Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis

A. K. Waljee[1,2] (iD) | B. Liu[3] | K. Sauder[2] | J. Zhu[3] | S. M. Govani[2] |
R. W. Stidham[2] (iD) | P. D. R. Higgins[2] (iD)

[1]VA Center for Clinical Management Research, VA Ann Arbor Health Care System, Ann Arbor, MI, USA

[2]Division of Gastroenterology and Hepatology, Department of Internal Medicine, University of Michigan Health System, Ann Arbor, MI, USA

[3]Department of Statistics, University of Michigan, Ann Arbor, MI, USA

Correspondence
Dr. A K Waljee, Gastroenterology, Ann Arbor, MI, USA.
Email: awaljee@med.umich.edu

Funding information
United States Department of Veterans Affairs, Grant/Award Number: CDA 11-217; National Institutes of Health

Summary

Background: Vedolizumab is an effective therapy for ulcerative colitis (UC), but costly and slow to work. New clinical responses occur after 30 weeks of therapy.

Aims: To enable physicians, patients, and insurers to predict whether a patient with UC will respond to vedolizumab at an early time point after starting therapy.

Methods: The clinical study data request website provided the phase 3 clinical trial data for vedolizumab. Random forest models were trained on 70% and tested on 30% of the data to predict corticosteroid-free endoscopic remission at week 52. Models were constructed using baseline data, or data through week 6 of vedolizumab therapy from 491 subjects.

Results: The AuROC for prediction of corticosteroid-free endoscopic remission at week 52 using baseline data was only 0.62 (95% CI: 0.53-0.72), but was 0.73 (95% CI: 0.65-0.82) when using data through week 6. A total of 47% of subjects were predicted to be remitters, and 59% of these subjects achieved corticosteroid-free endoscopic remission, in contrast to 21% of the predicted non-remitters. A week 6 prediction using FCP ≤ 234 μg/g was nearly as accurate.

Conclusions: A machine learning algorithm using laboratory data through week 6 of vedolizumab therapy was able to accurately identify which UC patients would achieve corticosteroid-free endoscopic remission on vedolizumab at week 52. Application of this algorithm could have significant implications for clinical decisions on whom to continue on this costly medication when the benefits of the vedolizumab are not clinically apparent in the first 6 weeks of therapy.

1 | BACKGROUND

Ulcerative colitis (UC) affects over 700 000 people in the United States.[1] For many patients with moderate to severe disease, the gut-selective alpha-4-beta-7 integrin therapy, vedolizumab (VDZ), has proven effective, but is expensive and relatively slow to produce remission[2-4]

Given the high cost of VDZ, insurers are often reluctant to pay for this therapy, which produces remission in roughly one-third of patients with UC in clinical trials. Patients treated with VDZ often do not respond immediately, and additional new clinical responses continue to accumulate even after 30 weeks of therapy.[5] This encourages physicians and patients to continue VDZ therapy at great expense in the hope of a late remission, though the likelihood of this outcome is low. Physicians, patients and insurers would like to be able to predict whether a given patient with UC will respond to VDZ at baseline, or at some early time point after starting therapy, rather than waiting up to 30 weeks (seven doses, or $43 782.76 in Average Wholesale Price drug costs)[6] to determine whether a satisfactory clinical response will occur.

With a pipeline of new therapies in IBD,[7] the growing number of treatment options for each patient raises the question of which therapy is most likely to work for each patient, and how long should we try a new, slow-acting therapy before switching drugs, increasing dose intensity, or adding a combination therapy. There is an increasing need to target therapies to the patients most likely to respond, given the high cost and growing number of therapeutic options in inflammatory bowel disease.

Leveraging clinical trial data may help elucidate which UC patients are most likely to benefit from each therapy, and make clinical decisions about starting or continuing therapies. The Clinical Study Data Request (CSDR) website[8] was implemented to provide researchers the opportunity to conduct further analysis with anonymised data from previously completed clinical studies. Through the CSDR, we obtained access to the phase 3 patient-level clinical trial data for the induction and maintenance of UC using VDZ. These data were used to predict whether baseline data, or data through week 6, could be predictive of week 52 corticosteroid-free endoscopic remission in UC patients treated with VDZ.

2 | METHODS

2.1 | Overview

We obtained confirmation from our Institutional Review Board (IRB) that IRB approval was not necessary to evaluate previously collected and de-identified clinical trial data from the CSDR for the phase 3 clinical trial data for the induction and maintenance of UC using VDZ (HUM00118527). Predictors and outcomes from the clinical trial dataset were used to develop and test predictive models of the outcome.

2.2 | Cohort and demographics

Our initial cohort consisted of 895 subjects. Subjects were excluded from model development if they were on placebo (N = 275), leaving 620 subjects. Additional subjects were excluded if they had missing predictor variables (N = 125) or were missing the outcome (N = 4). A final dataset of 491 de-identified subjects was used for modelling. Table 1 shows the demographics of the final cohort, as compared to the original cohort in the two arms that received VDZ.

2.3 | Predictor variables

Baseline model predictor variables included patient age, gender, race, height, weight, VDZ Interval (VDZ Interval: dosing every 4 or 8 weeks), Immunomodulator use at the start of the trial (ImmAtStart), Steroid use at the start of the trial (SteroidAtStart), previous exposure to anti-TNF therapy (PriorTNF) and all available quantitative laboratory tests at baseline.

A model at week 6 included the baseline variables mentioned above with one exception: quantitative laboratory test results were included from week 6 (or nearest earlier date if week 6 results were not available), rather than from baseline. The VDZ drug level at week 6, as well as calculated longitudinal variables, were also included as predictors.

Longitudinal variables that were calculated included the slope of Faecal Calprotectin (FCP) ((week 6 FCP minus FCP prior to initiation of medication)/6), and the slope of the VDZ drug level ((VDZ drug level at week 6-VDZ drug level at week 2)/4). The slope, acceleration, mean and maximum of each of the other laboratory predictors were tested, but provided no added improvement to the AuROC of the week 6 model, and were removed.

Additional clinical predictors, including disease extent, baseline sigmoidoscopic severity, and disease duration, were tested, but added no improvement to the AuROC, and were removed. A model including values from week 0 through week 14 was also tested, but was only slightly better than the week 6 model, so this model is not reported here. A list of all predictors used for each model, as well as variables that were tested and removed due to no improvement of AuROC, can be found in Supplemental Table 1.

2.4 | Definition of outcomes

The primary outcome was corticosteroid-free endoscopic remission at week 52, defined by no use of corticosteroid medications (including prednisone and budesonide) at week 52, and a Mayo Sigmoidoscopy Score of 0 or 1 at week 52. Subjects without a visit at week 52, having a sigmoidoscopy score greater than 1, or using steroids (Prednisone EQ Dose or Budesonide EQ Dose > 0 at week 52) were defined as failures. Patients who had a visit but who did not have a sigmoidoscopy score at week 52 were defined as a missing outcome, and were removed from the cohort (N = 4).

TABLE 1 Subject demographics

Variable	Original cohort (N = 620)	Final modelling cohort (N = 491)
Mean age (y)	40.1 ± 13.1	40.2 ± 13.4
Percentage male sex	58.7%	57.4%
Percentage white race	83.5%	83.1%
Mean body weight (kg)	73.4 ± 18.3	73.0 ± 18.5
Percentage current smoker	5.8%	5.9%
Mean duration of disease (y)	6.7 ± 6.0	6.5 ± 5.9
Mean mayo clinic score at baseline	8.6 ± 1.8	8.5 ± 1.7
Median faecal calprotectin at baseline	844 (346-1727)	872 (371-1727)
Site of disease (%)		
Rectum and sigmoid colon only	13.7%	14.1%
Left colon	36.6%	37.7%
Proximal to splenic flexure	11.9%	11.4%
Pancolitis	37.7%	36.9%
Median prednisone equivalent dose in people who were on prednisone (mg)	20.0 (10.0-30.0) (N = 316)	20.0 (10.0-30.0) (N = 261)
Median budesonide equivalent dose in people who were on budesonide (mg)[a]	5.0 (5.0-9.0) (N = 11)	5.0 (5.0-9.0) (N = 6)
Percentage with prior anti-TNF therapy	50.2%	49.1%
Baseline concomitant medications (%)		
Glucocorticoids only	36.5%	37.5%
Immunosuppressants only	18.4%	17.3%
Glucocorticoids and immunosuppressants	16.0%	16.9%
No glucocorticoids or immunosuppressants	29.2%	28.3%

This table compares the subject demographics of the original clinical trial cohort (which combines both VDZ arms) and the final modelling cohort (where any subject with missing data or missing outcomes was removed) to demonstrate that the modelling cohort is similar to the original clinical trial cohort.
[a]Budesonide prescriptions of 5 mg were reported in this dataset.

2.5 | Statistical analysis and model development

Random forest (RF) machine learning[9] was used to construct these algorithms. This method of prediction uses decision trees[9,10] to classify a new observation. Each observation is run through each of the trees in the forest and each tree gives a classification (which can also be thought of as a vote). The forest combines the votes from all of the trees to compute a predicted score of the outcome. By choosing a value of the predicted score as a cut-off, one can obtain the desired balance of sensitivity and specificity for the outcome. We developed two random forest models using the predictor variables, one using only baseline data and one using data through week 6. To validate the predictive models, each dataset was split into a 70% training dataset for model development, and a 30% testing dataset. This split method was preferred over an out-of-bag validation to show a true training and validation cohort, and to generate misclassification tables.

A baseline random forest model of 1000 trees with baseline variables was fit on the training set, and was used to produce the predictions on the test set. An additional week 6 random forest model of 1000 trees using cross sectional data at the final time point and longitudinal predictor variables was fit, the predictions tested on the test set, and the results compared across models.

2.5.1 | Training and testing cohorts

Training and testing datasets were derived by splitting the data randomly into 70% and 30% subsets. This was done 50 times, and the random forest model was fit on the training dataset and tested on the testing dataset each time. The AuROC (area under the receiver operating characteristic curve) values from the 50 replication test sets were averaged to obtain a mean AuROC. Subsequently, the one split of the 50 that produced an AuROC closest to the average AuROC for the week 6 model was selected as a representative split (training and testing cohorts) for both the week 6 and baseline models, which was used for ROC (receiver operating characteristic) plots, representative AuROCs, cut-off point selection, and misclassification tables. We then built baseline and week 6 random forest models on the entire dataset to have the most accurate models for future use, and we calculated variable importance and produced partial dependence plots based on these models.

2.5.2 | Model performance

The AuROC was used to evaluate the performance of each model. All statistical analyses were performed using the statistical language R (version 3.3), using the packages randomForest,[9] and pROC.

2.5.3 | Cut-off point selection

The optimal cut-off was defined as the point on the ROC plot that is closest to the perfect point where both sensitivity and specificity are 1. In other words, we minimise the following criterion: $(1\text{-sensitivity})^2 + (1\text{-specificity}).^2$

2.5.4 | Variable importance

We evaluated the importance of predictors based on a random forest model built on the entire dataset. The relative importance of each predictor variable was determined by identifying nodes in the ensemble of trees in which the individual predictor variable appeared and summing the relative information content provided by all of the nodes containing that variable. Predictor variables that provided the greatest combined discrimination have higher importance.

2.5.5 | Partial dependence plots

Partial dependence plots were constructed to demonstrate how individual predictors can affect the probability of success, that is, corticosteroid-free endoscopic remission at week 52. For each predictor, we used the random forest model on the entire dataset to predict the probability of success for each patient with that predictor set to a given value, and averaged over all patients to produce a mean probability of success. We repeated this procedure for each value of that predictor in the data, or for its 50 quantiles if more than 50 values were available.

2.5.6 | Simpler pragmatic models

Given the complexity of these models, simpler models were considered more attractive for routine clinical use, if accurate. Informed by the Variable Importance Plots, we hypothesised that two simplified models for week 6 might be helpful, one with the week 6 faecal calprotectin alone, and the second using the FCP/VDZ level ratio at week 6. We also hypothesised that the baseline FCP might be helpful as a single predictor. We evaluated these single variable predictor models on the entire (100%) dataset to determine whether they could predict the week 52 corticosteroid-free endoscopic remission outcome.

2.5.7 | Reproducible research code repository

The code used to produce this analysis in R is available in a public Github repository at: https://github.com/higgi13425/vedoUC. Note that access to the trial data can only be obtained through the CSDR website at https://clinicalstudydatarequest.com/.

3 | RESULTS

3.1 | Predicting week 52 steroid free endoscopic remission at baseline

The average AuROC for the baseline model over 50 replications is 0.63. The AuROC for the baseline model under the selected training and testing split is 0.62 (95% CI: 0.53-0.72), as shown in Figure 1A. The variable importance plot for the baseline model is shown in Figure 1B. The 5 strongest baseline predictors of corticosteroid-free endoscopic remission at week 52 were: faecal calprotectin, albumin, neutrophils, white blood cell count, and absolute lymphocyte count. Notably weak baseline predictors of this outcome were the interval between VDZ doses, prior use of anti-TNF therapy, use of immunomodulators or corticosteroids at baseline, race, and sex. The best cut-off, number of predicted successes and failures, and the proportion of subjects with true success in the testing set (N = 148) within these two predicted classes are displayed in Figure 2A and Table 2. True positives and true negatives are also displayed in Table 2, which provides the sensitivity of 0.63 and specificity of 0.62.

3.2 | Effect of predictor variables at baseline

The effects of the predictors on the outcome of corticosteroid-free endoscopic remission at week 52 can be illustrated by partial dependence plots. Many of these are not linear. These are displayed in multiple panels in Figure S1. Faecal calprotectin levels ≤811.5 μg/g at baseline predict success, and the rate of success increases steeply with lower FCP levels. Lower uric acid levels, perhaps a marker of ongoing bowel damage and cell death, also predict higher rates of success with VDZ in UC.

3.3 | Predicting week 52 steroid free endoscopic remission at week 6

The average AuROC for the week 6 model over 50 replications is 0.73. The AuROC for the week 6 model under the selected representative training and testing split is 0.73 (95% CI: 0.65-0.82), as shown in Figure 3A and Table 2. The P-value for the AuROC of the week 6 model vs. an AuROC of 0.5 is 1.23×10^{-6}. The variable importance graph for the week 6 model is shown in Figure 3B. The five strongest predictors of corticosteroid-free endoscopic remission at week 52 were faecal calprotectin at week 6, the slope of VDZ level, the slope of FCP, albumin at week 6, and VDZ level at week 6. Notably weak predictors at week 6 for this outcome were the interval between VDZ doses, prior use of anti-TNF therapy, use of immunomodulators or corticosteroids at baseline, race and sex. The best cut-off, number of predicted successes and failures, and the proportion of subjects of true success within these two predicted classes are displayed in Table 2 and Figure 2B. This model classifies subjects into two groups: those likely to succeed (47%) and those likely to fail (53%). Those classified as likely to succeed achieve corticosteroid-free endoscopic remission in 58.6%, while those classified as likely to fail only achieve this outcome in 20.5%. The true positives and true negatives are also displayed in Table 2, which provides the sensitivity of 0.72 and specificity of 0.68.

3.4 | Effect of predictor variables at week 6

The effects of the week 6 predictors on the outcome of corticosteroid-free endoscopic remission at week 52 can be illustrated in
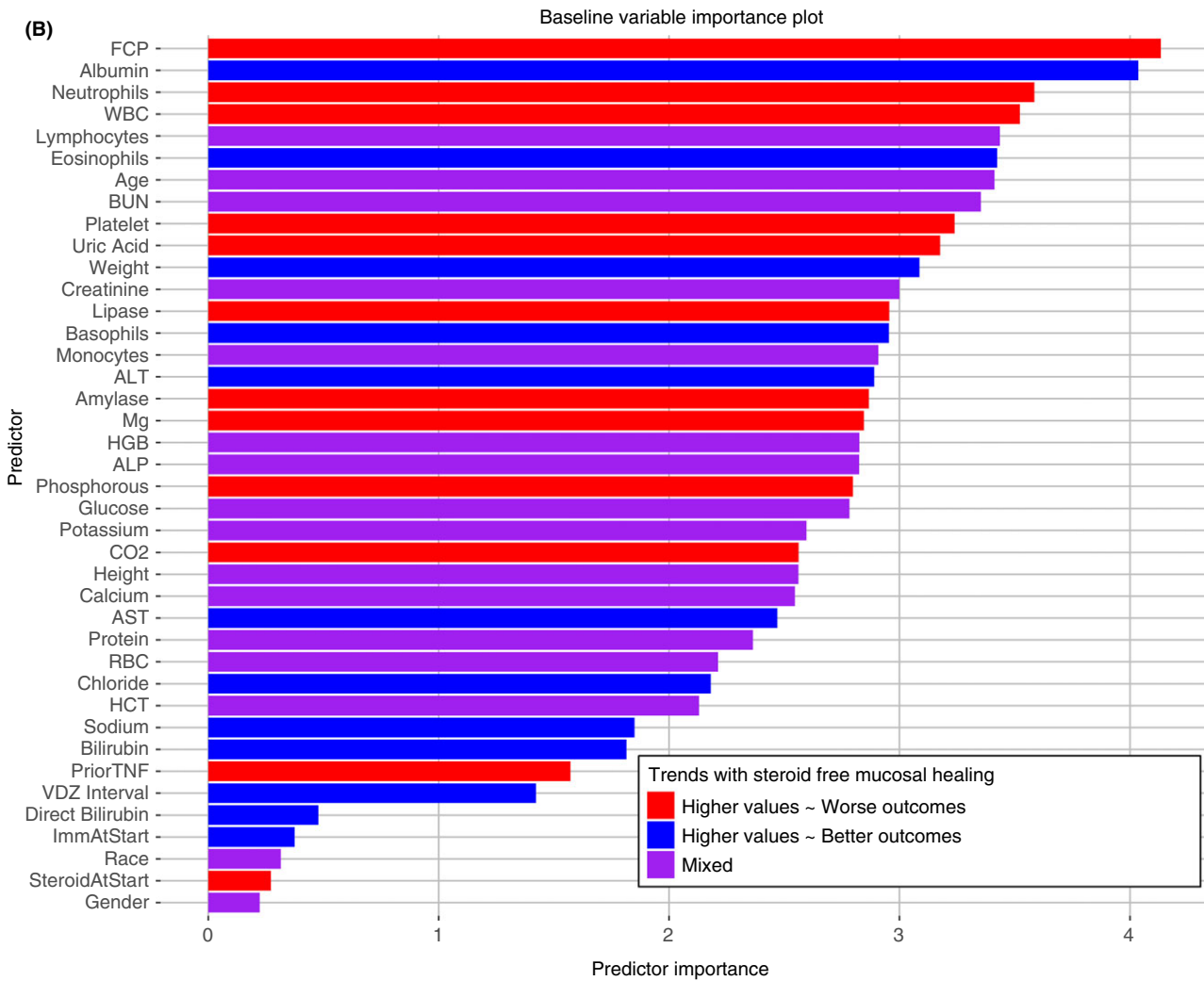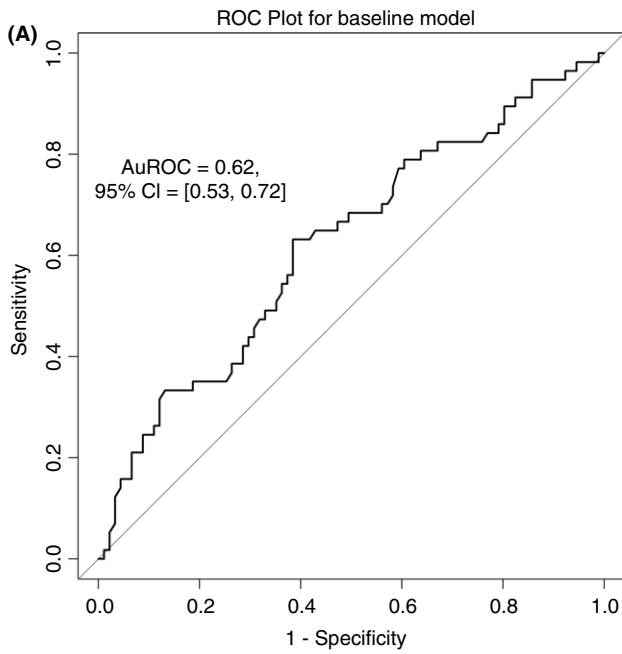
FIGURE 1 Baseline model ROC plot and variable importance plot. A, Receiver operating characteristic (ROC) plot for the baseline model. B, Variable importance plot for the baseline model shows the relative importance of each predictor variable
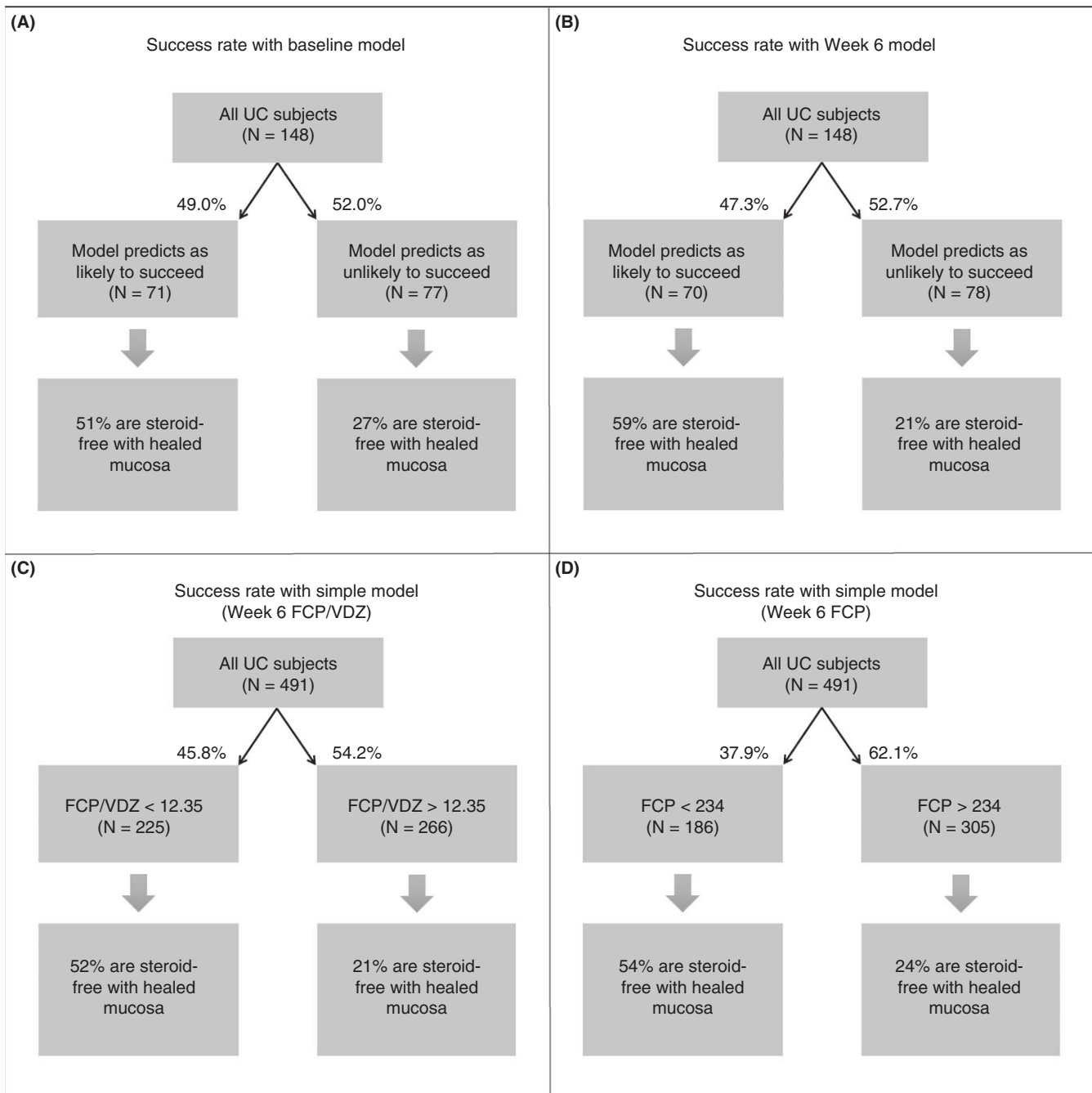
Endpoint Success

(A) Success rate with baseline model

All UC subjects
(N = 148)

49.0% ↙   ↘ 52.0%

Model predicts as likely to succeed (N = 71)   Model predicts as unlikely to succeed (N = 77)

51% are steroid-free with healed mucosa   27% are steroid-free with healed mucosa

(B) Success rate with Week 6 model

All UC subjects
(N = 148)

47.3% ↙   ↘ 52.7%

Model predicts as likely to succeed (N = 70)   Model predicts as unlikely to succeed (N = 78)

59% are steroid-free with healed mucosa   21% are steroid-free with healed mucosa

(C) Success rate with simple model (Week 6 FCP/VDZ)

All UC subjects
(N = 491)

45.8% ↙   ↘ 54.2%

FCP/VDZ < 12.35 (N = 225)   FCP/VDZ > 12.35 (N = 266)

52% are steroid-free with healed mucosa   21% are steroid-free with healed mucosa

(D) Success rate with simple model (Week 6 FCP)

All UC subjects
(N = 491)

37.9% ↙   ↘ 62.1%

FCP < 234 (N = 186)   FCP > 234 (N = 305)

54% are steroid-free with healed mucosa   24% are steroid-free with healed mucosa

FIGURE 2 Endpoint success. Each flow diagram shows the percentages of endpoint success (corticosteroid-free and endoscopically healed subjects) at week 52 within the predicted success and predicted failure groups. A, Success rate with baseline model. B, Success rate with week 6 model. C, Success rate with simplified week 6 model-week 6 FCP/VDZ. D, Success rate with simplified week 6 model-week 6 FCP

partial dependence plots. These are displayed in multiple panels in Figure S2. Faecal calprotectin levels ≤233.67 μg/g at week 6 predict success, and the rate of success increases steeply with lower FCP levels. A significant fall in FCP between week 0 and week 6 (steeper negative slope) is a good prognostic marker. Lower albumin levels, perhaps a marker of chronicity and severity of ulcers, and colonic leak of VDZ, predict lower rates of success. As expected, higher VDZ levels predict success, but a falling VDZ level between week 2 and week 6 is a negative prognostic marker.

3.5 | Simpler pragmatic models

On the basis of the variable importance plots, we tested simpler models using a single predictor on the entire dataset, to see if simplified models could have comparable predictive performance. A simple model composed of very few variables would be easier to acquire and use in clinical practice. For comparison to the baseline model, we tested a simpler model using faecal calprotectin measured before the first dose of VDZ (FCP baseline). A baseline FCP of

TABLE 2 Model performance

Model	Validation sample size	AuROC (95% CI)	Best cut-off	Prediction category	Predicted cases	True success rate (%)	Sensitivity	Specificity
Baseline model	148	0.62 (0.53, 0.72)	0.35	Predicted success ($\geq$ cut-off)	71	50.7	0.63	0.62
				Predicted failure ($<$ cut-off)	77	27.3		
Faecal Calprotectin before 1st dose	491	0.58 (0.52, 0.63)	811.50	Predicted success ($\leq$ cut-off)	236	41.9	0.57	0.57
				Predicted failure ($>$ cut-off)	255	29.0		
Week 6 model	148	0.73 (0.65, 0.82)	0.32	Predicted success ($\geq$ cut-off)	70	58.6	0.72	0.68
				Predicted failure ($<$ cut-off)	78	20.5		
Week 6 FCP/VDZ level ratio	491	0.71 (0.67, 0.76)	12.35	Predicted success ($\leq$ cut-off)	225	52.0	0.68	0.66
				Predicted failure ($>$ cut-off)	266	21.1		
Week 6 Faecal calprotectin	491	0.71 (0.66, 0.76)	233.67	Predicted success ($\leq$ cut-off)	186	54.3	0.58	0.73
				Predicted failure ($>$ cut-off)	305	23.6		

This table presents the performance details for each model. The area under the receiver operating characteristic curve (AuROC) provides the discriminative power of each model. The best cut off for each model was based on the ROC plot to optimise the balance between the sensitivity and specificity. The number of cases predicted to be a success or failure is listed for each model, along with the respective true success rate, sensitivity and specificity.

$\leq$811.5 µg/g was a weak predictor of success. This model was less accurate than the full baseline model, with an AuROC of 0.58 (95% CI: 0.52-0.63).

For comparison to the week 6 model, we tested two simpler models. A week 6 FCP/VDZ ratio $\leq$12.35 predicted success reasonably well, and this single predictor had an AuROC of 0.71 (95% CI: 0.67-0.76). A single predictor of week 6 FCP $\leq$233.67 µg/g predicted success, and this predictor had a similar AuROC of 0.71 (95% CI: 0.66-0.76). Other single predictors were not as strong (Slope of VDZ AuROC: 0.65, Slope of FCP AuROC: 0.57, Week 6 Albumin AuROC: 0.65, Week 6 VDZ AuROC: 0.63). The best cut-off, number of predicted successes and failures, and the proportion of patients of true success within each predicted class for these simplified models are displayed in Figures 3C and D, and Table 2. The true positives and true negatives are also displayed in Table 2, which provides the sensitivity and specificity.

4 | DISCUSSION

Leveraging access to the GEMINI 1 clinical trial data[2] through the CSDR, we were able to apply machine learning tools to develop and validate predictive models of corticosteroid-free endoscopic remission in response to VDZ in UC. While the baseline model is relatively inaccurate, with a sensitivity of 63% and a specificity of 62%, it is helpful for clinicians to know that patients with a very elevated baseline faecal calprotectin (>811.5 µg/g) are at higher risk of VDZ failure.

The week 6 model is numerically more accurate, with an AuROC of 0.73 (95% CI: 0.65-0.82). This model incorporates the change in faecal calprotectin over time, VDZ levels, and the slope of the VDZ concentration along with laboratory values at week 6. This model can classify patients into two distinct groups, who achieve corticosteroid-free endoscopic remission at week 52 in very different proportions of 59% and 21%. Simpler models are not quite as accurate, but it can be helpful to clinicians to know that values of either a Week 6 FCP/VDZ ratio of $\leq$12.35, or a week 6 FCP $\leq$233.67 µg/g, predict success. Takeda has provided various laboratories with VDZ level results and serum samples for validation of external assays, making the Takeda assay used in the clinical trials the *de facto* gold standard.

These models can be implemented as a cloud-based service for laboratories with HL7-compatible laboratory information systems,[11] which can securely submit lab values and accept returned calculated algorithmic results without exposing protected health information. This study is limited to the data available from GEMINI 1, and is only as generalisable as these data which resulted in the FDA approval of VDZ for use in the treatment of UC.

Given the expense of biological therapies, being able to identify patients with a low rate of success after a short trial of therapy is valuable, as these patients can then move on to a different therapy, or a different mechanism of action in a timely fashion. This would reduce the time period of their active symptoms, likely reduce their exposure to steroids, and reduce the expense of a biological therapy that is likely to be futile in a given patient.
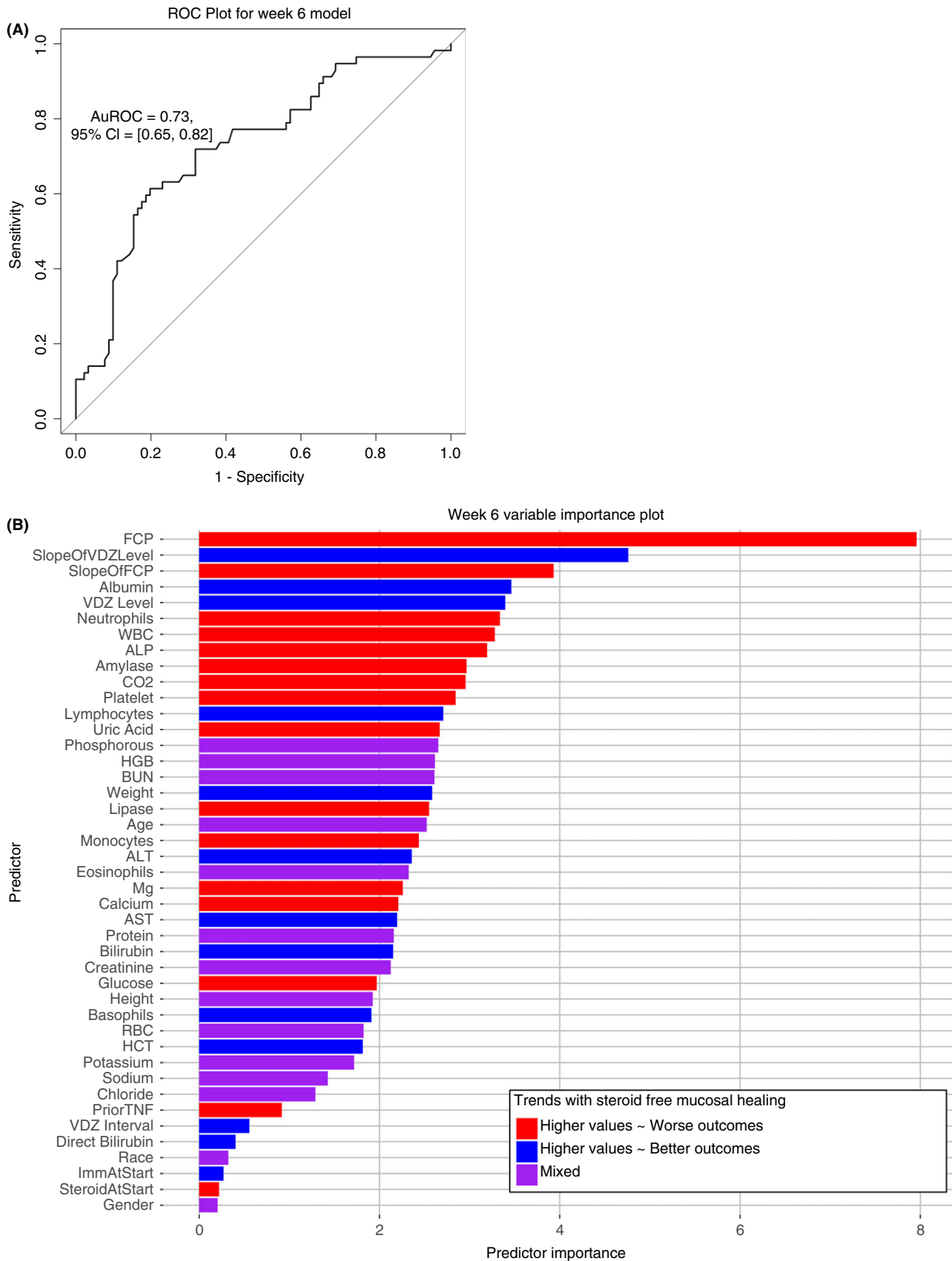
FIGURE 3 Week 6 longitudinal model ROC plot and variable importance plot. A, Receiver operating characteristic (ROC) plot for the week 6 longitudinal model. B, Variable importance plot for the longitudinal model shows the relative importance of each predictor variable

Being able to identify UC patients who are at higher risk of failure with VDZ at baseline or at week 6 also provides an opportunity to improve the outcomes for these patients. Patients who are predicted as likely to fail VDZ might benefit from addition of a "booster" therapy after a model prediction of high risk of VDZ failure at week 6, possibly including an anti-TNF therapy,[12,13] a JAK inhibitor,[14] or an anti-IL23 therapy.[15] These severe UC patients, once they achieve low faecal calprotectin levels and adequate VDZ levels, may well be able to continue on VDZ maintenance monotherapy successfully. Future randomised clinical trials in this subset of patients are needed.

The limitations of this study include: (1) 224 of the 491 patients were already corticosteroid-free at baseline. However, to enter into the trial, subjects had to have active disease, defined by a Mayo score of at least 6 (with an endoscopy score of 2-3). Achieving steroid-free remission in these subjects is an important clinical endpoint, whether they started on steroids or not; (2) patients included in all clinical trials are not necessarily representative of the general UC patient population and (3) the simple predictors (Week 6 FCP/VDZ ratio, and Week 6 FCP) presented here were developed post hoc, and should be evaluated in an external dataset to determine their external validity.

Strengths of this study include: (1) the inclusion of patients from multiple sites and across multiple countries; (2) the data used in this study were the same data used for FDA approval of VDZ in UC and 3) our internal validation used a model developed on a randomly selected 70% of subjects, and validated this model by testing it on the remaining 30% of subjects.; (3) It should also be noted that the simpler pragmatic model of a week 6 FCP/VDZ ratio ≤ 12.35 provides a convenient measure for point-of-care use as its accuracy is close to the full model. Because this simple ratio is quite accurate, independent of previous anti-TNF use and body weight, this suggests that any patient can be treated effectively with Vedolizumab if sufficient drug for their inflammatory load is provided. These results suggest that titration of VDZ dosing during induction to produce a low FCP/VDZ level ratio could significantly increase the efficacy of VDZ induction in ulcerative colitis.

5 | CONCLUSION

A machine learning algorithm was able to distinguish IBD patients who are highly likely to achieve week 52 corticosteroid-free endoscopic remission with VDZ from those who are likely to fail VDZ using data collected by week 6 of therapy. A baseline model was numerically less accurate in predicting this outcome with VDZ. The ability to make early and accurate predictions of outcomes could help reduce costs by targeting this expensive therapy to the UC patients most likely to benefit, or could target additional interventions to patients who are likely to fail VDZ. While these algorithms are imperfect, they are arguably better than the current practice.

AUTHORSHIP

Guarantor of the article: Akbar K. Waljee, the submission's guarantor, takes responsibility for the integrity of the work as a whole, from inception to published article.

Author contributions: Akbar K. Waljee, Concept and design, Data interpretation, Writing, Figures, Critical revision of the manuscript, Final approval; Boang Liu, Data Analysis, Data interpretation, Critical revision of the manuscript, Final approval; Kay Sauder, Data collection, Figures, Critical revision of the manuscript, Final approval; Ji Zhu, Data Analysis, Data interpretation, Critical revision of the manuscript, Final approval; Shail M. Govani, Critical revision of the manuscript, Final approval; Ryan W. Stidham, Critical revision of the manuscript, Final approval; Peter D.R. Higgins, Concept and design, Data interpretation, Figures, Critical revision of the manuscript, Final approval. All authors approved the final version of the manuscript.

ORCID

A. K. Waljee http://orcid.org/0000-0003-1964-8790
R. W. Stidham http://orcid.org/0000-0001-9638-2186
P. D. R. Higgins http://orcid.org/0000-0003-1602-4341

REFERENCES

1. Loftus EV. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology.* 2004;126:1504-1517.

2. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med*. 2013;369:699-710.

3. Stallmach A, Langbein C, Atreya R, et al. Vedolizumab provides clinical benefit over 1 year in patients with active inflammatory bowel disease – a prospective multicenter observational study. *Aliment Pharmacol Ther*. 2016;44:1199-1212.

4. Amiot A, Serrero M, Peyrin-Biroulet L, et al. One-year effectiveness and safety of vedolizumab therapy for inflammatory bowel disease: a prospective multicentre cohort study. *Aliment Pharmacol Ther*. 2017;46:310-321.

5. Dulai PS, Singh S, Jiang X, et al. The real-world effectiveness and safety of vedolizumab for moderate-severe Crohn's disease: results from the US VICTORY consortium. *The American journal of gastroenterology*. 2016;111:1147-1155.

6. ENTYVIO (VEDOLIZUMAB) FOR INJ 300MG SDV. metromedicalorder.com. https://www.metromedicalorder.com/entyvio-vedolizumab-for-inj-300mg-sdv.html. Accessed March 20, 2017.

7. Cohen BL, Sachar DB. Update on anti-tumor necrosis factor agents and other new drugs for inflammatory bowel disease. *BMJ*. 2017;357:j2505. https://doi.org/10.1136/bmj.j2505

8. CSDR: Clinical Study Data Request. https://clinicalstudydatarequest.com/.

9. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.

10. Liaw A, Wiener M. Classification and regression by random forest. *R News*. 2002;2:18-22.

11. Romero J, Lopez P, Vazquez Noguera JL, Cappo C, Pinto-Roa PD, Villalba C. Integrated, reliable and cloud-based personal health record: a scoping review. *HIIJ*. 2016;5:01-20.

12. Stidham RW, Lee TCH, Higgins PDR, et al. Systematic review with network meta-analysis: the efficacy of anti-tumour necrosis factor-alpha agents for the treatment of ulcerative colitis. *Aliment Pharmacol Ther*. 2014;39:660-671.

13. Stidham RW, Lee TCH, Higgins PDR, et al. Systematic review with network meta-analysis: the efficacy of anti-TNF agents for the treatment of Crohn's disease. *Aliment Pharmacol Ther*. 2014;39:1349-1362.

14. Olivera P, Danese S, Peyrin-Biroulet L. JAK inhibition in inflammatory bowel disease. *Exp Rev Clin Immunol*. 2017;13:693-703.

15. Furfaro F, Gilardi D, Allocca M, et al. IL-23 Blockade for Crohn s disease: next generation of anti-cytokine therapy. *Exp Rev Clin Immunol*. 2017;13:457-467.

SUPPORTING INFORMATION

Additional Supporting Information will be found online in the supporting information tab for this article.

How to cite this article: Waljee AK, Liu B, Sauder K, et al. Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis. *Aliment Pharmacol Ther*. 2018;47:763–772. https://doi.org/10.1111/apt.14510