

# Head-to-head comparison of biological drugs for inflammatory bowel disease: from randomized controlled trials to real-world experience

Fabio Salvatore Macaluso<sup>1</sup>, Marcello Maida, Mauro Grova, Federica Crispino, Giulia Teresi, Adele Orlando and Ambrogio Orlando<sup>2</sup>

**Abstract:** During past years, the increasing knowledge of molecular mechanisms of inflammatory bowel disease (IBD) have led to the development of several targeted biological therapies. This great expansion of available medical options has prompted the need for comparative data between drugs. For years, given that most randomized controlled trials (RCTs) were performed only *versus* placebo, this demand has clashed with the absence of head-to-head trials comparing two or more treatments. The quality of evidence coming from real-world experience was low overall, so it was extremely difficult to clarify the correct positioning of the biologicals inside the therapeutic algorithms for IBD. Fortunately, times are changing: head-to-head comparative RCTs have been conducted or are ongoing, and the methodological quality of real-world studies is gradually increasing, mainly thanks to a higher rate of application of statistical methods capable of reducing the selection bias, such as the propensity score. In this evolving scenario, the increasing number of comparative RCTs is providing high-quality data for a correct drug positioning in IBD. In parallel, real-world observational studies are supporting the data coming from RCTs, and covering those comparisons not performed in the RCT setting. We believe that there is moderate evidence already available to support clinicians in the correct choice between different biologicals, and data will certainly be more robust in the near future.

**Keywords:** biologicals, propensity score analysis, randomized controlled trials, real-world experience

Received: 31 December 2020; revised manuscript accepted: 29 March 2021.

## Introduction

Inflammatory bowel disease (IBD) – a term including Crohn's disease (CD), ulcerative colitis (UC), and unclassified colitis – are chronic conditions that require lifelong management, as the precise etiology is unknown and a causal therapy does not exist.<sup>1</sup> Considerable improvement in the understanding of molecular mechanisms of these diseases has been achieved over the past two decades,<sup>2</sup> leading to the introduction of novel targeted therapies, namely biologicals – that is, monoclonal antibodies that selectively block key

mediators of inflammation – and novel small molecule drugs – that is, compounds with a molecular weight <1 kDa able to diffuse through cell membranes and then fit for the oral route of administration.<sup>3</sup> The expansion of the therapeutic options – coupled with the intention to cure beyond symptoms, that is, to achieve ambitious therapeutic targets such as mucosal healing,<sup>4</sup> deep remission,<sup>5</sup> and histological healing<sup>6</sup> – makes IBD management more and more complex. In this scenario, all physicians dealing with IBD would like to know which drugs are most effective

*Ther Adv Gastroenterol*

2021, Vol. 14: 1–11

DOI: 10.1177/  
17562848211010668

© The Author(s), 2021.  
Article reuse guidelines:  
sagepub.com/journals-  
permissions

Correspondence to:

**Ambrogio Orlando**  
Department of Medicine,  
IBD Unit, "Villa Sofia-  
Cervello" Hospital, Viale  
Strasburgo 233, Palermo  
90146, Italy  
[ambrogiorlando@gmail.com](mailto:ambrogiorlando@gmail.com)

**Fabio Salvatore Macaluso**  
**Giulia Teresi**  
**Adele Orlando**  
Department of Medicine,  
IBD Unit, "Villa Sofia-  
Cervello" Hospital,  
Palermo, Italy

**Marcello Maida**  
Gastroenterology and  
Endoscopy Unit, "S.  
Elia-M. Raimondi"  
Hospital, Caltanissetta,  
Italy

**Mauro Grova**  
**Federica Crispino**  
Department of Medicine,  
IBD Unit, "Villa Sofia-  
Cervello" Hospital,  
Palermo, Italy  
Department of Health  
Promotion, Sciences,  
Maternal and Infant  
Care, Section of  
Gastroenterology and  
Hepatology, Internal  
Medicine and Medical  
Specialties, PROMISE,  
University of Palermo, Italy

for their patients, in order to make the correct therapeutic choices. In other words, there is a great demand for comparative data between drugs, a need that for years has clashed with the absence of head-to-head trials comparing two (or more) treatments and with the availability of randomized controlled trials (RCTs) analyzing the efficacy of active compounds only *versus* placebo. Furthermore, the quality of evidence arising from real-world experience (RWE) was low overall – with few exceptions – so that it was extremely difficult to clarify the correct positioning of drugs inside the therapeutic algorithms for IBD. Fortunately, times are changing: head-to-head RCTs have been conducted or are ongoing, and the methodological quality coming from RWE is gradually increasing following the application of statistical methods capable of reducing the selection bias, such as the propensity score.

This review aims to summarize the current available evidence on head-to-head comparisons of biological drugs for IBD, based on data from RCTs and RWE.

### **RCTs *versus* RWE: and the winner is. . . ?**

When a treatment is clearly effective, no sophisticated statistical analysis is necessary. For instance, no complex study design is needed to demonstrate the efficacy of paracetamol in reducing the body temperature of patients with fever. Unfortunately, this is not the rule, as most drugs are unlikely to cause significant and positive effects in all patients, particularly in heterogeneous diseases such as IBD. As a consequence, we often need large sample sizes to demonstrate mild effects. This is also true when a study is designed to compare two treatments. In this context, RCTs are at the top of the hierarchy of evidence-based medicine. This is clearly expressed by the two methodologies most frequently used to grade the quality of evidence and the strength of recommendations, key points for the development of clinical guidelines: the Oxford levels of evidence and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria. The Oxford levels of evidence rank systematic reviews (with homogeneity) of RCTs and individual RCTs (with narrow confidence intervals) at the top of the evidence (1a and 1b, respectively), while the evidence arising from cohort and case-control studies is positioned lower (2a: systematic reviews of cohort studies; 2b: individual cohort study; 2c:

'outcomes' research/ecological studies; 3a: systematic reviews with homogeneity of case-control studies; 3b: individual case-control study; 4: case-series and poor quality cohort and case-control studies). The GRADE approach is a system for rating the quality of a body of evidence in systematic reviews in which the evidence from RCTs is graded as high/moderate quality, while that arising from RWE is graded as low/very low quality. The high level of evidence of RCTs is due to their rigorous data reporting and to the strength of randomization, a procedure that eliminates known and unknown confounders of treatment effect. In other words, RCTs are designed to explore (or compare) the intrinsic efficacy of drugs in ideal circumstances. However, RCTs are not perfect tools. The major drawback of RCTs lies in their poor external validity, as there is a well-known sharp discrepancy between the patients enrolled in clinical trials and those encountered in clinical practice.<sup>7</sup> Pair-wise or network meta-analyses of RCTs are often used to compare different treatments, and their overall quality of evidence is judged to be high. Even though pair-wise and network meta-analyses evaluate data from direct comparisons in RCTs (drug *versus* drug or drug *versus* placebo), their final comparisons should be considered as indirect. Keeping in mind that indirect evidence can be helpful in the absence of head-to-head RCTs, meta-analyses should be interpreted with caution, as they may suffer from the same limitations as the RCTs they pool (poor external validity) and from big differences in the study designs.<sup>8</sup> As a consequence of these issues and the current paucity of head-to-head RCTs, the direct comparisons between drugs in IBD mostly rely on observational studies. RWE ranks one step lower in the hierarchy of evidence-based medicine due to numerous factors: observational studies are often retrospective, data collection is not rigorous, they suffer from selection bias, the sample size and length of follow-up are often too small to give reliable results.<sup>9</sup> In addition, the definitions of study outcomes are often not rigorous: response and remission are usually based on clinical judgement, and not on validated indexes or objective measures of inflammation such as C-reactive protein or fecal calprotectin, while endoscopic data are usually limited or absent. Over the past few years, their overall quality has increased. This improvement is mainly due to two factors: (a) the implementation of large, multi-center cohorts able to create considerable sample sizes; (b) the application of rigorous statistical

methodologies able to reduce the selection bias. In this context, the use of the propensity score has become increasingly popular in recent years.<sup>10</sup> This is a score built according to the likelihood that a treatment has been administered to a patient taking into account all covariates that may influence this choice. It thus acts as a proxy between treatment and confounders in order to simulate a sort of randomization. It should be noted that this score cannot create a real randomization, but only an adjustment for known confounders, while the adjustment for unknown confounders can be obtained only with the randomization.<sup>11</sup> It is still unclear whether the propensity score is superior to standard multivariable analyses, such as simple linear models or complex hierarchical models taking into account multiple levels – methods aiming to explore the relationship between a dependent variable and two or more independent variables (covariates) by concomitantly adjusting for the contributions of all the other covariates present in the predictive model.

On the basis of all the aforementioned considerations, it would be incorrect to put RCTs and RWE on two opposite levels. They should be seen as complementary settings, in which the qualitatively high evidence obtained by RCTs can be confirmed – that is, externally validated – by the data obtained by real world studies, particularly when they are conducted with rigorous statistical methods such as the propensity score analysis. In addition, other relevant applications of RWE can be emphasized, such as investigation on disease subtypes, outcomes, or populations that are usually excluded from clinical trials (e.g. assessment of the effectiveness of a drug on extra-intestinal manifestations, or analysis of safety and effectiveness on elderly or frail patients), as well as the evaluation of the long-term safety of drugs using prospective observational registries.

## Head-to-head comparative RCTs of biologicals in IBD

### *The VARSITY study*

The VARSITY study<sup>12</sup> is considered a milestone for scientific literature of IBD, because it is the first head-to-head comparative RCT between two drugs. In detail, it was a multicenter, phase IIb, double-blind, double-dummy, randomized trial which compared vedolizumab (VDZ) with adalimumab (ADA)

in 729 patients with moderately to severely active UC. Previous exposure to tumor necrosis factor (TNF)- $\alpha$  inhibitors was allowed in up to 25% of patients. Dose escalation was not permitted in either treatment group. The primary outcome was clinical remission at week 52 (defined with the full Mayo score), while endoscopic improvement and corticosteroid-free remission at week 52 were secondary endpoints. At week 52, clinical remission was observed in a higher percentage in the VDZ group than in the ADA group (31.3% *versus* 22.5%;  $p=0.006$ ), as was endoscopic improvement (39.7% *versus* 27.7%;  $p<0.001$ ). Conversely, corticosteroid-free clinical remission occurred more frequently in the ADA group (12.6% *versus* 21.8%), although this difference was not statistically significant.

While recognizing the courageous pioneering approach in comparing two drugs for the first time in IBD history, the VARSITY study had some important limitations. First of all, it included mostly TNF- $\alpha$  inhibitor naive patients, while VDZ is used, in clinical practice, mostly as a second-line drug, after failure of TNF- $\alpha$  inhibitor(s). This issue, coupled with the lack of treatment optimization, represents a clear gap between the setting of this RCT and clinical practice. Finally, evidence of a greater rate of corticosteroid-free clinical remission in patients with ADA highlights a certain degree of variability in reduction or withdrawal of steroids that may have influenced all clinical endpoints.

### *Other head-to-head comparative RCTs*

To date, the VARSITY study is the only head-to-head comparative RCT to have been fully published, but other RCTs have recently been completed, or are nearing completion, or are in progress (Table 1). Some of them compare directly two different biologicals, while others have a placebo-controlled arm plus an active comparator arm. These studies are mainly testing etrolizumab (ETRO) – an anti-integrin that selectively inhibits  $\alpha 4\beta 7$  and  $\alpha E\beta 7$  to control trafficking of immune cells into the gut and their inflammatory effects on the intestinal lining<sup>13</sup> – and several selective interleukin (IL)-23 inhibitors, *versus* ustekinumab (UST) or VDZ. There are no planned head-to-head RCTs involving Janus kinase inhibitors (JAK) inhibitors or modulators of the sphingosine-1-phosphate receptor.

**Table 1.** Comparative head-to-head or placebo and active-controlled trials in inflammatory bowel disease.

Study name or NCT number	Drugs	Disease – patients	Status	Estimated completion date
VARSITY, <sup>12</sup>	Vedolizumab <i>versus</i> adalimumab	UC – 75% Bio-naive	Completed, data published	–
EFFICACI	Vedolizumab <i>versus</i> infliximab	UC – Failure to subcutaneous anti-TNFs	Recruiting	4 January 2022
SEAVUE	Ustekinumab <i>versus</i> adalimumab	CD – Bio-naive patients	Active, not recruiting	4 June 2021
HIBISCUS I, <sup>14</sup>	Etrolizumab <i>versus</i> adalimumab <i>versus</i> placebo	UC – Anti-TNF-naive patients	Completed, data presented	–
HIBISCUS II, <sup>15</sup>	Etrolizumab <i>versus</i> adalimumab <i>versus</i> placebo	UC – Anti-TNF-naive patients	Completed, data presented	
GARDENIA, <sup>16</sup>	Etrolizumab <i>versus</i> infliximab	UC – Anti-TNF-naive patients	Completed, data presented	
EXPEDITION	Brazikumab <i>versus</i> vedolizumab <i>versus</i> placebo	UC	Recruiting	4 December 2022
NCT03759288	Brazikumab <i>versus</i> adalimumab <i>versus</i> placebo	CD	Recruiting	4 December 2022
LUCENT-ACT	Mirikizumab <i>versus</i> vedolizumab <i>versus</i> placebo	UC	Not yet recruiting	5 June 2024
VIVID-1	Mirikizumab <i>versus</i> ustekinumab <i>versus</i> placebo	CD	Recruiting	25 July 2023
NCT04524611	Risankizumab <i>versus</i> ustekinumab	CD – Failure to anti-TNFs	Recruiting	11 June 2023
GALAXI	Guselkumab <i>versus</i> ustekinumab <i>versus</i> placebo	CD	Recruiting	15 October 2024
NCT03558152	UTTR1147A <i>versus</i> vedolizumab <i>versus</i> placebo	UC	Recruiting	12 January 2022
TRIDENT	JNJ-64304500 <i>versus</i> ustekinumab <i>versus</i> placebo	CD	Active, not recruiting	10 September 2022

CD, Crohn's disease; TNF, tumor necrosis factor; UC, ulcerative colitis.

Results of the HIBISCUS I<sup>14</sup> and II<sup>15</sup> studies have recently been disclosed. These are two identical, randomized, double-blind, double-dummy, placebo-controlled, multicenter studies evaluating the efficacy and safety of ETRO *versus* ADA and placebo in patients with moderately to severely active UC who have not been previously treated with TNF- $\alpha$  inhibitors. In the HIBISCUS I induction study, ETRO met the primary endpoint (percentage of participants in clinical remission with ETRO compared to placebo at week 10), while the HIBISCUS II induction study did

not meet the primary endpoint, which was the same as HIBISCUS I.

GARDENIA was a randomized, double-blind, double-dummy study evaluating the efficacy and safety of ETRO *versus* infliximab (IFX) in patients with moderately to severely active UC who have not been previously treated with TNF- $\alpha$  inhibitors.<sup>16</sup> The primary endpoint was a combination of clinical response at week 10 and clinical remission at week 54, which was similar for IFX (19.7%) and ETRO (18.6%). Similarly, relevant

secondary endpoints – including endoscopic improvement, endoscopic remission, and clinical remission at week 54 – were not different between the two drugs. On the basis of these results, the future of ETRO as an effective treatment for UC is currently uncertain.

### Head-to-head comparisons of biologicals in IBD arising from RWE

There is a huge body of RWE on the comparison of effectiveness between different drugs in IBD. In order to provide a summary of the more relevant studies, we selected those with a higher methodological quality, with a particular focus on those based on the application of the propensity score (Table 2).

#### *Comparisons between different TNF- $\alpha$ inhibitors*

As IFX and ADA are the main TNF- $\alpha$  inhibitors used in patients with CD, the absence of comparative RCTs was compensated by several real-world studies aiming at comparing the two biologicals. Overall, the two drugs appeared to be equally effective in patients with CD. A 2014 Dutch study reported no significant difference in 1-year and 2-year rates of steroid-free clinical response between ADA and IFX-treated TNF- $\alpha$  inhibitor-naïve patients.<sup>17</sup> Similar findings were observed in a consecutive series of 362 naïve patients with CD in Austria.<sup>18</sup> A prospective, registry-based study reported similar rates of clinical response and drug survival in ADA and IFX-treated patients at 6 months and at 2 years.<sup>19</sup> The Sicilian Network for Inflammatory Bowel Diseases (SN-IBD) – a group composed of 16 centers licensed to prescribe biologicals in Sicily, Italy – performed this comparison in TNF- $\alpha$  inhibitor-naïve and non-naïve patients, showing comparable efficacy and safety of the two drugs in both subgroups.<sup>20</sup> Other studies were conducted on administrative databases. Osterman *et al.*<sup>21</sup> showed no significant difference in the risk of hospitalization and abdominal surgery between ADA and IFX-treated patients. An elegant nationwide, register-based, propensity score-matched cohort study performed in Denmark among 827 biological-naïve patients with CD confirmed the equivalence between the two drugs, in terms of the percentage of CD-related hospitalizations, major abdominal surgery and serious infections requiring hospitalization.<sup>22</sup>

In UC, the most interesting comparison among TNF- $\alpha$  inhibitors, in our opinion, is between the two subcutaneous agents, ADA and golimumab (GOL). A direct comparison with IFX is limited by the fact that this drug is generally used when the disease activity is more severe, while the other two biologicals are chosen more often in patients with less severe (i.e. moderate) activity. Nevertheless, some comparative studies involving IFX have also been published. A retrospective, single-center, French study compared treatment persistence rates of IFX *versus* ADA as first and second-line TNF- $\alpha$  inhibitors, showing overall similar rates of persistence between the two drugs.<sup>23</sup> A recent small Korean study compared IFX and ADA among patients naïve to TNF- $\alpha$  inhibitors, confirming no significant differences between the two groups in the rates of clinical remission and clinical response at 8 or 52 weeks.<sup>24</sup>

Focusing on the comparison between ADA and GOL, the SN-IBD performed a comparative real-world study between the two biologicals.<sup>25</sup> Clinical benefit was reported in 79% of patients in the ADA group and 63% in the GOL group ( $p=0.026$ ) after 8 weeks, and in 67% in the ADA group and 47% in the GOL group ( $p=0.008$ ) at the end of follow-up. These results were confirmed by propensity score analysis. A further analysis considering ADA optimization as treatment failure showed that the difference in clinical benefit between ADA and GOL was not significant. This finding suggests that the difference between the two drugs is affected by the fact that GOL could not be optimized at the time in Italy.

#### *Vedolizumab versus TNF- $\alpha$ inhibitors*

The approval of VDZ for IBD treatment by the US Food and Drug Administration (FDA) and in Europe in 2014 opened the possibility of choosing an alternative target to TNF- $\alpha$  inhibition. Consequently, the interest in comparing the efficacy of different mechanisms of action in RWE was unavoidable.

The SN-IBD performed a comparison of the effectiveness of VDZ and ADA in CD through a propensity score weighted cohort study.<sup>26</sup> Five hundred and eighty-five treatments (VDZ:  $n=277$ ; ADA:  $n=308$ ) were included. Despite rates of clinical response being numerically higher for ADA at 12 and 52 weeks, the difference with VDZ was not statistically significant. Cox survival

**Table 2.** Real-world studies on the comparison of effectiveness between biologics included in the review.

Authors	Biologics	Disease/patients	Propensity score	Main study outcomes	Main findings
Kestens <i>et al.</i> <sup>17</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive	No	Steroid-free clinical response, defined in absence of these points: hospitalization or CD-related surgery, discontinuation of anti-TNF, steroid-dependency	No difference between IFX and ADA after 1 and 2 years of follow-up
Narula <i>et al.</i> <sup>18</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive	No	Clinical remission, clinical response, steroid-free remission (graded with HBI)	No difference between IFX and ADA after 12 weeks and 1 year
Cosnes <i>et al.</i> <sup>19</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive	No	6-month and 2-year response rates and drug survival. Clinical response defined as clinical improvement without need for steroids, retrospectively assisted evaluation of CDAI <150, and absence of surgery	No difference between IFX and ADA
Macaluso <i>et al.</i> <sup>20</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive and experienced	Yes, Matching	Steroid-free remission (resolution of abdominal pain and normalization of bowel habit without steroids), and clinical response (mild or no abdominal pain plus a reduction of at least 50% of the number of bowel movements compared with baseline)	No difference between IFX and ADA after 12 weeks and 1 year. Lower response rates among anti-TNF experienced compared with naive.
Osterman <i>et al.</i> <sup>21</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive	Yes, Matching	Persistence on therapy at week 26, surgery, and hospitalization for CD	No difference between IFX and ADA
Singh <i>et al.</i> <sup>22</sup>	IFX <i>versus</i> ADA	CD/Anti-TNFs naive	Yes, Matching	All-cause hospitalisation, CD-related hospitalisation, major abdominal surgery and serious infections	No difference between IFX and ADA
Pouillon <i>et al.</i> <sup>23</sup>	IFX <i>versus</i> ADA	UC/Anti-TNFs naive and experienced	No	Persistence on therapy at the end of follow-up	No difference in treatment persistency between IFX and ADA
Lee <i>et al.</i> <sup>24</sup>	IFX <i>versus</i> ADA	UC/Anti-TNFs naive	No	Clinical remission and response at 8 and 52 weeks (graded with partial Mayo score)	No difference between IFX and ADA
Renna <i>et al.</i> <sup>25</sup>	ADA <i>versus</i> GOL	UC/Anti-TNFs naive and experienced	Yes, weighting	Steroid-free clinical remission and response at 8 weeks and at the end of follow-up (graded with partial Mayo score)	Higher rate of clinical benefit with ADA compared with GOL
Macaluso <i>et al.</i> <sup>26</sup>	VDZ <i>versus</i> ADA	CD/Anti-TNFs naive and experienced	Yes, weighting	Clinical response, steroid-free clinical remission (graded with HBI) at 12 and 52 weeks, and failure-free survival at the end of follow-up	No difference between VDZ and ADA
Bohm <i>et al.</i> <sup>27</sup>	VDZ <i>versus</i> Anti-TNFs	CD/Anti-TNFs naive and experienced	Yes, weighting	Risk for infections or non-infectious serious adverse events, clinical remission (resolution of CD-related symptoms), steroid-free clinical remission, and endoscopic remission (absence of ulcers/erosions)	Lower risk of non-infectious serious adverse events, but not serious infections, with VDZ, with no significant difference for achieving disease remission

(Continued)

**Table 2.** (continued)

Authors	Biologicals	Disease/patients	Propensity score	Main study outcomes	Main findings
Macaluso <i>et al.</i> , <sup>28</sup>	VDZ <i>versus</i> ADA <i>versus</i> GOL	UC/Anti-TNFs naive and experienced	Yes, weighting	Clinical response, steroid-free clinical remission (graded with partial Mayo score) at 12 and 52 weeks, and treatment persistence at the end of follow-up	VDZ was superior to ADA and GOL at 52 weeks and as treatment persistence, while ADA had superior treatment persistence compared to GOL
Lukin <i>et al.</i> , <sup>29</sup>	VDZ <i>versus</i> Anti-TNFs	UC/Anti-TNFs naive and experienced	Yes, weighting	Time to serious adverse events and serious infections, time to clinical remission (resolution of UC-related symptoms), steroid-free clinical remission, deep remission (achieving both clinical and endoscopic remission)	Higher rates of remission with VDZ than anti-TNFs, and lower rates of serious adverse events in anti-TNFs-naïve patients
Favale <i>et al.</i> , <sup>30</sup>	VDZ <i>versus</i> ADA	UC/Failure with IFX	No	Treatment failure at week 52 (therapy discontinuation due to adverse events or inadequate clinical improvement based on PGA or clinical relapse requiring steroids or need for colectomy)	Failure rate was significantly higher in the ADA group as compared with VDZ group among IFX secondary failures
Rundquist <i>et al.</i> , <sup>31</sup>	VDZ <i>versus</i> Anti-TNFs	CD and UC/Anti-TNFs experienced	Yes, matching	Drug survival at 12 months, survival without IBD-related hospitalisation, IBD-related surgery, antibiotics, or hospitalisation because of infection, corticosteroid exposure	No difference between VDZ and anti-TNFs
Biemans <i>et al.</i> , <sup>32</sup>	UST <i>versus</i> VDZ	CD/Anti-TNFs experienced	Yes, matching	Corticosteroid-free clinical remission (graded with HBI), biochemical remission (C-reactive protein $\leq 5$ mg/L and fecal calprotectin $\leq 250$ $\mu$ g/g), combined corticosteroid-free clinical and biochemical remission, and safety outcomes after 52 weeks	UST was associated with superior effectiveness outcomes when compared to VDZ, while safety outcomes were comparable
Alric <i>et al.</i> , <sup>33</sup>	UST <i>versus</i> VDZ	CD/Anti-TNFs experienced	Yes, weighting	Clinical remission, steroid-free clinical remission (graded with HBI) and treatment persistence at week 48	UST was associated with a higher rate of clinical remission and treatment persistence than VDZ after 48 weeks
Townsend <i>et al.</i> , <sup>34</sup>	UST <i>versus</i> VDZ	CD/Anti-TNFs experienced	Yes, weighting	Steroid-free remission at end of induction (2 months) and at 12 months (graded with HBI), clinical response and remission, treatment persistence, surgery and adverse events	UST was associated with superior effectiveness outcomes when compared to VDZ
ADA, adalimumab; CD, Crohn's disease; GOL, golimumab; HBI, Harvey-Bradshaw Index; IFX, infliximab; TNF, tumor necrosis factor; UC, ulcerative colitis; UST, ustekinumab; VDZ, vedolizumab.					

analysis weighted for propensity score showed no significant difference in the probability of failure-free survival between the two drugs. These findings were overall confirmed by the data from the VICTORY Consortium, in which the propensity score weighted comparison between VDZ and anti-TNF $\alpha$  agents in patients with CD showed no significant difference between the two treatments in terms of clinical remission, steroid-free clinical remission or endoscopic remission at 1 year. Therapy with TNF- $\alpha$  inhibitors was associated with higher treatment persistence compared to VDZ, while rates of non-infectious serious adverse events, but not serious infections, were significantly lower with VDZ compared to TNF- $\alpha$  inhibitors.<sup>27</sup>

In an attempt to confirm the results of the VARSITY study, the SN-IBD recently performed a multicenter, real-world comparison of the efficacy of VDZ and ADA – with the addition of GOL as a third arm – in UC through a propensity score weighted cohort study.<sup>28</sup> Four hundred and sixty-three patients (VDZ:  $n=187$ ; ADA:  $n=168$ ; GOL:  $n=108$ ) were included, with a median follow-up of 48 weeks. While no significant difference was found between the three biologicals after 12 weeks of treatment, steroid-free remission was reported in 51% of patients in the VDZ group, 31% of patients in the ADA group, and in 29% of patients in the GOL group ( $p=0.002$  for VDZ *versus* ADA,  $p=0.001$  for VDZ *versus* GOL,  $p=n.s.$  for ADA *versus* GOL) at 52 weeks. Patients treated with VDZ had reduced probability of treatment discontinuation compared to those treated with ADA [hazard ratio (HR): 0.42,  $p<0.001$ ] and GOL (HR: 0.30,  $p<0.001$ ), while patients treated with ADA had reduced risk of treatment discontinuation compared to those treated with GOL (HR: 0.71,  $p=0.048$ ). Overall, this study was able to confirm the superiority of VDZ over ADA – extending this finding also over GOL – thus helping to translate the results derived from a RCT to everyday practice. These findings have been confirmed by the recently published study of the VICTORY Consortium, which analyzed the effect of VDZ and TNF- $\alpha$  inhibitors in UC using propensity score weighted comparisons.<sup>29</sup> VDZ-treated patients were more likely to achieve clinical remission, steroid-free clinical remission, and steroid-free deep remission than those treated with TNF- $\alpha$  inhibitors. Results were consistent across subgroup analyses in TNF-antagonist naive and exposed patients,

and for VDZ *versus* IFX and *versus* subcutaneous TNF- $\alpha$  inhibitors separately. There were no statistically significant differences in the risk of serious adverse events or serious infections between the two groups.

Finally, focusing on the comparison between TNF- $\alpha$  inhibitors and VDZ as second-line biologicals, at least two studies should be mentioned. Favale *et al.*<sup>30</sup> demonstrated higher effectiveness of VDZ over ADA in patients with UC who had experienced a secondary failure with IFX (failure rate: 48.0% for ADA *versus* 22.4% for VDZ,  $p=0.035$ ). A recent Swedish study with propensity score matching showed that, in patients with a previous exposure to a first-line treatment with TNF- $\alpha$  inhibitors, drug survival was comparable in VDZ and TNF- $\alpha$  inhibitors as second-line biologicals in both UC and CD. The only difference in the reported outcomes was that VDZ-treated patients had lower survival without IBD-related hospitalization compared to TNF- $\alpha$  inhibitor treated patients (82% *versus* 93%,  $p=0.02$ ).<sup>31</sup>

#### *Vedolizumab versus ustekinumab*

The recent approval of ustekinumab (UST) for the treatment of moderately to severely active CD led to the introduction of a biological with a novel mechanism of action – the inhibition of IL-12 and IL-23 axis<sup>35</sup> – which has particular relevance, similarly to VDZ, for patients who fail TNF- $\alpha$  inhibitor treatment(s). Observational cohort studies have shown that 85–100% of patients receiving VDZ or UST in daily practice had previously been exposed to TNF- $\alpha$  inhibitors.<sup>36,37</sup> As a consequence, the absence of RCTs comparing VDZ and UST in anti-TNF refractory patients has been compensated by real-world studies exploring differences between the two drugs in this setting. Three studies have recently been published. Biemans *et al.*<sup>32</sup> compared 128 VDZ and 85 UST-treated patients with previous failure to anti-TNF treatment, showing that UST-treated patients were more likely to achieve corticosteroid-free clinical remission [odds ratio (OR): 2.58,  $p=0.004$ ], biochemical remission (OR: 2.34,  $p=0.027$ ), and combined corticosteroid-free clinical and biochemical remission (OR: 2.74,  $p=0.014$ ) after 52 weeks, while safety outcomes were comparable between the two drugs. These findings were confirmed after propensity score matched analysis. Similar results were obtained by a French study comparing 107

**Table 3.** Head-to-head comparisons between biologics arising from RCTs and RWE in inflammatory bowel disease: summary of evidence.

<p>Crohn's disease</p> <ol style="list-style-type: none"> <li>1. IFX and ADA seem to have similar effectiveness.</li> <li>2. VDZ and TNF-inhibitors seem to have similar effectiveness.</li> <li>3. There are signals of higher effectiveness of UST over VDZ as a second-line biological after a previous failure with TNF-<math>\alpha</math> inhibitors.</li> </ol>
<p>Ulcerative colitis</p> <ol style="list-style-type: none"> <li>1. The current evidence is not conclusive about differences of effectiveness between IFX and ADA (or GOL).</li> <li>2. There is superiority of VDZ over ADA. It is possible that this superiority may be extended also to the other TNF-<math>\alpha</math> inhibitors.</li> <li>3. As a second-line biological after a failure with a TNF-<math>\alpha</math> inhibitor, there are controversial data on potential differences between VDZ and another TNF-<math>\alpha</math> inhibitor.</li> </ol>
<p>ADA, adalimumab; GOL, golimumab; IFX, infliximab; RCT, randomized controlled trial; RWE, real-world experience; TNF, tumor necrosis factor; UST, ustekinumab; VDZ, vedolizumab.</p>

UST-treated patients and 132 VDZ-treated patients with intention-to-treat analysis and propensity scores weighted comparison.<sup>33</sup> At week 48, UST achieved a higher clinical remission rate (OR=1.92) and treatment persistence (OR=2.54) than VDZ, although the rate of corticosteroid-free clinical remission did not differ significantly between the two drugs. Finally, the study by Townsend *et al.*,<sup>34</sup> comparing 85 VDZ-treated patients and 45 UST-treated patients, confirmed the possibility of a higher efficacy of UST in this setting, even though methodological issues can be raised regarding this study, and caution is still needed before declaring the superiority of UST over VDZ in anti-TNF- $\alpha$  failure patients.<sup>38</sup> One must keep in mind that the incorrect application of the propensity score can lead to distorted estimates, and sophisticated statistical methods cannot replace randomization, because they cannot completely adjust for all the differences in risk factors at baseline.

## Conclusions

Table 3 summarizes the current evidence on direct comparisons between biologics arising from RCTs and RWE. The first completed head-to-head comparative RCTs and the increasing number of ongoing RCTs will provide high-quality data for a correct drug positioning in IBD. In parallel, real-world observational studies, with improved methodological quality, will support the data coming from RCTs and will cover comparisons not performed in the RCT setting. We believe that there is already moderate evidence to support clinicians in the correct choice of the

most appropriate drug and that in the near future the data will be even more robust.

## Acknowledgements

The authors thank Mariella Mannino for the extensive English language editing.

## Conflict of interest statement

Fabio Salvatore Macaluso served as an advisory board member and/or received lecture grants from AbbVie, Biogen, Janssen, MSD, Samsung Biosipis, and Takeda Pharmaceuticals. Ambrogio Orlando served as an advisory board member for AbbVie, MSD, Janssen, Pfizer, Takeda Pharmaceuticals, and received lecture grants from AbbVie, MSD, Sofar, Chiesi, Janssen, Pfizer, and Takeda Pharmaceuticals. The other authors have nothing to disclose.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Fabio Salvatore Macaluso  <https://orcid.org/0000-0001-5128-3846>

Ambrogio Orlando  <https://orcid.org/0000-0002-5298-2638>

## References

1. Abraham C and Cho JH. Inflammatory bowel disease. *N Engl J Med* 2009; 361: 2066–2078.

2. Hold GL, Smith M, Grange C, *et al.* Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we learnt in the past 10 years? *World J Gastroenterol* 2014; 20: 1192–1210.
3. Olivera P, Danese S and Peyrin-Biroulet L. Next generation of small molecules in inflammatory bowel disease. *Gut* 2017; 66: 199–209.
4. Shah SC, Colombel JF, Sands BE, *et al.* Systematic review with meta-analysis: mucosal healing is associated with improved long-term outcomes in Crohn's disease. *Aliment Pharmacol Ther* 2016; 43: 317–333.
5. Ungaro RC, Yzet C, Bossuyt P, *et al.* Deep remission at 1 year prevents progression of early Crohn's disease. *Gastroenterology* 2020; 159: 139–147.
6. Gupta A, Yu A, Peyrin-Biroulet L, *et al.* Treat to target: the role of histologic healing in inflammatory bowel diseases: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. Epub ahead of print 30 September 2020. DOI: 10.1016/j.cgh.2020.09.046
7. Ha C, Ullman TA, Siegel CA, *et al.* Patients enrolled in randomized controlled trials do not represent the inflammatory bowel disease patient population. *Clin Gastroenterol Hepatol* 2012; 10: 1002–1007.
8. Cipriani A, Higgins JP, Geddes JR, *et al.* Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013; 159: 130–137.
9. Blonde L, Khunti K, Harris SB, *et al.* Interpretation and impact of real-world clinical data for the practicing clinician. *Adv Ther* 2018; 35: 1763–1774.
10. Biondi-Zoccai G, Romagnoli E, Agostoni P, *et al.* Are propensity scores really superior to standard multivariable analysis? *Contemp Clin Trials* 2011; 32: 731–740.
11. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011; 46: 399–424.
12. 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.
13. Sandborn WJ, Vermeire S, Tyrrell H, *et al.*; Etrolizumab Global Steering Committee. Etrolizumab for the treatment of ulcerative colitis and Crohn's disease: an overview of the phase 3 clinical program. *Adv Ther* 2020; 37: 3417–3431.
14. ClinicalTrials.gov. HIBISCUS I, <https://clinicaltrials.gov/ct2/show/NCT02163759> (accessed 29 December 2020).
15. ClinicalTrials.gov. HIBISCUS II, <https://clinicaltrials.gov/ct2/show/NCT02171429> (accessed 29 December 2020).
16. ClinicalTrials.gov. GARDENIA, <https://clinicaltrials.gov/ct2/show/NCT02136069> (accessed 29 December 2020).
17. Kestens C, van Oijen MG, Mulder CL, *et al.*; Dutch Initiative on Crohn and Colitis [ICC]. Adalimumab and infliximab are equally effective for Crohn's disease in patients not previously treated with anti-tumor necrosis factor- $\alpha$  agents. *Clin Gastroenterol Hepatol* 2013; 11: 826–831.
18. Narula N, Kainz S, Petritsch W, *et al.* The efficacy and safety of either infliximab or adalimumab in 362 patients with anti-TNF- $\alpha$  naïve Crohn's disease. *Aliment Pharmacol Ther* 2016; 44: 170–180.
19. Cosnes J, Sokol H, Bourrier A, *et al.* Adalimumab or infliximab as monotherapy, or in combination with an immunomodulator, in the treatment of Crohn's disease. *Aliment Pharmacol Ther* 2016; 44: 1102–1113.
20. Macaluso FS, Fries W, Privitera AC, *et al.*; Sicilian Network for Inflammatory Bowel Diseases [SN-IBD]. A propensity score-matched comparison of infliximab and adalimumab in tumour necrosis factor- $\alpha$  inhibitor-naïve and non-naïve patients with Crohn's disease: real-life data from the Sicilian Network for Inflammatory Bowel Disease. *J Crohns Colitis* 2019; 13: 209–217.
21. Osterman MT, Haynes K, Delzell E, *et al.* Comparative effectiveness of infliximab and adalimumab for Crohn's disease. *Clin Gastroenterol Hepatol* 2014; 12: 811–817.e3.
22. Singh S, Andersen NN, Andersson M, *et al.* Comparison of infliximab with adalimumab in 827 biologic-naïve patients with Crohn's disease: a population-based Danish cohort study. *Aliment Pharmacol Ther* 2018; 47: 596–604.
23. Pouillon L, Baumann C, Rousseau H, *et al.* Treatment persistence of infliximab versus adalimumab in ulcerative colitis: a 16-year single-center experience. *Inflamm Bowel Dis* 2019; 25: 945–954.
24. Lee YI, Park Y, Park SJ, *et al.* Comparison of long-term outcomes of infliximab versus adalimumab treatment in biologic-naïve patients with ulcerative colitis. *Gut Liver* 2021; 15: 232–242.

25. Renna S, Mocciaro F, Ventimiglia M, *et al.* A real life comparison of the effectiveness of adalimumab and golimumab in moderate-to-severe ulcerative colitis, supported by propensity score analysis. *Dig Liver Dis* 2018; 50: 1292–1298.
26. Macaluso FS, Ventimiglia M, Fries W, *et al.*; Sicilian Network for Inflammatory Bowel Disease (SN-IBD). A propensity score weighted comparison of vedolizumab and adalimumab in Crohn's disease. *J Gastroenterol Hepatol* 2021; 36: 105–111.
27. Bohm M, Xu R, Zhang Y, *et al.*; VICTORY Collaboration. Comparative safety and effectiveness of vedolizumab to tumour necrosis factor antagonist therapy for Crohn's disease. *Aliment Pharmacol Ther* 2020; 52: 669–681.
28. Macaluso FS, Ventimiglia M, Fries W, *et al.*; Sicilian Network for Inflammatory Bowel Disease (SN-IBD). A propensity score weighted comparison of Vedolizumab, Adalimumab, and Golimumab in patients with ulcerative colitis. *Dig Liver Dis* 2020; 52: 1461–1466.
29. Lukin D, Faleck D, Xu R, *et al.* Comparative safety and effectiveness of vedolizumab to tumor necrosis factor antagonist therapy for ulcerative colitis. *Clin Gastroenterol Hepatol*. Epub ahead of print 8 October 2020. DOI: 10.1016/j.cgh.2020.10.003
30. Favale A, Onali S, Caprioli F, *et al.*; Italian Group for the study of Inflammatory Bowel Disease (IG-IBD). Comparative efficacy of vedolizumab and adalimumab in ulcerative colitis patients previously treated with infliximab. *Inflamm Bowel Dis* 2019; 25: 1805–1812.
31. Rundquist S, Sachs MC, Eriksson C, *et al.*; SWIBREG Study Group. Drug survival of anti-TNF agents compared with vedolizumab as a second-line biological treatment in inflammatory bowel disease: results from nationwide Swedish registers. *Aliment Pharmacol Ther* 2021; 53: 471–483.
32. Biemans VBC, van der Woude CJ, Dijkstra G, *et al.*; Dutch Initiative on Crohn and Colitis (ICC). Ustekinumab is associated with superior effectiveness outcomes compared to vedolizumab in Crohn's disease patients with prior failure to anti-TNF treatment. *Aliment Pharmacol Ther* 2020; 52: 123–134.
33. Alric H, Amiot A, Kirchgesner J, *et al.* The effectiveness of either ustekinumab or vedolizumab in 239 patients with Crohn's disease refractory to anti-tumour necrosis factor. *Aliment Pharmacol Ther* 2020; 51: 948–957.
34. Townsend T, Razanskaite V, Dodd S, *et al.* Comparative effectiveness of ustekinumab or vedolizumab after one year in 130 patients with anti-TNF-refractory Crohn's disease. *Aliment Pharmacol Ther* 2020; 52: 1341–1352.
35. Macaluso FS, Orlando A and Cottone M. Anti-interleukin-12 and anti-interleukin-23 agents in Crohn's disease. *Expert Opin Biol Ther* 2019; 19: 89–98.
36. Engel T, Ungar B, Yung DE, *et al.* Vedolizumab in IBD-lessons from real-world experience; a systematic review and pooled analysis. *J Crohns Colitis* 2018; 12: 245–257.
37. Macaluso FS, Maida M, Ventimiglia M, *et al.* Effectiveness and safety of Ustekinumab for the treatment of Crohn's disease in real-life experiences: a meta-analysis of observational studies. *Expert Opin Biol Ther* 2020; 20: 193–203.
38. Macaluso FS and Orlando A. Letter: propensity score-handle with care. *Aliment Pharmacol Ther* 2021; 53: 360–361.

Visit SAGE journals online  
journals.sagepub.com/  
home/tag

 SAGE journals