1 Title

- 2 Large-scale clustering of longitudinal faecal calprotectin and C-reactive protein
- 3 profiles in inflammatory bowel disease

4 Short title

5 Longitudinal biomarker clustering in inflammatory bowel disease

6 Authors

- 7 Nathan Constantine-Cooke^{1,2}, Nikolas Plevris^{1,3}, Karla Monterrubio-Gómez^{1,2}, Clara
- 8 Ramos Belinchón^{1,3}, Solomon Ong³, Alexander T. Elford^{3,4}, Beatriz Gros^{3,5}, Gareth-
- 9 Rhys Jones^{3,6}, Charlie W. Lees^{† 1,3} and Catalina A. Vallejos^{† 2}.
- 10 [†] Joint senior authors
- 11 Affiliations
- Centre for Genomic and Experimental Medicine, Institute of Genetics and
 Cancer, University of Edinburgh, Edinburgh, UK
- MRC Human Genetics Unit, Institute of Genetics and Cancer, University of
 Edinburgh, Edinburgh, UK
- 16 3. Edinburgh IBD Unit, Western General Hospital, Edinburgh, UK
- 17 4. Faculty of Medicine, The University of Melbourne, Melbourne, Australia
- 18 5. Reina Sofía University Hospital, Gastroenterology and Hepatology. IMIBIC.
- 19 University of Cordoba. Cordoba, Spain
- Centre for Inflammation Research, The Queen's Medical Research Institute,
 University of Edinburgh, Edinburgh, UK

22 Abstract

- 23 Background
- 24 Crohn's disease (CD) and ulcerative colitis (UC) are highly heterogeneous, dynamic
- 25 and unpredictable, with a marked disconnect between symptoms and intestinal NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

inflammation. Attempts to classify inflammatory bowel disease (IBD) subphenotypes
to inform clinical decision making have been limited. We aimed to describe the latent
disease heterogeneity by modelling routinely collected faecal calprotectin (FC) and
CRP data, describing dynamic longitudinal inflammatory patterns in IBD.

30 Methods

31 In this retrospective study, we analysed patient-level longitudinal measurements of FC 32 and CRP recorded within seven years since diagnosis. Latent class mixed models 33 (LCMMs) were used to cluster individuals with similar longitudinal FC or CRP profiles. 34 Associations between cluster assignment and information available at diagnosis (e.g. 35 age, sex, and Montreal classification) were quantified using multinomial logistic 36 regression. Differences in advanced therapy use across clusters were also explored 37 using cumulative distributions over time. Finally, we considered uncertainty in cluster 38 assignments with respect to follow-up length and explored the overlap between 39 clusters identified based on FC and CRP.

40 Findings

41 We included 1036 patients (544 CD, 380 UC, 112 IBD-unclassified (IBDU)) in the FC 42 analysis with a total of 10545 FC observations (median 9 per subject, IQR 6–13). The 43 CRP analysis consisted of 1838 patients (805 CD, 847 UC, 186 IBDU) with 49364 44 CRP measurements (median 20 per subject: IQR 10-36). Eight distinct clusters of 45 inflammatory behaviour over time were identified by LCMM in each analysis. The 46 clusters, FC1-8 and CRP1-8, were ordered from the lowest cumulative inflammatory 47 burden to the highest. The clusters included groups with high diagnostic levels of 48 inflammation which rapidly normalised, groups where high inflammation levels 49 persisted throughout the full seven years of observation, and a series of intermediates 50 including delayed remitters and relapsing remitters.

51 CD and UC patients were unevenly distributed across the clusters. In CD, whilst 52 patients with upper GI involvement (L4) were less likely to be in FC1 and FC2, there 53 was no impact on ileal versus colonic disease on cluster assignment. In UC, male sex 54 was associated with the poorest prognostic cluster (FC8). The use and timing of 55 advanced therapy was associated with cluster assignment, with the highest use of 56 early advanced therapy in FC1. Of note, FC8 and CRP8 captured consistently high

57 patterns of inflammation despite a high proportion of patients receiving advanced 58 therapy, particularly for CD individuals (56.8% and 33.3%, respectively). We observed 59 that uncertainty in cluster assignments was higher for individuals with short 60 longitudinal follow-up, particularly between clusters capturing similar earlier 61 inflammation patterns. There was broadly poor agreement between FC and CRP 62 clusters in keeping with the need to monitor both in clinical practice.

63 Interpretation

Distinct patterns of inflammatory behaviour over time are evident in patients with IBD.
Cluster assignment is associated with disease type and both the use and timing of
advanced therapy. These data pave the way for a deeper understanding of disease
heterogeneity in IBD and enhanced patient stratification in the clinic.

68

69 Introduction

70 Inflammatory bowel disease (IBD), an umbrella term for Crohn's disease (CD), 71 ulcerative colitis (UC) and inflammatory bowel disease unclassified (IBDU), has a prevalence of almost 1% in the UK population.^{1,2} The condition is characterised by 72 73 chronic relapsing and remitting inflammation of the gastrointestinal (GI) tract that 74 confers a host of debilitating symptoms, negatively impacting quality of life.^{3,4} Studies 75 have clearly demonstrated that uncontrolled inflammation of the GI tract increases the risk of disease progression and complications including the development of colorectal 76 77 cancer,⁵ stricturing/penetrating complications and surgery.^{6,7} However, IBD is highly 78 heterogeneous with respect to symptoms, inflammatory burden, treatment response, 79 and long-term outcomes.

80 Current IBD classification methods are mostly based on historic nomenclature which 81 utilise baseline phenotypic characteristics and do not take into account the dynamic 82 and unpredictable nature of the disease. Attempts at developing prediction tools to identify high risk patients have been made but again these use static clinical 83 84 parameters, dismissing the changing nature of the disease.⁸ Furthermore, they do not take into account the influence of other factors, such as advanced therapy timing, on 85 86 the disease course. In the wake of the increasing prevalence of IBD and associated 87 healthcare burden, new methods to characterise the dynamic disease course and help

88 identify at-risk individuals who require aggressive therapy with close monitoring versus

89 those that need less intense input are essential.

90 The original IBSEN studies provided the first data on the clinical course of patients 91 with UC and CD during the first 10-years of diagnosis.^{9,10} However, these data were 92 based on predefined disease patterns, rather than data driven, utilising symptoms 93 alone. It is now widely accepted that there is a clear disconnect between inflammation 94 and symptoms in IBD,¹¹ therefore it is imperative characterisations of disease course 95 include objective parameters of inflammation.

- 96 C-reactive protein (CRP) and faecal calprotectin (FC) are well established tools for 97 monitoring patients with IBD, but are typically interpreted in terms of the most recent 98 measurement or short-term trends. Interrogation of long-term trends of inflammation 99 could greatly assist clinical decision making and improve prediction of future events. 100 Moreover, modelling inflammatory behaviour over time might be a key tool for 101 characterising the largely unexplained heterogeneity seen in IBD, and provide new tools for disease sub-phenotyping beyond the current Montreal classification.¹² 102 103 In this study, we aimed to 1) identify groups of IBD patients with similar longitudinal
- 104 patterns of inflammation 2) determine if these groupings were associated with age.
- 105 sex, IBD type, Montreal classification, or early advanced therapy, and 3) explore
- 106 whether subjects with similar longitudinal FC profiles also share similar CRP profiles.

107 Methods

- 108 Ethics
- 109 This project was approved by the local Caldicott Guardian (Project ID: CRD18002,
- 110 registered NHS Lothian information asset #IAR-954). Patients or the public were not
- 111 involved in the design, conduct, reporting or dissemination plans of our research.

112 Study design

This was a retrospective cohort study. Patients with a confirmed diagnosis of IBD (as per Lennard-Jones criteria)¹³ were followed up for a period of seven years from the date of diagnosis. Baseline phenotype data (sex, age at diagnosis, IBD type, date of diagnosis) were obtained from the Lothian IBD registry (LIBDR), a retrospective cohort

117 of patients receiving IBD care in Lothian, Scotland. The LIBDR is estimated to have 118 identified 94.3% of all true IBD patients in the area using a capture-recapture 119 approach.¹ Using a population level cohort reduces potential biases associated with cohort recruitment.¹⁴ When available, additional phenotyping information was 120 121 extracted by the clinical team from electronic health records (TrakCare: InterSystems. 122 Cambridge, MA). This includes smoking and Montreal classification for disease 123 location, behaviour and extent, all recorded as per patient status at the time of 124 diagnosis. Data on prescribing of all advanced therapies, including start/stop dates, 125 were extracted from both TrakCare and NHS Lothian pharmacy databases. Primary 126 care prescribing data were not available. See Supplementary Note 1 for more detailed 127 definitions.

128 Inclusion/exclusion criteria

Subjects were required to have a confirmed diagnosis of IBD at any age and receive secondary care for their condition from the NHS Lothian health board. Only subjects with a recorded date of IBD diagnosis between 2005 and 2019 were included. The lower bound for this criteria was established as FC testing was not routinely performed prior to this date. The upper bound ensured subjects had the possibility of having at least five years of follow-up at the time of data extraction.

135 For subjects which met the above requirements, the following criteria was applied to 136 their FC and/or CRP longitudinal measurements. Subjects were required to have a 137 diagnostic measurement (± three months of diagnosis, Figures S1 and S2) and have 138 a further two observations available within seven years of diagnosis. If any biomarker 139 measurements were observed within three months prior to the recorded diagnosis 140 date, measurement time scales were realigned with respect to diagnosis (Figure S3). 141 Only non-censored observations were considered in this calculation for FC. For CRP, 142 this filtering was applied after preprocessing (see "Longitudinal measurements and preprocessing" section), and subjects with constant biomarker measurements over 143 144 time were excluded. As the clustering analyses were performed separately for FC and 145 CRP, subjects did not need to meet the criteria for both biomarkers.

146 Statistical analysis

- 147 Cohort description
- 148 Continuous variables were summarised as their median and interquartile range (IQR).
- 149 Categorical variables were summarised as counts and percentages.
- 150 Longitudinal measurements and preprocessing

151 FC and CRP measurements were obtained from an extract by the local biochemistry 152 team describing tests recorded up to August 13, 2024. For each individual, all 153 measurements made within seven years from diagnosis were considered. Failed tests, 154 for example due to contamination, were discarded. All FC tests were performed from 155 stool samples using the same ELISA technology.¹⁵ Due to limits of detection, 156 observations < 20 μ g/g were recorded to 20 μ g/g whilst observations > 1250 μ g/g were 157 mapped to 1250 µg/g. Such observations were treated as censored when applying the 158 inclusion exclusion criteria described above. CRP was measured from blood samples; 159 observations for which only an upper bound was available, e.g. < 1 mg/L, were 160 mapped to the corresponding upper bound.

Further processing was applied to CRP data to smooth out short-term fluctuations. Measurements were grouped into intervals of t: [0, 0.5), [0.5, 1.5), [1.5, 2.5), [2.5, 3.5),[3.5, 4.5), [4.5, 5.5), [5.5, 7], where t = 0 (years) is the time of diagnosis. The median CRP for each interval was calculated for each subject and used as input for subsequent analyses. The centre of each interval was used as the corresponding observation time.

167 Longitudinal biomarker clustering

168 Prior to model fitting, FC and CRP observations were log-transformed. FC and CRP trajectories were modelled separately using latent class mixed models (LCMMs),¹⁶ an 169 170 extension of linear mixed effects models that enables clustering of individuals that 171 share similar longitudinal biomarker trajectories. LCMM consists of two submodels: 172 one which captures the longitudinal behaviour of the biomarker, and one which 173 captures cluster assignment. Fixed effects for the longitudinal submodel were 174 specified using natural cubic splines with three interior knots placed at quantiles with 175 respect to observation times. An alternative specification was considered leading to

similar results. Only the intercept was treated as a random effect. The cluster
assignment submodel used IBD type (CD, UC, and IBDU) as a covariate. Formal
definitions of the LCMM models considered here and the associated hyper-parameter

179 choices are provided in Supplementary Note 2.

As the number of clusters is not known a priori, the optimal model was found using a grid search approach. We considered models with 2–10 clusters for both FC and CRP. The likelihood based statistics,¹⁷ Akaike information criterion (AIC) and Bayesian information criterion (BIC), and visual inspection of cluster trajectories were used to compare models with different specifications and determine the most appropriate number of assumed clusters.

LCMM calculates, for each individual, the probability of being assigned to each cluster. In subsequent analysis, each individual was assigned to the cluster with the highest probability. The distribution of cluster assignment probabilities was used to assess uncertainty in these allocations with respect to follow-up length. This was defined as the time difference between diagnosis and the last available biomarker measurement (FC or CRP, depending on the analysis). For each cluster, the average probability of individual-specific probabilities of cluster assignment were reported.

To avoid displaying potentially identifiable individual-level data, exemplar trajectories within each cluster are visualised as aggregated trends, where measurements were summarised as the median across six randomly selected individuals. Cluster labels were ordered based on the area under the overall biomarker trend inferred for each cluster (a proxy for cumulative inflammatory burden).

198 Associations with respect to cluster assignments

199 To facilitate the interpretation of each cluster, we considered potential associations 200 between cluster assignments and patient-level information. Violin plots and 201 percentage bar plots were used as a visual summary when considering continuous 202 (age) and discrete (sex, IBD type, additional phenotyping) patient-level covariates, 203 respectively. Associations with respect to additional phenotyping were only explored 204 after stratifying by IBD type (CD and UC only). Multinomial logistic regression was 205 used to quantify associations between cluster assignments and patient-level 206 covariates. The cluster used as the reference class was chosen to closely resemble 207 the overall distribution of IBD types within the corresponding cohort. Univariate and

208 multivariate models were considered, and the associated 95% confidence intervals 209 are reported. Individuals with missing covariate values were excluded when fitting

- 210 each model. Due to small cluster sizes and low frequency of some covariate levels.
- 211 the analysis was repeated after merging clusters with similar cumulative inflammatory
- 212 burden. Effect sizes were largely consistent (data not shown).

213 Advanced therapy use

214 To compare patterns of advanced therapy (AT) across clusters, the cumulative 215 distribution of first-line advanced therapy use was calculated. Results are reported 216 stratified by IBD type (CD and UC only).

Comparison between FC and CRP cluster assignments 217

218 For subjects meeting the criteria for both FC and CRP analyses, the relationship 219 between FC and CRP cluster assignment was visualised using alluvial plots and side-220 by-side comparisons of mean cluster trajectories for the optimal models. Stratified 221 results for CD and UC subjects are also reported.

Software 222

R (v. $4 \cdot 4 \cdot 0$), extended using the lcmm (v. $2 \cdot 1 \cdot 0$),¹⁸ ggalluvial (v. $0 \cdot 12 \cdot 5$),¹⁹ and datefixR 223 224 (v.1.6.1)²⁰ packages, was used for all analyses. Analytical reports have been 225 generated using the Quarto scientific publishing system and are hosted online 226 (https://vallejosgroup.github.io/Lothian-IBDR/). An R package, libdr (v.1.0.0), has also 227 been produced, supporting the reuse of our R code with other datasets.

- 228 Role of the funding source
- 229 Funders were not involved in the study design, collection, analysis, or interpretation
- 230 of the data, writing, or decision to submit the paper for publication.

231 Results

232 Cohort derivation and description

Of the 10153 subjects with a confirmed IBD diagnosis, 5508 were reported as being diagnosed between 2005 and 2019 (Figure 1). Of these subjects, 1036 and 1838 subjects were included in the FC and CRP analysis, respectively. We identified 808 subjects which met the inclusion criteria for both biomarkers. <u>Table 1</u> describes key demographic factors for subjects included in the FC and CRP analyses.



Figure 1. Derivation of the study cohorts based on separate faecal calprotectin (FC) and C-reactive
 protein (CRP) inclusion/exclusion criteria.

241

	FC analysis (n = 1036)	CRP analysis (n = 1838)	Overlap (n = 808)
Age at diagnosis	33 (25–75)	37 (24–55)	32 (20–50)
Male sex	503 (48.6%)	956 (52-0%)	397 (49-1%)
IBD type			
Crohn's disease (CD)	544 (52·5%)	805 (43-8%)	451 (55-8%)
Ulcerative colitis (UC)	380 (36.7%)	847 (46-1%)	276 (34-2%)

IBDU	112 (10-8%)	186 (10.1%)	81 (10-0%)
Baseline FC (µg/g)	740 (320–1070)		760 (330–1081)
FC observations	9 (6–13)		9 (6–14)
Baseline CRP (µg/mL)		8 (3–24)	8 (3–20)
Total unprocessed CRP observations		20 (10–36)	26 (14–36)
Total CRP observations after pre-processing		6 (4–7)	6 (5–7)
Advanced therapies (AT)			
AT within seven years for CD	270 (49-6%)	374 (46-5%)	
AT within one year for CD	129 (23-7%)	167 (20.7%)	
AT within seven years for UC	108 (28-4%)	85 (10-0%)	
AT within one year for UC	42 (11·1%)	35 (4.1%)	
Crohn's disease			
Montreal location			
L1	168 (30-9%)	249 (30-9%)	136 (30-2%)
L2	200 (36-8%)	287 (35.7%)	157 (34-8%)
L3	169 (31-1%)	246 (30.6%)	153 (33-9%)
Missing	7 (1·3%)	23 (2-9%)	5 (1.1%)
Upper GI inflammation			
Present	77 (14·2%)	87 (10-8%)	68 (15-1%)
Not Present	467 (85.8.4%)	718 (89-2%)	383 (84-9%)
Montreal behaviour			
B1	443 (81-4%)	607 (75-4%)	364 (80.7%)
B2	63 (11-5%)	110 (13.7%)	54 (12.0%)
B3	28 (5-1%)	61 (7.6%)	25 (5.5%)

Missing	10 (1.8%)	27 (3-4%)	8 (1-8%)
Perianal disease			
Present	79 (14·5%)	113 (14-0%)	68 (15-1%)
Not Present	459 (84-4%)	673 (83-6%)	379 (84.0%)
Missing	6 (1·1%)	19 (2-4%)	4 (0-9%)
Smoking status			
Current or previously	166 (30-5%)	274 (34-0%)	138 (30-6%)
Never	352 (64.7%)	484 (60-1%)	293 (65-0%)
Missing	26 (4-8%)	47 (5-8%)	20 (4-4%)

242Table 1. Demographic and clinical data at diagnosis for subjects meeting the faecal calprotectin (FC)243or C-reactive protein (CRP) study inclusion criteria. Continuous data are presented as median and244interquartile range. Categorical data are presented as counts and percentages. Missingness is only245directly reported if values were missing. The column labelled as "Overlap" denotes subjects which met246the inclusion criteria for both FC and CRP modelling. Missing observations for upper gastrointestinal247inflammation were assumed to be "not present" (Supplementary Note 1). Missingness was not248inferred to be a value for any other variable.

249 Longitudinal measurements of FC and CRP

For subjects in the FC analysis, 10545 FC observations were available (median 9 per subject, IQR 6–13). Prior to processing 49364 CRP observations were available (median 20 per subject, IQR 10–36). Following the pre-processing of CRP observations, there were 9898 data points (median 6 per subject, IQR 4–7). The distribution of log-transformed FC and CRP values is shown in <u>Figure S4</u>.

255 Modelling of FC trajectories

AIC suggested the 10-cluster model (<u>Figure S5</u>), whilst BIC suggested the 9-cluster model (<u>Figure S6</u>) was more appropriate (<u>Table S1</u>, <u>Figure S7A</u>). However, the 8cluster model was chosen as a parsimonious choice, as it captures the main observed longitudinal patterns without generating very small clusters (<50 individuals) which could be difficult to interpret.

261 Figure 2 shows representative cluster profiles for the 8-cluster model, ordered from
 262 lowest (FC1) to highest (FC8) cumulative inflammatory burden. FC2 (n=67; ~6%)

263 represents low FC values throughout the whole observation period. Instead, FC1 264 (n=140; ~14%) and FC3 (n=157; ~15%) and FC7 (n=244; ~24%) were characterised 265 by initially high FC values (>250 µg/g) which decreased over time at different rates. 266 Whilst FC1 exhibited a sharp decrease within the first year, the decrease was more 267 gradual for FC3 and FC7, where FC was normalised (<250 µg/g) approximately 268 around two and five years post diagnosis, respectively. Furthermore, FC4 (n=103; 269 ~10%), FC5 (n=67; ~6%) and FC6 (n=64; ~6%) capture relapsing and remitting 270 patterns of gastrointestinal inflammation. Finally, FC8 (n=194; ~19%) represented 271 individuals with consistently high FC values.

272

273 Associations with respect to FC cluster assignments

Associations with respect to age, sex and IBD type were first considered. Whilst the effect of age and sex was not substantial (<u>Figure S8</u>), this was not the case for IBD type (<u>Figure S9</u>). For example, IBDU and UC patients were less likely to be assigned to FC6 (1.56% IBDU and 28.1% UC vs 11.4% and 37.3% respectively elsewhere).

278 A similar analysis was performed after stratifying by IBD type (UC and CD only) and 279 considering additional phenotyping (Figures S10 - S16). In most cases, effect sizes 280 were not statistically significant (in some cases this was due to low counts and small 281 cluster sizes). However, amongst CD patients, those without upper GI inflammation 282 were more likely to be assigned to FC1 or FC2 (2.4% and 3.03% L4 vs 18.1% 283 elsewhere). In UC, and to some extent CD, males were more likely to be assigned to 284 FC8 (72% male versus 49.7% elsewhere). Finally, UC patients with ulcerative proctitis 285 were found to be less likely to be assigned to FC3 (2.44% E1 versus 14.2% 286 elsewhere).

FC cluster assignment and advanced therapy usage

In total, 270 (49-6%) CD and 108 (28-4%) UC subjects received an advanced therapy within the seven year observation period (<u>Table 1</u>). Overall AT rates and the distribution of time to first AT were not homogeneous across clusters (<u>Figure 3</u>). For example, whilst AT prescription rates in FC1 largely matched the overall FC cohort, prescriptions were generally earlier in this cluster, especially in CD patients. It was noteworthy that patients in FC8, with consistently high FC levels and later onset of AT,

- had cumulative AT rates of 56.8% in CD and 37.9% in UC by the end of the seven
- 295 year observation period.



296

Figure 2. Cluster trajectories obtained from LCMM assuming eight clusters fitted to FC data (logtransformed). Red lines indicate predicted mean cluster profiles with 95% confidence intervals. The
blue dotted lines indicate log(250µg/g). For visualisation purposes, pseudo subject-specific
trajectories have been generated by amalgamating observations from randomly selected groups of six
subjects. Clusters are ordered from lowest (FC1) to highest (FC8) cumulative inflammatory burden.
Cluster sizes are shown as panel titles.

304 Modelling of CRP trajectories

AIC and BIC both suggested the 8-cluster model was the most appropriate (<u>Table S4</u>). Visual inspection also supported this finding as the 9-cluster model did not identify new trajectories when compared to the 8-cluster model, producing two trajectories with consistently low CRP (<u>Figure S8</u>). In contrast, the 7-cluster model (<u>Figure S9</u>) lacks one of the clinically interesting trajectories, characterised by an initially elevated CRP which then decreases to slightly above biochemical remission after one year, when compared to the 8-cluster model.

312 Figure 4 presents exemplar cluster profiles for the 8-cluster model. Over a third of 313 subjects (n=702; ~38%) were assigned to CRP1 which was defined by consistently 314 low CRP. CRP2 (n=225; ~12%) was characterised by high CRP at diagnosis which 315 rapidly decreased shortly thereafter remaining low. CRP3 (n=51; ~3%) and CRP4 316 (n=60; ~3%) are both small clusters with the former described by low CRP until the 317 last year of follow-up and the latter presenting as low inflammation within the first year 318 of diagnosis before increasing until the third year where the inflammation then 319 decreases again. CRP5 (n=110; 6%) is characterised by elevated CRP at diagnosis 320 which then decreases gradually over time. CRP6 (n=434; 24%) consists of trajectories 321 which are elevated at diagnosis which then falls slightly for the first two years after 322 diagnosis, remaining elevated across the remaining duration of follow-up. CRP8 323 (n=172; 9%) is consistently elevated and does not change over time.

324 Associations with CRP cluster assignments

325 Figure S19 and S20 visualise the distribution of age, sex and IBD type within each 326 CRP cluster. Older patients were more likely to be assigned to a CRP cluster with 327 higher cumulative inflammatory burden. IBD type was not evenly distributed among 328 CRP clusters, with CRP1, CRP3, CRP4 and CRP7 enriched for UC patients (59.3%, 329 56.8%, 61.7% and 72.6% UC vs 32.3% elsewhere). Amongst CD patients, higher 330 smoking rates were generally observed for CRP clusters with higher inflammatory 331 burden (e.g. 26.9% in CRP1, 54.0% in CRP8; Figure S21), but there were not 332 substantial differences when considering Montreal location and behaviour, or upper 333 gastrointestinal inflammation (Figures S22 - S24). AT prescribing patterns in CRP 334 clusters are shown in Figure S25.



336 Figure 3. FC cluster-specific cumulative distribution for first-line advanced therapy prescribing for 337 Crohn's disease (red) and ulcerative colitis (teal) subjects. Clusters are ordered from lowest (FC1) to 338 highest (FC8) cumulative inflammatory burden. The number of CD and UC subjects present in each 339 cluster is displayed as panel titles. Total advance therapy prescribing (as a percentage of the 340 corresponding group) within seven years from diagnosis is shown next to each distribution curve. 341 Curves which would describe fewer than five subjects are not shown.



342

343 Figure 4. Cluster trajectories obtained from LCMM assuming eight clusters fitted to processed CRP 344 data (log-transformed). Red lines indicate predicted mean cluster profiles with 95% confidence 345 intervals. The blue dotted lines indicate log(5µg/mL). For visualisation purposes, pseudo subject-346 specific trajectories have been generated by amalgamating observations from randomly selected

347 groups of six subjects. Clusters are ordered from lowest (CRP1) to highest (CRP8) cumulative 348 inflammatory burden. Cluster sizes are shown as panel titles.



349 350

351

Figure 5. Exploration of cluster assignment uncertainty for A) faecal calprotectin (FC) and B) CRP clusters. LCMM assigns individuals to the cluster with the highest estimated probability. For

352 individuals assigned to a given cluster, bars show the average probability of cluster assignment to 353 each possible cluster. Results are stratified according to follow-up length, defined as the time 354 difference between diagnosis and the last available biomarker measurement (FC or CRP for A) and 355 B), respectively). Clusters are ordered from lowest (FC1 and CRP1) to highest (FC8 and CRP8) 356 cumulative inflammatory burden, with adjacent clusters coloured sequentially in the plots for FC (black 357 to yellow) and CRP (blue to yellow).

Uncertainty in cluster assignments 358

In the FC analysis, with the exception of FC2, cluster assignments were on average 359 360 more uncertain for subjects with a short follow-up (Figure 5 (A)). This is particularly the 361 case for FC clusters that share similar earlier trends. For example, individuals 362 assigned to FC1 (rapid FC normalisation) had a low average probability of being 363 assigned to FC8 (consistently high FC) and vice-versa, even for those with a short 364 follow-up. This is not the case when comparing FC3 and FC6, both of which capture 365 similar FC trajectories within the first two years. Indeed, those assigned to FC6 with 366 less than two years of follow-up also have, on average, a high probability of being 367 assigned to FC3. On average, cluster assignments were less uncertain in the CRP 368 analysis, even for individuals with a short follow-up (Figure 5 (B)). This is expected as 369 CRP clusters are associated with more distinct early trajectories.

Comparison of FC and CRP clustering 370

371 Overall, all FC clusters were well represented amongst the 808 subjects included in both analyses (overlap cohort), but CRP8 was underrepresented (~27% of CRP8 was 372 373 in the overlap cohort vs ~46% elsewhere; Figure S26). When comparing FC and CRP 374 clustering in the overlap cohort, there was little agreement (Figure 6 (A)). Whilst most 375 subjects in FC1 (71.6%) were also in CRP1 or CRP2, this relationship was not mirrored 376 as 81.2% of subjects in the latter two clusters were assigned to substantially different 377 FC clusters. However, CRP8 did overlap with elevated FC as the vast majority of 378 subjects in this cluster (80.8%) were assigned to either FC7 or FC8. For the 451 CD 379 subjects in the overlap cohort, largely similar patterns were observed (Figure 6 (B)). 380 For the 276 UC subjects, a substantial proportion of FC7 (45.1%) were assigned to 381 CRP1 (Figure 6 (C)), perhaps reflecting the characteristic where UC patients are less 382 likely to mount an abnormal CRP response than CD patients²¹.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.08.24316916; this version posted November 11, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in

perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .





384 Figure 6. Comparison between faecal calprotectin (FC) and processed CRP for models with chosen 385 specification (three NCS) assuming eight clusters. Results are reported based on the overlap cohort, 386 consisting of 808 subjects included in both the FC and CRP analysis. (A) all subjects; (B) Crohn's 387 disease; and (C) ulcerative colitis. Each segment denotes the size of the cluster whilst the alluvial 388 segments connecting the nodes visualises the number of subjects shared between clusters.

Discussion 389

390 We have characterised IBD behaviour using long-term longitudinal trends of objective 391 inflammatory markers routinely collected for clinical care. Our model has uncovered, 392 in a large IBD cohort, eight clusters with distinct inflammatory profiles based on FC and CRP, respectively. Our data highlights the heterogeneity of the disease course, 393

and present a novel approach for understanding real-world patterns of inflammatory activity in IBD patients. For the first time, we are capturing the dynamic nature of the disease in a more biologically nuanced way than traditional behaviour endpoints such as treatment escalation, surgery and "Montreal progression". Moreover, this represents a marked shift from the traditional symptom-based behaviour profiles exemplified by the IBSEN cohorts over a decade ago.^{9,10}

400

401 This study builds on our earlier proof-of-concept work, where we first demonstrated 402 the feasibility of using long-term individualised profiles of FC to cluster IBD patients 403 with CD.¹⁵ In our previous analysis we identified four distinct FC clusters: one cluster 404 with persistently high FC (non-remitters) and three clusters with different downward 405 longitudinal trends. Here, we expand upon this by considering a substantially larger 406 IBD cohort (FC cohort, n=1036; CRP cohort, n=1838), with longer follow-up, not 407 excluding patients based on indicators of disease severity, and also including patients 408 with UC or IBDU. In addition, we modelled CRP. This has generated results with 409 greater representativeness across IBD phenotypes, uncovering more granular 410 structure with eight distinct FC clusters rather than four. Whilst the modelling 411 suggested further partitioning of the FC data was possible (Figure S5, Figure S7 (A)), 412 we selected the eight-cluster model as a parsimonious choice that captured the key 413 inflammatory patterns (Figure 2). Notably, the CRP data also partitioned into eight 414 distinct clusters. Together, the clusters broadly fall into four patterns of inflammatory 415 behaviour mirroring those recognised by gastroenterologists managing IBD patients 416 (i) rapid remitters (FC1 and CRP2), (ii) delayed remitters (FC3, FC7 and CRP5), (iii) 417 relapsing-remitters (FC4-6 and CRP4 and CRP7), and (iv) non-remitters (FC8 and 418 CRP6 and CRP8). These classifications provide new insights into the inflammatory 419 course of IBD patients.

420

We observed broadly poor agreement between FC and CRP clusters, although there was overlap amongst those at the varying ends of the inflammatory spectrum with FC1 and FC2 correlating with CRP1 and CRP2 (rapid remitters with low cumulative inflammatory burden), and FC7 and FC8 correlating with CRP6-8 (non-remitters with high cumulative inflammatory burdens). Although, the lack of overlap between the two is not surprising. CRP is a marker of systemic inflammation, whilst FC is more specific for detecting inflammation at a mucosal level. Hence, why both biomarkers are

428 complimentary when monitoring patients with IBD. Studies have also shown that a
429 proportion of IBD patients will have lower CRP values at diagnosis,²² as is seen in
430 CRP1 which is significantly enriched for UC cases.

431

432 We observed several important observations regarding cluster assignment. All IBD 433 subtypes featured in each cluster, however cluster assignment was unevenly 434 distributed across CD, UC and IBDU (Figures S9 and S20). In CD, although patients 435 with L4 disease were less represented in FC1 and FC2 there was no association with 436 ileal versus colonic disease location. Males were slightly overrepresented in the FC8 437 cluster, which is characterised by persistently elevated values. Interestingly, in the 438 recent SEXEII study, they observed male patients with UC were more likely to have 439 extensive colonic involvement and abdominal surgery, which may account for this 440 finding.²³ CRP cluster membership was also associated with smoking and older age, 441 with both more likely to have higher inflammatory burdens.

442

443 Multiple lines of evidence, including the recent PROFILE study,²⁴ have clearly 444 demonstrated improved disease control with and outcomes in Crohn's patients 445 receiving early advanced therapy. As such, we anticipated that the biggest driver of 446 inflammatory patterns over time would be advanced therapies and that this effect 447 would be more pronounced in CD versus UC especially given the era of our cohort. 448 Indeed, rates of advanced therapy use were not homogenous across clusters or IBD 449 subtypes. Overall, in the FC cohort, advanced therapies were used in 49.6% of CD 450 patients of which 47.7% started AT in the first year. In FC1, similar rates of advanced 451 therapies were used for Crohn's patients, however they were used earlier, suggesting 452 their positive benefit in rapidly inducing and maintaining remission. Rates of advanced 453 therapy prescriptions for patients with CD in FC8 (persistently high levels of 454 inflammatory behaviour) were similar, but started later in the disease course, which 455 may have negatively affected the ability to bring about remission of disease. This 456 cluster may also represent a more refractory group of patients, with a higher risk 457 phenotype. Whilst this may provide additional support for the use of early advanced 458 therapy, a causal interpretation of these effects is not possible in this study, which is 459 based on observational data. Additional work with alternative cohorts where treatment 460 assignment is randomised is planned.

461 Our study also highlights the importance of using a probabilistic approach to account 462 for uncertainty in cluster assignments. Indeed, we observed higher uncertainty for 463 subjects with a shorter longitudinal follow-up, particularly for FC clusters (Figure 5). In 464 such cases, we cannot confidently assign individuals to a specific cluster. Instead, 465 cluster assignment probabilities are sometimes evenly split across multiple clusters. 466 mostly between those with similar earlier behaviour and inflammatory burdens. This 467 effect is less prominent in CRP cluster assignments, partly due to the more distinct early trajectories observed across CRP clusters. As such, we anticipate that a 468 469 multivariate approach which simultaneously considers other biomarkers of disease 470 activity, such as haemoglobin, albumin, and platelet count, may further increase the 471 robustness of cluster assignments. Such analysis may also consider pre-diagnostic 472 biomarker measurements, following the recent observations in Danish registry data,²⁵ 473 as well as metabolomics, genetics or microbiome data to inform cluster assignments. 474 Critically, this study is limited to data generated in a single health board in Scotland. 475 As such, independent validation of the identified longitudinal clusters will be necessary 476 before clinical implementation and to better understand the associations between 477 cluster assignment and disease phenotyping. Furthermore, there are limitations 478 associated with the use of observational data. Beyond advanced therapy use, the 479 frequency and amount of biomarker measurements is likely to be higher for those with 480 a more severe disease, and very mild cases may be excluded from our cohort. 481 Moreover, FC and CRP clusters cannot be directly used in a prognostic way. Indeed, 482 future work is needed to assess associations between clusters and IBD related 483 complications (such as steroid use, hospitalisations and surgery) but also non-484 conventional complications related to a high cumulative inflammatory burden, such as 485 major cardiovascular adverse events, neuropsychiatric illness, and malignancy. Such 486 analyses will ultimately support the development of a low cost clinical support tool to

487 help deliver precision medicine to patients with IBD.

Classifying patients by their inflammatory behaviour may better inform therapy decisions, including timing of advanced therapy, as well as monitoring and follow-up requirements. This is a paradigm shift in thinking about disease behaviour compared with the previous symptom based profiles first reported by the IBSEN cohorts. Moreover, this approach - rooted in data that is widely available in clinical settings²⁶ and probabilistic modelling- paves the way for predictive analytics integrated into a

494 clinical support tool for population wide risk stratification and individual patient level495 prognostication and treatment.

496 Data sharing statement

- 497 As the data collected for this study has been derived from unconsented patient data,
- it is not possible to share subject-level data with external entities. Detailed summary
- 499 level data is available online at <u>https://vallejosgroup.github.io/Lothian-IBDR</u>. The
- 500 code used to conduct the analysis is also publicly available
- 501 (https://github.com/VallejosGroup/Lothian-IBDR).

502 Author contributions

503 CAV, CWL, and NCC were involved in conceptualising the study. CRB, SO, ATE, GRJ, 504 and NP curated the data. NCC and CAV implemented computer code, tested existing 505 code components, and conducted the formal analysis and visualisation. The original 506 draft was written by NCC, NP, ATE, CMR, CWL and CAV with all authors involved in 507 reviewing and editing the manuscript. CWL and CAV provided supervisory support.

508 Competing interests

509 NP has served as a speaker for Janssen, Takeda and Pfizer. BG has acted as 510 consultant to Galapagos and Abbyie and as speaker for Abbyie, Jansen, Takeda, 511 Pfizer and Galapagos. G-RJ has served as a speaker for Takeda, Janssen, Abbvie, Fresnius and Ferring. CWL has acted as a speaker and/or consultant to AbbVie, 512 513 Janssen, Takeda, Pfizer, Galapagos, GSK, Gilead, Vifor Pharma, Ferring, Dr Falk, 514 BMS, Boehringer Ingelheim, Eli Lilly, Merck, Novartis, Sandoz, Celltrion, Cellgene, 515 Amgen, Samsung Bioepis, Fresenius Kabi, Tillotts, Kuma Health, Trellus Health and 516 Iterative Health. None of the other authors report any conflicts of interest.

517 Funding

518 CWL is funded by a UKRI (UK Research and Innovation) Future Leaders Fellowship 519 'Predicting outcomes in IBD' (MR/S034919/1). G-RJ is funded by a Wellcome Trust 520 Clinical Research Career Development Fellowship. NC-C was partially supported by 521 the Medical Research Council and The University of Edinburgh via a Precision 522 Medicine PhD studentship (MR/N013166/1).

References 523

- 525 1. Jones GR, Lyons M, Plevris N, et al. IBD prevalence in Lothian, Scotland, 526 derived by capture-recapture methodology. Gut. 2019;68(11):1953-1960. 527 doi:10.1136/gutinl-2019-318936
- 528 2. Hamilton B, Green H, Heerasing N, et al. Incidence and prevalence of 529 inflammatory bowel disease in Devon, UK. Frontline Gastroenterol. 530 2021;12(6):461-470. doi:10.1136/flgastro-2019-101369
- 531 3. Cushing K, Higgins PDR. Management of Crohn disease: A review. JAMA. 532 2021;325(1):69. doi:10.1001/jama.2020.18936
- 533 4. Gros B, Kaplan GG. Ulcerative colitis in adults: A review. JAMA. 534 2023;330(10):951. doi:10.1001/jama.2023.15389
- 535 5. Zhou RW, Harpaz N, Itzkowitz SH, Parsons RE. Molecular mechanisms in colitis-536 associated colorectal cancer. Oncogenesis. 2023;12(1):48. doi:10.1038/s41389-537 023-00492-0
- 538 6. Kennedy NA, Jones GR, Plevris N, Patenden R, Arnott ID, Lees CW. Association 539 between level of fecal calprotectin and progression of Crohn's disease. Clin 540 Gastroenterol Hepatol. 2019;17(11):2269-2276.e4. 541 doi:10.1016/j.cgh.2019.02.017
- 542 7. Plevris N, Fulforth J, Lyons M, et al. Normalization of fecal calprotectin within 12 543 months of diagnosis is associated with reduced risk of disease progression in 544 patients with Crohn's disease. Clin Gastroenterol Hepatol. 2021;19(9):1835-545 1844.e6. doi:10.1016/j.cgh.2020.08.022
- 546 8. Verstockt B, Noor NM, Marigorta UM, et al. Results of the Seventh Scientific 547 Workshop of ECCO: Precision medicine in IBD—Disease outcome and response 548 to therapy. J Crohns Colitis. 2021;15(9):1431-1442. doi:10.1093/ecco-jcc/jjab050
- 549 9. Solberg IC, Vatn MH, Høie O, et al. Clinical course in Crohn's disease: Results of 550 a Norwegian population-based ten-year follow-up study. Clin Gastroenterol 551 Hepatol. 2007;5(12):1430-1438. doi:10.1016/j.cgh.2007.09.002
- 552 10. Solberg IC, Lygren I, Jahnsen J, et al. Clinical course during the first 10 years of 553 ulcerative colitis: Results from a population-based inception cohort (IBSEN 554 Study). Scand J Gastroenterol. 2009;44(4):431-440. 555 doi:10.1080/00365520802600961
- 556 11. Peyrin-Biroulet L, Reinisch W, Colombel JF, et al. Clinical disease activity, C-557 reactive protein normalisation and mucosal healing in Crohn's disease in the 558 SONIC trial. Gut. 2014;63(1):88-95. doi:10.1136/gutjnl-2013-304984
- 559 12. Silverberg MS, Satsangi J, Ahmad T, et al. Toward an integrated clinical, 560 molecular and serological classification of inflammatory bowel disease: Report of

- 561 a working party of the 2005 Montreal World Congress of Gastroenterology. Can J 562 Gastroenterol Hepatol. 2005;19:5A-36A. doi:10.1155/2005/269076
- 563 13. Lennard-Jones JE. Classification of inflammatory bowel disease. Scand J 564 Gastroenterol. 1989;24(sup170):2-6. doi:10.3109/00365528909091339
- 565 14. Sorlie P, Wei GS. Population-based cohort studies: Still relevant? J Am Coll 566 Cardiol. 2011;58(19):2010-2013. doi:10.1016/j.jacc.2011.08.020
- 567 15. Constantine-Cooke N, Monterrubio-Gómez K, Plevris N, et al. Longitudinal fecal 568 calprotectin profiles characterize disease course heterogeneity in Crohn's 569 disease. Clin Gastroenterol Hepatol. 2023;21(11):2918-2927.e6. 570 doi:10.1016/j.cgh.2023.03.026
- 571 16. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using 572 latent classes and latent processes: The R package lcmm. J Stat Softw. 573 2017;78(2):1-56. doi:10.18637/jss.v078.i02
- 574 17. Stoica P, Selen Y. Model-order selection: A review of information criterion rules. 575 IEEE Signal Process Mag. 2004;21(4):36-47. doi:10.1109/MSP.2004.1311138
- 576 18. Proust-Lima C, Philipps V, Diakite A, Liquet B. Icmm: extended mixed models 577 using latent classes and latent processes. Published online 2021. https://cran.r-578 project.org/package=lcmm
- 579 19. Brunson JC. ggalluvial: layered grammar for alluvial plots. J Open Source Softw. 580 2020;5(49):2017. doi:10.21105/joss.02017
- 581 20. Constantine-Cooke N. datefixR: fix really messy dates in R. Published online 582 2024. doi:10.5281/zenodo.5655311
- 583 21. Clough J, Colwill M, Poullis A, Pollok R, Patel K, Honap S. Biomarkers in 584 inflammatory bowel disease: A practical guide. Ther Adv Gastroenterol. 585 2024;17:17562848241251600. doi:10.1177/17562848241251600
- 586 22. Henriksen M, Jahnsen J, Lygren I, et al. C-reactive protein: A predictive factor 587 and marker of inflammation in inflammatory bowel disease. Results from a 588 prospective population-based study. Gut. 2008;57(11):1518-1523. 589 doi:10.1136/gut.2007.146357
- 590 23. Gargallo-Puyuelo CJ, Ricart E, Iglesias E, et al. Sex-related differences in the 591 phenotype and course of inflammatory bowel disease: SEXEII study of ENEIDA. 592 Clin Gastroenterol Hepatol. 2024;22(11):2280-2290. 593 doi:10.1016/j.cgh.2024.05.013
- 594 24. Noor NM, Lee JC, Bond S, et al. A biomarker-stratified comparison of top-down 595 versus accelerated step-up treatment strategies for patients with newly 596 diagnosed Crohn's disease (PROFILE): A multicentre, open-label randomised 597 controlled trial. Lancet Gastroenterol Hepatol. 2024;9(5):415-427. 598

- 599 25. Vestergaard MV, Allin KH, Poulsen GJ, Lee JC, Jess T. Characterizing the pre-
- 600 clinical phase of inflammatory bowel disease. Cell Rep Med. 2023;4(11):101263. 601 doi:10.1016/j.xcrm.2023.101263
- 602 26. Markowetz F. All models are wrong and yours are useless: Making clinical 603 prediction models impactful for patients. Npj Precis Oncol. 2024;8(1):54.
- doi:10.1038/s41698-024-00553-6 604

Supplemental display items 606

Table of contents 607

Supplemental figures 608 609 Figure S1. Distribution of observation times for diagnostic FC • 610 Figure S2. Distribution of observation times for diagnostic CRP • 611 Figure S3. Illustration of how biomarker (faecal calprotectin or CRP) 612 observation times were adjusted based on time of diagnostic biomarker 613 observation 614 Figure S4. Distribution of diagnostic biomarker measurements across the 615 study cohort 616 Figure S5. Cluster trajectories obtained from LCMM assuming ten clusters fitted to FC data 617 618 Figure S6. Cluster trajectories obtained from LCMM assuming nine clusters 619 fitted to FC data 620 Figure S7. Alluvial plot visualising cluster assignment as the number of 621 assumed clusters increases 622 Associations with FC cluster membership 623 Figure S8. Age and sex 0 624 Figure S9. IBD type 0 625 Figure S10. Smoking for CD subjects 0 626 Figure S11. Montreal location for CD subjects 0 627 Figure S12. L4 for CD subjects 0 628 Figure S13. Montreal behaviour for CD subjects 0 629 Figure S14. Perianal disease for CD subjects 0 630 Figure S15. Smoking for UC subjects 0 631 Figure S16. Montreal extent for UC subjects 0 632 Figure S17. Cluster trajectories obtained from model assuming nine clusters 633 fitted to CRP data 634 Figure S18. Cluster trajectories obtained from model assuming seven clusters fitted to CRP data 635

636 Associations with CRP cluster membership

637	 Figure S19. Age and sex
638	 Figure S20. IBD type
639	 Figure S21. Smoking for CD subjects
640	 Figure S22. Montreal location for CD subjects
641	 Figure S23. L4 for CD subjects
642	 Figure S24. Montreal behaviour for CD subjects
643	• Figure S25. Time to first-line advanced therapy by CRP cluster assignment
644	• Figure S26. Proportion of individuals included in the overlap cohort by FC and
645	CRP cluster assignment
646	Supplemental tables

- 647 Table S1. Model statistics for models fitted to FC data •
- Table S2. Model statistics for models fitted to CRP data 648 •

Supplemental figures 649



650

651 Figure S1. Distribution of observation times for diagnostic faecal calprotectin (FC) relative to reported 652 date of diagnosis for the study cohort. Stratified by FC cluster assignment. Diagnostic FC was defined as the first FC test within ±90 days of diagnosis.







655 Figure S2. Distribution of observation times for diagnostic C-reactive protein (CRP) relative to 656 reported date of diagnosis for the study cohort. Stratified by CRP cluster assignment. Diagnostic CRP 657 was defined as the first CRP test within ±90 days of diagnosis.





calprotectin after applying a logarithmic transformation; (B) pre-processed (grouped into time intervals with the median used for multiple measurements) CRP after logarithmic transformation.



667 Figure S5. Cluster trajectories obtained from LCMM assuming ten clusters fitted to faecal calprotectin 668 data. Red lines indicate predicted mean cluster profiles with 95% confidence intervals. Dotted 669 horizontal lines indicate log(250µg/g). For visualisation purposes, pseudo subject-specific trajectories 670 have been generated by amalgamating observations from groups of six subjects.



671

672 Figure S6. Cluster trajectories obtained from LCMM assuming nine clusters fitted to faecal 673 calprotectin data. Red lines indicate predicted mean cluster profiles with 95% confidence intervals. 674 Dotted horizontal lines indicate log(250µg/g). For visualisation purposes, pseudo subject-specific

675 trajectories have been generated by amalgamating observations from groups of six subjects.



677 Figure S7. Alluvial plot demonstrating how cluster assignment changes as the number of assumed 678 clusters increases for the chosen models for (A) faecal calprotectin and (B) C-reactive protein. The 679 clusters found by the 8-cluster models are labelled.



681	Figure S8. (A) For each FC cluster, violin plots show the distribution of age at diagnosis across
682	subjects, highlighting median and interquartile ranges. (B) Forest plot showing the estimated effect
683	sizes and associated 95% confidence intervals for age in a multinomial logistic regression model that
684	uses FC cluster assignment as outcome. (C) For each FC cluster, panels show the proportion of
685	individuals with female and male sex. The dashed horizontal line represents overall proportions
686	across the entire FC cohort. (D) Forest plot showing the estimated effect sizes and associated 95%
687	confidence intervals for male sex versus females (baseline category) in a multinomial logistic
688	regression model that uses FC cluster assignment as outcome. In (B) and (C), effect sizes are with
689	respect to the reference cluster (in this case FC4). In both cases, the multivariate model includes age,
690	sex and IBD type as covariates. The dashed vertical lines are used as a reference to indicate no
691	effect.





692

693 Figure S9. (A) For each FC cluster, panels show the proportion of individuals with Crohn's disease, 694 ulcerative colitis and IBDU respectively. The dashed horizontal line represents overall proportions 695 across the entire FC cohort. (B) Forest plot showing the estimated effect sizes and associated 95% 696 confidence intervals for IBD type: ulcerative colitis and IBDU versus Crohn's disease (baseline 697 category) in a multinomial logistic regression model that uses FC cluster assignment as outcome. 698 Effect sizes are with respect to the reference cluster (in this case FC4). In both cases, the multivariate 699 model includes age, sex and IBD type as covariates. The dashed vertical lines are used as a 700 reference to indicate no effect.









713 714

715 individuals with Montreal Location recorded as L1, L2 and L3 respectively. The dashed horizontal line 716 represents overall proportions across the entire FC cohort. (B) Forest plot showing the estimated 717 effect sizes and associated 95% confidence intervals for Montreal Location: L2 and L3 versus L1 718 (baseline category) in a multinomial logistic regression model that uses FC cluster assignment as 719 outcome. Effect sizes are with respect to the reference cluster (in this case FC4). The multivariate 720 model includes age, sex, smoking, Montreal location (L1, L2, L3), upper gastrointestinal inflammation 721 (L4), perianal disease (yes, no) and Montreal behaviour (B1, B2/B3) as covariates. The dashed 722 vertical lines are used as a reference to indicate no effect.



723	
724	
725	
700	

Figure S12 Crohn's disease patients only. (A) For each FC cluster, panels show the proportion of individuals with Montreal L4 (upper gastrointestinal inflammation), recorded as present or non present. The dashed horizontal line represents overall proportions across the entire FC cohort. (B) 726 727 Forest plot showing the estimated effect sizes and associated 95% confidence intervals for L4: 728 present versus not present (baseline category) in a multinomial logistic regression model that uses FC 729 cluster assignment as outcome. Effect sizes are with respect to the reference cluster (in this case 730 FC4). The multivariate model includes age, sex, smoking, Montreal location (L1, L2, L3), upper 731 gastrointestinal inflammation (L4), perianal disease (yes, no) and Montreal behaviour (B1, B2/B3) as 732 covariates. The dashed vertical lines are used as a reference to indicate no effect.





734	Figure S13. Crohn's disease patients only. (A) For each FC cluster, panels show the proportion of
735	individuals with Montreal behaviour recorded as B1 or B2/B3 respectively. The dashed horizontal line
736	represents overall proportions across the entire FC cohort. (B) Forest plot showing the estimated
737	effect sizes and associated 95% confidence intervals for Montreal behaviour: B2/B3 versus B1
738	(baseline category) in a multinomial logistic regression model that uses FC cluster assignment as
739	outcome. Effect sizes are with respect to the reference cluster (in this case FC4). The multivariate
740	model includes age, sex, smoking, Montreal location (L1, L2, L3), upper gastrointestinal inflammation
741	(L4), perianal disease (yes, no) and Montreal behaviour (B1, B2/B3) as covariates. The dashed
742	vertical lines are used as a reference to indicate no effect.



743 744 Figure S14 Crohn's disease patients only. (A) For each FC cluster, panels show the proportion of 745 individuals with perianal disease recorded as present or not present. The dashed horizontal line 746 represents overall proportions across the entire FC cohort. (B) Forest plot showing the estimated 747 effect sizes and associated 95% confidence intervals for perianal disease: present versus not present 748 (baseline category) in a multinomial logistic regression model that uses FC cluster assignment as 749 outcome. Effect sizes are with respect to the reference cluster (in this case FC4). The multivariate 750 model includes age, sex, smoking, Montreal location (L1, L2, L3), upper gastrointestinal inflammation 751 (L4), perianal disease (yes, no) and Montreal behaviour (B1, B2/B3) as covariates. The dashed 752 vertical lines are used as a reference to indicate no effect.



753 754 755 Figure S15. Ulcerative colitis patients only. (A) For each FC cluster, panels show the proportion of 756 individuals with smoking behaviour recorded as "no" (no and never) and "yes" (current or previously 757 smoked) at diagnosis respectively. The dashed horizontal line represents overall proportions across 758 the entire FC cohort. (B) Forest plot showing the estimated effect sizes and associated 95% 759 confidence intervals for smoking: yes versus no (baseline category) in a multinomial logistic 760 regression model that uses FC cluster assignment as outcome. Effect sizes are with respect to the 761 reference cluster (in this case FC4). The multivariate model includes age, sex, smoking and Montreal 762 extent (E1, E2, E3). The dashed vertical lines are used as a reference to indicate no effect.



763 764 765 766 767 768 769 770

Figure S16. Ulcerative colitis patients only. (A) For each cluster, panels show the proportion of individuals with Montreal extent recorded as E1, E2 and E3 respectively. The dashed horizontal line represents overall proportions across the entire FC cohort. (B) Forest plot showing the estimated effect sizes and associated 95% confidence intervals for Montreal extent: E2 and E3 versus E1 (baseline category) in a multinomial logistic regression model that uses FC cluster assignment as outcome. Effect sizes are with respect to the reference cluster (in this case FC4). The multivariate model includes age, sex, smoking and Montreal extent (E1, E2, E3). The dashed vertical lines are 771 used as a reference to indicate no effect.



773 Figure S17. Cluster trajectories obtained from LCMM assuming nine clusters fitted to C-reactive 774 protein data. Red lines indicate predicted mean cluster profiles with 95% confidence intervals. Dotted 775 horizontal lines indicate log(5µg/mL). For visualisation purposes, pseudo subject-specific trajectories 776 have been generated by amalgamating observations from groups of six subjects.



777

778 Figure S18. Cluster trajectories obtained from LCMM assuming seven clusters fitted to C-reactive 779 protein data. Red lines indicate predicted mean cluster profiles with 95% confidence intervals. Dotted 780 horizontal lines indicate log(5µg/mL). For visualisation purposes, pseudo subject-specific trajectories 781 have been generated by amalgamating observations from groups of six subjects.



784 Figure S19. (A) For each CRP cluster, violin plots show the distribution of age at diagnosis across 785 subjects, highlighting median and interquartile ranges. (B) Forest plot showing the estimated effect 786 sizes and associated 95% confidence intervals for age in a multinomial logistic regression model that 787 uses CRP cluster assignment as outcome. (C) For each CRP cluster, panels show the proportion of 788 individuals with female and male sex. The dashed horizontal line represents overall proportions 789 across the entire CRP cohort. (D) Forest plot showing the estimated effect sizes and associated 95% 790 confidence intervals for male sex versus females (baseline category) in a multinomial logistic 791 regression model that uses CRP cluster assignment as outcome. In (B) and (C), effect sizes are with 792 respect to the reference cluster (in this case CRP8). In both cases, the multivariate model includes 793 age, sex and IBD type as covariates. The dashed vertical lines are used as a reference to indicate no 794 effect.





795

796 Figure S20. (A) For each CRP cluster, panels show the proportion of individuals with Crohn's disease, 797 ulcerative colitis and IBDU respectively. The dashed horizontal line represents overall proportions 798 across the entire CRP cohort. (B) Forest plot showing the estimated effect sizes and associated 95% 799 confidence intervals for IBD type: ulcerative colitis and IBDU versus Crohn's disease (baseline 800 category) in a multinomial logistic regression model that uses CRP cluster assignment as outcome. 801 Effect sizes are with respect to the reference cluster (in this case CRP8). In both cases, the 802 multivariate model includes age, sex and IBD type as covariates. The dashed vertical lines are used 803 as a reference to indicate no effect.





805	Figure S21. Crohn's disease patients only. (A) For each CRP cluster, panels show the proportion of
806	individuals with smoking behaviour recorded as "no" (no and never) and "yes" (current or previously
807	smoked) at diagnosis respectively. The dashed horizontal line represents overall proportions across
808	the entire CRP cohort. (B) Forest plot showing the estimated effect sizes and associated 95%
809	confidence intervals for smoking: yes versus no (baseline category) in a multinomial logistic
810	regression model that uses CRP cluster assignment as outcome. Effect sizes are with respect to the
811	reference cluster (in this case CRP8). The multivariate model includes age, sex, smoking, Montreal
812	location (L1, L2, L3), upper gastrointestinal inflammation (L4), and Montreal behaviour (B1, B2/B3) as
813	covariates. The dashed vertical lines are used as a reference to indicate no effect.



814

815 Figure S22. Crohn's disease patients only. (A) For each CRP cluster, panels show the proportion of 816 individuals with Montreal Location recorded as L1, L2 and L3 respectively. The dashed horizontal line 817 represents overall proportions across the entire CRP cohort. (B) Forest plot showing the estimated 818 effect sizes and associated 95% confidence intervals for Montreal Location: L2 and L3 versus L1 819 (baseline category) in a multinomial logistic regression model that uses CRP cluster assignment as 820 outcome. Effect sizes are with respect to the reference cluster (in this case CRP8). The multivariate 821 model includes age, sex, smoking, Montreal location (L1, L2, L3), upper gastrointestinal inflammation 822 (L4), and Montreal behaviour (B1, B2/B3) as covariates. The dashed vertical lines are used as a 823 reference to indicate no effect.





825	Figure S23. Crohn's disease patients only. (A) For each CRP cluster, panels show the proportion of
826	individuals with Montreal L4 (upper gastrointestinal inflammation), recorded as present or non
827	present. The dashed horizontal line represents overall proportions across the entire CRP cohort. (B)
828	Forest plot showing the estimated effect sizes and associated 95% confidence intervals for L4:
829	present versus not present (baseline category) in a multinomial logistic regression model that uses
830	CRP cluster assignment as outcome. Effect sizes are with respect to the reference cluster (in this
831	case CRP8). The multivariate model includes age, sex, smoking, Montreal location (L1, L2, L3), upper
832	gastrointestinal inflammation (L4), and Montreal behaviour (B1, B2/B3) as covariates. The dashed
833	vertical lines are used as a reference to indicate no effect.





835	Figure S24. Crohn's disease patients only. (A) For each CRP cluster, panels show the proportion of
836	individuals with Montreal behaviour recorded as B1 or B2/B3 respectively. The dashed horizontal line
837	represents overall proportions across the entire CRP cohort. (B) Forest plot showing the estimated
838	effect sizes and associated 95% confidence intervals for Montreal behaviour: B2/B3 versus B1
839	(baseline category) in a multinomial logistic regression model that uses CRP cluster assignment as
840	outcome. Effect sizes are with respect to the reference cluster (in this case CRP8). The multivariate
841	model includes age, sex, smoking, Montreal location (L1, L2, L3), upper gastrointestinal inflammation
842	(L4), and Montreal behaviour (B1, B2/B3) as covariates. The dashed vertical lines are used as a
843	reference to indicate no effect.



844

845 Figure S25. CRP cluster-specific cumulative distribution for first-line advanced therapy prescribing for 846 Crohn's disease (red) and ulcerative colitis (teal) subjects. Clusters are ordered from lowest (CRP1) to

- 847 highest (CRP8) cumulative inflammatory burden. The number of CD and UC subjects present in each
- 848 cluster is displayed as panel titles. Curves which would describe fewer than five subjects are not
- 849

shown.



850

851 Figure S26. (A) For each FC cluster, panels show the proportion of individuals included in the overlap 852 cohort, which consists of subjects included in both the FC and CRP analysis. The dashed horizontal 853 line represents overall proportions across the entire FC cohort. (B) As in (A), but focusing on CRP 854 clusters instead. The dashed horizontal line represents overall proportions across the entire CRP 855 cohort.

Supplemental tables 856

Clusters	Maximum log- likelihood	AIC	BIC
2	-16532-10	33094-20	33168.35
3	-16354-39	32754.78	32868-47
4	-16260-21	32582.42	32735.66
5	-16201.86	32481.71	32674-49
6	-16141.82	32377.63	32609-96
7	-16109-77	32329.54	32601.42
8	-16076-89	32279.77	32591.19
9	-16039.66	32221.33	32572.29
10	-16015-29	32188-59	32579.09

857

Table S1. Likelihood-based model statistics for LCMMs fitted to faecal calprotectin data across 2–10

858

clusters using the chosen model specification. Boldface font indicates optimal values.

Clusters	Maximum log- likelihood	AIC	BIC
2	-15451-41	30932.82	31015-57
3	-15314.76	30675.52	30802-40
4	-15240-15	30542.29	30713-30
5	-15173-04	30424.09	30639-23
6	-15106-25	30306-50	30565-78
7	-15051-43	30212.86	30516-27
8	-15016-96	30159.91	30507-45
9	-15014.76	30171.53	30563-20
10	-15014-40	30186-81	30622-61

859 860 Table S2. Likelihood-based model statistics for LCMMs fitted to processed CRP data across 2-10 clusters using the chosen model specification. Boldface font indicates optimal values.