n (%)	Patients, n
Infections, n (%)	120 (38.0)	31
Upper respiratory tract	32	29
Lower respiratory tract	29	29
Gastrointestinal tract	30	18
Skin and mucosa infections	18	7
Urinary tract	7	4
Other	4	
Cancer or dysplasia, n (%)	26 (8.1)	26
Arthralgia or arthritis, n (%)	47 (15.0)	46
Skin reaction, n (%)	23 (7.2)	22
Cholestasis and/or hepatitis, n (%)	11 (3.0)	10
Infusion reations, n (%)	17 (5.3)	17
Neurological symptoms and deseases, n (%)	15 (4.7)	13
Cardiological deseases, n (%)	4 (1.2)	4
Pancreatitis, n (%)	2 (0.6)	2
Others, n (%)	53 (16.6)	50

For each AE subtype, it is reported the number of AEs occurred (left column, expressed as percentage of the total number of AEs in the cohort) and the number of patients experiencing an AE (right column).

unfortunately, data on this variable were not available to explore a potential correlation.

In our machine learning model, bionaïve status was the main predictor of vedolizumab effectiveness, in line with previous findings [16,21–23]. A shorter disease duration was also associated with increased effectiveness: long-standing IBD is more likely to be complex and, therefore, resistant to medical therapies (at least, for CD) [24], and our finding might support the notion that early treatment might be associated with more favourable outcomes. We observed that female sex was associated with a reduced effectiveness of vedolizumab. The impact of gender on the treatment course of patients affected by IBD is an emerging concept observed for several drugs, possibly related to the effect of sexual hormones, gender-specific differences in drug volume distribution/clearance [25], and intestinal microbiota composition [26]. Recently, Macaluso et al. performed a comprehensive meta-analysis of observational studies on the effectiveness and safety of vedolizumab in patients with IBD. They report rates of SFCR during maintenance comparable to ours; interestingly, they observe a positive association between male sex and vedolizumab effectiveness in the CD cohort [16]. Similarly, Coletta et al. also reported that female patients treated with vedolizumab, regardless of IBD type, have lower rates of clinical remission [27]. Of note, in a 2015 study the volume of distribution and clearance of vedolizumab was observed to be lower in female patients, but with no clinical relevance [28]. Surprisingly, the machine learning model showed that older age is associated with a higher likelihood of achieving SFCR at 12 and 24 months, which is not in line with our previous findings [20]. However, in our previous study, elderly and non-elderly patients were propensity matched to account for potential selection biases: it is possible that the finding from our present study might reflect that, in the overall cohort, elderly patients tended to have a milder disease course and, hence, better outcomes; nevertheless, this represents one possible explanation, and reinforces the idea that the impact of age on treatment response needs to be further investigated.

Our data confirm the overall favourable safety profile of vedolizumab. Only 63 patients had to discontinue the treatment due to an AE. In our cohort, 90 patients had an history of previous cancer and 26 developed a new or recurrent dysplasia/cancer. This figure is in line with previous studies showing absence of increased risk of new or recurrent cancer in vedolizumab-treated patients, compared to anti TNF-α or conventional drugs [29–31]. Looking at intestinal cancers, we reported an incidence rate of 1.3 per 1000 patient-year. A large Scandinavian population-based cohort study reported an incidence rate of 0.82 cases of colorectal cancer (CRC) per 1000 patient-year in CD [32] and of 1.29 cases of CRC per 1000 patient-year in UC [33]. In our cohort, we also report a case of small bowel cancer: when considering only CRC, we had an incidence rate of 0.9 cases per 1000 patient-year, which is in line with previous findings. In our machine learning model, higher baseline CRP level was the main features associated with an increased risk of AEs, reflecting that a higher inflammatory burden tends to be associated with a reduced safety of treatments. The presence of comorbidities also contributed to the risk of AEs. Indeed, comorbidities contribute to determining a "frailty phenotype", which is associated with an increased vulnerability to adverse health outcomes, irrespective of patients' chronological age [34]. Finally, the machine learning model also showed that elderly patients appear to have an increased risk for AEs, which might be related to the fact that elderly patients are more likely to have comorbidities and to be frail.

Dulai et al. developed a scoring system for outcome prediction in vedolizumab-treated patients, by using logistic regression analysis on the data from the GEMINI program, and they validated it on patients from the VICOTRY cohort [22,23]. In those works, no previous anti-TNFα exposure was the most relevant predictor of SFCR at 12 months, which is in line with our finding; conversely, other variables deemed relevant in their models (*i.e.*, baseline endoscopic activity for UC, and absence of prior fistulizing disease and baseline CRP concentration for CD) did not share the same importance in ours. In their work, their model showed a slightly lower discriminative ability compared to ours (AUC-ROC for SFCR prediction: 0.66 in CD, 0.64 in UC), which might depend on the overall better performance of AI compared to traditional statistics in risk-stratification and outcome prediction [35,36]. A direct comparison between the two models would have been ideal; however, our study was initiated prior to the publication of the abovementioned manuscripts and baseline albumin levels (one of the five parameters used in their model) were not collected in our dataset: as a result, a formal comparison was not feasible in our cohort.

As we approach the 'era of precision medicine', machine learning will likely represent a non-renounceable tool. The application of XGBoost and SHAP ensures both a good performance and the explainability of the model. Three studies [37–39] applied AI to develop prediction models using data from the registration trials of biologic therapies. In regard to vedolizumab, Waljee et al. observed that a model constructed using clinical and laboratory data through week 6 of therapy could successfully predict steroid-free clinical (AUC-ROC of 0.75) and free endoscopic remission (AUC-ROC of 0.73) at 52 weeks. Despite their sound methodology, these studies risk suffering from a low external validity owing to the strict inclusion/exclusion criteria that patients need to meet for enrolment [40].

The main limitations of our study include: 1) the imbalance between bionaïve and bioexperinced patients, which might depend on the fact that our study enrolled the first patients who received vedolizumab in Italy; 2) the scarcity of endoscopic data, with a potential bias toward more severe cases being more likely to undergo endoscopy; 3) the lack of external validation for the machine learning model, which was both trained and validated on two internal subsets (*i.e.*, training and testing datasets) of our cohort; 4) the inability to perform AI analyses separately for CD and UC due to insufficient sample size, which would have compromised the performance of the – of note, SHAP analysis did not identify 'disease type' (*i.e.*, CD *vs.* UC) as an important contributor to any of the outcomes investigated, consistently ranking it among the least influential variables.

On the other hand, major strengths are represented by the overall large sample size, the inclusion of real-life patients and the good performance of the adopted machine learning model. More-

over, our study reports data that are relevant from several perspectives, including the best positioning of vedolizumab in the therapeutic algorithms, the importance of gender medicine and the role of AI in precision medicine.

In conclusion, vedolizumab is confirmed as an effective and safe therapy in patients with IBD. With machine learning, we identified the most important predictors of SFCR and development of AEs.

Authorship statement

Daniela Pugliese, Giuseppe Privitera and Alessandro Armuzzi are responsible for the original planning of the study, drafting of the article, statistical analysis, and interpretation of data. Nicola Cersullo and Harsh Bordekar performed machine learning-based data analysis and contributed to the manuscript drafting. All authors contributed to the design of the study, performed data collections and critical revision of article for important intellectual content. All authors approved the final version of the manuscript including authorship list. Alessandro Armuzzi oversaw the entire project and approved the final version of the manuscript.

Financial support

This work was supported by Takeda, which gave unconditional financial support.

Data transparency statement

Data can be made available upon request from third parties.

Conflict of interest

Daniela Pugliese received speaker fees and/or advisory board from AbbVie, Galapagos, MSD, Takeda and Janssen, Pfizer. Giuseppe Privitera received consultancy fees from Alphasigma and Janssen. Alessandro Armuzzi: consulting and/or advisory board fees from AbbVie, Amgen, Biogen, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Celltrion, Eli-Lilly, Ferring, Galapagos, Gilead, Janssen, MSD, Mylan, Pfizer, Protagonist Therapeutics, Roche, Samsung Bioepis, Sandoz, Takeda; lecture and/or speaker bureau fees from AbbVie, Amgen, Arena, Biogen, Bristol-Myers Squibb, Eli-Lilly, Ferring, Galapagos, Gilead, Janssen, MSD, Mitsubishi-Tanabe, Novartis, Pfizer, Roche, Sandoz, Samsung Bioepis, Takeda; and research grants from MSD, Pfizer, Takeda and Biogen. The remaining authors declare no competing interests.

Acknowledgments

The Authors wish to thank Ennio Sarli for performing statistical analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dld.2025.04.021.

References

[1] Burisch J, Jess T, Martinato M, Lakatos PL. The burden of inflammatory bowel disease in Europe. J Crohns Colitis 2013;7(4):322–37. doi:10.1016/J.CROHNS.2013.01.010.
[2] Kaplan GG, Windsor JW. The four epidemiological stages in the global evolution of inflammatory bowel disease. Nat RevGastroenterol Hepatol 2020;18(1):56–66. doi:10.1038/s41575-020-00360-x.
[3] Neurath MF. Targeting immune cell circuits and trafficking in inflammatory bowel disease. Nat Immunol 2019;20(8):970–9. doi:10.1038/s41590-019-0415-0.
[4] Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med 2013;369(8):699–710. doi:10.1056/NEJMoa1215734.
[5] Sandborn WJ, Feagan BG, Rutgeerts P, Hanauer S, Colombel JF, Sands BE, et al. Vedolizumab as induction and maintenance therapy for Crohn's disease. N Engl J Med 2013;369(8):711–21. doi:10.1056/NEJMoa1215739.
[6] Sands BE, Feagan BG, Rutgeerts P, Colombel JF, Sandborn WJ, Sy R, et al. Effects of vedolizumab induction therapy for patients with Crohn's disease in whom tumor necrosis factor antagonist treatment failed. Gastroenterology 2014;147(3):618–627.e3. doi:10.1053/j.gastro.2014.05.008.
[7] Privitera G., Pugliese D., Rapaccini G.L., Gasbarrini A., Armuzzi A., Guidi L. Predictors and early markers of response to biological therapies in inflammatory bowel diseases. 2021;10(4). https://pubmed.ncbi.nlm.nih.gov/33669579/. Accessed March 12, 2021
[8] Takayama T, Okamoto S, Hisamatsu T, Naganuma M, Matsuoka K, Mizuno S, et al. Computer-aided prediction of long-term prognosis of patients with ulcerative colitis after cytoapheresis therapy. PLoS One 2015;10(6):e0131197. doi:10.1371/JOURNAL.PONE.0131197.
[9] Weng F, Meng Y, Lu F, Wang Y, Wang W, Xu L, et al. Differentiation of intestinal tuberculosis and Crohn's disease through an explainable machine learning method. Sci Rep 2022;12(1):1–12. doi:10.1038/s41598-022-05571-7.
[10] Waljee AK, Joyce JC, Wang S, Saxena A, Hart M, Zhu J, et al. Algorithms outperform metabolite tests in predicting response of patients with inflammatory bowel disease to thiopurines. Clin Gastroenterol Hepatol 2010;8(2):143–50. doi:10.1016/j.cgh.2009.09.031.
[11] Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med 2001;23(1):89–109. doi:10.1016/S0933-3657(01)00077-X.
[12] Bhavsar KA, Abugabah A, Singla J, AlZubi AA, Bashir AK, Nikita. A comprehensive review on medical diagnosis using machine learning. Comput, Mater Cont 2021;67(2):1997–2014. doi:10.32604/cmc.2021.014943.
[13] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57. doi:10.1613/JAIR.953.
[14] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining; 2016. Published online. doi:10.1145/2939672.2939785.
[15] Feng DC, Wang WJ, Mangalathu S, Taciroglu E. Interpretable XGBoost-SHAP machine-learning model for shear strength prediction of squat RC walls. J Struct Eng 2021;147(11):04021173. doi:10.1061/(ASCE)ST.1943-541X.0003115.
[16] Macaluso FS, Ventimiglia M, Orlando A. Effectiveness and safety of Vedolizumab in inflammatory bowel disease: a comprehensive meta-analysis of observational studies. J Crohns Colitis 2023 Published online March 13,. doi:10.1093/ECCO-JCC/JJAD043.
[17] Narula N, Peerani F, Meserve J, Kochhar G, Chaudrey K, Hartke J, et al. Vedolizumab for ulcerative colitis: treatment outcomes from the VICTORY Consortium. Am J Gastroenterol 2018;113(9):1345–54. doi:10.1038/S41395-018-0162-0.
[18] Dulai PS, Singh S, Jiang X, Peerani F, Narula N, Chaudrey K, et al. The real-world effectiveness and safety of vedolizumab for moderate-severe Crohn's disease: results from the US VICTORY consortium. Am J Gastroenterol 2016;111(8):1147–55. doi:10.1038/ajg.2016.236.
[19] Faleck DM, Winters A, Chablaney S, Shashi P, Meserve J, Weiss A, et al. Shorter disease duration is associated with higher rates of response to Vedolizumab in patients with Crohn's disease but not ulcerative colitis. Clin Gastroenterol Hepatol 2019;17(12):2497–2505.e1. doi:10.1016/J.CGH.2018.12.040.
[20] Pugliese D, Privitera G, Crispino F, Mezzina N, Castiglione F, Fiorino G, et al. Effectiveness and safety of vedolizumab in a matched cohort of elderly and nonelderly patients with inflammatory bowel disease: the IG-IBD LIVE study. Aliment Pharmacol Ther 2022;56(1):95–109. doi:10.1111/APT.16923.
[21] Attauabi M, Madsen GR, Bendtsen F, Seidelin JB, Burisch J. Vedolizumab as the first line of biologic therapy for ulcerative colitis and Crohn's disease – a systematic review with meta-analysis. Digest Liver Dis 2022;54(9):1168–78. doi:10.1016/j.dld.2021.11.014.
[22] Dulai PS, Boland BS, Singh S, Chaudrey K, Koliani-Pace JL, Kochhar G, et al. Development and validation of a scoring system to predict outcomes of Vedolizumab treatment in patients with Crohn's disease. Gastroenterology 2018;155(3):687–695.e10. doi:10.1053/j.gastro.2018.05.039.
[23] Dulai PS, Singh S, Vande CN, Meserve J, Winters A, Chablaney S, et al. Development and validation of clinical scoring tool to predict outcomes of treatment with Vedolizumab in patients with ulcerative colitis. Clin Gastroenterol Hepatol 2020;18(13):2952–2961.e8. doi:10.1016/j.cgh.2020.02.010.
[24] Ben-Horin S, Zhao Y, Guo J, Mao R, Novack L, Sergienko R, et al. Efficacy of biological drugs in short-duration *versus* long-duration inflammatory bowel disease: a protocol for a systematic review and an individual-patient level meta-analysis of randomised controlled trials. BMJ Open 2019;9(1):1–6. doi:10.1136/bmjopen-2018-024222.
[25] Rosario M, Dirks NL, Milch C, Parikh A, Bargfrede M, Wyant T, et al. A review of the clinical pharmacokinetics, pharmacodynamics, and immunogenicity of Vedolizumab. Clin Pharmacokinet 2017;56(11):1287–301. doi:10.1007/S40262-017-0546-0/FIGURES/5.
[26] Goodman WA, Erkkila IP, Pizarro TT. Sex matters: impact on pathogenesis, presentation and treatment of inflammatory bowel disease. Nat Rev Gastroenterol Hepatol 2020;17(12):740–54. doi:10.1038/s41575-020-0354-0.
[27] Coletta M, Paroni M, Alvisi MF, De Luca M, Rulli E, Mazza S, et al. Immunological variables associated with clinical and endoscopic response to

Vedolizumab in patients with inflammatory bowel diseases. J Crohns Colitis 2020;14(9):1190–201. doi:10.1093/ECCO-JCC/JJAA035.
[28] Rosario M, Dirks NL, Gastonguay MR, Fasanmade AA, Wyant T, Parikh A, et al. Population pharmacokinetics-pharmacodynamics of vedolizumab in patients with ulcerative colitis and Crohn's disease. Aliment Pharmacol Ther 2015;42(2):188–202. doi:10.1111/apt.13243.
[29] Vedamurthy A, Gangasani N, Ananthakrishnan AN. Vedolizumab or tumor necrosis factor antagonist use and risk of new or recurrent cancer in patients with inflammatory bowel disease with prior malignancy: a retrospective cohort study. Clin Gastroenterol Hepatol 2022;20(1):88–95. doi:10.1016/j.cgh.2020.10.007.
[30] Poullenot F, Getaid AA, Getaid NM, Getaid, et al. Comparative risk of incident cancer in patients with inflammatory bowel disease with prior non-digestive malignancy according to Immunomodulator: a multicentre cohort study. J Crohns Colitis 2022;16(10):1523–30. doi:10.1093/ECCO-JCC/JJAC061.
[31] Hong SJ, Zenger C, Pecoriello J, Pang A, Vallely M, Hudesman DP, et al. Ustekinumab and Vedolizumab are not associated with subsequent cancer in IBD patients with prior malignancy. Inflamm Bowel Dis 2022;28(12):1826–32. doi:10.1093/IBD/IZAC035.
[32] Olén O, Erichsen R, Sachs MC, Pedersen L, Halfvarson J, Askling J, et al. Colorectal cancer in Crohn's disease: a Scandinavian population-based cohort study. Lancet Gastroenterol Hepatol 2020;5(5):475–84. doi:10.1016/S2468-1253(20)30005-4.
[33] Olén O, Erichsen R, Sachs MC, Pedersen L, Halfvarson J, Askling J, et al. Colorectal cancer in ulcerative colitis: a Scandinavian population-based cohort study. The Lancet 2020;395(10218):123–31. doi:10.1016/S0140-6736(19)32545-0.
[34] Fons A, Kalisvaart K, Maljaars J. Frailty and inflammatory bowel disease: a scoping review of current evidence. J Clin Med 2023;12(2):533. doi:10.3390/JCM12020533.
[35] Alkhamis MA, Al Jarallah M, Attur S, Zubaid M. Interpretable machine learning models for predicting in-hospital and 30 days adverse events in acute coronary syndrome patients in Kuwait. Sci Rep 2024;14(1):1–13. doi:10.1038/s41598-024-51604-8.
[36] Singal AG, Mukherjee A, Joseph Elmunzer B, Higgins PDR, Lok AS, Zhu J, et al. Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. Am J Gastroenterol 2013;108(11):1723. doi:10.1038/AJG.2013.332.
[37] Waljee AK, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, et al. Development and validation of machine learning models in prediction of remission in patients with moderate to severe Crohn disease. JAMA Netw Open 2019;2(5):e193721. doi:10.1001/JAMANETWORKOPEN.2019.3721.
[38] Waljee AK, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, et al. Predicting corticosteroid-free biologic remission with Vedolizumab in Crohn's Disease. Inflamm Bowel Dis 2018;24(6):1185–92. doi:10.1093/IBD/IZY031.
[39] Waljee AK, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, et al. Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis. Aliment Pharmacol Ther 2018;47(6):763–72. doi:10.1111/APT.14510.
[40] Pugliese D, Onali S, Privitera G, Armuzzi A, Papi C. Comparative effectiveness research: a roadmap to sail the seas of IBD therapies. J Clin Med 2022;11(22):6717. doi:10.3390/JCM11226717.