Development and validation of peripheral blood DNA methylation signatures to predict response to biological therapy in adults with Crohn's disease (EPIC-CD): an epigenome-wide association study

Vincent W Joustra, Andrew Y F Li Yim*, Peter Henneman*, Ishtu Hageman, Tristan de Waard, Evgeni Levin, Alexandra J Noble, Thomas P Chapman, Femke Mol, Sarah van Zon, Donghyeok Lee, Colleen G C McGregor, Melanie S Hulshoff, Jack J Satsangi†, Wouter J de Jonge†, Geert R D'Haens†, EPIC-CD Consortium‡*

Summary

Background Biological therapeutics are widely used in Crohn's disease, with evidence of efficacy from randomised trials and real-world experience. Primary non-response is a common, poorly understood problem. We aimed to assess blood methylation as a predictor of response to adalimumab, vedolizumab, or ustekinumab in patients with Crohn's disease.

Methods This epigenome-wide association study used data from two ongoing biobanks (one from the Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands [discovery cohort] and the other from the John Radcliffe Hospital, Oxford, UK [validation cohort]) that recruited patients between Oct 1, 2009, and June 17, 2022. Adult participants (age ≥18 years) with active symptomatic and endoscopic Crohn's disease who were scheduled to start adalimumab, vedolizumab, or ustekinumab treatment were included. Patients with ongoing malignancy or serious concomitant inflammatory diseases were excluded. Treatment response was assessed after a median of 28 weeks of treatment (IQR 18–36). Response was defined as a combination of endoscopic criteria (50% or more reduction in the Simple Endoscopic Score for Crohn's Disease) with either clinical or biochemical criteria (corticosteroid-free clinical response: ≥3 point decrease in Harvey–Bradshaw Index [HBI] score or remission [HBI ≤4] and no systemic steroids at follow up; biochemical response: C-reactive protein reduction ≥50% or ≤5 mg/L and faecal calprotectin reduction ≥50% or ≤250 µg/g) compared with baseline. Epigenome-wide DNA methylation and transcriptome-wide gene expression analyses were done on whole peripheral blood leukocyte samples that were collected before the start of treatment. To identify baseline DNA methylation markers associated with response or non-response to treatment, we performed supervised machine learning through stability selected gradient boosting. In a post-hoc analysis, we compared our DNA methylation-based prediction model with clinical decision support tools (CDSTs).

Findings We profiled the peripheral blood DNA methylome of 273 adults with Crohn's disease scheduled to start adalimumab, vedolizumab, or ustekinumab in the discovery (Amsterdam, n=183; 108 [59·0%] female and 75 [41·0%] male) and the validation cohort (Oxford, n=90; 46 [51·1%] female and 44 [48·9%] male). In the discovery cohort, we defined a panel of DNA methylation biomarkers that were associated with combined endoscopic and clinical or biochemical response to adalimumab (18 markers), vedolizumab (25 markers), or ustekinumab (68 markers), with an area under the curve (AUC) of 0·86 (95% CI 0·58–0·97) for adalimumab, 0·87 (0·67–0·98) for vedolizumab, and 0·89 (0·76–1·00) for ustekinumab. Validation in the Oxford cohort yielded an AUC of 0·25 (0·10–0·35) for adalimumab, 0·75 (0·65–0·85) for vedolizumab, and 0·75 (0·65–0·87) for ustekinumab. In comparison, implementing the CDSTs in the validation cohort yielded an AUC of 0·56 (0·44–0·68) for vedolizumab and 0·66 (0·54–0·77) for ustekinumab. Previous anti-TNF exposure was associated with a reduction in accuracy of the methylation models for vedolizumab (0·66 [0·55–0·73]) and ustekinumab (0·63 [0·52–0·70]) when analysed in the validation cohort.

Interpretation Our findings provide evidence for the potential use of DNA methylation as a modality for personalised medicine for Crohn's disease by predicting response to vedolizumab and ustekinumab. The models were more accurate in biologically naive patients and outperform available vedolizumab and ustekinumab CDSTs. We were unable to predict response to adalimumab. The vedolizumab and ustekinumab prediction models are currently being tested in a multicentre randomised clinical trial.

Funding The Leona M and Harry B Helmsley Charitable Trust.

Lancet Gastroenterol Hepatol 2025

Published Online July 1, 2025 https://doi.org/10.1016/S2468-1253(25)00102-5

See Online/Comment https://doi.org/10.1016/S2468-1253(25)00138-4

*These authors share first authorship

†These authors are shared senior authors

‡Members are listed in the appendix (p 1)

Department of Gastroenterology and Hepatology, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands (V W Joustra MD, I Hageman MD, S van Zon MD, M S Hulshoff PhD, Prof W J de Jonge PhD, Prof G D'Haens MD); **Tytgat Institute for Liver and Intestinal Research, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands** (A Y F Li Yim PhD, I Hageman MD, F Mol MSc, Prof W J de Jonge PhD); **Genome Diagnostics Laboratory, Department of Human Genetics, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands** (A Y F Li Yim PhD, P Henneman PhD); **Horaizon, Delft, Netherlands** (E Levin PhD, Tr de Waard MSc, D Lee BSc); **Translational Gastroenterology and Liver Unit, Nuffield Department of Medicine, University of Oxford and NIHR Oxford Biomedical Research Centre, Oxford, UK** (A J Noble PhD, T P Chapman DPhil, C G C McGregor DPhil, Prof J J Satsangi DPhil);

Department of Gastroenterology, St Richard's and Worthing Hospitals, University Hospitals Sussex NHS Foundation Trust, West Sussex, UK (T P Chapman DPhil); **Department of Surgery, University of Bonn, Bonn, Germany** (Prof W J de Jonge PhD)

Correspondence to: Prof Geert D'Haens, Department of Gastroenterology and Hepatology, Amsterdam University Medical Centre, University of Amsterdam, 1105 AZ Amsterdam, Netherlands
g.dhaens@amsterdamumc.nl

See Online for appendix

Research in context

Evidence before this study

The anti-TNF antibody adalimumab, the anti-α4β7 integrin antibody vedolizumab, and the anti-IL12/23p40 antibody ustekinumab are biologicals used to treat Crohn's disease. Despite the established efficacy of these biological treatments to induce corticosteroid-free clinical remission in patients with Crohn's disease, sustained endoscopic remission is observed in less than a third of patients after 1 year of treatment. To date, treatment selection has been based on a trial-and-error approach, with response probabilities estimated at 37% for adalimumab, 45% for vedolizumab, and 42% for ustekinumab. Most patients are therefore at risk of receiving ineffective therapy, necessitating drug switching. The development of strategies that allow selection of treatment based on the likelihood of response is an important unmet need. We searched PubMed using the terms "response prediction" and "prognosis" and "Crohn's disease" and "adalimumab" or "vedolizumab" or "ustekinumab" from database inception to June 25, 2024, with no restrictions for article type. Only English articles were considered. Of the 5038 results, 201 were randomised controlled trials (RCTs), none of which represented an RCT on response prediction to adalimumab, vedolizumab, or ustekinumab. Clinical decision support tools (CDSTs) are available for vedolizumab and ustekinumab based on clinical parameters; however, no RCTs have validated these tools. To our knowledge, there are, to date, no positive RCTs of biomarkers to use in response prediction for biologicals in Crohn's disease.

Added value of this study

This study shows that DNA methylation in peripheral blood taken before treatment can predict response to treatment with vedolizumab and ustekinumab in patients with Crohn's disease, yielding an area under the curve (AUC) of 0·75 for both in the external validation cohort and outperforming available CDSTs. For adalimumab, we were unable to accurately predict response, yielding an AUC of 0·25 in the external validation cohort. Calculating the potential impact of our prediction models on clinical practice showed a post-test probability of response of 65% for vedolizumab and 66% for ustekinumab, compared with pre-test probabilities of 45% and 42%, respectively, indicating an improvement of 20 percentage points for vedolizumab and 24 percentage points for ustekinumab over the current standard of care, which could substantially affect health-care costs and disease burden in these patients.

Implications of all the available evidence

Our results show that response to vedolizumab and ustekinumab can be predicted before starting therapy, suggesting that DNA methylation biomarkers could have a role in personalising treatment.

Introduction

Crohn's disease is an incurable, chronic, relapsing inflammatory bowel disease (IBD) caused by a complex interplay between the environment, gut microbiome, and a dysregulated immune system in genetically susceptible patients.[1] Accessibility to high-throughput omics technology enhanced the understanding of the underlying molecular pathogenesis of Crohn's disease, leading to the development of several monoclonal antibodies or biologics that target specific inflammatory pathways in an effort to suppress inflammation and to induce or maintain a state of clinical and endoscopic remission.[2] Currently, the repertoire of approved biologics in Crohn's disease includes anti-TNF antibodies (adalimumab and infliximab), the anti-α4β7 integrin antibody vedolizumab, and the anti-IL12/23p40 antibody ustekinumab; specific IL23p19 antibodies and JAK inhibitors have also been approved for clinical use.[3,4] Despite the established efficacy of these biological treatments to induce corticosteroid-free clinical remission in up to 65% of patients with Crohn's disease, sustained endoscopic remission is observed in no more than a third of patients after 1 year of treatment.[5,6] This finding creates a clinical challenge because therapeutic guidelines suggest the use of endoscopic remission as a target.

To date, treatment selection is based on a trial-and-error approach. Many patients are provided with an insufficiently effective treatment, increasing the chance of disease progression, which is associated with a higher risk of complications (stenosis, fistula, abscesses, and nutritional deficiencies) and surgery. The development of strategies that allow selection of treatment based on the likelihood of response is an important unmet need. Although various efforts using clinical,[7,8] transcriptomic,[9,10] proteomic,[11] or microbial[12] technologies have been investigated, only a clinical decision support tool (CDST) for vedolizumab has been approved as a medical device for patients with Crohn's disease in Europe,[7] with an equivalent tool being investigated for ustekinumab.[8]

DNA methylation is one of the most studied epigenetic features and is characterised by the covalent binding of methyl groups to nucleotides, most often a cytosine in a cytosine-phosphate-guanine (CpG) sequence in humans. DNA methylation is believed to play an essential role in the regulation of gene expression, thereby determining cellular phenotype and behaviour without altering the DNA sequence itself. Within the context of IBD, interest has grown in DNA methylation due to its potential capability of interfacing between the host and the environment, such as the microbiome.[13] Some studies showed differential DNA methylation profiles associated with the presence of Crohn's disease or specific Crohn's disease-phenotypes[14] in peripheral blood leukocytes,[15] circulating $CD8^+$ T cells,[16] or intestinal mucosa.[17] Most studies proposed a potential role of the

DNA methylome in diagnostics and prediction of treatment response.

Here we report the results of EPIC-CD, an epigenome-wide association study in which we aimed to identify and validate prognostic DNA methylation signatures in peripheral blood of adults with Crohn's disease that were associated with objective therapeutic response to adalimumab, vedolizumab, and ustekinumab.

Methods

Study design

EPIC-CD was an epigenome-wide association study conducted at the Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands (discovery cohort) and the John Radcliffe Hospital, Oxford, UK (validation cohort). The study was initiated in 2019 to identify biomarkers for adalimumab, vedolizumab, or ustekinumab response. In Amsterdam, patient material was collected from an ongoing biobank with enrolments for adalimumab dating back to 2009, with participants included between Oct 1, 2009, and Jan 1, 2022. Patients in the Oxford biobank (also ongoing) were recruited from March 1, 2019, to June 17, 2022.

The study was approved by the medical ethics committee of the Academic Medical Centre, Amsterdam (NL53989.018.15) and the National Health Service Research Ethics committee (21/PR/0010; protocol number 14833; and Integrated Research Application System project identification 266041). Written informed consent was obtained from all participants on entry into the biobanks before sampling. Patients in the UK were consented under the ethics of the Translational Gastrointestinal Unit Biobank IBD cohort (09/H1204/30) and the Gastro-Intestinal cohort (16/YH/0247 and 21/YH/0206).

Participants

We used data from adult patients (age ≥18 years) with Crohn's disease who presented with endoscopic disease activity (Simple Endoscopic Score for Crohn's Disease [SES-CD] ≥3) at ileo-colonoscopy and either clinical activity (Harvey–Bradshaw Index [HBI] ≥4), or biochemical activity (C-reactive protein [CRP] ≥5 mg/L or faecal calprotectin ≥250 µg/g) and were scheduled to start adalimumab, vedolizumab, or ustekinumab treatment. All patients had not previously received the biological therapy they were scheduled to start but could have received previous biologics. Patients self-reported their sex at birth (male or female). Concomitant use of immunomodulators and prednisolone taper scheme (40 mg starting dose with 5 mg per week tapering) at initiation of treatment was permitted. Patients with ongoing malignancy or serious concomitant inflammatory diseases that might impair the interpretability of the biomarker analysis, per investigator's interpretation, were excluded. Patients that developed anti-drug antibodies over the course of the treatment, presented without a measurable serum drug concentration, or stopped treatment due to adverse events (ie, side-effects, infections, or non-compliance) without objective response assessment were also excluded.

Patients in both biobanks were prospectively followed up through standard of care and were selected from the biobanks for this study if they met the strict criteria for either response or non-response. Patients without sufficient objective data to strictly categorise into responder or non-responder with a degree of certainty, or who were not able to provide blood samples for DNA methylation analysis at the required timepoint, were not selected.

Procedures

Patients were treated according to standard-of-care protocols. For adalimumab, patients received 160 mg subcutaneous injections at week 0, 80 mg at week 2, and 40 mg at week 4, followed by 40 mg every other week. For vedolizumab, patients received 300 mg infusions at weeks 0, 2, and 6 followed by infusions every 8 weeks. For ustekinumab, patients received a single intravenous infusion (6 mg/kg rounded to 260 mg, 390 mg, or 520 mg) at week 0 and subsequent 90 mg subcutaneous injections every 8 weeks. Interval intensification to weekly injections for adalimumab and interval intensification to infusions either every 6 weeks or every 4 weeks for both vedolizumab and ustekinumab (as well as an extra week 10 infusion for vedolizumab or extra intravenous boost infusion for ustekinumab) were allowed at the treating physicians' discretion. Endoscopic assessment of response was typically done between 26–52 weeks of treatment.

A full description of the DNA methylation and gene expression experimental procedures can be found in the appendix (pp 3–4). We analysed epigenome-wide DNA methylation and transcriptome-wide gene expression in whole peripheral blood leukocyte samples that had been collected before the start of treatment, either before the baseline endoscopy or the first infusion (timepoint 1) and after a median of 28 weeks into treatment (IQR 18–36; timepoint 2) at response assessment. At both timepoints, HBI, CRP, faecal calprotectin, and SES-CD were also determined. To limit the probability of pharmacokinetic failures of each treatment, only data from patients with measurable serum drug concentrations without anti-drug antibodies at response assessment were used for methylation analyses.

The quality of the collected data and study procedures were assessed by an independent monitor.

Outcomes

At response assessment (timepoint 2), patients were classified as responders to treatment if they had an endoscopic response (≥50% reduction in SES-CD score) together with either a clinical response (≥3 point decrease

in HBI score or HBI ≤4, both with no systemic steroids) or biochemical response (CRP reduction ≥50% or ≤5 mg/L and faecal calprotectin reduction ≥50% or ≤250 µg/g) compared with baseline.

Modified criteria for response were used for patients whose endoscopic assessment during follow-up was not possible in the COVID-19 pandemic. Modified response was defined as a combination of corticosteroid-free clinical remission (HBI ≤4) and biochemical remission (CRP ≤5 mg/L or faecal calprotectin ≤250 µg/g or both) between week 26 and week 52 without treatment change up until week 52.

DNA methylation analysis

Peripheral blood genomic DNA was extracted and bisulfite converted, then the DNA methylome was quantified using the HumanMethylation EPIC BeadChip array (Illumina, San Diego, CA, USA) at Core Facility Genomics, Amsterdam University Medical Centre, Amsterdam for the discovery cohort and at UCL Genomics, University College London, London, UK for the validation cohort. The raw methylation data was imported into the R statistical environment (v4.3.1) using the Bioconductor minfi package (v1.44). Raw signals were normalised using functional normalisation. Probe and sample level quality control was performed, resulting in the removal of two patient samples from the vedolizumab discovery cohort and two patient samples from the ustekinumab discovery cohort. Technical artifacts because of batch, plate, and plate position were removed using ComBat as implemented in the sva package (v3.50.0). Probes hybridising to allosomes were removed to identify sex-independent differences. Moreover, probes hybridising to known and tentative genetic variants were removed to identify true methylation signals. To train the prediction models, 808 329 CpGs were used for adalimumab, 806 308 CpGs were used for vedolizumab, and 808 815 CpGs were used for ustekinumab. To identify baseline DNA methylation markers associated with response or non-response to treatment, we implemented a supervised machine learning approach that is detailed later. For the validation of the prediction models, raw methylation data from the validation cohort was preprocessed together with the discovery cohort using functional normalisation and ComBat to mitigate batch effects introduced by the different experimental setup. From the combined dataset, the predictor CpGs were extracted, and the prediction model was recalibrated against the discovery dataset then predictions were made on the validation dataset. A full description of the DNA methylation analysis can be found in the appendix (pp 3–4).

For more on the **Bioconductor minfi package** see https://www.bioconductor.org/packages/release/bioc/html/minfi.html

For more on the **sva package** see https://bioconductor.org/packages/release/bioc/html/sva.html

Machine learning analysis

The machine learning modelling was divided into two steps: feature selection and validation (appendix p 5). We performed feature selection on the discovery cohort and validation was performed on the validation cohort. Baseline epigenetic markers were associated with response or non-response to treatment using stability selected gradient boosting combined with covered information disentanglement to capture linear, non-linear, and interaction effects. Gradient boosting is an algorithm for supervised learning, which operates through stepwise improvement of weak learners. Covered information disentanglement, in conjunction with stability selected gradient boosting, represents an approach for feature selection that assigns permutation-based feature importance that, unlike other methods for assigning feature importance, is unbiased by multicollinearity. We used stability selection to identify reliable biomarkers by randomly splitting the discovery data into an 80% training data and a 20% test data using a stratified shuffle split and repeating this process 100 times to mitigate overfitting. During each split, we computed the covered information disentanglement for each CpG by randomly permuting it 100 times and calculating the mean feature importance and assessing the average effect of permutation on the model's performance (ie, predicted *vs* true outcome). After all 100 iterations, the mean feature importance per iteration was compared against a randomly generated variable that was included throughout the entire modelling process, whereby CpGs with an aggregated feature importance ranked above the random variable were termed predictor CpGs. Having determined the predictor CpGs, we subsequently validated the predictive performance internally and externally. We performed internal validation by extracting the predictor CpGs, training an ensemble of 100 gradient boost models on the 80% discovery training data, and using all models to predict against the withheld 20% discovery test data. We also performed external validation. First, as noted earlier, the discovery and validation cohorts were merged and normalised to prevent in-silico batch effects, and predictor CpGs were extracted. An ensemble of 100 models were trained on the discovery cohort samples only, and these models were subsequently used to predict against the external validation cohort. The resultant models were used to predict the response in the external validation cohort. In both internal and external validation, the output of each model yielded a prediction score on a scale of 0 to 1 per sample, as returned by the XGBoost classifier's default prediction function. The prediction scores were aggregated by calculating the mean prediction score per sample, representing the final, ensembled output of the model. This final prediction score was used to calculate the receiver operator characteristic together with bootstrapped 95% CIs to assess performance. The resultant prediction scores were subsequently converted into classes by freezing the model at the Youden index, thereby balancing the true positive rate relative to the false positive rate. Resultant performances were visualised using receiver operator characteristic curves.

Results for the vedolizumab and ustekinumab cohort were visualised as a whole and separately for patients that were previously treated with anti-TNF medication. Further details of the machine learning processes can be found in the appendix (pp 3–4).

Clinical decision support tool

In a post-hoc analysis we compared the DNA methylation-based prediction model with previously published CDST scores for vedolizumab[7] and ustekinumab.[8] The vedolizumab CDST score was calculated using the following five variables: no previous anti-TNF exposure (+3 points), no previous bowel surgery (+2 points), no previous fistulising disease at baseline (+2 points), baseline albumin level (+0·4 points for every g/L), and baseline CRP concentration (≥3 mg/L to ≤10 mg/L [–0·5 points], and >10 mg/L [–3 points]). Patients with a vedolizumab CDST score of more than 19 were classified as high probability of response to vedolizumab. The ustekinumab CDST was calculated using the following five variables: no previous anti-TNF exposure (+2 points), no previous bowel surgery (+2 points), no active fistulising disease at baseline (+1 point), no current or previous smoking history (+1 point), and baseline albumin concentration (≤2·5 g/dL [–3 points], >2·5 g/dL to ≤3·2 g/dL [–1 point], >3·2 g/dL to ≤3·9 g/dL [0 point], >3·9 g/dL to ≤4·3 g/dL [+1 point], and >4·3 g/dL [+3 points]). Patients with an ustekinumab CDST score of more than 4 were classified as high probability of response to ustekinumab. Patients that did not have all measurements within 3 months before the start of the study were excluded from these analyses.

Statistical analysis

We based our sample size on an initial pilot experiment and on calculations performed by Tsai and Bell,[18] in which we used a nominal two-tailed p value threshold of 0·05. The pilot experiment, which included a subset of vedolizumab-treated patients (seven responders and five non-responders), indicated on average that the most differentially methylated CpGs presented a 10% mean difference in percentage methylation when comparing responders with non-responders. Assuming an approximately equal number of responders and non-responders and a mean difference in percentage methylation across all CpGs assessed of at least 10% at a nominal two-tailed p value threshold of 0·05, a statistical power of at least 80% would be achieved if we included 40 patients (20 responders and 20 non-responders) per drug. We conducted a secondary power calculation by simulating a set of 865 859 predictive and background features, mimicking the total number of CpGs located on the Illumina HumanMethylation EPIC BeadChip array. Our parameter of interest was the sample size, which we varied by increasing the sample size from 30 to 70 in steps of ten per iteration with each iteration being repeated 20 times. Features were modelled to follow a beta distribution, where each simulated sample was obtained from a uniform distribution bounded between 0 and 1. A set of 50 predictive features were defined by separating the responders and non-responders with a 0·01–0·20 mean difference sampled from a uniform distribution, representing mean difference in methylation. Background features were not subjected to the simulated mean difference in methylation and reflect random variation. Using this simulated set, we used our supervised machine learning approach described earlier to classify responders from non-responders and calculate the AUC. We found that at a minimum sample size of 40, we would achieve an expected AUC of 0·80, matching the observations by Tsai and Bell.[18] To further mitigate the possibility of being underpowered, we aimed to collect at least 60 patients per drug for the discovery cohort, which was supplemented by at least 20 patients for the validation cohort.

We summarised baseline characteristics of all included patients using descriptive statistics. Categorical variables are presented as percentages, and continuous variables as either mean (SD) for normal distributions or median (IQR) for skewed distributions. We assessed differences in distribution between responders, non-responders, and the different cohorts using either a χ^2 test or Fisher's exact test for categorical variables, the latter being used for comparisons where 20% or more cells had expected counts less than five. For continuous variables, we used Mann–Whitney *U* tests. We used two-tailed probabilities where p≤0·05 was considered statistically significant. We analysed clinical data using SPSS (v26).

We performed differential methylation analyses on the DNA methylation data using R using limma (v3.46) and eBayes, which yielded residuals and p values per CpG measured. Specifically, we conducted the following comparisons: (1) a case-control analysis comparing responders versus non-responders pretreatment, analysing without covariates and subsequently with the covariates prior anti-TNF use (yes *vs* no), sex (male *vs* female), age (continuous, years), smoking status (active, former, never, or unknown), and estimated cellular composition (continuous, CD8 T cell, CD4 T cell, NK cell, B cell, Mono, and Neu), and (2) a longitudinal analysis of the predictor CpGs by comparing treated samples versus pretreatment samples, using a paired analysis by including patient as random effects. p values are nominal in nature and were not adjusted for multiplicity. Residuals for the predictor CpGs from the differential methylation analyses are in the appendix (p 6). We conducted analyses to assess the correlation between pretreatment and on-treatment differences between responders and non-responders and DNA methylation and gene expression correlations analyses using the cor. test function in R set to calculate the Pearson correlation coefficient. We performed classification analyses using gradient boosting. Full details of the machine learning models are in the appendix (pp 3–4).

For more on **limma and eBayes** see https://bioconductor.org/packages/release/bioc/html/limma.html

To estimate the probability of a patient responding to either vedolizumab or ustekinumab after being predicted to be a responder, we calculated the post-test probability using the the pre-test probability of responses for vedolizumab of 45%[5] and ustekinumab of 42%,[6] and the likelihood ratios observed in this study. Sensitivity and specificity were calculated by determining the number of true positives, true negatives, false positives, and false negatives when predicting response in the validation cohort. These analyses were also done for patients who had previously received anti-TNF medication.

For more on **DESeq2** see https://bioconductor.org/packages/release/bioc/html/DESeq2.html

For our transcriptome analyses, differential expression analysis was done within R using the Bioconductor package DESeq2 (v1.38.3). We specifically focused on genes associated with the predictor CpG loci based on the latter's location in either promoter or enhancer regions. A full description of the preprocessing of the transcriptomic data analysis is in appendix (p 4).

	Discovery cohort (n=183)	Validation cohort (n=90)
Sex		
Female	108 (59·0%)	46 (51·1%)
Male	75 (41·0%)	44 (48·9%)
Age, years	35 (26–52)	40 (27–53)
Disease duration, years	11 (5–21)	16 (3–27)
Ethnicity		
White European	140 (76·5%)	79 (87·8%)
Non-White European	43 (23·5%)	11 (12·2%)
C-reactive protein, mg/L	6·3 (2·1–14·7)	7·2 (2·3–22·2)
Faecal calprotectin, µg/g	828 (267–1800)	209 (100–800)
Total baseline HBI	6 (4–10)	6 (3–9)
Total baseline SES-CD	8 (6–13)	11 (6–16)
Endoscopic evaluation at follow-up	172 (94·0%)	28 (31·1%)
Disease location		
Ileal disease (L1)	53 (29·0%)	25 (27·8%)
Colonic disease (L2)	39 (21·3%)	27 (30·0%)
Ileocolonic disease (L3)	91 (49·7%)	36 (40·0%)
Upper gastrointestinal involvement (L4)	3 (1·6%)	1 (1·1%)
Disease behaviour		
Non-stricturing non-penetrating (B1)	71 (38·8%)	68 (75·6%)
Stricturing (B2)	64 (35·0%)	19 (21·2%)
Penetrating (B3)	48 (26·2%)	3 (3·3%)
Perianal disease (p)	58 (31·7%)	25 (27·8%)
Previous inflamatory bowel disease-related surgery	105 (57·4%)	33 (36·7%)
Concomitant medication		
Immunomodulators*	18 (9·8%)	14 (15·6%)
Prednisone taper scheme	15 (8·2%)	1 (1·1%)
Previous treatment exposure		
Immunomodulators*	158 (86·3%)	64 (71·1%)
Anti-TNFs†	139 (76·0%)	36 (40%)
Vedolizumab	33 (18·0%)	5 (5·6%)
Ustekinumab	14 (7·7%)	7 (7·8%)
Active smoking	33 (18·0%)	14 (15·6%)

Data are n (%) or median (IQR). HBI=Harvey–Bradshaw Index. SES-CD=Simple Endoscopic Score for Crohn's Disease. *Azathioprine, mercaptopurine, thioguanine, or methotrexate. †Infliximab, adalimumab, or golimumab. Golimumab was provided to a single patient who was initially diagnosed with ulcerative colitis but was later reclassified as having Crohn's disease due to inflammation in the terminal ileum, confirmed by the presence of granulomas measured through histology.

Table 1: **Baseline characteristics for the discovery and validation cohorts**

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The discovery cohort consisted of 183 adults (108 [59·0%] female and 75 [41·0%] male) starting adalimumab (n=57), vedolizumab (n=64), or ustekinumab (n=62) at the IBD Centre of Amsterdam University Medical Centre, Amsterdam, Netherlands. At baseline, all patients had documented active clinical (median HBI 6 [IQR 4–10]), biochemical (median CRP 6·3 mg/L [2·1–14·7] and median faecal calprotectin 828 µg/g [267–1800]), and endoscopic (median SES-CD 8 [6–13]) disease activity (table 1). The external validation cohort consisted of 90 adults (46 [51·1%] female and 44 [48·9%] male) with Crohn's disease starting adalimumab (n=32), vedolizumab (n=25), or ustekinumab (n=33) at the John Radcliffe Hospital, Oxford, UK. A detailed overview of the clinical characteristics across the different cohorts and treatments is in table 1 and the appendix (pp 2–3, 25–26).

The response-predicting models based on baseline epigenetic markers yielded an AUC of 0·86 (95% CI 0·58–0·97) for adalimumab, 0·87 (0·67–0·98) for vedolizumab, and 0·89 (0·76–1·00) for ustekinumab when testing in the discovery cohort (figure 1A). The models comprised 18 differentially methylated CpGs for adalimumab, 25 differentially methylated CpGs for vedolizumab, and 68 differentially methylated CpGs for ustekinumab (figure 1B; appendix pp 7–19, 27–32). Validating our models in the external validation cohort yielded an AUC of 0·25 (95% CI 0·10–0·35) for adalimumab, 0·75 (0·65–0·85) for vedolizumab, and 0·75 (0·65–0·87) for ustekinumab (figure 1A). Given the inability to predict the adalimumab response in the validation cohort, we continued our subsequent analyses with the vedolizumab and ustekinumab models only.

We compared our DNA methylation-based prediction model with the CDSTs for vedolizumab[7] and ustekinumab.[8] For vedolizumab, we obtained all CDST measurements for 40 patients from the discovery cohort and 15 patients from the validation cohort. For ustekinumab, we obtained all CDST measurements for

57 patients in the discovery cohort and 26 patients from the validation cohort (three patients had missing data). Predicting high probability of vedolizumab response using the CDST yielded an AUC of 0·56 (95% CI 0·42–0·70) in the discovery cohort and 0·56 (0·44–0·68) in the validation cohort (appendix p 20). By contrast,

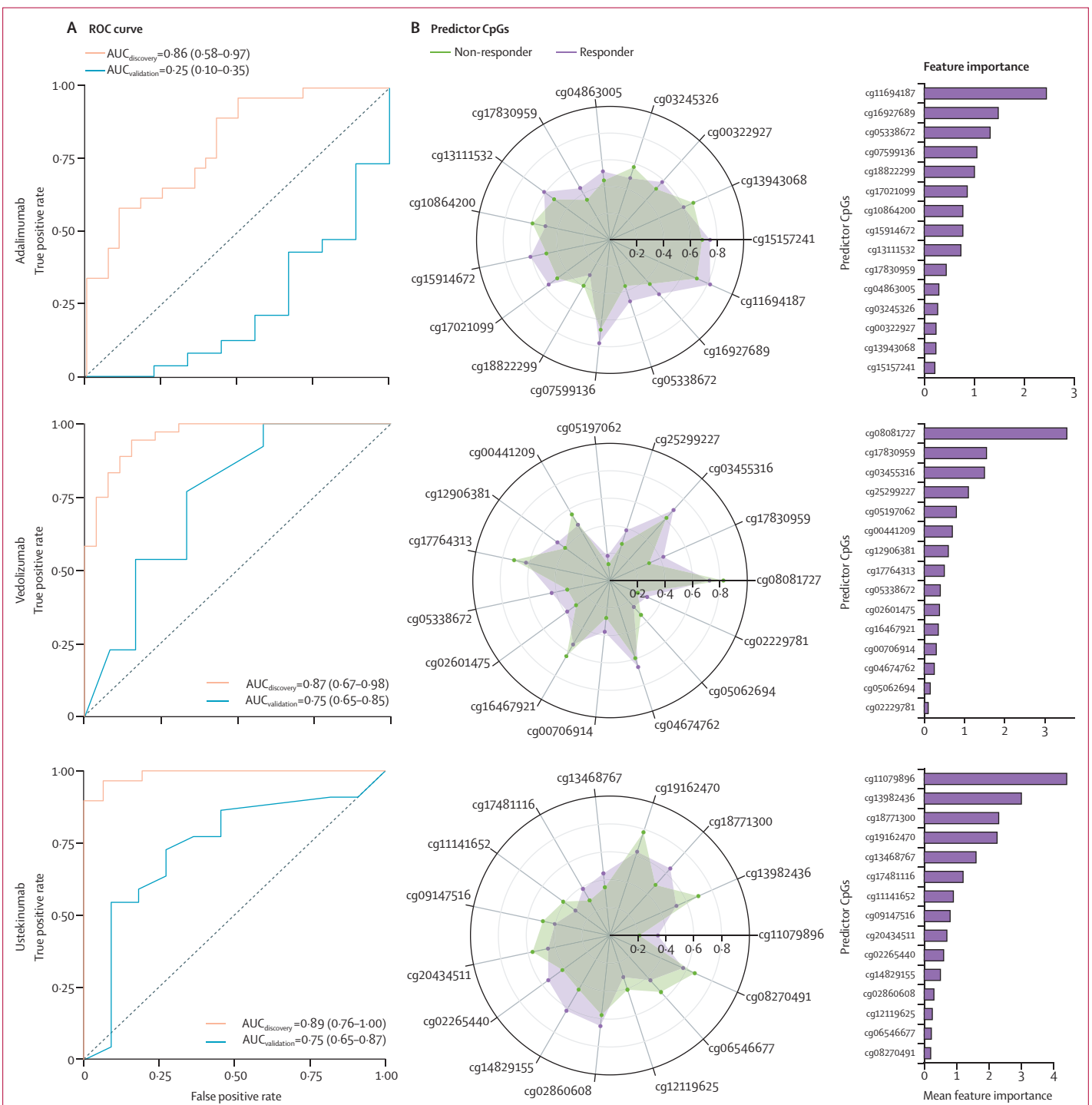

Figure 1: **Predictive model using stability selected gradient boosting for response to therapy**
(A) ROC plots showing the mean AUC performance of the models in the discovery (n=183) and validation (n=90) cohorts. (B) Left: radar plots presenting the standardised difference in methylation between responders (purple) and non-responders (green) for the top 15 predictor CpGs. Right: aggregated feature importance of the top 15 predictor CpGs. AUC=area under the curve. CpG=cytosine-phosphate-guanine. ROC=receiver operator characteristic.

	Adalimumab discovery cohort (n=57)	Adalimumab validation cohort (n=32)	Vedolizumab discovery cohort (n=62)*	Vedolizumab validation cohort (n=25)	Ustekinumab discovery cohort (n=60)†	Ustekinumab validation cohort (n=33)
True positive‡	17 (29·8%)	10 (31·3%)	33 (53·2%)	10 (40·0%)	27 (45·0%)	16 (48·5%)
True negative‡	25 (43·9%)	2 (6·3%)	22 (35·5%)	8 (32·0%)	29 (48·3%)	8 (24·2%)
False positive‡	3 (5·3%)	7 (21·9%)	4 (6·5%)	4 (16·0%)	2 (3·3%)	3 (9·1%)
False negative‡	12 (21·1%)	13 (40·6%)	3 (4·8%)	3 (12·0%)	2 (3·3%)	6 (18·2%)
AUC	0·86 (0·58–0·97)	0·25 (0·10–0·35)	0·87 (0·69–0·98)	0·75 (0·65–0·85)	0·89 (0·76–1·00)	0·75 (0·65–0·87)
Sensitivity	0·58 (0·50–1·00)	0·44 (0·17–0·58)	0·92 (0·56–1·00)	0·77 (0·62–0·85)	0·93 (0·58–1·00)	0·73 (0·61–0·91)
Specificity	0·89 (0·50–1·00)	0·22 (0·00–0·56)	0·85 (0·60–1·00)	0·67 (0·58–0·92)	0·94 (0·67–0·94)	0·73 (0·55–0·82)
Precision	0·85 (0·66–1·00)	0·59 (0·43–0·67)	0·89 (0·80–1·00)	0·71 (0·69–0·83)	0·93 (0·75–1·00)	0·84 (0·79–0·90)
F1-score	0·69 (0·60–0·95)	0·50 (0·25–0·60)	0·90 (0·69–0·96)	0·74 (0·67–0·82)	0·93 (0·72–1·00)	0·78 (0·72–0·86)

Data are n/N (%) or estimate (95% CI). AUC=area under the receiver operator characteristic curve. *62 patients in the vedolizumab discovery cohort were included in the analysis because two patients did not pass the methylation quality control. †60 patients in the ustekinumab discovery cohort were included in the analysis because two patients did not pass the methylation quality control. ‡Percentages are calculated relative to all samples in the cohort.

Table 2: **Predictive performance metrics in the discovery (n=183) and validation (n=90) cohorts**

predicting high probability of ustekinumab response using the CDST yielded an AUC of 0·53 (0·37–0·57) in the discovery cohort and 0·66 (0·54–0·77) in the validation cohort (appendix p 20).

Because a substantially lower proportion of participants in Oxford underwent endoscopy at follow-up due to COVID-19-related restrictions, we investigated whether discrepancies in performance between the discovery and validation cohorts were partly due to differences in the means of response assessment. This stratification indicated that accuracy was optimised by using the strictly defined combination of clinical and endoscopic endpoints for both vedolizumab (AUC_{strict} 0·83 [95% CI 0·68–0·93] *vs* $AUC_{modified}$ 0·66 [0·60–0·70]) and ustekinumab (AUC_{strict} 0·83 [0·60–1·00] *vs* $AUC_{modified}$ 0·72 [0·64–0·76]; appendix p 21).

Additionally, we hypothesised that the performance of our models might be affected by previous exposure to anti-TNF medication. We observed significantly better performance of our models among anti-TNF naive compared with anti-TNF exposed patients for both vedolizumab ($AUC_{non\text{-}exposed}$ 0·85 [95% CI 0·80–0·90] *vs* $AUC_{exposed}$ 0·66 [0·55–0·73]; p<0·0001) and ustekinumab ($AUC_{non\text{-}exposed}$ 0·97 [0·78–1·00] *vs* $AUC_{exposed}$ 0·63 [0·52–0·70]; p<0·0001; appendix p 21).

In the validation cohort, the models had a sensitivity of 0·77 (95% CI 0·62–0·85) and a specificity of 0·67 (95% CI 0·58–0·92) for vedolizumab and a sensitivity of 0·73 (0·61–0·91) and specificity of 0·73 (0·55–0·82) for ustekinumab (table 2). Based on a pre-test probability of response of 45%, the calculated sensitivity of 0·77 and specificity of 0·67, the likelihood ratio of response was 2·31. The post-test probability of response was therefore 65% for vedolizumab, indicating a 20 percentage point increase compared with current clinical practice. Similarly, for ustekinumab, with a pre-test probability of response of 42%, sensitivity and specificity of 0·73, the likelihood ratio was 2·67 and the post-test probability of response was 66%, indicating a 24 percentage point increase compared with current clinical practice. Because previous anti-TNF exposure affected the predictive performance of our model, we interrogated the effect on the post-test probability. We identified 49 of the 61 anti-TNF exposed patients in our vedolizumab cohort that responded to treatment, suggesting a pre-test probability of response of 51% and a post-test probability of response of 59%, suggesting an 8 percentage point increase. Similarly, we identified 37 of the 82 anti-TNF exposed patients in our ustekinumab cohort that responded to treatment, suggesting a pre-test probability of response of 45% and a post-test probability of response of 55%, suggesting a 10 percentage point increase.

Because baseline samples were acquired before treatment, we sought to understand whether exposure to either biological treatment affected the methylation status of the predictor CpGs by comparing samples obtained at the time of response assessment (timepoint 2) with samples obtained before treatment (timepoint 1) through multiple linear regression analysis. Interrogating the residuals indicated a largely normal distribution for most predictor CpGs for both vedolizumab and ustekinumab (appendix p 6). We identified one vedolizumab predictor CpG (cg09659072) that presented a statistically significant increase of 4·1% (95% CI –0·3 to 8·5; p=0·021) between pretreatment and response assessment (figure 2A; appendix pp 9–11). By contrast, ustekinumab presented no statistically significant differences over the course of treatment, with the most substantial difference being an increase of 3·5% (–1·3 to 8·4) at response assessment over pretreated for cg17037048 (p=0·074; figure 2A; appendix pp 12–19). Comparing the differences over time suggested that the mean difference between responders and non-responders was similar both pretreatment and at response assessment (figure 2B). Furthermore, a two-way, mixed, consistency intraclass correlation

analysis indicated highly stable DNA methylation over time with 24 of 25 vedolizumab and 62 of 68 ustekinumab predictor CpGs presenting intraclass correlation values of 0·75 or more (figure 2C; appendix pp 9–19). This observation was corroborated by interrogating our previous longitudinal consistency analysis of peripheral blood DNA methylation from 46 adults with IBD collected at two timepoints with a median of 7 years (range 2–9) in between.[20] Here, we observed that the majority (16 of 25 vedolizumab and 52 of 68 ustekinumab) of the predictor CpGs presented good (intraclass correlation 0·75 to <0·9) to excellent (intraclass correlation ≥0·9) stability[21] over a median span of 7 years (figure 2C).

Through multiple linear regression analyses we observed that 22 (88%) of the 25 vedolizumab response-associated CpGs and 38 (55%) of the 68 ustekinumab predictor CpGs presented statistically significant differences when comparing responders against non-responders. The smallest significant response-associated effect sizes were –6·7% (95% CI –12·8 to –0·51) for vedolizumab and –5·4% (–11·2 to 0·41) for ustekinumab (figure 3A; appendix pp 9–19). The large discrepancy between the number of ustekinumab predictor CpGs that differed significantly between responders and non-responders in the linear regression analyses and the ustekinumab response-associated predictor CpGs indicates that a more complex non-linear relationship exists among the response-associated predictor CpGs and underscores that statistical p values might not equal biological relevance, functional relevance, or importance at an individual level.

Because it has been established that the peripheral blood DNA methylome is associated with baseline steroid medication, previous anti-TNF, sex, age, smoking status, and underlying cellular composition,[21–24] we included these variables as covariates in linear regression analyses. We observed that 11 (50%) of 22 markers remained significantly associated with response for vedolizumab and 28 (74%) of 38 markers remained significantly associated with response for ustekinumab (figure 3A). In terms of effect size, the mean percentage methylation difference between responders and non-responders of the predictor CpGs on average decreased by 16% for vedolizumab and increased by 6% for ustekinumab (figure 3B). To understand whether the covariates can predict response, we constructed a prediction model solely based on the covariates using the discovery cohort and tested this on the validation cohort. The covariate model yielded an AUC of 0·57 (0·42–0·70) for vedolizumab and 0·64 (0·56–0·71) for ustekinumab in the validation cohort (figure 3C).

We next investigated whether the predictor CpGs were significantly associated with severity of systemic and intestinal inflammation at baseline measured using CRP and faecal calprotectin, as previously reported.[25] For vedolizumab, five predictor CpGs significantly associated with CRP whereas only one CpG was associated with faecal calprotectin (appendix p 22). For ustekinumab, we

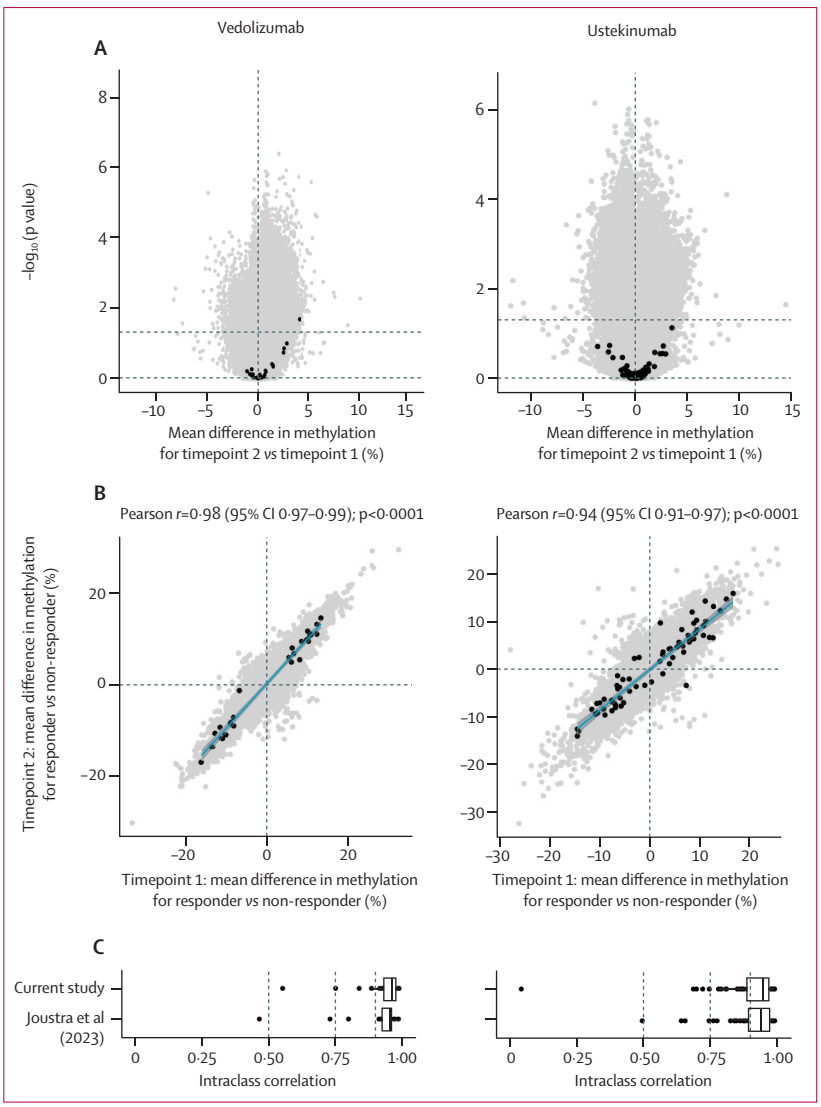

***Figure 2*: Longitudinal stability analyses**

(A) Volcano plot representing the differential methylation analyses when comparing treatment (timepoint 2) with pretreatment (timepoint 1) whereby grey dots represent CpG loci located on the Illumina HumanMethylation EPIC BeadChip array and black dots represent response-associated predictor CpGs. The x-axis represents mean difference in percentage methylation, the y-axis represents statistical significance as calculated using limma, and the dashed horizontal line is p=0·05. (B) Scatterplot showing the correlation of differential DNA methylation between responders and non-responders pretreatment (timepoint 1) and treatment (timepoint 2). Grey dots represent CpG loci located on the Illumina HumanMethylation EPIC BeadChip array and black dots represent response-associated predictor CpGs. Pearson correlation coefficients, 95% CIs, and the two-tailed p values were calculated using the cor.test function in R. The blue line represents the (linear) regression line. (C) Boxplot of the two-way consistency of the predictor CpGs calculated when comparing pretreatment and treatment as well as the intraclass correlation coefficients of the predictor CpGs obtained from a previous study on long-term stability of DNA methylation in patients with inflammatory bowel disease.[19] The vertical dashed grey lines represent classification boundaries introduced by Koo and Li,[20] with intervals representing poor (intraclass correlation coefficient <0·5), moderate (intraclass correlation coefficient 0·5 to <0·75), good (intraclass correlation coefficient 0·75 to <0·9), and excellent (intraclass correlation coefficient ≥0·9). Outlier dots are data points that fall beyond the whiskers of the boxplot (1·5 × IQR). The midline of the boxplot represents the median, the upper bound represents the third quartile, and the lower bound represents the first quartile. CpG=cytosine-phosphate-guanine.

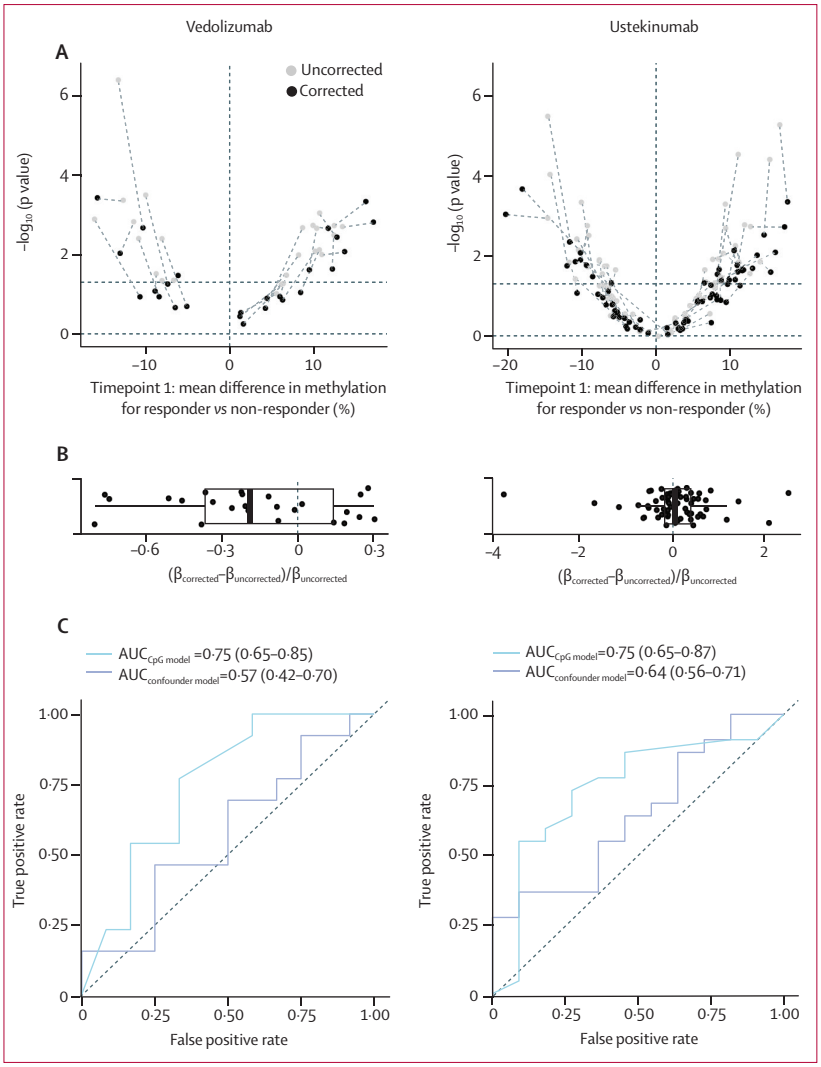

Figure 3: **Analyses of potential covariates**
(A) Volcano plot representing the change in response-associated differential methylation when correcting for the covariates of baseline steroid medication, previous anti-TNF usage, age, sex, and estimated cellular composition (black) or without any correction for covariates (grey). The x-axis represents mean difference in percentage methylation, the y-axis represents statistical significance as calculated using limma, and the dashed horizontal line represents a threshold set at p=0·05. (B) Boxplot of the change in effect size after correcting for the covariates as calculated by ($\beta_{corrected} - \beta_{uncorrected}$) divided by $\beta_{uncorrected}$. The midline of the boxplot represents the median, the upper bound represents the third quartile, and the lower bound represents the first quartile. (C) Receiver operator characteristic curve comparing the response-prediction model for the CpG model (grey) with the covariate model (blue). AUC=area under the curve. CpG=cytosine-phosphate-guanine.

observed nine predictor CpGs associated with CRP and two predictor CpGs with faecal calprotectin (appendix p 22). In all cases, mean differences in methylation between responders and non-responders were less than 0·5%.

Interrogating the blood expression of the predictor CpG-associated genes showed a significant differential expression when comparing responders with non-responders before treatment for *TULP4* (log_2[fold change] 0·38 [95% CI 0·011 to 0·76]; $p_{timepoint\ 1}$=0·044) and *RFPL2* (log_2[fold change] –0·77 [–1·5 to –0·0093]; $p_{timepoint\ 1}$=0·047) for vedolizumab (appendix pp 23–24, 33), with *RFPL2* having an inverse correlation between DNA methylation and gene expression (Pearson *r* –0·65 [95% CI –0·87 to –0·32]; p=0·0014; appendix pp 23–24). For ustekinumab, we observed significant differences in the expression of predictor CpG-associated genes *MRC1* (log_2[fold change] 1·3 [0·58 to 2·0]; $p_{timepoint\ 1}$=0·0004) and *TMEM191B* (log_2fold change –1·9 [–3·2 to –0·50]; $p_{timepoint\ 1}$=0·0073; appendix pp 23–24, 34–35) with *TMEM191B* presenting a significant positive correlation between DNA methylation and gene expression (Pearson *r* 0·57 [0·22 to 0·79]; p=0·0031; appendix pp 23–24). We investigated whether predictor CpG-associated genes presented differential expression at response assessment by comparing responders with non-responders at response assessment. Vedolizumab predictor CpG-associated genes *MCM2* (log_2[fold change] 1·3 [0·44 to 2·1]; $p_{timepoint\ 2}$=0·0024) and *RFPL2* (log_2[fold change] –1·6 [–2·6 to –0·50]; $p_{timepoint\ 2}$=0·0039) were differentially expressed at response assessment (appendix pp 23–24, 33) with *RFPL2*, once again, presenting a significant inverse correlation between DNA methylation and expression (Pearson *r* –0·55 [–0·81 to –0·12]; p=0·017; appendix pp 23–24). For ustekinumab, predictor CpG-associated genes *POTEF* (log_2[fold change] –0·87 [–1·5 to –0·18]; $p_{timepoint\ 2}$=0·015), *HDAC4* (log_2[fold change] –0·55 [–0·98 to –0·11]; $p_{timepoint\ 2}$=0·022), *PARP4* (log_2[fold change] –0·40 [–0·75 to –0·046]; $p_{timepoint\ 2}$=0·035) and *MARK3* (log_2[fold change] –0·45 [–0·85 to –0·054]; $p_{timepoint\ 2}$=0·029) presented significant, but limited, differential expression at response assessment (appendix pp 23–24, 34–35), but did not show any significant correlation with DNA methylation. Most response-associated differences in methylation occured in the absence of differential gene expression or were associated with limited log_2 fold changes.

Discussion

This study identified methylation signatures composed of 18, 25, and 68 markers that were associated with the combination of endoscopic response with either biochemical or clinical response to adalimumab, vedolizumab, and ustekinumab, respectively, in a cohort of adult patients with Crohn's disease. We built models in the discovery cohort with significant predictive performance for adalimumab (AUC 0·86 [95% CI 0·58–0·97]), vedolizumab (0·87 [0·67–0·98]), and ustekinumab (0·89 [0·76–1·00]). The 95% CIs show reasonably accurate prediction of response in the discovery cohort, with the lower bound exceeding 0·5. Testing the vedolizumab and ustekinumab models in an independent, external validation cohort demonstrated similar performance, with an AUC of 0·75 (0·65–0·85) and 0·75 (0·65–0·87) for vedolizumab and ustekinumab, respectively. By contrast, the adalimumab model failed in the validation cohort, yielding an AUC of 0·25 (0·10–0·35), consistently predicting the opposite to the

observed outcome. The reason for the lack of performance of the adalimumab model remains unclear and requires further investigation. However, our observations corroborate earlier attempts by Mishra and colleagues, in which the authors were unable to identify anti-TNF response-associated DNA methylation biomarkers before treatment in a prospective two-cohort approach.[26] Interestingly, Lin and colleagues identified 48 primary response-associated loci, which they associated with anti-TNF drug concentration levels at week 14 into treatment, suggesting a link between dosage and primary non-response.[27] Importantly, both Mishra and colleagues and Lin and colleagues defined response on biochemical or clinical endpoints only. The methodological differences and the absence of endoscopic data therefore preclude direct comparison with our study, in which outcome assessments were more stringent. We further note that anti-TNF exposure was associated with a reduction in the accuracy of both the vedolizumab and ustekinumab prediction models with an AUC of 0·66 (95% CI 0·55–0·73) for vedolizumab and 0·63 (0·52–0·70) for ustekinumab when considering only patients exposed to anti-TNF. Taken together, while we demonstrate the use of DNA methylation in predicting objective response to vedolizumab and ustekinumab, use of these predictors appears to have the most value in anti-TNF naive patients. Predicting response to anti-TNF using DNA methylation presents further complexities, which necessitate further investigation.

Recent real-world data and (post-hoc) findings from both the GEMINI and UNITI trials indicate superior response to both vedolizumab and ustekinumab in anti-TNF naive patients.[28] In our discovery cohorts, 77% of vedolizumab-treated and 98% of the ustekinumab-treated patients were previously exposed to anti-TNF drugs. Stratifying the patients in the validation cohort by previous anti-TNF exposure showed that both models performed noticeably better in anti-TNF naive rather than exposed patients. However, the number of patients included in both subset comparisons are relatively small and further exploration using larger groups of patients are needed. First-line treatment with anti-TNF agents has largely been driven by cost considerations. However, this economic background is rapidly changing with the advent of biosimilars for vedolizumab and ustekinumab. The American Gastroenterological Association currently recommends and suggests the use of ustekinumab and vedolizumab as first-line biologics for treating moderate-to-severe Crohn's disease as alternatives to anti-TNF therapy;[29] in clinical practice, these drugs are preferred over anti-TNF therapy in older patients or those with comorbidities on safety grounds.

Investigating the utility of the CDST for vedolizumab[7] and ustekinumab[8] in both discovery and validation cohorts indicated the superiority of the DNA methylation-based approach over these predictive tools in our cohorts. We hypothesise that this difference might in part be due to the differences in endpoints used to train the models. Where CDST was optimised for predicting clinical remission, our DNA methylation-based models were optimised for predicting endoscopic response. Dulai and colleagues showed that although 38% of patients classified as high probability of response reached clinical remission, only 19% reached clinical remission and mucosal healing in the absence of corticosteroids at week 26.[7]

Through two separate stability analyses, we showed both short-term and long-term stability of the majority of our identified CpG markers, indicating their independence of treatment and the resultant difference in inflammation. The latter is supported by the absence of correlation between the methylation status of the predictor CpGs and both baseline CRP and faecal calprotectin. Others have shown an absence of correlation between methylation status and therapy switch and even Crohn's disease-related surgery,[19] suggesting that the CpGs are response predictors but do not directly reflect inflammation. Nonetheless, when including baseline corticosteroid medication, previous anti-TNF usage, age, sex, smoking status, and blood cell distribution in the linear regression analyses, half of the vedolizumab predictor CpGs and about a quarter of the ustekinumab CpGs were no longer significantly associated with response. However, we observed that prediction modelling using the covariates only performed worse than the CpG model, indicating that the covariates we assessed alone do not contribute substantially to the predictive performance.

Our study represents the largest epigenome-wide association study on predicting biological response assessment in IBD and derives its main strength from the stringent sampling and objective endpoint assessment strategy. Nonetheless, there are some limitations to this study. First, patients with anti-drug antibodies or without a measurable serum drug concentration and patients that stopped treatment due to adverse events were excluded before the selection of this cohort to mitigate pharmacokinetic failure or intolerance. However, we did not set a minimum threshold on drug concentration to ensure sufficient exposure. Accordingly, we cannot fully exclude pharmacodynamic failure. Although achieving sufficient exposure is a well-documented issue with anti-TNFs, most of the adalimumab-treated patients presented with a median trough level above 5·85 µg/mL (appendix pp 25–26), a concentration associated with remission with adalimumab treatment.[30] Second, in the external cohort recruited in Oxford, response assessment was less stringent in approximately 70% of the patients due to the COVID-19 pandemic as a result of the restricted non-essential endoscopies in the UK. Therefore, the modified response criteria reflect clinical parameters that physicians commonly use in daily practice. Taking this pragmatic approach enhances the overall generalisability of our results to a larger IBD population. Notably, the analysis in

which we included the available endoscopic outcomes in the subset of UK patients for whom these data were available gave an AUC of 0·83 for both the vedolizumab and ustekinumab models, reinforcing the validity of both models in identifying objective responders to these biological therapies and indicating that our model is likely more capable at predicting response as defined using combined clinical, endoscopic, and biochemical evaluations than response defined by less robust means. While we acknowledge this limitation, we believe that our comprehensive approach provides valuable insights into the predictive capabilities of the models under diverse clinical scenarios. Third, although we purposely used peripheral blood leukocyte samples because these samples are minimally invasive and easily obtained during daily clinical practice, peripheral blood leukocytes represent a mixed cellular population. Therefore, the specific cell types responsible for the observed predictive signal remain unidentified. It should be noted that after correcting for covariates, including broad leukocyte subsets, several predictor CpGs remained statistically significant, suggesting independence of cellular composition. Fourth, although several of the predictor CpGs annotate to genes that encode proteins whose function can be related to IBD or general immunological functions (appendix p 36), most of the predictor CpG loci are situated within gene introns. Accordingly, beyond the utility of the predictor CpGs in classifying response to therapy, identifying their role in the pathogenesis and cause of non-response remains challenging at present and hence a subject for future studies. Finally, although we strictly removed most catalogued and predicted genetic variant-binding probes, we acknowledge that there was still a possibility that underlying genetic differences could have influenced our outcome. Nonetheless both models performed effectively in the discovery and the validation cohorts.

In summary, prognostic response-associated predictor CpGs for vedolizumab and ustekinumab could pave the way towards personalised medicine for Crohn's disease. We acknowledge that clinical validation of our findings in a randomised trial is needed, comparing our method of pretreatment selection with current clinical practice, to demonstrate both clinical and economic benefit. To this end, the Omicrohn trial, as part of the ongoing Horizon Europe funded METHYLOMIC project, has been launched and is currently underway.

Contributors

GRD, JJS, and WJdJ were the chief investigators of the EPIC-CD study. GRD, JJS, WJdJ, and PH conceptualised and developed the protocol. VWJ, IH, TPC, SvZ, and CGCM acquired the patient material. AYFLY, PH, FM, and AJN processed the material and acquired the data. AYFLY, PH, TdW, AJN, DL, and EL analysed the data. TdW, DL, and EL developed the algorithms for epigenetic biomarker discovery and performed the predictive machine learning analysis. VWJ, AYFLY, and MSH prepared the first draft of the manuscript. All authors critically reviewed and approved the final version before submission. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. AYFLY, AJN, and JJS have accessed and verified the data.

Declaration of interests

VWJ received speaker fees from Janssen-Cilag and Galapagos Biopharma; and received support for attending meetings or travel from Dr Falk Pharma Benelux. AYFLY received honoraria from Janssen, Johnson & Johnson, and DeciBio; was employed by GSK, and is shareholder of GSK. PH received grant support from Horizon Europe, TKI Health Holland, and the Leona M and Harry B Helmsley Charitable Trust. TdW was employed by Horaizon. EL is a cofounder of Horaizon. TPC received consulting fees from Eli Lilly and Takeda; speaker fees from Eli Lilly, Janssen, Takeda, and Tillotts Pharma; educational support from Dr Falk Pharma UK, Janssen, Abbvie, and Tillotts Pharma; and served as advisory board member for Eli Lilly and Dr Falk Pharma UK. DL is employed by Horaizon. JSJ received grant support from the Leona M and Harry B Helmsley Charitable Trust, the European Crohn's and Colitis Organisation, Crohn's And Colitis UK, UK Research and Innovation, Action Medical Research, and the European Commission's Innovative Health Initiative and Horizon 2020 programmes; travel support form Takeda and Janssen; is director of the UK IBD Registry; and has preliminary patent applications to develop non-immunogenic anti-TNF therapy and to develop diagnostic epigenetic biomarkers in inflammatory bowel disease. WJdJ received grant support from Leona M and Harry B Helmsley Charitable Trust, TKI Health Holland, and the European Commission; speaker fees from Janssen Cilag, Alimentiv, and IBD Canada; is scientific board member of the MDL fund; and is a board member and shareholder of AIBiomics. GRD served as a consultant for AbbVie, Alimentiv, AstraZeneca, Bristol Myers Squibb, Celltrion, Eli Lilly, Exeliom Biosciences, Johnson & Johnson, Pfizer, and Takeda; and has received speakers bureau fees from AbbVie, Eli Lilly, Pfizer, Bristol Myers Squibb, and Takeda. All other authors declare no competing interests. The Amsterdam University Medical Centre has a patent pending for the vedolizumab and ustekinumab response prediction models presented in this manuscript.

Data sharing

The raw DNA methylation and gene expression data alongside the de-identified patient metadata as reported on in this study have been published under controlled access for research purposes at the European Genome-phenome Archive. The accession number for the DNA methylation data for the adalimumab discovery cohort is EGAD00010002720, for the adalimumab validation cohort is EGAD00010002721, for the vedolizumab discovery cohort is EGAD00010002651, the vedolizumab validation cohort is EGAD00010002652, the ustekinumab discovery cohort is EGAD00010002649, and the ustekinumab validation cohort is EGAD00010002650. The accession number for the RNA-sequencing data for the vedolizumab discovery cohort is EGAD50000000385 and for the ustekinumab discovery cohort is EGAD50000000386. All Snakemake, bash calls, and R scripts have been made available on GitHub (https://github.com/ND91/HGPRJ0000008_EPICCD_multi_drug.git). The machine learning modelling was performed using proprietary algorithms founded on the same statistical principles as those of gradient boosting, permutation importance, and covered information disentanglement. The code for these techniques is openly available at https://xgboost.ai and https://github.com/JBPereira/CID or https://scikit-learn.org/stable.

Acknowledgments

This project was supported by a grant from The Leona M and Harry B Helmsley Charitable Trust (2019-1179]). We thank Floris de Voogd for assistance in evaluation of ultrasound images and Alex T Adams for assistance in the initial planning and setting up of this study.

References

1 de Souza HS, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nat Rev Gastroenterol Hepatol* 2016; **13:** 13–27.
2 Alsoud D, Vermeire S, Verstockt B. Biomarker discovery for personalized therapy selection in inflammatory bowel diseases: challenges and promises. *Curr Res Pharmacol Drug Discov* 2022; **3:** 100089.
3 Loftus EV Jr, Panés J, Lacerda AP, et al. Upadacitinib induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2023; **388:** 1966–80.

4 D'Haens G, Panaccione R, Baert F, et al. Risankizumab as induction therapy for Crohn's disease: results from the phase 3 ADVANCE and MOTIVATE induction trials. *Lancet* 2022; **399:** 2015–30.

5 Löwenberg M, Vermeire S, Mostafavi N, et al. Vedolizumab induces endoscopic and histologic remission in patients with Crohn's disease. *Gastroenterology* 2019; **157:** 997–1006.

6 Sands BE, Irving PM, Hoops T, et al, and the SEAVUE Study Group. Ustekinumab versus adalimumab for induction and maintenance therapy in biologic-naive patients with moderately to severely active Crohn's disease: a multicentre, randomised, double-blind, parallel-group, phase 3b trial. *Lancet* 2022; **399:** 2200–11.

7 Dulai PS, Boland BS, Singh S, et al. Development and validation of a scoring system to predict outcomes of vedolizumab treatment in patients with Crohn's disease. *Gastroenterology* 2018; **155:** 687–95.

8 Dulai P, Guizzetti L, Ma T, et al. Clinical prediction model and decision support tool for ustekinumab in Crohn's disease. *Am J Gastroenterol* 2019; **114:** S373.

9 Verstockt B, Verstockt S, Veny M, et al. Expression levels of 4 genes in colon tissue might be used to predict which patients will enter endoscopic remission after vedolizumab therapy for inflammatory bowel diseases. *Clin Gastroenterol Hepatol* 2020; **18:** 1142–51.

10 Noor NM, Lee JC, Bond S, et al, and the PROFILE Study Group. 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 Gastroenterol Hepatol* 2024; **9:** 415–27.

11 Soendergaard C, Seidelin JB, Steenholdt C, Nielsen OH. Putative biomarkers of vedolizumab resistance and underlying inflammatory pathways involved in IBD. *BMJ Open Gastroenterol* 2018; **5:** e000208.

12 Ananthakrishnan AN, Luo C, Yajnik V, et al. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel diseases. *Cell Host Microbe* 2017; **21:** 603–10.

13 Peery RC, Pammi M, Claud E, Shen L. Epigenome – a mediator for host-microbiome crosstalk. *Semin Perinatol* 2021; **45:** 151455.

14 Li Y, Wang Z, Wu X, et al. Intestinal mucosa-derived DNA methylation signatures in the penetrating intestinal mucosal lesions of Crohn's disease. *Sci Rep* 2021; **11:** 9771.

15 Ventham NT, Kennedy NA, Adams AT, et al, and the IBD BIOM consortium, and the IBD CHARACTER consortium. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat Commun* 2016; **7:** 13507.

16 Gasparetto M, Payne F, Nayak K, et al. Transcription and DNA methylation patterns of blood-derived $CD8^+$ T cells are associated with age and inflammatory bowel disease but do not predict prognosis. *Gastroenterology* 2021; **160:** 232–44.

17 Howell KJ, Kraiczy J, Nayak KM, et al. DNA methylation and transcription patterns in intestinal epithelial cells from pediatric patients with inflammatory bowel diseases differentiate disease subtypes and associate with outcome. *Gastroenterology* 2018; **154:** 585–98.

18 Tsai PC, Bell JT. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *Int J Epidemiol* 2015; **44:** 1429–41.

19 Joustra V, Li Yim AYF, Hageman I, et al. Long-term temporal stability of peripheral blood DNA methylation profiles in patients with inflammatory bowel disease. *Cell Mol Gastroenterol Hepatol* 2023; **15:** 869–85.

20 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; **15:** 155–63.

21 Tsai PC, Glastonbury CA, Eliot MN, et al. Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. *Clin Epigenetics* 2018; **10:** 126.

22 Houseman EA, Kelsey KT, Wiencke JK, Marsit CJ. Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective. *BMC Bioinformatics* 2015; **16:** 95.

23 Solomon O, Huen K, Yousefi P, et al. Meta-analysis of epigenome-wide association studies in newborns and children show widespread sex differences in blood DNA methylation. *Mutat Res-Rev Mutat* 2022; **789:** 108415.

24 Dobbs KR, Embury P, Koech E, et al. Age-related differences in monocyte DNA methylation and immune function in healthy Kenyan adults and children. *Immun Ageing* 2021; **18:** 11.

25 Somineni HK, Venkateswaran S, Kilaru V, et al. Blood-derived DNA methylation signatures of Crohn's disease and severity of intestinal inflammation. *Gastroenterology* 2019; **156:** 2254–65.

26 Mishra N, Aden K, Blase JI, et al, and the SYSCID Consortium. Longitudinal multi-omics analysis identifies early blood-based predictors of anti-TNF therapy response in inflammatory bowel disease. *Genome Med* 2022; **14:** 110.

27 Lin S, Hannon E, Reppell M, et al. Whole blood DNA methylation changes are associated with anti-TNF drug concentration in patients with Crohn's disease. *J Crohns Colitis* 2024; **18:** 1190–201.

28 Yang H, Huang Z, Li M, et al. Comparative effectiveness of ustekinumab vs. vedolizumab for anti-TNF-naïve or anti-TNF-exposed Crohn's disease: a multicenter cohort study. *EClinicalMedicine* 2023; **66:** 102337.

29 Feuerstein JD, Ho EY, Shmidt E, et al, and the American Gastroenterological Association Institute Clinical Guidelines Committee. AGA clinical practice guidelines on the medical management of moderate to severe luminal and perianal fistulizing Crohn's disease. *Gastroenterology* 2021; **160:** 2496–508.

30 Mazor Y, Almog R, Kopylov U, et al. Adalimumab drug and antibody levels as predictors of clinical and laboratory response in patients with Crohn's disease. *Aliment Pharmacol Ther* 2014; **40:** 620–28.