Inflammatory Bowel Diseases, 2022, **28**, 1555–1564
https://doi.org/10.1093/ibd/izab310
Advance access publication 30 December 2021
Original Research Articles - Clinical

Decision Support Tool Identifies Ulcerative Colitis Patients Most Likely to Achieve Remission With Vedolizumab vs Adalimumab

Parambir S. Dulai, MD,[*] Emily C.L. Wong, BS,[†] Walter Reinisch, MD,[‡] Jean-Frederic Colombel, MD,[§] John K. Marshall, MD, MSc,[†] and Neeraj Narula, MD[†]

From the [*]Division of Gastroenterology, University of California San Diego, La Jolla, CA, USA;
[†]Division of Gastroenterology, Department of Medicine and Farncombe Family Digestive Health Research Institute, McMaster University, Hamilton ON, Canada;
[‡]Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Vienna, Austria; and
[§]Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

Address correspondence to: Parambir S. Dulai, MD, University of California San Diego, Division of Gastroenterology, 9500 Gilman Drive, La Jolla, CA 92093, USA (pdulai@health.ucsd.edu).

Background & Aims: We have previously validated a clinical decision support tool (CDST) (vedolizumab CDST [VDZ-CDST]) for clinical and endoscopic remission with VDZ in ulcerative colitis (UC). We aim to expand the validation for predicting histoendoscopic mucosal improvement (HEMI) with VDZ vs adalimumab (ADA).

Methods: In a post hoc analysis of a clinical trial for VDZ vs ADA in moderate to severe UC (VARSITY trial; NCT02497469), comparative accuracy was evaluated for the VDZ-CDST among an external validation cohort of VDZ- and ADA-treated patients for week 52 HEMI (Mayo endoscopic subscore 0-1 and Geboes score <3.2). Comparative effectiveness of VDZ and ADA was assessed after stratifying the cohort by baseline probability of response to VDZ using the VDZ-CDST.

Results: A total of 419 patients were included. The majority of patients enrolled in the VARSITY trial had a high (61%) or intermediate (29%) baseline predicted probability of response to VDZ. The baseline VDZ-CDST score was significantly more likely to predict week 52 HEMI for VDZ (area under the curve , 0.712; 95% confidence interval, 0.636-0.787) relative to ADA-treated patients (area under the curve, 0.538; 95% confidence interval, 0.377-0.700; $P < .001$ for AUC comparison). A significant ($P < .001$) association was observed between the VDZ-CDST and measured VDZ drug exposure over 52 weeks. Superiority of VDZ to ADA was only observed in patients with a high baseline predicted probability of response to VDZ.

Conclusions: Superiority of VDZ to ADA is dependent on baseline probability of response, and a VDZ-CDST is capable of identifying UC patients most appropriate for VDZ vs ADA.

Key Words: precision medicine, therapeutic drug monitoring, comparative effectiveness

Introduction

Clinical practice guidelines have recommended vedolizumab (VDZ) or tumor necrosis factor (TNF) antagonist therapy as first-line biologics for moderate to severe ulcerative colitis (UC).[1,2] A recently completed head-to-head clinical trial demonstrated VDZ to be superior to adalimumab (ADA) for the achievement of clinical, endoscopic, and histologic endpoints after 1 year of therapy.[3] Within the subgroup analyses of this trial, the superiority of VDZ to ADA was driven by the TNF antagonist–naïve subgroup, suggesting that VDZ might be the preferred first-line biologic for use in UC. Caution must be taken when interpreting subgroup analyses, and some degree of personalization is still required when choosing between VDZ and ADA for the treatment of UC. There remains no guidance on how to consider these data and best personalize this decision.

Using the phase 3 registration trial programs for VDZ in UC, we have previously built a prediction model and clinical decision support tool (CDST) from the GEMINI clinical trial datasets that was observed to have good prognostic performance for predicting VDZ concentrations and identifying UC patients more likely to achieve clinical and endoscopic remission with VDZ.[4] This CDST for VDZ (VDZ-CDST) could therefore be used to help identify UC patients most appropriate for VDZ therapy relative to TNF-antagonist therapy in routine practice; however, several limitations in the prior validation study need be addressed prior to integration. In particular, the routine practice validation cohort used was not prospectively recruited, blinded readers did not centrally assess disease activity measures, and an assessment for evolving histologic endpoints was not made. Finally, the prognostic comparison for predicting response to VDZ relative to TNF-antagonist therapy may have been limited by selection biases inherent to routine practice.

In the current study we addressed these gaps by further validating the VDZ-CDST within an external cohort of a head-to-head clinical trial of VDZ vs ADA. By studying the relative prognostic value of the VDZ-CDST for predicting treatment outcomes within a randomized clinical trial we

Received for publication: June 26, 2021. Editorial Decision: November 16, 2021

Table 1. Baseline Demographics

Variable	Entire Included Population			Low Predicted Probability Response to VDZ		Intermediate Predicted Probability Response to VDZ		High Predicted Probability Response to VDZ	
	Overall (N = 419)	VDZ (n = 207)	ADA (n = 212)	VDZ (n = 10)	ADA (n = 11)	VDZ (n = 66)	ADA (n = 77)	VDZ (n = 131)	ADA (n = 124)
Age, y	39.9 ± 13.8	29.9 ± 14.2	39.9 ± 13.5	42.2 ± 14.5	43.3 ± 15.5	37.6 ± 12.4	38.8 ± 14.4	40.9 ± 14.9	40.2 ± 12.7
Male	243 (58.0)	125 (60.4)	118 (55.7)	5 (50.0)	5 (45.5)	38 (57.6)	41 (53.3)	82 (62.6)	72 (58.1)
Current smoking	18 (4.3)	7 (3.4)	11 (5.2)	1 (10.0)	1 (9.1)	5 (7.6)	2 (2.6)	1 (0.8)	8 (6.5)
Disease duration, y	3.6 (1.5-7.2)	3.3 (1.5-7.1)	3.9 (1.6-7.7)	6.2 (1.7-11.7)	1.2 (0.4-3.9)	1.9 (0.9-4.0)	2.9 (1.4-6.3)	4.0 (2.1-7.1)	4.8 (2.5-8.6)
Mayo endoscopic score at baseline									
2	153 (36.5)	77 (37.2)	76 (35.9)	0	0	8 (12.1)	12 (15.6)	69 (52.7)	64 (51.6)
3	266 (63.5)	130 (62.8)	136 (64.2)	10 (100.0)	11 (100.0)	58 (87.9)	65 (84.4)	62 (47.3)	60 (48.4)
Anti-TNF naïve	333 (79.5)	159 (76.8)	174 (82.1)	2 (20.0)	8 (72.7)	45 (68.2)	55 (71.4)	112 (85.5)	111 (89.5)
Concomitant immunomodulator use	99 (23.6)	53 (26.1)	46 (21.7)	1 (10.0)	1 (9.1)	17 (25.8)	19 (24.7)	35 (26.7)	26 (21.0)
Concomitant corticosteroid use	154 (36.8)	84 (40.6)	70 (33.0)	7 (70.0)	4 (36.4)	27 (40.9)	27 (35.1)	50 (38.2)	39 (31.5)
Baseline fecal calprotectin, µg/g	2647.9 ± 5417.0	2606.4 ± 6671.4	2691.0 ± 3709.1	4585.5 ± 3759.9	2475.2 ± 1038.5	2488.2 ± 2367.9	2663.0 ± 3223.7	2501.5 ± 8143.0	2726.9 ± 4122.6
Baseline C-reactive protein, mg/L	9.7 ± 16.3	9.8 ± 16.4	9.4 ± 16.3	25.2 ± 25.6	35.7 ± 33.2	15.3 ± 21.1	11.8 ± 17.8	5.8 ± 10.3	5.7 ± 9.4
Baseline albumin, mg/L	42.4 ± 4.4	42.5 ± 4.2	42.2 ± 4.5	34.1 ± 3.4	31.5 ± 2.5	39.5 ± 3.3	39.5 ± 2.7	44.7 ± 2.6	44.9 ± 2.8
Total Mayo score	8.9 ± 1.5	8.9 ± 1.4	8.8 ± 1.6	9.8 ± 0.5	10.3 ± 1.0	9.5 ± 1.3	9.5 ± 1.3	8.6 ± 1.4	8.4 ± 1.6

Values are mean ± SD, n (%), or median (interquartile range).
Abbreviations: ADA, adalimumab; TNF, tumor necrosis factor; VDZ, vedolizumab.

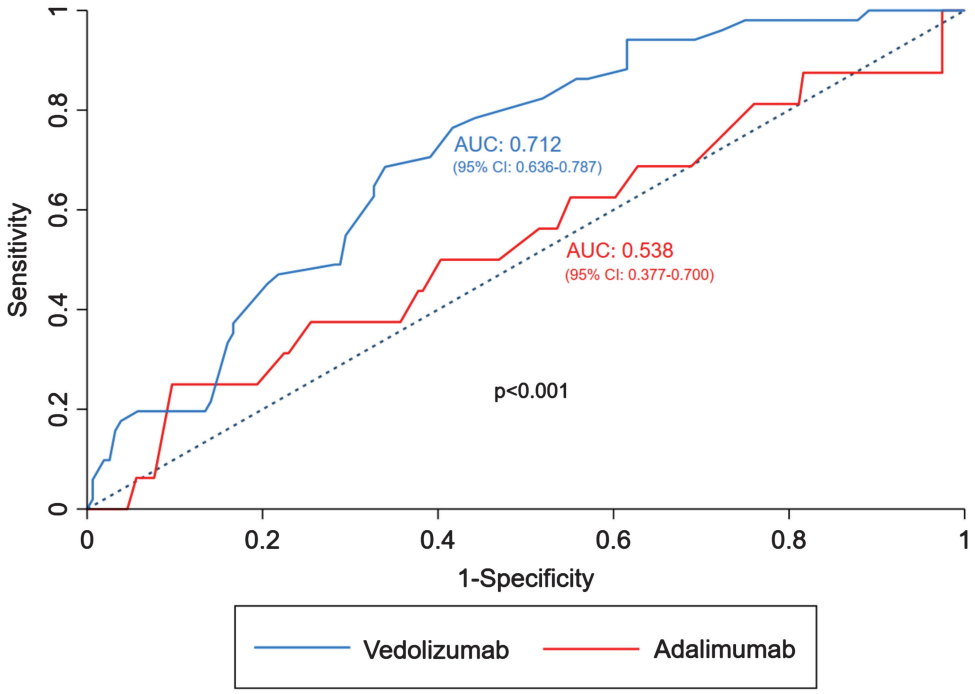


Figure 1. Comparative accuracy of the vedolizumab clinical decision support tool for histoendoscopic mucosal improvement with vedolizumab vs adalimumab in ulcerative colitis. AUC, area under the curve; CI, confidence interval.

were able to address biases related to both measurable and unmeasurable confounders, and we were able to assess the prognostic value of the VDZ-CDST for endoscopic and histologic endpoints using blinded centralized scoring. This validation therefore provides high-quality evidence for the use of this VDZ-CDST to guide biologic treatment selection for UC patients in routine practice.

Methods

Study Aim and Hypothesis

The proposed study aimed to determine whether the VDZ-CDST can predict both response to VDZ using objective measures of disease activity that include histology and exposure profiles to assess exposure-response relations. We further aimed to assess whether this tool can identify patients who would achieve higher rates of treatment response with VDZ as compared with ADA. Our a priori hypothesis was that the VDZ-CDST specifically predicts exposure-response to VDZ in UC, and that among the high probability of response group we would see significantly higher rates of remission for VDZ vs ADA. Among the intermediate or low probability of response groups, in which VDZ was predicted to have lower exposure profiles and be less effective, we would see comparable outcomes between VDZ and ADA and possibly superiority of ADA to VDZ in patients in the low probability of response group.

Study Design

Publicly available data were used from the double-blind, head-to-head, randomized clinical trial program for VDZ vs ADA for the treatment of moderate to severe UC (VARSITY trial; NCT02497469).[3] Data access and permission for analyses were obtained through the open access data sharing platform VIVLI (Cambridge, MA; Protocol #00006686). Institutional review board and ethics review determined that approval was not needed given that data had been previously collected and because only deidentified data were being used for analyses.

Participants and Assessments

The original trial design, inclusion and exclusion criteria, and endpoint measurements have been previously reported for the VARSITY trial in the primary publication.[3] In brief, UC patients with a total Mayo score of 6 to 12 (Mayo endoscopic subscore of 2-3) were enrolled, with a cap of 25% for the recruitment of patients with prior TNF-antagonist failure. Enrolled patients were randomized to receive VDZ or ADA with follow-up assessments at weeks 14 and 52 for clinical, endoscopic, and histologic endpoints. The study design was a double blind without rerandomization or consideration for response status at week 14. Blinded central readers using the Mayo endoscopic subscore, Robarts Histopathology Index (RHI), and Geboes score assessed endoscopy and histology endpoints. Drug exposure assessments for serum concentrations of VDZ were made at baseline, and predose at weeks 6, 14, 22, 30, 38, and 52. Drug exposure assessments for serum concentrations of ADA were not available in the public dataset.

Data Variables and VDZ-CDST Calculation

Of the 769 patients enrolled in the phase 3 clinical trial program, 419 had all baseline data needed for calculating the VDZ-CDST and follow-up data for both endoscopy and histology endpoint measures. The reason for exclusion of the remaining patients was an absence of disease duration assessments at baseline, which limited the ability to calculate the VDZ-CDST score. For the 419 patients included, the baseline VDZ-CDST was then calculated as follows:[4]

Table 2. Diagnostic Performance of VDZ-CDST for VDZ-Treated vs ADA-Treated Patients

Week 52 Outcome	Sensitivity (95% CI) (%)		Specificity (95% CI) (%)		PLR (95% CI)		NLR (95% CI)	
	VDZ	ADA	VDZ	ADA	VDZ	ADA	VDZ	ADA
Clinical remission	83.9 (72.3-92.0)	71.4 (55.4-84.3)	45.5 (37.2-54.0)	41.3 (33.5-49.3)	1.54 (1.28-1.85)	1.22 (0.96-1.53)	0.35 (0.20-0.64)	0.69 (0.41-1.16)
Endoscopic remission	83.0 (70.2-91.9)	74.1 (53.7-88.9)	43.5 (35.6-51.72)	43.8 (36.5-51.3)	1.47 (1.22-1.77)	1.32 (1.02-1.70)	0.39 (0.21-0.73)	0.59 (0.31-1.14)
Histoendoscopic mucosal improvement	86.3 (73.7-94.3)	62.5 (35.4-84.8)	44.2 (36.3-52.4)	41.8 (34.9-40.1)	1.55 (1.30-1.85)	1.07 (0.72-1.60)	0.31 (0.15-0.63)	0.90 (0.47-1.72)

Performance is presented for high-probability cutoff of CDST—32 points.
Abbreviations: ADA, adalimumab; CDST, clinical decision support tool; CI, confidence interval; NLR, negative likelihood ratio; PLR, positive likelihood ratio; VDZ, vedolizumab.

- Absence of prior anti-TNF exposure: + 3 points
- Disease duration of 2 years or more:+ 3 points
- Moderate (Mayo 2) as opposed to severe (Mayo 3) baseline endoscopic activity: + 2 points
- Baseline albumin: + 0.65 point per 1 g/L

This resulted in the generation of a baseline VDZ-CDST score as a continuous variable for all patients (VDZ- and ADA-treated UC patients). This continuous variable was also further transformed into a categorical variable using previously identified and validated cutoffs to categorize patients as low (26 points or less), intermediate (27-32 points), or high (33 points or higher) probability of response to VDZ.[4] This was done for VDZ- and ADA-treated patients to generate comparative subgroups within predicted probability groupings (ie, VDZ vs ADA in the low probability of response to VDZ group; VDZ vs ADA in the intermediate probability of response to VDZ group; and VDZ vs ADA in the high probability of response to VDZ group).

Disease Activity Measures and Endpoints

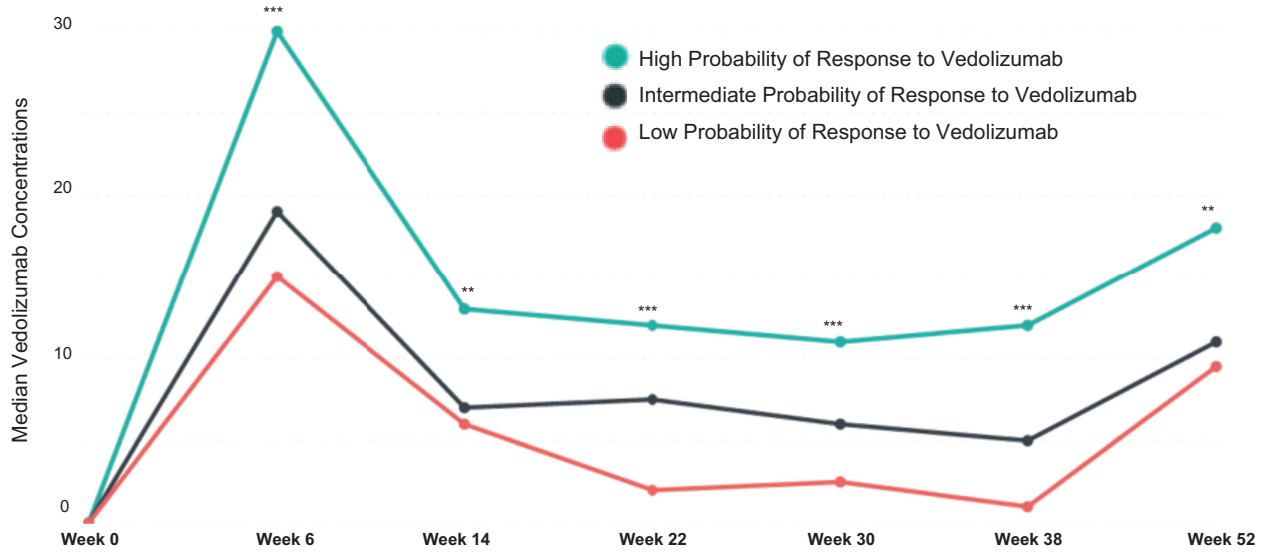
Within the original trial program, blinded assessments were made for the full Mayo score (stool frequency, rectal bleeding, physician global assessment, endoscopic subscore; range 0-12) and for the RHI and Geboes histology subscores. The primary disease activity endpoint for this analysis was week 52 histoendoscopic mucosal improvement (HEMI) defined as the achievement of a Mayo endoscopic subscore of 0-1 and a Geboes highest grade <3.2. This was chosen as the primary endpoint given the evolving integration of this endpoint in clinical trials and clinical practice, and the association between endoscopy and histology with longer-term disease outcomes in UC.[5] Secondary outcomes included week 14 HEMI, week 14 and 52 clinical remission (full Mayo score 0-2), endoscopic improvement (Mayo endoscopic subscore 0-1), endoscopic remission (Mayo endoscopic subscore 0), and histologic remission (Geboes highest grade <3.2 or RHI <5). Safety outcomes included adverse events as defined by the original trial, and included all reported adverse events ranging from mild to severe.

Drug Exposure

The baseline VDZ-CDST was previously observed to predict VDZ drug exposure profiles over 52 weeks in the original derivation cohort from the GEMINI trial programs.[4] To fully validate the ability of the VDZ-CDST to predict exposure-response relationships in this validation cohort, a secondary assessment was also made for the association between the baseline VDZ-CDST score assessed prior to VDZ treatment (as a linear score and categorical classification of low, intermediate, and high probability of response) and measured VDZ drug exposure over 52 weeks.

Statistical Analyses

Descriptive statistics were used to summarize baseline characteristics (eg, disease activity and patient demographics) as well as outcomes among patients after stratifying them based on the VDZ-CDST. Dichotomous variables are presented as proportions or percentages. Continuous variables are reported as mean ˆ SD or medians (interquartile range [IQR]). Using the VDZ-CDST as a continuous baseline variable, we assessed the area under the curve (AUC) of the receiver-operating



	Week 0	Week 6	Week 14	Week 22	Week 30	Week 38	Week 52
Low Probability		14.8 (5.3-21.0)	6.0 (4.0-20.7)	1.7 (0.5-7.8)	2.5 (0.7-3.8)	1.1 (0.6-5.6)	9.5 (4.2-13.0)
Intermediate Probability		18.9 (13.3-33.2)	7.4 (3.6-13.7)	7.7 (3.5-12.4)	5.8 (3.1-13.1)	4.7 (2.6-12.4)	11.3 (6.2-23.3)
High Probability		29.7 (23.4-39.6)	13.3 (7.2-20.4)	12.3 (7.0-17.9)	11.4 (7.9-17.4)	12.4 (8.2-16.1)	17.7 (10.1-25.0)
p value		**0.0001**	**0.0048**	**0.0001**	**0.0001**	**0.0001**	**0.0050**

Comparison of Median (IQR) Vedolizumab Concentrations between Groups Based on Probability of Response to Vedolizumab

Figure 2. Measured drug exposure over 52 weeks in vedolizumab-treated ulcerative colitis based on the baseline vedolizumab clinical decision support tool score. $^{**}P < .01$; $^{***}P < .001$; numeric values represent medians with interquartile ranges (IQRs) in parentheses. μg per milliliter.

characteristic assessments for our primary and secondary outcomes, and these were done separately for VDZ- and ADA-treated patients. AUCs for the VDZ-CDST among VDZ- and ADA-treated patients were then compared using generalized U-statistics as described by DeLong et al.[6] Subsequently, after categorization into low, intermediate, and high probability of response to VDZ, binary comparisons were made for the primary and secondary endpoints between VDZ- and ADA-treated patients within these subgroups. Sensitivity, specificity, positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were calculated, in addition to number needed to treat (NNT) and number needed to harm (NNH) to help inform shared decision-making discussions. For the comparison of drug concentrations across the 3 probability of response groups at different time points, we used the nonparametric Kruskal-Wallis test with closed-loop testing as previously done in our original GEMINI derivation cohort.[4] Confidence intervals (CIs) were calculated at the 95% interval, and *P* value comparisons were 2-tailed and held to a .01 significance threshold to account for multiple hypothesis testing. Data were analyzed using STATA version 16.0 (StataCorp, College Station, TX, USA).

Results

Demographics

Table 1 outlines demographics for the 419 patients included in this post hoc analysis, further stratified by their baseline predicted probability of response to VDZ using the VDZ-CDST. A minority (n = 21, 5%) of patients included in the VARSITY trial had a low baseline predicted probability of response to VDZ, and the majority of patients recruited had a high (n = 255, 61%) or intermediate (n = 143, 29%) baseline predicted probability of response to VDZ. Notably, TNF antagonist–naïve patients were present in all subgroups classified using the VDZ-CDST, demonstrating that not all TNF antagonist–naïve patients have an increased probability of response to VDZ and that the baseline predicted probability of response to VDZ is not entirely driven by prior TNF-antagonist exposure status.

Diagnostic Accuracy of VDZ-CDST

The baseline VDZ-CDST score demonstrated significantly better accuracy for predicting week 52 HEMI in VDZ-treated patients (AUC, 0.712; 95% CI, 0.636-0.787) relative to ADA-treated patients (AUC, 0.538; 95% CI, 0.377-0.700; $P < .001$ for comparison of AUC between VDZ- and ADA-treated patients) (Figure 1). A baseline VDZ-CDST score of 32 points or higher, which delineates low-intermediate baseline predicted probability of response to VDZ from high baseline predicted probability of response to VDZ, had good sensitivity for identifying VDZ-treated patients who achieved clinical remission, endoscopic remission, or HEMI by week 52 but had fair to poor sensitivity for identifying ADA-treated patients who achieved these endpoints by week 52 (Table 2). Notably, the PLR and NLR among ADA-treated patients were not significant as CIs crossed 1.0, but the PLR was consistently >1.4 and the NLR was consistently <0.4 among VDZ-treated patients, and the CIs did not cross 1.0.

Drug Exposure Predictions of VDZ-CDST

Measured VDZ concentrations were significantly and incrementally higher among baseline VDZ-CDST strata (Figure 2). Week 6 trough VDZ concentrations were 2-fold higher in

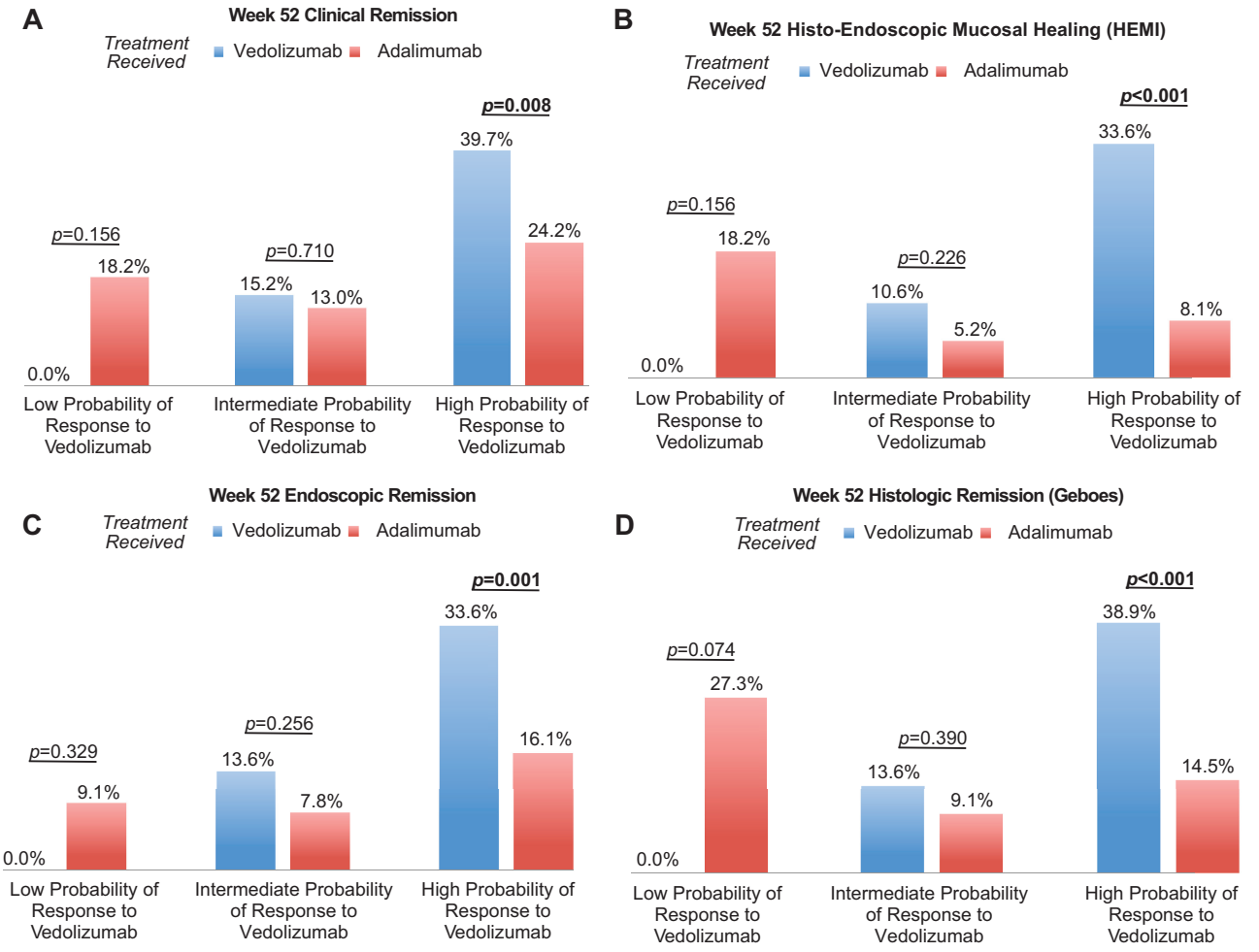


Figure 3. Comparative effectiveness of vedolizumab (VDZ) vs adalimumab (ADA) after patient stratification using the VDZ clinical decision support tool (VDZ-CDST). The y-axis represents proportion of patients achieving this outcome. Patients in the clinical trial program were stratified using the VDZ-CDST calculation at baseline prior to treatment exposure, and then comparisons were made between VDZ- and ADA-treated patients at week 52 within each baseline VDZ response probability subgroup. A, Week 52 clinical remission (Full Mayo score 0-2) comparison of VDZ vs ADA after stratification based on the VDZ-CDST. B, Week 52 histoendoscopic mucosal improvement (HEMI) comparison of VDZ vs ADA after stratification based on the VDZ-CDST. C, Week 52 endoscopic remission (Mayo endoscopic subscore 0) comparison of VDZ vs ADA after stratification based on the VDZ-CDST. D, Week 52 histologic remission (Geboes score of <3.2) comparison of VDZ vs ADA after stratification based on the VDZ-CDST.

the high probability of response to VDZ group (29.7 [IQR, 23.4-39.6] µg/mL) vs the low probability of response to VDZ group (14.8 [IQR, 5.3-21.0] µg/mL; $P < .001$ for comparison across groups), with nonoverlapping IQR intervals. During maintenance, week 30 trough VDZ concentrations were over 4-fold higher in the high probability of response to VDZ group (11.4 [IQR, 7.9-17.4] µg/mL) vs the low probability of response to VDZ group (2.5 [IQR, 0.7-3.8] µg/mL; $P < .001$ for comparison across groups), again with nonoverlapping IQR intervals.

Comparative Effectiveness of VDZ and ADA Based on VDZ-CDST Classification

Upon stratifying the entire cohort into strata with low, intermediate, and high VDZ-CDST baseline predicted probability of response, significant differences in treatment outcomes between VDZ and ADA were only observed in the patients with high baseline predicted probability of response to VDZ (Figure 3; Supplementary Tables 1-3). Numeric rates of treatment outcomes were comparable between VDZ and ADA among the subgroup classified as having an intermediate baseline predicted probability of response to VDZ. They were numerically better, but not statistically significant, for ADA vs VDZ among the subgroup classified as having a low baseline predicted probability of response to VDZ. Based on these comparative estimates within the VDZ-CDST subgroups, NNT estimates are provided in Table 3. The NNT favored VDZ over ADA in the entire cohort and ranged from 5.8 to 9.8. Among the low baseline predicted probability of response to VDZ subgroup, the NNT favored ADA and ranged from 3.7 to 11. Among the intermediate baseline predicted probability of response to VDZ subgroup, the NNT favored VDZ and ranged from 11.2 to 45.5. The overall NNT estimate for VDZ vs ADA were driven by the patients in the high baseline predicted probability of response to VDZ subgroup, in which NNT favored VDZ and ranged from 3.9 to 6.5.

Upon limiting the cohort to TNF antagonist–naïve patients and then stratifying by the VDZ-CDST baseline predicted probability of response, we observed no significant differences between VDZ- and ADA-treated patients

Table 3. NNT or NNH for Vedolizumab vs Adalimumab in Ulcerative Colitis

Week 52 Outcome	NNT (Which Drug Is Favored for Effectiveness)	NNH (Which Drug iIs Favored for Safety)[a]
Entire cohort (without stratification using the decision support tool)		
Clinical remission	9.8 (vedolizumab)	5.7 (vedolizumab)
Histoendoscopic mucosal improvement	5.8 (vedolizumab)	
Endoscopic improvement	7.4 (vedolizumab)	
Endoscopic remission	7.8 (vedolizumab)	
Histologic remission (Geboes)	6.8 (vedolizumab)	
Histologic remission (RHI)	7.9 (vedolizumab)	
Low probability of response to vedolizumab		
Clinical remission	5.5 (adalimumab)	22.6 (adalimumab)
Histoendoscopic mucosal improvement	5.5 (adalimumab)	
Endoscopic improvement	5.5 (adalimumab)	
Endoscopic remission	11 (adalimumab)	
Histologic remission (Geboes)	3.7 (adalimumab)	
Histologic remission (RHI)	5.5 (adalimumab)	
Intermediate probability of response to vedolizumab		
Clinical remission	45.5 (vedolizumab)	160.5 (adalimumab)
Histoendoscopic mucosal improvement	18.5 (vedolizumab)	
Endoscopic improvement	17.9 (vedolizumab)	
Endoscopic remission	17.2 (vedolizumab)	
Histologic remission (Geboes)	22.2 (vedolizumab)	
Histologic remission (RHI)	11.2 (vedolizumab)	
High probability of response to vedolizumab		
Clinical remission	6.5 (vedolizumab)	3.6 (vedolizumab)
Histoendoscopic mucosal improvement	3.9 (vedolizumab)	
Endoscopic improvement	5.2 (vedolizumab)	
Endoscopic remission	5.7 (vedolizumab)	
Histologic remission (Geboes)	4.1 (vedolizumab)	
Histologic remission (RHI)	6.3 (vedolizumab)	

Comparisons between vedolizumab and adalimumab were only statistically significantly different in the high probability of response group, and other NNT and NNH estimates should be interpreted with caution.
Abbreviations: NNH, number needed to harm; NNT, number needed to treat; RHI, Robarts Histopathology Index.
[a]Calculated using total number of adverse events and duration of exposure.

classified as being low or intermediate probability of response to VDZ across all outcomes, and again observed that superiority of VDZ to ADA was only observed in patients classified as having a high baseline probability of response to VDZ (Supplementary Tables 4-6). Rates of HEMI at 52 weeks for VDZ- and ADA-treated patients who were TNF antagonist naïve but classified as having an intermediate baseline probability of response according to the VDZ-CDST (n = 45 VDZ, n = 55 ADA) were low at 11.1% and 5.5%, respectively (P = .30). Rates of HEMI at 52 weeks for VDZ- and ADA-treated patients who were TNF antagonist naïve and classified as having a high baseline probability of response according to the VDZ-CDST (n = 112 VDZ, n = 111 ADA) were 33.9% and 8.1%, respectively (P < .001).

Comparative Safety of VDZ and ADA Based on VDZ-CDST

The proportions of patients who experienced adverse events were not significantly different between VDZ- and ADA-treated patients within the subgroups of low (20% vs 9.1%; P = .476), intermediate (10.6% vs 18.2%; P = .202), and high (14.5% vs 17.7%; P = .350) baseline predicted probability of response to VDZ. The total number of adverse events were also not significantly different between VDZ- and ADA-treated patients within the subgroups of low (n = 2 vs n = 1; P = .476) or intermediate (n = 17 vs n = 16; P = .806) baseline predicted probability of response to VDZ; however, the total number of adverse events was significantly lower for VDZ-treated patients compared with ADA-treated patients within the subgroup classified as having a high baseline predicted probability of response to VDZ based on the VDZ-CDST (n = 25 vs n = 117; P < .001). NNH estimates are provided in Table 3.

Discussion

In this post hoc analysis of trial VARSITY trial we have made several key observations that will better inform the positioning and selection of biologics for UC. First, in this external validation cohort a CDST built for VDZ predicted clinical, endoscopic, and histologic remission with VDZ but not with ADA in UC. The drug specificity of prediction for this CDST and the extended validation for histologic outcomes

improved its clinical utility. Second, most VARSITY trial patients had a high baseline predicted probability of response to VDZ, and superiority of VDZ to ADA was only observed in these patients with a high baseline predicted probability of response to VDZ. In contrast, ADA was favored in patients with a low baseline predicted probability of response to VDZ. This subcohort of patients with a high probability of response to VDZ included both TNF antagonist–naïve and TNF antagonist–exposed individuals, and a subset of TNF antagonist–naïve patients were identified to have a low or intermediate baseline predicted probability of response to VDZ. Finally, we have further validated our VDZ-CDSTs ability to predict measured VDZ exposure and added a new validation for safety in which it was observed that our VDZ-CDST tool identifies patients who are at an increased risk for adverse events.

Prior TNF-antagonist exposure has been associated with reduced effectiveness for VDZ,[7] and the VARSITY trial observed that superiority of VDZ to ADA was primarily seen in the TNF antagonist–naïve subgroup.[3] Our VDZ-CDST incorporates prior TNF-antagonist exposure status within the scoring; however, it also takes into consideration other variables found to be predictive of response to VDZ including disease duration, baseline endoscopic severity, and baseline albumin. This allows for the identification of TNF antagonist–exposed patients who have a high probability of response to VDZ and TNF antagonist–naïve patients who have a low or intermediate probability of response to VDZ. Combined with the extended validation of our VDZ-CDST for endoscopic and histologic endpoints, and comparative validation for VDZ vs ADA-treated patients, this suggests that the VDZ-CDST be considered first when choosing between VDZ and ADA therapy in practice.

In our original derivation cohort, we observed that the baseline VDZ-CDST score was able to predict measured VDZ concentrations over 52 weeks, with significant ($P < .001$) differences in VDZ concentrations across the 3 predicted probability of response groups. We have now validated this observation in the current study, thereby confirming that our VDZ-CDST is capable of predicted exposure-response relations. Notably, the measured VDZ concentrations were 2- to 5-fold higher in the high probability of response group compared with the low probability of response group, and median concentrations for VDZ in the high probability of response group during maintenance therapy were approximately 12 µg/mL. A systematic review has suggested that VDZ concentrations >12 µg/mL during maintenance may be associated with better clinical outcomes,[8] which is in keeping with the observed median VDZ concentration during maintenance among the highest-responding subgroup of VDZ-treated UC patients. Future consideration might therefore be given to whether the VDZ-CDST could be used to identify patients at risk for lower drug concentrations who might benefit from early therapeutic drug monitoring and dose optimization for persistent disease activity. The ongoing VDZ dose optimization trial in UC (ENTERPRET trial; NCT03029143) will determine whether dose optimization is of value for VDZ therapy, and future post hoc analyses of our VDZ-CDST in that dataset might allow for validation of this hypothesis.

Shared decision making requires a discussion of both effectiveness and safety, and ideal decision support tools are capable of differentiating therapies that are more effective and less risky. We observed that UC patients with a high baseline predicted probability of response to VDZ were more likely to achieve key effectiveness outcomes with VDZ but were less likely to experience adverse events with VDZ therapy compared with ADA. Furthermore, UC patients with a low baseline predicted probability of response to VDZ were more likely to achieve key effectiveness outcomes with ADA and also less likely to experience adverse events with ADA therapy, confirming the relationship between treatment response and treatment safety. A similar observation has been made by our group for the VDZ-CDST in Crohn's disease, in which patients with a high baseline predicted probability of response to VDZ were less likely to experience an adverse event during the long-term safety extension studies.[9,10] A consideration could be given to whether treatment safety is the inverse of treatment effectiveness, and the safest drug for an individual is the one that is likely to be most effective and result in tapering of steroids, resolution of symptoms, avoidance of narcotics, and healing of inflammation, all of which are known risk factors for adverse events in inflammatory bowel disease.[11,12]

Our study has several strengths including the use of a prospective randomized comparative effectiveness trial, blinded assessments for endoscopy and histology, and consistency of observations for this clinical trial validation cohort compared with the prior clinical trial derivation cohort and the routine practice validation cohort. The study is not without limitations, which include the ability to use only a subset of the original VARSITY trial program cohort, the small subcohort of patients who had a low baseline predicted probability of response to VDZ, and the post hoc nature of these analyses. Furthermore, none of the patients in the VARSITY trial received dose optimization, and no standardized steroid tapering was done, which may create some inherent biases between study groups. Notably, in the original model development and validation process from the GEMINI trial, concomitant immunomodulators or steroids did not predict treatment response, however, further validation work will need to take these factors into consideration, recognizing the practice variability in use and prescribing patterns. We also did not have access to ADA concentrations to confirm the drug specificity of this prediction for VDZ exposure profiles, or reason for prior TNF-antagonist failure to account for mechanism of failure when making comparisons in TNF antagonist–exposed patients. Prospective validation of our VDZ-CDST would help to address these remaining limitations and determine whether the tool can be used to guide not only biologic selection, but also considerations for therapeutic drug monitoring and dose optimization.

In conclusion, we have demonstrated that a CDST built for VDZ in UC is capable of predicting measured VDZ concentrations and identifying patients most likely to achieve clinical, endoscopic, and histologic remission with VDZ but not with ADA. Furthermore, we have determined that the superiority of VDZ over ADA in UC is predominately due to the inclusion of patients in the trial who had a high baseline predicted probability of response to VDZ, and this includes both TNF antagonist–naïve and TNF antagonist–exposed individuals. Our VDZ-CDST should therefore be considered for use in routine practice to guide biologic treatment selection for UC. To aid in the uptake of this tool, it is freely available to providers (https://via.juxlyapps.com/pathway/archemedx/

ibd-cdst/index.html#/disease-selection) and has recently gained sponsorship from the American Gastroenterological Association as an accredited educational activity.[13] This helps to bring us 1 step closer to precision medicine, and further augmentation and refinement of our tool with translational parameters may help to ultimately bridge the gap of personalized care.

Supplementary data

Supplementary data is available at *Inflammatory Bowel Diseases* online.

Acknowledgments

This publication (Vivli protocol #00006686) is based on research using data from Takeda that have been made available through Vivli. Vivli has not contributed to or approved, and is not in any way responsible for, the contents of this publication.

Author Contributions

Study concept and design: P.S.D., N.N.; acquisition and compilation of data: E.C.L.W., N.N.; statistical analyses: E.C.L.W.; data interpretation: all authors; drafting and final approval of manuscript: all authors.

Funding

P.S.D. and the work outlined in this manuscript was directly supported by an American Gastroenterology Association Research Scholar Award and National Institutes of Health Digestive Diseases Research Center grant DK120515. N.M. holds a McMaster University Department of Medicine Internal Career Award. P.S.D. and the University of California San Diego hold a provisional patent for the prediction modeling and decision support tool outlined herein (U.S. 16/792,090).

Conflicts of Interest

P.S.D. has received research support and consulting for Takeda, AbbVie, Janssen, and Pfizer. W.R. has received support for the following: speaker for Abbott Laboratories, AbbVie, Aesca, Aptalis, Astellas, Centocor, Celltrion, Danone Austria, Elan, Falk Pharma GmbH, Ferring, Immundiagnostik, Mitsubishi Tanabe Pharma Corporation, MSD, Otsuka, PDL, Pharmacosmos, PLS Education, Schering-Plough, Shire, Takeda, Therakos, Vifor, and Yakult; consultant for Abbott Laboratories, AbbVie, Aesca, Algernon, Amgen, AM Pharma, AMT, AOP Orphan, Arena Pharmaceuticals, Astellas, AstraZeneca, Avaxia, Roland Berger GmBH, Bioclinica, Biogen IDEC, Boehringer Ingelheim, Bristol-Myers Squibb, Cellerix, Chemocentryx, Celgene, Centocor, Celltrion, Covance, Danone Austria, DSM, Elan, Eli Lilly, Ernest & Young, Falk Pharma GmbH, Ferring, Galapagos, Genentech, Gilead, Grünenthal, ICON, Index Pharma, Inova, Janssen, Johnson & Johnson, Kyowa Hakko Kirin Pharma, Lipid Therapeutics, LivaNova, Mallinckrodt, Medahead, MedImmune, Millenium, Mitsubishi Tanabe Pharma Corporation, MSD, Nash Pharmaceuticals, Nestle, Nippon Kayaku, Novartis, Ocera, OMass, Otsuka, Parexel, PDL, Periconsulting, Pharmacosmos, Philip Morris Institute, Pfizer, Procter & Gamble, Prometheus, Protagonist, Provention, Robarts Clinical Trials, Sandoz, Schering-Plough, Second Genome, Seres Therapeutics, Setpointmedical, Sigmoid, Sublimity, Takeda, Therakos, Theravance, Tigenix, UCB, Vifor, Zealand, Zyngenia, and 4SC; and advisory board member for Abbott Laboratories, AbbVie, Aesca, Amgen, AM Pharma, Astellas, AstraZeneca, Avaxia, Biogen IDEC, Boehringer Ingelheim, Bristol-Myers Squibb, Cellerix, Chemocentryx, Celgene, Centocor, Celltrion, Danone Austria, DSM, Elan, Ferring, Galapagos, Genentech, Grünenthal, Inova, Janssen, Johnson & Johnson, Kyowa Hakko Kirin Pharma, Lipid Therapeutics, MedImmune, Millenium, Mitsubishi Tanabe Pharma Corporation, MSD, Nestle, Novartis, Ocera, Otsuka, PDL, Pharmacosmos, Pfizer, Procter & Gamble, Prometheus, Sandoz, Schering-Plough, Second Genome, Setpointmedical, Takeda, Therakos, Tigenix, UCB, Zealand, Zyngenia, and 4SC. J.-F.C. reports receiving research grants from AbbVie, Janssen Pharmaceuticals, and Takeda; receiving payment for lectures from AbbVie, Amgen, Allergan, Ferring Pharmaceuticals, Shire, and Takeda; and receiving consulting fees from AbbVie, Amgen, Arena Pharmaceuticals, Boehringer Ingelheim, Celgene Corporation, Celltrion, Eli Lilly, Enterome, Ferring Pharmaceuticals, Genentech, Janssen Pharmaceuticals, Landos, Ipsen, Medimmune, Merck, Novartis, Pfizer, Shire, Takeda, Tigenix, Viela bio; and holds stock options in Intestinal Biotech Development and Genfit. J.K.M. has received honoraria from Janssen, AbbVie, Allergan, Bristol-Myers Squibb, Ferring, Janssen, Lilly, Lupin, Merck, Pfizer, Pharmascience, Roche, Shire, Takeda, and Teva. N.N. has received honoraria from Janssen, AbbVie, Takeda, Pfizer, Merck, Sandoz, Novartis, and Ferring. No other authors have any relevant conflicts of interest. No authors have received support for the submitted manuscript. All authors approved the final version of the manuscript.

REFERENCES

1. Rubin DT, Ananthakrishnan AN, Siegel CA, et al. ACG clinical guideline: ulcerative colitis in adults. *Am J Gastroenterol*. 2019;114:834-413.
2. Feuerstein JD, Isaacs KL, Schneider Y, et al. AGA clinical practice guidelines on the management of moderate to severe ulcerative colitis. *Gastroenterology*. 2020;158:1450-1461.
3. Sands BE, Peyrin-Biroulet L, Loftus EV Jr, et al.; VARSITY Study Group. Vedolizumab versus adalimumab for moderate-to-severe ulcerative colitis. *N Engl J Med*. 2019;381:1215-1226.
4. Dulai PS, Singh S, Casteele NV, et al. Development and validation of clinical scoring tool to predict outcomes of treatment with vedolizumab in patients with ulcerative colitis. *Clin Gastroenterol Hepatol*. 2020;18:2952-2961.e8.
5. Dulai PS, Levesque BG, Feagan BG, et al. Assessment of mucosal healing in inflammatory bowel disease: review. *Gastrointest Endosc*. 2015;82:246-255.
6. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.
7. Narula N, Peerani F, Meserve J, et al. Vedolizumab for ulcerative colitis: treatment outcomes from the VICTORY consortium. *Am J Gastroenterol*. 2018;113:1345.
8. Singh S, Dulai PS, Vande Casteele N, et al. Systematic review with meta-analysis: association between vedolizumab trough concentration and clinical outcomes in patients with inflammatory bowel diseases. *Aliment Pharmacol Ther*. 2019;50:848-857.
9. Dulai PS, Peyrin-Biroulet L, Demuth D, et al. Early intervention with vedolizumab on longer term surgery rates in crohn's disease: post hoc analysis of the GEMINI phase 3 and long-term safety programs. *J Crohns Colitis*. 2020;15:195-202.

10. Dulai P, Feng D, Lasch K. 688 Probability of vedolizumab response as defined by clinical decision support tool is associated with lower risk of serious infection in patients with inflammatory bowel disease. *Am J Gastroenterol*. 2019;114: S403-S404.
11. Lichtenstein GR, Feagan BG, Cohen RD, et al. Serious infection and mortality in patients with Crohn's disease: more than 5 years of follow-up in the TREAT™ registry. *Am J Gastroenterol*. 2012;107:1409-1422.
12. Dulai PS, Siegel CA. Optimization of drug safety profile in inflammatory bowel disease through a personalized approach. *Curr Drug Targets*. 2018;19:740-747.
13. American Gastroenterological Association endorses the IBD Clinical Decision Support Tool (CDST) developed by the US IBD Health Outcomes Consortium. *AP News*, February 18, 2021. Accessed February 21, 2021. https://apnews.com/press-release/business-wire/technology-business-science-corporate-news-products-and-services-253efecbd0194bcd8ad1a07073b141fb