

# Supplementary Information

## Supplementary Methods

### Study design

#### The PANTS study

PANTS is a UK-wide, prospective, observational cohort study of response to anti-TNF therapy in CD patients, previously described in detail by Kennedy et al.<sup>1</sup> and Sazonovs et al.<sup>27</sup> The study is registered with ClinicalTrials.gov identifier NCT03088449, and the protocol is available at <https://www.ibdresearch.co.uk/pants/>. Total enrolment was 1610 patients, who were at least 6 years old, had active luminal CD, and were naive to anti-TNF therapy.

The anti-TNF drugs evaluated were adalimumab (ADA), infliximab (IFX), and infliximab biosimilars. Patients received treatment according to standard dosing intervals (8 weeks for infliximab, 2 weeks for adalimumab), and were invited to attend up to ten major study visits over the course of the study, or until anti-TNF drug withdrawal. These visits were timed to allow sampling immediately prior to drug doses, and to allow the same visit structure to be used for all patients. Additional visits were scheduled in case of secondary LOR or premature study exit due to drug withdrawal, usually replacing the next scheduled major visit.

The target timings for the four major visits in the first year of follow-up in the PANTS study were week 0, week 14, week 30, and week 54. The week 0 major visit is the visit prior to the first dose of drug. Week 0 to week 14 is the induction period. After week 14, patients continued to take their drug for maintenance according to the standard dosing interval. For RNA-seq, whole blood samples at these four major visits were taken prior to the scheduled drug doses that aligned with those visits, labelled with the visit name, and preserved in Tempus Blood RNA Tubes.

## **Definition of primary response and primary non-response**

The definition of primary response and non-response was based on the clinical decision tree from Kennedy et al.<sup>1</sup>, as detailed in the main text (“RNA-seq sample selection”). Primary response and non-response could be assessed from week 12, with a final classification made by the scheduled week 14 visit. As PANTS was an observational study that continued until drug withdrawal, a patient’s clinician may have decided to continue anti-TNF therapy even if a patient demonstrated primary non-response, so it was possible for primary non-responders to remain in the study past week 14.

## **Selection of samples for RNA-seq**

Additional selection criteria were used to choose the subcohort of PANTS patients put forward for RNA-seq. Patients were required to be at least 16 years old, on either infliximab or adalimumab (not on an infliximab biosimilar), and to have an available baseline serum sample. Equal numbers of responders and non-responders to each drug were selected, excluding grey zone patients that only partially met the criteria for either primary non-response or response. Within the patients on infliximab, there was propensity score matching between primary non-responders and other patients based on baseline immunomodulator use, baseline steroid use, age, sex, and body mass index (BMI). Furthermore, primary non-responders were selected excluding patients known to be in remission at week 54, and primary responders were selected from patients known to be in remission by week 30 or week 54. The primary non-responders and responders in the RNA-seq subcohort thus represent phenotypic extremes of response.

It could not be guaranteed that study visits occurred on the target day specified in the protocol, thus samples were mapped to timepoints using windows around major visits defined by Kennedy et al.<sup>1</sup> Relative to the first dose of anti-TNF (day 0), these windows were: week 0 (week -4-0), week 14 (week 10-20), week 30 (week 22-38), and week 54 (week 42-66). Only a small minority of labelled major visit samples fell outside their respective windows, mostly for later timepoints where there was more variation around the target day. Samples taken at additional (LOR or exit) visits falling within one of the windows were mapped to that timepoint.

## Library preparation and sequencing

Total RNA was extracted following the Qiagen QIA Symphony instrument protocol (RNA Isolation PAX RNA CR22332 ID 2915). RNA was quantified with the ThermoFisher QuBit BR RNA (Q10211), and RNA integrity assessed with the Agilent RNA ScreenTape assay (5067-5579, 5067-5577, 5067-5576) on the Agilent 4200 TapeStation. Library preparation was performed using the Kapa mRNA HyperPrep Kit, including enrichment for messenger RNA (mRNA) using magnetic oligo-dT beads, depletion of ribosomal RNA (rRNA) and globin mRNA using the QIAseq FastSelect RNA Removal Kit, and adapter ligation with IDT xGEN Dual Index UMI adapters. Libraries were sequenced on the Illumina HiSeq 4000 with 75 bp paired-end reads.

## RNA-seq quantification and preprocessing

A total of 1141 samples from 396 patients were sequenced to a target minimum depth of 20 million total read pairs before deduplication. Sequencing data was demultiplexed with Picard.<sup>2</sup> Sequence quality, overrepresented sequences, adapter content, and sequence duplication rates were checked using FastQC.<sup>3</sup> Reads were mapped to GRCh38 using STAR (v2.6.1d)<sup>4</sup> and deduplicated to unique reads using UMI-tools.<sup>5</sup> Gene expression was quantified against the Ensembl 96 gene annotation with featureCounts (v1.6.4).<sup>6</sup>

Samples were filtered to remove outliers (>2 standard deviations from the mean) according to percentage of aligned reads in coding regions reported by Picard, percentage of unique reads, and number of unique reads. Samples that could not be mapped to a timepoint were removed. Samples with sex mismatch were removed. Samples from patients with grey zone primary response were removed. Samples for which there was missingness in the data matrix for clinical and cell proportion were removed. A total of 814 samples remained after filtering (Fig. S1a). The number of samples per patient ranged from one to four, with a median of three (Fig. S1b).

The Ensembl 96 gene annotation contains 58 884 genes, many of which are not expressed in whole blood. Effective library sizes were computed using the trimmed mean of M-values

(TMM) method in edgeR (v3.28.1),<sup>7</sup> then between-sample normalisation for library size was performed using edgeR::cpm, converting counts to counts per million (CPM). Genes with low expression were filtered, requiring >1.25 CPM in >10% of samples (1.25 CPM being approximately 10 counts at the median library size of 8 million unique mapped read pairs) and non-zero expression in >90% of samples. Globin genes and short non-coding RNAs (ncRNAs) were removed. A total of 15 511 genes remained after filtering. Finally, CPMs were converted to the log<sub>2</sub> scale, and precision weights to account for the expression mean-variance relationship were computed for each gene and sample using variancePartition::voomWithDreamWeights.<sup>8</sup>

## Differential gene expression

### Variable selection by variance components analysis

For each gene, the DGE model was a regression expressing the response variable (gene expression), as a linear function of predictor variables of interest (primary response status, drug, timepoint), and other selected predictor variables. Fig. S10 shows the correlation matrix of variables considered for selection. These included three variables from Kennedy et al.<sup>1</sup> associated with primary response: baseline immunomodulator use, smoking, and BMI. Also available were proportions of six common cell types in whole blood (CD4+ T cells, CD8+ T cells, B cells, NK cells, monocytes, granulocytes), estimated from whole blood Illumina MethylationEPIC methylation array data collected from the same patients and timepoints.<sup>9</sup> Estimates were computed using the Houseman method, which uses differentially methylated regions between immune cell types as cell type markers,<sup>10</sup> implemented in Minfi (minfi::estimateCellCounts).<sup>11</sup>

A variance components analysis was performed to quantify the proportion of expression variance explained by each variable for each gene using variancePartition,<sup>8</sup> as variables that do not explain much variation in the response would be unlikely to improve statistical efficiency if conditioned on in DGE models. The variance components analysis model was a linear mixed effects regression model with variables in Fig. S10 included as predictors. Additional categorical variables were included for patient and RNA-seq library preparation

plate. An additional continuous variable consisting of random numbers drawn from the standard normal distribution was included as a null (random\_numbers). Granulocyte proportion estimates were dropped to relieve perfect multicollinearity. Categorical variables were coded as random intercepts, and continuous variables as fixed effects—simulations from Hoffman et al.<sup>8</sup> showed variance proportion estimates were unbiased even when coding categorical variables with as few as two categories as random effects, as long as model parameters were estimated using maximum likelihood (ML) rather than restricted maximum likelihood (REML). It was also shown this approach avoids overestimates of variance proportions that occur if categorical variables with many levels are treated as fixed.

As downstream DGE methods required the same set of predictors for all genes, we aimed to select variables that explained substantial variance for many genes. Variables that explained the most variance on average were patient, cell proportions, and RNA-seq plate (Fig. S2). Some variables that did not explain more variance on average than the null nevertheless had high maximum values, indicating their importance for a relatively small number of genes. These included sex, library preparation protocol version (always the same within RNA-seq plates), and smoking status. However primary response status—a variable of interest—also fell into this group, so it was difficult to justify excluding all variables with lower median variance explained than the null. Consequently, all non-null variables in Fig. S2 were selected as predictors in downstream models apart from Ever\_Immunomodulator (whether the patient had ever had immunomodulator treatment), as that variable had both low median variance explained and was correlated with baseline immunomodulator use. The sample size is large compared to the number of degrees of freedom (df) lost by including predictors that may not be relevant for some genes.

As expected, cell proportions were among the biological factors that explained the most variance on average; they are one of the largest sources of variation in bulk blood expression data, and are a major driver of transcriptional response to immune perturbations.<sup>12</sup> In DGE analyses, two sets of separate models were fit including and excluding cell proportions as predictors, but otherwise identical. In models without cell proportions included, differential expression after drug perturbation could represent up or downregulation on a per-cell basis, but could also come from differences in cell proportions

induced by anti-TNF treatment. The estimates from models adjusted for cell proportions are more likely to reflect up or downregulation on a per-cell basis.

### Linear contrasts for pairwise group comparisons

Per-gene DGE models were fit in dream.<sup>13</sup> Like the variance components analysis models, these DGE models were linear mixed models:

$$y = 0 + \beta_{G_{trd}} G_{trd} + \sum^9 \beta_Z Z + \left( \sum^5 \beta_C C \right) + u + v + \varepsilon$$

where:

- The response variable is normalised gene expression  $y$ .
- 0 indicates a group means parameterisation, where the fixed intercept term is omitted to avoid over-parameterisation.
- $G_{trd}$  is a fixed effect for experimental group defined by combinations of the predictors of interest: timepoint  $t$  (week 0, 14, 30, 54), response  $r$  (PR, PNR), and drug  $d$  (infliximab, adalimumab). This is equivalent to having an intercept term and a three-way interaction between visit, response, and drug, including all lower order terms, but is more convenient for testing pairwise expression differences between groups, as the coefficient for each term is the estimate of mean expression for that group.
- $\sum^9 \beta_Z Z$  are the nine non-cell proportion fixed effects from variable selection (Fig. S2): sex (Sex), age of disease onset (Age\_of\_Onset), disease duration (Disease\_Duration), smoking history (Smoking\_History: ex, current or never), whether the patient has had surgery for CD (Crohns\_Surgery), whether the patient was on immunomodulator at baseline (On\_Immunomodulator\_At\_Baseline), whether the patient was on steroids at baseline (On\_Steroids\_At\_Baseline), BMI at baseline (Earliest\_BMI), and library preparation protocol version (Library\_Prep\_Protocol).
- $(\sum^5 \beta_C C)$  are the five cell proportion fixed effects from variable selection for NK cells, monocytes, B cells, CD4+ T cells, and CD8+ T cells.
- $u$  is a random intercept for RNA-seq plate (RNA\_Plate).
- $v$  is a random intercept for patient (PANTS.ID), nested inside RNA-seq plate.

Four sets of per-gene models were fit, with or without the cell proportion terms ( $\sum^5 \beta_c C$ ), and replacing  $\beta_{G_{trd}} G_{trd}$  (separate drug models) with  $\beta_{G_{tr}} G_{tr} + \beta_d d$  (pooled drug models) or not. Unlike with variance components analysis, to avoid small-sample bias in estimates of fixed effect standard errors, REML was used for estimation.<sup>14</sup> Specific hypotheses were tested using sum-to-zero contrasts, which are linear combinations of model coefficients with weights summing to zero. For example, to test for DGE between responders and non-responders to infliximab at baseline in the non-pooled model, we used a contrast where the weight for the week 0/responder/infliximab group coefficient was 1, the weight for the week 0/non-responder/infliximab group coefficient was -1, and all other coefficient weights were zero. To compute p-values, the contrast divided by its standard error was compared to the *t*-distribution using the Satterthwaite approximation for df. False discovery rate (FDR) was controlled with the Benjamini-Hochberg (BH) method,<sup>15</sup> with threshold set at 0.05, computed separately for each contrast.

### **Spline model comparing primary responders and non-responders over time**

The aim was to use expression data from all four timepoints to find genes associated with response, while avoiding a large number of pairwise comparisons. We fit a natural cubic spline (splines::ns, R v3.6.2)<sup>16</sup> to the study day to allow for non-linear trajectories of expression over time. We set two inner knots at week 14 (day 98) and week 30 (day 210) representing the centre of the study day windows for those two visits, as expression was expected to change after each drug dose. To include all data within the boundaries, the two boundary knots were set at the minimum and maximum values of study day rather than week 0 and week 54. A basis matrix was computed with `ns(Study_Day, knots = 7*c(14, 30))`, giving a matrix with three columns, each column being a transformation of study day. The columns were fit in the regression model in place of study day to allow for non-linear effects of study day on expression. The model form used was the same as that used for pairwise contrasts, except with  $\beta_{G_{trd}} G_{trd}$  replaced by  $\beta_r r + \sum^3 \beta_b b + \sum^3 \beta_{rb} rb + \beta_d d$ , where *r* is response status, *d* is drug,  $\sum^3 \beta_b b$  are the three columns of the basis matrix, and  $\sum^3 \beta_{rb} rb$  are the second-order interaction terms between response status and the basis matrix

columns. Separate sets of per-gene models were again fit with and without cell proportion terms ( $\sum^5 \beta_c C$ ).

When testing for response-associated differences in the spline parameters, the predictors of interest were the interaction terms  $\sum^3 \beta_{rb} r b$ . The three terms were tested jointly using an F-test, with the FDR threshold set at 0.05 (BH method). A significant result indicates a significant difference in the trajectory of expression over study day between responders and non-responders.

## **Clustering expression over all timepoints**

We clustered genes by their expression trajectories to define sets of genes with similar trajectories over time. This was done to aid the interpretation of significant genes from the cell proportion-adjusted spline model using gene set enrichment analysis. Expression data was converted to the CPM scale using TMM normalisation factors, then regressed against cell proportions. Residuals were centred and scaled per gene. A distance matrix was computed using  $1 - r$  as the distance metric, where  $r$  is the Pearson correlation. Hierarchical clustering was performed with complete agglomeration for inter-cluster distance (`fastcluster::hclust(method = "complete")`).<sup>17</sup> The optimal number of clusters was assessed by the gap statistic (`factoextra::fviz_nbclust(method = "gap_stat", nboot = 500)`), which determines when the change in within-cluster dispersions are no longer significantly improved by increasing the number of clusters.<sup>18</sup> The default "firstSEmax" criteria was used to choose the optimal number of clusters  $k$ , which finds the first local maximum at  $m$  clusters where  $Gap(m) \geq Gap(m + 1)$ , then finds the smallest  $k: 1 \leq k \leq m$  such that  $Gap(k)$  is not less than  $Gap(m)$  minus the bootstrapped standard error of  $Gap(m)$ . The hierarchical clustering tree was then cut into  $k$  clusters.

## **Gene set enrichment analyses**

### **Ranked gene set enrichment**

The gene sets used for ranked gene set enrichment analyses were blood transcription modules (BTMs) available in the `tmod` package.<sup>19</sup> Modules are "specific to the context of



immune responses in blood tissue”: sets of genes with transcriptional and functional similarities across a variety of healthy, diseased, and stimulated conditions.<sup>20</sup> The 260 modules from Chaussabel et al.<sup>21</sup> (prefixed “DC”) were constructed by unsupervised clustering of 239 PBMC transcriptomes from multiple disease datasets, then annotated by data mining of gene names in PubMed abstracts. The 346 modules from Li et al.<sup>20</sup> (prefixed “LI”) were constructed from coexpression analysis of approximately 30 000 blood transcriptomes, then annotated making use of Gene Ontology (GO) terms, cell type-specific markers, pathway databases, and manual literature searches. Li et al. modules are better annotated and were used as the default module set for gene set enrichment.

Ranked gene set enrichment analyses were conducted using `tmod::tmodCERNOtest`, which is a non-parametric test for enrichment of small ranks within specific sets of genes compared to all genes, after the genes are ranked by some metric. The CERNO statistic for a gene set is:

$$-2 \sum_{i=1}^n \ln \frac{r_i}{N} \sim \chi^2(2n)$$

where  $n$  is the number of genes in the set,  $N$  is the number of measured genes in the experiment, and  $r_i$  is the rank of the  $i$ th gene in the set. For each contrast, as the t-statistics are not comparable between genes due to the use of approximate df, genes were ranked by the signed z-score reported by `dream`, which is a monotonic transformation of the p-value. Similarly, moderated F-statistics from the spline model are not comparable between genes, so we used the signed F-statistics computed by `dream` from the transformation of the p-value. FDR control for the number of gene sets tested was performed using the BH procedure, separately for each tested contrast. `tmod::tmodCERNOtest` is one-sided and only considers enrichment of small ranks when computing significance. As genes can be down or upregulated, separate tests were performed sorting genes in ascending and descending order, and the more significant result was used to determine the overall direction of effect for each gene set. The effect size of gene set enrichments can be quantified with the area under the curve (AUC), computed from  $U$ , the test statistic from a Mann-Whitney U test (a.k.a. Wilcoxon rank-sum test):

$$U = n(N - n) + \frac{n(n + 1)}{2} - \sum_{i=1}^n r_i$$

Then  $AUC = U/(n(N - n)) \in [0, 1]$ , with significant results from the one-sided `tmod::tmodCERNOtest` having  $AUC > 0.5$ .

## Gene set overrepresentation

Gene set overrepresentation analyses for BTMs were run with `tmod::tmodHGtest`, which implements the hypergeometric test for enrichment of a query gene set within a gene module, controlling the FDR at 0.05 using the BH method. The 15 511 genes assayed by RNA-seq were used as a custom background set.

Gene set overrepresentation analyses were also run using `gprofiler2::gost`, which implements a similar test using publicly-available gene set databases including Gene Ontology, KEGG Reactome, WikiPathways, miRNA targets from miRTarBase, and regulatory motif matches from TRANSFAC. Multiple testing was controlled at  $\alpha = 0.05$  using the `g:SCS` method (`correction_method = "g_SCS"`), with genes assayed by RNA-seq provided as a custom background (`domain_scope = "custom_annotated"`).<sup>22</sup>

## Prediction of primary response status from expression

### Computing module scores by single-sample GSEA

The single-sample gene set enrichment analysis method (ssGSEA 2.0, <https://github.com/broadinstitute/ssGSEA2.0/>, commit `b3d035ed31043b277512fd4f08e90f4c1728beec`) was used to compute an expression score for each module in each RNA-seq sample. This module score is a measure of coordinate up or downregulation of the genes in the module in a particular sample as compared to other samples based on expression ranks. Combining the Li et al.<sup>20</sup> and Chaussabel et al.<sup>21</sup> datasets, 427 modules had sufficient overlap with the 15 511 genes in the filtered expression matrix for ssGSEA scores to be calculated.

## Model training and internal validation

Models were built to predict binary primary response status based on baseline expression and clinical variables. The predictors included the 427 module expression scores, as well as clinical variables (drug, sex, age of onset, disease duration, smoking status, previous CD surgery, immunomodulator use at baseline, steroid use at baseline, BMI) and cell proportions (CD4+ T cells, CD8+ T cells, B cells, NK cells, monocytes). At the week 0 timepoint, complete data was available for 152 responders and 116 non-responders; at week 14, complete data was available for 142 responders and 104 non-responders.

Model training and evaluation was performed using caret (v6.0-8).<sup>23</sup> Pre-processing was performed to remove zero variance predictors, then predictors were centred and scaled. To avoid lack of power during model development, no data splitting was performed (TRIPOD model type 1b).<sup>24</sup> The predictive methods used were: penalised and regularised logistic regression methods (plr, regLogistic, glmnet), parallel random forest (parRF), eXtreme Gradient Boosting (xgbTree), support vector machines with a radial basis (svmRadial), k-nearest neighbours (knn), naive Bayes (naive\_bayes), and Gaussian process models (gaussprLinear). To evaluate the incremental advantage of adding gene expression information, each method was trained with each of three sets of predictors. Initially only clinical variables were used as predictors, then cell proportions were added to the predictor set, and finally ssGSEA module scores were added. Model parameters were tuned by resampling (50 bootstraps per method, predictor set, and parameter combination; trainControl(method = "boot")), optimising using the AUC metric (train(method = "ROC")). Each bootstrap involves resampling the dataset with replacement to obtain a bootstrap dataset of the same size (on average 63.2% of observations will be sampled one or more times), training a predictive model on the bootstrap dataset, then evaluating the model AUC on the observations that were not sampled (the "out-of-bag" observations, on average 36.8% of observations). The parameter combination with the highest mean AUC over 50 bootstraps was selected for the final model. The resampling metrics (AUCs) from the final model per method and predictor set were compared to identify the best performing models. Bootstrapping provides internal estimates of model performance without data splitting, which is undesirable given the relatively small sample size.<sup>25</sup>

A single AUC for each final tuned model was computed after merging all 50 bootstraps. ROC curves were plotted with pROC (v1.18.0).<sup>26</sup> Pairwise tests for differences in AUC between ROC curves were performed with `pROC::roc.test(method = "bootstrap", boot.n = 2000)`, which uses bootstrapping to estimate the standard deviation of the difference in AUCs.

## Supplementary Results

### Evaluating possible contributors to heterogeneity of effects between drug subgroups

We considered factors that may contribute to the heterogeneity observed for baseline models unadjusted for cell composition: 859 differentially expressed genes for infliximab patients versus only one for adalimumab patients. Power differences related to sample size were not the sole contributor. Although there was a greater sample size for the infliximab subgroup (Fig. S1a), heterogeneity remained after randomly downsampling the infliximab subgroup to the same size as the adalimumab subgroup (Fig. S5). To examine the effect of adjusting for clinical variables on the association between response and expression, a set of models was run excluding all clinical covariates, such that the only predictor terms in the model were grouping variables (timepoint, response status, drug), technical terms (library preparation protocol and plate), and the random intercept for patient. A set of models was run including additional two-way interaction terms between drug and components of the propensity score used to match IFX patients during RNA-seq sample selection: sex, age, immunomodulator usage at baseline, steroid usage at baseline, and BMI. This allows coefficient estimates for the propensity score terms to vary between drugs. These two sets of models were run per drug subgroup, adjusted or unadjusted for cell composition. Heterogeneity remained regardless of model form, with a high number of single-gene associations uniquely observed in the IFX subgroup unadjusted for cell composition (Fig. S4). We also considered adjusting for week 14 drug level due to its association with PNR,<sup>1</sup> but this was precluded by substantial missingness in week 14 drug level data (196/814 samples with missing data). Inclusion of post-treatment variables in the model is also undesirable

when testing for baseline associations with response. Finally, we directly tested for differences in cell composition at baseline between study groups. Infliximab responders had significantly lower proportions of CD4+ and CD8+ T cells, and higher proportions of granulocytes compared to infliximab non-responders at baseline (Mann-Whitney test, Bonferroni-corrected  $p = 0.028$ ). These differences were not observed in adalimumab patients (Fig. S11). Thus associations between baseline expression of immune gene modules and post-induction primary response may be partially mediated by baseline differences in cell composition, and particularly in the infliximab subgroup. In the main manuscript, we focus on models adjusting for cell composition, which allows a more sensible interpretation for analyses where expression data for both drugs are pooled to increase sample size.

## Supplementary References

1. Kennedy, N.A., Heap, G.A., Green, H.D., Hamilton, B., Bewshea, C., Walker, G.J., Thomas, A., Nice, R., Perry, M.H., Bouri, S., et al. (2019). Predictors of anti-TNF treatment failure in anti-TNF-naive patients with active luminal Crohn's disease: a prospective, multicentre, cohort study. *The Lancet Gastroenterology & Hepatology* 4, 341–353. 10.1016/S2468-1253(19)30012-3.
2. Broad Institute (2019). Picard toolkit. <https://broadinstitute.github.io/picard/>.
3. Simon Andrews (2010). FastQC: A Quality Control tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
4. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. 10.1093/bioinformatics/bts635.
5. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* 27, 491–499. 10.1101/gr.209601.116.
6. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. 10.1093/bioinformatics/btt656.

7. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* *11*, R25. 10.1186/gb-2010-11-3-r25.
8. Hoffman, G.E., and Schadt, E.E. (2016). variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* *17*. 10.1186/s12859-016-1323-z.
9. Lin, S., and Hannon, E.J. (2022). Understanding the molecular mechanisms of anti-TNF treatment failure: Whole blood DNA methylation changes are associated with primary non-response to anti-TNF treatment in patients with Crohn's disease. Manuscript in preparation.
10. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., and Kelsey, K.T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* *13*, 86. 10.1186/1471-2105-13-86.
11. Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., and Irizarry, R.A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* *30*, 1363–1369. 10.1093/bioinformatics/btu049.
12. Piasecka, B., Duffy, D., Urrutia, A., Quach, H., Patin, E., Posseme, C., Bergstedt, J., Charbit, B., Rouilly, V., MacPherson, C.R., et al. (2018). Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proceedings of the National Academy of Sciences* *115*, E488–E497. 10.1073/pnas.1714765115.
13. Hoffman, G.E., and Roussos, P. (2020). dream: Powerful differential expression analysis for repeated measures designs. *Bioinformatics*. 10.1093/bioinformatics/btaa687.
14. McNeish, D. (2017). Small Sample Methods for Multilevel Modeling: A Colloquial Elucidation of REML and the Kenward-Roger Correction. *Multivariate Behavioral Research* *52*, 661–670. 10.1080/00273171.2017.1344538.
15. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* *57*, 289–300.
16. R Core Team (2019). R: A language and environment for statistical computing.

17. Müllner, D. (2013). **fastcluster** : Fast Hierarchical, Agglomerative Clustering Routines for *R* and *Python*. *J. Stat. Soft.* 53. 10.18637/jss.v053.i09.
18. Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63, 411–423. 10.1111/1467-9868.00293.
19. Weiner 3rd, J., and Domaszewska, T. (2016). tmod: an R package for general and multivariate enrichment analysis. 10.7287/peerj.preprints.2420v1.
20. Li, S., Roupahel, N., Duraisingham, S., Romero-Steiner, S., Presnell, S., Davis, C., Schmidt, D.S., Johnson, S.E., Milton, A., Rajam, G., et al. (2013). Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nature Immunology* 15, 195–204. 10.1038/ni.2789.
21. Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N., Stichweh, D., Blankenship, D., Li, L., Munagala, I., et al. (2008). A Modular Analysis Framework for Blood Genomics Studies: Application to Systemic Lupus Erythematosus. *Immunity* 29, 150–164. 10.1016/j.immuni.2008.05.012.
22. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* 47, W191–W198. 10.1093/nar/gkz369.
23. Kuhn, M. (2020). caret: Classification and regression training.
24. Collins, G.S., Reitsma, J.B., Altman, D.G., and Moons, K. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 13, 1. 10.1186/s12916-014-0241-z.
25. Steyerberg, E.W., and Jr, F.E.H. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 8.
26. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12, 77. 10.1186/1471-2105-12-77.

27. Sazonovs, A., Kennedy, N.A., Moutsianas, L., Heap, G.A., Rice, D.L., Reppell, M., Bewshea, C.M., Chanchlani, N., Walker, G.J., Perry, M.H., et al. (2019). HLA-DQA1\*05 Carriage Associated With Development of Anti-Drug Antibodies to Infliximab and Adalimumab in Patients With Crohn's Disease. *Gastroenterology*. 10.1053/j.gastro.2019.09.041.

## Supplementary Figure Captions

*See accompanying files for Supplementary Figures.*

**Figure S1. Distribution of PANTS RNA-seq samples over study timepoints. (a)** RNA-seq samples stratified by timepoint and study group. Windows from Kennedy et al. for the four major PANTS visits are coloured in grey. Samples mostly come from major visits, but a small number of LOR and exit visit samples were also included. **(b)** Numbers of patients with samples at multiple timepoints. Filled circles indicate the presence of at least one sample at the corresponding timepoint.

**Figure S2. Distributions of per-gene percentage of variance in expression explained by study variables.**

Variables are ordered by the median of per-gene variance explained estimates from variance components analysis. PANTS.ID = patient ID, NK = NK cell, Gran = granulocyte, Mono = monocyte, Bcell = B cell, CD4T = CD4<sup>+</sup> T cell, CD8 = CD8<sup>+</sup> T cell. random\_numbers is a null drawn from the standard normal distribution. Variables other than random\_numbers and Ever\_Immunomodulator were retained for use in downstream DGE and prediction models.

**Figure S3. Baseline expression associated with primary response, unadjusted for cell composition.**

**(a)** Volcano plots of DGE between responders (PR) and non-responders (PNR) at week 0; for infliximab (IFX), adalimumab (ADA), or with drug subgroups pooled. Annotated genes show significant associations from this study, and previously reported associations from the literature in both blood and gut biopsies. Dashed line shows significance threshold at FDR = 0.05. **(b)** Top gene modules differentially expressed between PR and PNR at week 0. Columns correspond to results for infliximab (IFX), adalimumab (ADA), difference between IFX and ADA (IFX – ADA, i.e. drug-by-response interaction), and pooled drug analyses. The top 30 modules ranked by minimum FDR in any column are shown. Dashed lines show significance thresholds at FDR = 0.05.

**Figure S4. Baseline gene expression differences between responders and non-responders with varying model specification.**



Volcano plots of DGE between responders (PR) and non-responders (PNR) at week 0; for adalimumab (ADA) and infliximab (IFX) subgroups. Columns correspond to cell proportion adjusted and cell proportion unadjusted models with either clinical variables excluded, or additional interaction terms between drug and variables used in propensity score matching during RNA-seq sample selection. Annotated genes show significant associations from this study, and previously reported associations from the literature in both blood and gut biopsies. Dashed line shows significance threshold at FDR = 0.05.

**Figure S5. Baseline gene expression differences between responders and non-responders, with downsampling to match baseline IFX and ADA subgroup sample sizes.**

Volcano plots of DGE between responders (PR) and non-responders (PNR) at week 0; for infliximab (IFX), adalimumab (ADA), or with drug subgroups pooled. Downsampling was performed, randomly selecting 66/86 IFX PR baseline samples and 57/59 IFX PNR baseline samples to match the corresponding ADA group sizes. Annotated genes show significant associations from this study, and previously reported associations from the literature in both blood and gut biopsies. Dashed line shows significance threshold at FDR = 0.05.

**Figure S6. Barcode plots showing selected gene modules with greater expression at baseline in primary responders compared to primary non-responders.** Genes were ranked in descending order by DGE z-statistic, with coloured bars indicating the rank of genes in a module. Curves show the cumulative fraction of genes in a module at a particular rank threshold. The area under the curve (AUC) reflects the effect size of the module association. Diagonal line shows the null of randomly-distributed ranks.

**Figure S7. Top modules differentially expressed between week 14 and week 0 in a) responders (PR) and b) non-responders (PNR), adjusted for cell composition.**

Columns correspond to results for infliximab (IFX), adalimumab (ADA), difference between IFX and ADA (IFX – ADA, i.e. drug-by-response interaction), and pooled drug analyses. The top 30 modules ranked by minimum FDR in any column are shown. Dashed lines show significance thresholds at FDR = 0.05.

**Figure S8. Distributions of model specificity, sensitivity and AUC (ROC) from resampling (50 bootstraps) for each combination of algorithm and predictor set at baseline (week 0).**

Metrics shown are the resampling metrics from the final tuned models. The prefixes “clinical”, “cellProps”, and “moduleScores” correspond to the inclusion of clinical variables, cell proportion estimates, and expression module scores in the predictor datasets respectively.

Suffixes indicate the method used (R 3.6, caret package).

**Figure S9. Distributions of model specificity, sensitivity and AUC (ROC) from resampling (50 bootstraps) for each combination of algorithm and predictor set at week 14.**

Metrics shown are the resampling metrics from the final tuned models. The prefixes “clinical”, “cellProps”, and “moduleScores” correspond to the inclusion of clinical variables, cell proportion estimates, and expression module scores in the predictor datasets respectively.

Suffixes indicate the method used (R 3.6, caret package).

**Figure S10. Correlation matrix of variables considered as potential predictor variables for DGE models.**

NK = NK cell, Gran = granulocyte, Mono = monocyte, Bcell = B cell, CD4T = CD4+ T cell, CD8 = CD8+ T cell.

**Figure S11. Differences in baseline cell proportion estimates between study groups.**

NK = NK cell, Gran = granulocyte, Mono = monocyte, Bcell = B cell, CD4T = CD4+ T cell, CD8 = CD8+ T cell. Nominal significance of pairwise comparisons using the Mann-Whitney test (R 3.6, wilcox.test(paired=F)) is shown: ns 0.05-1, \* 0.01-0.05, \*\* 0.001-0.01, \*\*\* 0.0001-0.001, \*\*\*\* <0.0001.

## Supplementary Tables

**Table S1. The UK Inflammatory Bowel Disease Pharmacogenetics Study Group.** All UK gastroenterologists were invited to participate in the PANTS study, which was promoted through the UK National Institute for Health Research and the British Society of Gastroenterology.

Hospital or Trust name	City	Name	Job Title
Tameside Hospital NHS Foundation Trust	Ashton U Lyne	Dr Vinod Patel	Consultant Gastroenterologist
Basildon and Thurrock University Hospitals NHS Foundation Trust	Basildon	Dr Zia Mazhar	Consultant Gastroenterologist
Hampshire Hospitals NHS Foundation Trust	Basingstoke	Dr Rebecca Saich	Consultant Gastroenterologist

Royal United Hospital	Bath	Dr Ben Colleypriest	Consultant Gastroenterologist
Ulster Hospital	Belfast	Dr Tony C Tham	Consultant Gastroenterologist
University Hospital's Birmingham NHS Foundation Trust	Birmingham	Dr Tariq H Iqbal	Consultant Gastroenterologist
East Lancashire NHS Teaching Trust	Blackburn	Dr Vishal Kaushik	Consultant Gastroenterologist
Blackpool Teaching Hospitals NHS Foundation Trust	Blackpool	Dr Senthil Murugesan	Consultant Gastroenterologist
Bolton NHS Trust	Bolton	Dr Salil Singh	Consultant Gastroenterologist
Royal Bournemouth Hospital	Bournemouth	Dr Sean Weaver	Consultant Gastroenterologist
Bradford Teaching Hospitals Foundation Trust - (St Lukes Hospital & Bradford Royal Infirmary)	Bradford	Dr Cathryn Preston	Consultant Gastroenterologist
Brighton and Sussex University Hospitals NHS Trust	Brighton	Dr Assad Butt	Paediatric Consultant Gastroenterologist
Brighton and Sussex University Hospitals NHS Trust	Brighton	Dr Melissa Smith	Consultant Gastroenterologist
University Hospitals Bristol NHS Foundation Trust	Bristol	Dr Dharamveer Basude	Consultant Paediatric Gastroenterologist
University Hospitals Bristol NHS Foundation Trust	Bristol	Dr Amanda Beale	Consultant Gastroenterologist
Frimley Park Hospital NHS Foundation Trust	Camberley	Dr Sarah Langlands	Consultant Gastroenterologist
Frimley Park Hospital NHS Foundation Trust	Camberley	Dr Natalie Direkze	Consultant gastroenterologist
Cambridge University Hospitals NHS Foundation Trust	Cambridge	Dr Miles Parkes	Consultant Gastroenterologist
Cambridge University Hospitals NHS Foundation Trust	Cambridge	Dr Franco Torrente	Consultant Paediatric Gastroenterologist
Cambridge University Hospitals NHS Foundation Trust	Cambridge	Dr Juan De La Revella Negro	Research fellow
North Cumbria University Hospitals NHS Trust	Carlisle	Dr Chris Ewen MacDonald	Consultant Gastroenterologist
Ashford & St Peter's Hospitals NHS Foundation Trust	Chertsey	Dr Stephen M Evans	Consultant Gastroenterologist
St Peter's Hospital	Chertsey	Dr Anton V J Gunasekera	Consultant Gastroenterologist
Ashford & St Peter's Hospitals NHS Foundation Trust	Chertsey	Dr Alka Thakur	Paediatric Consultant
Chesterfield Royal NHS Foundation Trust	Chesterfield	Dr David Elphick	Consultant Gastroenterologist
Colchester Hospital University NHS Foundation Trust	Colchester	Dr Achuth Shenoy	Consultant Gastroenterologist
University Hospitals Coventry and Warwickshire NHS Trust	Coventry	Prof Chuka U Nwokolo	Consultant Gastroenterologist
County Durham and Darlington NHS Foundation Trust	Darlington	Dr Anjan Dhar	Consultant Gastroenterologist & Hon. Clinical Lecturer
Derby Hospital NHS Foundation Trust	Derby	Dr Andrew T Cole	Consultant Gastroenterologist
Doncaster and Bassetlaw Hospitals NHS Foundation Trust	Doncaster	Dr Anurag Agrawal	Consultant Gastroenterologist
Dorset County Hospital NHS Foundation Trust	Dorchester	Dr Stephen Bridger	Consultant Gastroenterologist
Dorset County Hospitals Foundation Trust	Dorchester	Dr Julie Doherty	Paediatric Consultant
Dudley Group NHS Foundation Trust	Dudley	Dr Sheldon C Cooper	Consultant Gastroenterologist

Russells Hall Hospital, The Dudley Group NHS Foundation Trust	Dudley	Dr Shanika de Silva	Consultant Gastroenterologist
Ninewells Hospital & Medical School	Dundee	Dr Craig Mowat	Consultant Gastroenterologist
East Sussex Healthcare Trust	Eastborne	Dr Phillip Mayhead	Consultant Gastroenterologist
NHS Lothian	Edinburgh	Dr Charlie Lees	Consultant Gastroenterologist and Honorary Senior Lecturer
NHS Lothian	Edinburgh	Dr Gareth Jones	Research fellow
Royal Devon and Exeter NHS Foundation Trust	Exeter	Dr Tariq Ahmad	Consultant Gastroenterologist
Royal Devon and Exeter NHS Foundation Trust	Exeter	Dr James W Hart	Consultant Paediatrician
Glasgow Royal Infirmary	Glasgow	Dr Daniel R Gaya	Consultant Gastroenterologist
Royal Hospital for Children	Glasgow	Prof Richard K Russell	Consultant Paediatric Gastroenterologist
Royal Hospital for Children	Glasgow	Dr Lisa Gervais	Research fellow
Gloucestershire Hospitals NHS Trust	Gloucester	Dr Paul Duncley	Consultant Gastroenterologist
United Lincolnshire Hospitals NHS Trust	Grantham	Dr Tariq Mahmood	Consultant Gastroenterologist
James Paget University Hospitals NHS Foundation Trust	Great Yarmouth	Dr Paul J R Banim	Consultant Gastroenterologist
Calderdale and Huddersfield NHS Trust	Halifax	Dr Sunil Sonwalkar	Consultant Gastroenterologist
Princess Alexandra Hospital NHS Trust	Harlow	Dr Deb Ghosh	Consultant Gastroenterologist
Princess Alexandra Hospital NHS Trust	Harlow	Dr Rosemary H Phillips	Consultant Gastroenterologist
Hull and East Yorkshire NHS Trust	Hull	Dr Amer Azaz	Paediatric Consultant Gastroenterologist
Hull and East Yorkshire NHS Trust	Hull	Dr Shaji Sebastian	Consultant Gastroenterologist
Airedale NHS Foundation Trust	Keighley	Dr Richard Shenderey	Consultant Gastroenterologist
Crosshouse Hospital	Kilmarnock	Dr Lawrence Armstrong	Consultant Paediatrician
Crosshouse Hospital	Kilmarnock	Dr Claire Bell	Research fellow
The Queen Elizabeth Hospital NHS Foundation Trust	Kings Lynn	Dr Radhakrishnan Hariraj	Consultant Gastroenterologist
Kingston Hospital NHS Trust	Kingston upon Thames	Dr Helen Matthews	Consultant Gastroenterologist
NHS Fife	Kirkcaldy	Dr Hasnain Jafferbhoy	Consultant Gastroenterologist
Leeds Teaching Hospitals NHS Trust	Leeds	Dr Christian P Selinger	Consultant Gastroenterologist
Leeds Teaching Hospitals NHS Trust	Leeds	Dr Veena Zamvar	Paediatric Consultant Gastroenterologist
University Hospitals of Leicester NHS Trust	Leicester	Prof John S De Caestecker	Consultant Gastroenterologist
University Hospitals of Leicester NHS Trust	Leicester	Dr Anne Willmott	Paediatric Consultant Gastroenterologist
Mid Cheshire Hospitals NHS Foundation Trust	Leighton	Mr Richard Miller	Research Nurse
United Lincolnshire Hospitals NHS Trust	Lincoln	Dr Palani Sathish Babu	Consultant Gastroenterologist
Alder Hey Childrens Hospital	Liverpool	Dr Christos Tzivnikos	Consultant Paediatric Gastroenterologist
University College London Hospitals NHS Foundation Trust	London	Dr Stuart L Bloom	Consultant Gastroenterologist

Kings College Hospital NHS Foundation Trust	London	Dr Guy Chung-Faye	Consultant Gastroenterologist
Royal London Childrens Hospital, Barts Health NHS Trust	London	Prof Nicholas M Croft	Paediatric Consultant Gastroenterologist
Chelsea & Westminster Hospital	London	Dr John ME Fell	Consultant Paediatric Gastroenterologist
Chelsea and Westminster Hospital NHS Foundation	London	Dr Marcus Harbord	Consultant Gastroenterologist
North West London Hospitals NHS Trust	London	Dr Ailsa Hart	Consultant Gastroenterologist
Kings College Hospital NHS Foundation Trust	London	Dr Ben Hope	Consultant Paediatrician
Guys & St Thomas' NHS Foundation Trust	London	Dr Peter M Irving	Consultant Gastroenterologist
Barts and The London NHS Trust	London	Prof James O Lindsay	Consultant Gastroenterologist
Guy's and St Thomas' NHS trust	London	Dr Joel E Mawdsley	Gastroenterology Consultant
Lewisham and Greenwich Healthcare NHS Trust	London	Dr Alistair McNair	Consultant Gastroenterologist
Chelsea and Westminster Hospital NHS Foundation	London	Dr Kevin J Monahan	Consultant Gastroenterologist
Royal Free London NHS Foundation Trust	London	Dr Charles D Murray	Consultant Gastroenterologist
Imperial College Healthcare NHS Trust	London	Prof Timothy Orchard	Consultant Gastroenterologist
St George's Healthcare NHS Trust	London	Dr Thankam Paul	Paediatric Consultant Gastroenterologist
St George's Healthcare NHS Trust	London	Dr Richard Pollok	Reader and Consultant Gastroenterologist
Great Ormond Street Hospital for Children NHS Foundation Trust	London	Dr Neil Shah	Consultant Gastroenterologist
North West London Hospitals NHS Trust	London	Dr Sonia Bouri	Research fellow
The Luton & Dunstable University Hospital	Luton	Dr Matt W Johnson	Consultant Gastroenterologist
Luton and Dunstable Hospital Foundation Trust	Luton	Dr Anita Modi	Paediatric Consultant with Allergy and Gastroenterology interest
The Luton & Dunstable University Hospital	Luton	Dr Kasamu Dawa Kabiru	Research fellow
Maidstone and Tunbridge Wells NHS Trust	Maidstone	Dr B K Baburajan	Consultant Gastroenterologist
Maidstone and Tunbridge Wells NHS Trust	Maidstone	Prof Bim Bhaduri	Paediatric Consultant Gastroenterologist
Manchester University Hospitals NHS Foundation Trust	Manchester	Dr Andrew Adebayo Fagbemi	Consultant Gastroenterologist
Central Manchester University Hospitals NHS Foundation Trust	Manchester	Dr Scott Levison	Consultant Gastroenterologist
The Pennine Acute Hospitals NHS Trust	Manchester	Dr Jimmy K Limdi	Consultant Gastroenterologist
Manchester University NHS Foundation Trust, Wythenshawe Hospital	Manchester	Dr Gill Watts	Consultant Gastroenterologist
Sherwood Forest Hospitals NHS Foundation Trust	Mansfield	Dr Stephen Foley	Consultant Gastroenterologist
South Tees Hospital NHS Foundation Trust	Middlesbrough	Dr Arvind Ramadas	Consultant Gastroenterologist
Milton Keynes Hospital NHS Foundation Trust	Milton Keynes	Dr George MacFaul	Consultant Gastroenterologist

Newcastle Upon Tyne Hospital Trust	Newcastle	Dr John Mansfield	Consultant Gastroenterologist
Isle of Wight NHS Foundation Trust	Newport	Dr Leonie Grellier	Consultant Gastroenterologist
Norfolk & Norwich University Hospital NHS Foundation Trust	Norwich	Dr Mary-Anne Morris	Consultant Paediatric Gastroenterologist
Norfolk & Norwich University Hospital NHS Foundation Trust	Norwich	Dr Mark Tremelling	Consultant Gastroenterologist
Nottingham University Hospitals NHS Trust	Nottingham	Prof Chris Hawkey	Consultant Gastroenterologist
Nottingham University Hospitals NHS Trust	Nottingham	Dr Sian Kirkham	Consultant Paediatric Gastroenterologist
Nottingham University Hospitals NHS Trust	Nottingham	Dr Charles PJ Charlton	Consultant gastroenterologist
Oxford University Hospitals NHS Foundation Trust	Oxford	Dr Astor Rodrigues	Paediatric Consultant Gastroenterologist
Oxford University Hospitals NHS Trust	Oxford	Prof Alison Simmons	Consultant Gastroenterologist
Plymouth Hospitals NHS Trust	Plymouth	Dr Stephen J Lewis	Consultant Gastroenterologist
Poole Hospital NHS Foundation Trust	Poole	Dr Jonathon Snook	Consultant Gastroenterologist
Poole Hospital NHS Foundation Trust	Poole	Dr Mark Tighe	Paediatric Consultant with interest in Oncology and Gastroenterology
Portsmouth Hospitals NHS Trust	Portsmouth	Dr Patrick M Goggin	Consultant Gastroenterologist
Royal Berkshire NHS Foundation Trust	Reading	Dr Aminda N De Silva	Consultant Gastroenterologist
Salford Royal NHS Foundation Trust	Salford	Prof Simon Lal	Consultant Gastroenterologist
Shrewsbury and Telford Hospital NHS Trust	Shrewsbury	Dr Mark S Smith	Consultant Gastroenterologist
South Tyneside NHS Foundation Trust	South Shields	Dr Simon Panter	Consultant Gastroenterologist
Southampton University Hospitals NHS Trust	Southampton	Dr Fraser Cummings	Consultant Gastroenterologist
Southampton University Hospitals NHS Trust	Southampton	Dr Suranga Dharmisari	Research fellow
East and North Herts NHS Trust	Stevenage	Dr Martyn Carter	Consultant Gastroenterologist
NHS Forth Valley	Stirling	Dr David Watts	Consultant Gastroenterologist
Stockport NHS foundation Trust	Stockport	Dr Zahid Mahmood	Consultant Gastroenterologist
North Tees and Hartlepool NHS Foundation Trust	Stockton	Dr Bruce McLain	Paediatric Consultant Gastroenterologist
University Hospitals of North Staffordshire	Stoke-on Trent	Dr Sandip Sen	Consultant Gastroenterologist
University Hospitals of North Midlands NHS Trust	Stoke-on-Trent	Dr Anna J Pigott	Consultant Paediatric Gastroenterologist
City Hospitals Sunderland NHS Foundation Trust	Sunderland	Dr David Hobday	Consultant Gastroenterologist
Taunton and Somerset NHS Foundation Trust	Taunton	Dr Emma Wesley	Consultant Gastroenterologist
South Devon Healthcare NHS Foundation Trust	Torquay	Dr Richard Johnston	Consultant Gastroenterologist
South Devon Healthcare NHS Foundation Trust	Torquay	Dr Cathryn Edwards	Consultant gastroenterologist
Royal Cornwall Hospitals NHS Trust	Truro	Dr John Beckly	Consultant Gastroenterologist
Mid Yorkshire Hospitals NHS Trust	Wakefield	Dr Deven Vani	Consultant Physician & Gastroenterologist

Warrington& Halton NHS Foundation	Warrington	Dr Subramaniam Ramakrishnan	Consultant Gastroenterologist
West Hertfordshire Hospitals NHS Trust	Watford	Dr Rakesh Chaudhary	Consultant Gastroenterologist
Sandwell and West Birmingham Hospitals NHS Trust	West Bromwich	Dr Nigel J Trudgill	Consultant Gastroenterologist
Sandwell and West Birmingham Hospitals NHS Trust	West Bromwich	Dr Rachel Cooney	Consultant gastroenterologist
Weston Area Health NHS Trust	Weston-Super-Mare	Dr Andy Bell	Consultant Gastroenterologist
Royal Albert Edward Infirmary, Wrightington, Wigan & Leigh NHS Foundation Trust	Wigan	Dr Neeraj Prasad	Consultant Gastroenterologist
Hampshire Hospitals NHS Foundation Trust	Winchester	Dr John N Gordon	Consultant Gastroenterologist
Royal Wolverhampton Hospitals NHS Trust	Wolverhampton	Prof Matthew J Brookes	Consultant Gastroenterologist
Western Sussex Hospitals NHS Trust	Worthing	Dr Andy Li	Consultant Gastroenterologist
Yeovil District Hospital NHS Foundation Trust	Yeovil	Dr Stephen Gore	Consultant Gastroenterologist