**Pharmacoeconomics Review Article**

**Title page**

**Ustekinumab for Treating Moderately-to-Severely Active Crohn’s Disease After Prior Therapy: An Evidence Review Group Perspective of a NICE Single Technology Appraisal**

Robert Hodgson,1 Matthew Walton,1 Mousumi Biswas,1 Teumzghi Mebrahtu,1 Nerys Woolacott1

1Centre for Reviews and Dissemination (CRD), University of York, York, UK

**Corresponding author:** Matthew Walton, Centre for Reviews and Dissemination (CRD), University of York, York YO10 5DD, UK. Telephone: +44 (0)1904 321077, E-mail:matthew.walton@york.ac.uk

**Running title:** **Review of Ustekinumab for Treating Moderately-to-Severely Active Crohn’s Disease After Prior Therapy**

Word count: 6352

**Abstract**

As part of the single technology appraisal (STA) process, the National Institute for Health and Care Excellence (NICE) invited Janssen to submit evidence on the clinical and cost-effectiveness of their drug ustekinumab, an interleukin-12/23 inhibitor, for treating moderate-to-severe active Crohn’s disease (CD). The Centre for Reviews and Dissemination (CRD) and Centre for Health Economics (CHE) Technology Appraisal Group at the University of York was commissioned to act as the independent Evidence Review Group (ERG). This article provides a description of the Company’s submission, the ERG’s critical review of submitted evidence, and the resulting NICE guidance.

The main supporting clinical evidence was derived from four well-conducted randomised controlled trials, comparing ustekinumab with placebo in two sub-populations (conventional care failure and anti-TNFα failure patients) of adults with moderate-to-severe CD. Three trials assessed treatment induction over eight weeks, while the fourth recruited successfully induced patients into a maintenance trial for one year. These trials showed ustekinumab to be more effective than placebo in terms of its ability to induce and maintain clinical response and remission. In the absence of any direct head-to-head data, the Company conducted a network meta-analysis (NMA), which synthesised induction trial data on ustekinumab and relevant comparators; vedolizumab, adalimumab, and infliximab, using placebo data as a common comparator. This analysis found ustekinumab to be of comparable efficacy to previously approved biologics in treatment induction. A ‘treatment sequence analysis’ compared long-term treatment efficacy, finding ustekinumab to be comparable in maintaining treatment response and remission to the three other biologic therapies. However, the ERG had identified many limitations and potential bias in this analysis, and urged caution when interpreting the results.

The company’s economic model estimated ustekinumab to be dominant in both sub-populations compared to conventional care, however, the ERG’s preferred base-case estimated an ICER of £109,279 in the conventional care failure sub-population, and £110,967 in the anti-TNFα failure sub-population when compared to conventional care. However, the ERG identified significant failings in both the model structure and data inputs, which could not be addressed without complete restructuring. The ERG considered that the economic analysis presented by the company failed to adequately address the decision problem specified in NICE’s scope. The NICE Appraisal Committee recommended ustekinumab within its market authorisation, on the grounds of sufficiently similar efficacy and costs to previously recommended biologic therapies. However, the ERG’s analyses demonstrated that all currently recommended biologics are unlikely to be cost-effective relative to conventional care, raising broader questions regarding the appropriateness of cost-comparison exercises for decision-making.

**Key points for Decision Makers**

* Ustekinumab appears to be more effective than placebo in induction and maintenance treatment of patients with moderate-to-severe active Crohn’s disease in patients whose disease has responded inadequately to, or is no longer responding to, either conventional therapy, or a tumour necrosis factor-α inhibitor, or have medical contraindications to such therapy.
* The efficacy of ustekinumab is comparable to the previously approved biologics vedolizumab, adalimumab, and infliximab, but is subject to significant uncertainty.
* The Evidence Review Group (ERG) identified significant structural failings in the company’s economic model structure, which were too substantial to be addressed within the scope of the STA process. The quality of the model introduced a great deal of uncertainty in the results.
* The National Institute for Health and Care Excellence Appraisal Committee recommended ustekinumab as an option for treating moderately-to-severely active Crohn’s disease within its market authorisation. The Committee largely disregarded the company’s economic model, considering a cost-minimisation analysis more appropriate, which found ustekinumab to be sufficiently similar in costs and efficacy to previously approved biologics to warrant recommendation.
* The ERG’s analysis suggested that no currently recommended biologic treatments are likely to be cost-effective relative to conventional care, highlighting potential issues with the new NICE fast track appraisal process.

# **Introduction**

The National Institute for Health and Care Excellence (NICE) is an independent organisation whose remit includes provision of evidence-based recommendations and advice on which interventions are to be offered to patients through the National Health Service (NHS) in England. In order to gain a recommendation for use on the NHS, a health technology must be found to be both clinically effective and an efficient use of NHS resources through the NICE technology appraisal programme [1]. The single technology appraisal (STA) detailed here was commissioned to evaluate ustekinumab (Stelara®) in Crohn’s disease (CD). The company provided NICE with a submission detailing their estimates of the drug’s clinical and cost effectiveness [2], which was critiqued by the independent Evidence Review Group (ERG) [3]. A NICE appraisal committee then considered the evidence submitted by the company, in the context of evidence supplied by patient, clinical, and NHS commissioning experts, as well as the review conducted by the ERG.

The objective of the scope issued by NICE for this technology appraisal was to assess the clinical and cost effectiveness of ustekinumab within its market authorisation as described below. This article presents a summary of the ERG’s critique of the company’s submission (CS) in line with the NICE scope and the issues that arose during the review. The committee’s decision making process and final consultation results are also discussed. Full details of the appraisal and relevant supporting documents can be found on the NICE website [4].

# **The Decision Problem**

Crohn’s disease is an immune-mediated disease which causes inflammation and progressive damage to the gastrointestinal (GI) system, it is estimated to affect at least 115,000 people in the UK [5]. Its clinical manifestation is highly heterogeneous, determined partly by the location of the disease in the GI tract, and by the pattern of disease progression. Disease features generally include diarrhoea, abdominal pain, fatigue, unintended weight loss, and blood and mucus in stools [6-8]. Symptoms occur in a gradually worsening relapsing-remitting fashion, characterised by acute exacerbations between periods of lower activity or remission.

Crohn’s disease presents most commonly in adolescents and young adults, with incidence peaking between 15 to 30 years of age [9, 10], making control of the condition a significant and often life-long burden. Management can be complicated by development of strictures, obstructions, and fistulae, which require surgery to correct. Between 50 and 70% of patients will require surgery to remove diseased sections of bowel within 5 years of diagnosis [11]. Surgery is not a cure, however: relapse is inevitable, and inflammation and damage will go on to develop in previously unaffected areas of bowel.

Treatment of CD aims to control manifestations of active disease to improve quality of life through amelioration of symptoms, and if possible to induce and maintain remission while minimising treatment-related adverse events. Drugs currently recommended by NICE as first line treatments include glucocorticosteroids, 5-aminosalicylate (5-ASA), and various immunosuppressants such as methotrexate and azathioprine. For adult patients who either fail to respond to conventional therapy or relapse, the TNFα inhibitors infliximab and adalimumab are recommended. On failure or contraindication to these, vedolizumab is used as a final option. It is estimated that over 4000 patients in England and Wales have failed or become refractory to all therapies currently available in general clinical practice [12, 13].

Ustekinumab (brand name Stelara®, Janssen-Cilag International NV) is a humanised monoclonal antibody which inhibits inflammatory activity linked to interleukin-12 and interleukin-23; signalling molecules implicated in the pathogenesis of CD [14]. Ustekinumab has received the following market authorisation for CD: *“Stelara is indicated for the treatment of adult patients with moderately to severely active Crohn’s disease who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a TNFα antagonist or have medical contraindications to such therapies.*” [15] Ustekinumab is licensed to be administered initially as an intravenous (IV) induction dose of approximately 6mg/kg, followed by maintenance doses of 90 mg administered via subcutaneous injection. It should be noted that ustekinumab is also licensed and has a NICE recommendation for treatment of plaque psoriasis [16] and for psoriatic arthritis [17].

# **The Independent Evidence Review Group (ERG) Review**

The company submitted evidence to NICE on the clinical and cost effectiveness of ustekinumab for the treatment of previously treated moderate-to-severe active CD. The ERG performed a critical review of the evidence presented in the CS and associated documents. In line with established processes for STAs, the ERG had the opportunity to seek clarification on specific issues in the CS and request additional information pertinent to the decision problem. The ERG’s review of the CS included an assessment of the company’s adherence to NICE methodological guidelines, provided a critique of the company’s interpretation and analysis of the evidence, and checked for the existence of other evidence, or alternative interpretations of the evidence available. The following sections present a summary of the evidence presented in the CS and the ERG review of that evidence.

## **Clinical Evidence**

The company’s submission incorporated a systematic review of studies evaluating the efficacy and safety of ustekinumab for previously treated moderate to severe active Crohn’s disease. The evidence submitted in support of ustekinumab was based primarily upon four placebo-controlled, double-blind randomised controlled trials (RCTs). Three of these trials (UNITI-1 [18], UNITI-2 [19], and CERTIFI [20]), were designed to evaluate ustekinumab as an induction therapy where the objective is to induce disease remission. The fourth trial IM-UNITI [21], provided longer-term evidence on the efficacy of ustekinumab as a maintenance therapy which aims to maintain remission achieved by treatment induction.

The induction trials (UNITI-1, UNITI-2 and CERTIFI) randomised individuals to receive a single intravenous infusion (IV) dose of ustekinumab or an equivalent volume of placebo at week 0, with outcomes assessed over 8 weeks. Treatment dose was weight-based, the UNITI trials used 2, 3, or 4x130mg vials according to a patient’s weight category, at approximately 6mg/kg, while the CERTIFI trial used exactly 6mg of ustekinumab per kilogram bodyweight. The primary endpoint for these trials was clinical response at week 6, defined as a reduction in baseline CDAI score of ≥100 points. Treatment efficacy and safety was assessed in two distinct populations, representing the patients at two different positions in the NHS treatment pathway. The trials UNITI-1 and CERTIFI followed 496 and 263 individuals respectively, and recruited adults who had previously failed, or were intolerant to anti-TNFα therapy – in essence investigating ustekinumab as tertiary treatment option alongside vedolizumab. The third trial (UNITI-2) contained 419 patients who had failed conventional therapy only, positioning ustekinumab as second-line therapy alongside infliximab and adalimumab.

The population in the IM-UNITI trial was enrolled from those who completed the UNITI-1 and -2 trials, comprising 397 conventional care and anti-TNFα failure-patients who were treated and assessed for 44 weeks from the end of induction therapy. These patients received 90 mg ustekinumab or placebo subcutaneously every 8 or 12 weeks. This trial investigated the effect of treatment withdrawal in those who had previously exhibited a clinical response to ustekinumab induction therapy by re-randomising induction responders to either placebo or continued ustekinumab treatment. The primary outcome was clinical remission at week 44, defined as CDAI score <150.

Results from the three induction trials demonstrated patients treated with ustekinumab to be more likely to achieve clinical response and remission than those given placebo; this was the case for both the conventional care failure and anti-TNFα failure populations. In the anti-TNFα inhibitor therapy failure population (UNITI-1), ustekinumab induced response in 33.7% (~6 mg/kg) and 34.3% (130 mg) of patients vs 21.5% for placebo (p=0.003 and p=0.002). In the conventional care failure population (UNITI-2), ustekinumab induced response in 55.5% (~6 mg/kg) and 51.7% (130 mg) of patients vs 28.7% on placebo. In the maintenance trial IM-UNITI, a greater proportion of patients randomised to either of the two ustekinumab treatment regimens (90 mg SC UST every 8 weeks or 12 weeks) maintained a response than those re-randomised to placebo (59.4% and 58.1% versus 44.3% respectively). Remission at week 44 was also more likely in the ustekinumab arms than on placebo (53% and 49% versus 36% respectively).

The company’s systematic review identified all studies containing clinical data on patients with moderately to severely active CD treated with biologics. After application of eligibility criteria, 12 trials were selected for inclusion in the review (Targan (1997) [22], CLASSIC I [23], Watanabe (2012) [24], GAIN [25], GEMINI II [12], GEMINI III [26], UNITI-1, UNITI-2, CERTIFI, CHARM [27], ACCENT I [28], and IM-UNITI).

Network meta-analyses were conducted to allow comparison of all available biologic therapies as induction therapy in the absence of head-to-head trials, with separate analyses conducted on the conventional care and anti-TNFα failure populations for the following outcomes; clinical response (CDAI-70; CDAI-100) and clinical remission (CDAI≤150). Conventional care failure populations treated with ustekinumab had a statistically comparable probability of achieving a CDAI-100 response and clinical remission to those given vedolizumab: OR (95% CI) 1.85 (0.96 to 3.51) and 0.93 (0.39 to 2.08) respectively, to adalimumab 160/80mg: 1.03 (0.47 to 2.20) and 0.64 (0.25 to 1.53), and to adalimumab 80/40mg: 1.39 (0.64 to 2.97) and 1.14 (0.44 to 2.82). Sufficient data on infliximab was not available for directly comparable outcomes; however, an imputed measure of CDAI-70 response indicated that infliximab was statistically superior to ustekinumab at inducing clinical response in conventional care failure patients, and patients treated with ustekinumab were less likely to achieve remission (CDAI≤150) than those on infliximab (OR 0.08, 95% CI 0.01 to 0.59). The results of the NMAs performed on the anti-TNFα failure population were similar, with no significant differences in the probability of patients treated with ustekinumab achieving a CDAI-100 response or clinical remission versus vedolizumab: OR (95% CI) 1.10 (0.65 to 1.87) and 1.23 (0.60 to 2.53) respectively, adalimumab 160/80mg: 0.97 (0.55 to 1.70), and 0.51 (0.22 to 1.13), and adalimumab 80/40mg: 0.68 (0.19 to 2.39) and 1.80 (0.29 to 16.98).

The company considered that synthesis of the available trial evidence in the maintenance phase using traditional NMA techniques was inappropriate, as selection into the maintenance trials was conditional on patients successfully achieving remission in the induction trial. As such, the maintenance trials were not considered directly comparable. To overcome the inherent issues with synthesising the maintenance phase trial evidence, the company opted to conduct a ‘treatment sequence analysis’[29] [30] which attempted to account for the different treatment regimens and placebo arm compositions by linking patients’ maintenance phase results to their induction phase response status. The company also made a number of adjustments to their placebo data in order to reduce study heterogeneity and therefore increase the between-study comparability.

The results of the treatment sequence analysis estimated that in the anti-TNFα failure subpopulation, ustekinumab was associated with a higher probability of achieving clinical remission (80%, OR [Credible interval (CrI)]: 1.35 [0.66, 2.73]) compared to vedolizumab at 1 year. In the conventional care failure subpopulation, ustekinumab was similarly associated with a high probability of achieving clinical remission (69%, OR [CrI]: 1.26 [0.50, 3.07]) compared to adalimumab standard dose at 1 year. Infliximab was associated with the highest chance of being in clinical remission at 1 year in the conventional care subpopulation (29%, OR [CrI]: 0.60 [0.07, 324]).

## **Critique of Clinical Evidence and Interpretation**

The systematic review process as undertaken by the company was deemed satisfactory by the ERG, though the submissions lacked detail. However, despite some unclear reporting and some potential inconsistencies in application of eligibility criteria, the ERG was confident that the review as originally conducted was likely to have identified and included all relevant literature. There was no quality assessment of all 34 trials identified in the systematic literature review, nor was full methodology reported. There were full quality assessments performed for the ustekinumab trials, these were conducted in accordance with Centre for Reviews and Dissemination guidelines, and as such were deemed appropriate by the ERG.

The ERG highlighted a number of issues with the clinical evidence, and particularly with the way in which it was analysed and interpreted.

### **Trial evidence**

The UNITI and CERTIFI trials demonstrating the efficacy and safety of ustekinumab versus placebo were conducted to a good standard and had high internal validity; however, the ERG raised a number of concerns regarding their wider applicability to UK clinical practice and CD population. Firstly, the duration of induction phase follow-up (6 weeks) may have classed patients who were slower to respond as non-responders. Clinical experts advising the ERG suggested that the assessment of treatment response in UK practice is generally performed over a longer period, and may be based on multiple doses of ustekinumab, in contrast to the single dose received by patients in the UNITI and CERTIFI trials. Secondly, the reliance upon CDAI as the primary measure of treatment response does not reflect UK practice or opinion; clinicians consider CDAI to be an unreliable representation of disease activity and would avoid using it as a basis for clinical decisions. Treatment success in practice is instead based upon amelioration of symptoms and objective measures of response such as endoscopic outcomes. Thirdly, the conventional therapy failure trial population (UNITI-2) comprised a mixture of both anti-TNFα naïve and experienced patients (though none had previously failed an anti-TNFα). It is unclear whether this reflects the NHS population eligible for ustekinumab, or if this implies that the trial results may overestimate the likely benefit patients achieve in practice.

The assessment of ustekinumab against its comparators used evidence from 12 RCTs on all four biologics currently approved for CD treatment in the UK – adalimumab, infliximab, ustekinumab, and vedolizumab. These trials were judged by the ERG to be well conducted, and were methodologically similar to the UNITI trials for ustekinumab. As such they suffered from a number of the same weaknesses, namely the use of short induction phases and reliance upon CDAI to represent clinical improvement. Treatment history and prior anti-TNFα exposure differed substantially between trial populations; a significant proportion of participants in the UNITI-2 trial had a history of anti-TNFα treatment, yet were pooled with the anti-TNFα naïve populations of the other conventional care failure trials. There was additional heterogeneity within the anti-TNFα failure populations between trials; the GAIN trial of adalimumab recruited only secondary failure patients, i.e. patients who had exhibited a response to biologic therapy prior to treatment failure, while participants in the Watanabe trial of adalimumab had previously received, but not necessarily failed an anti-TNFα. At least four of the other trials synthesised in the NMA included a combination of both primary and secondary anti-TNFα failure patients. There was no available randomised evidence on infliximab in an anti-TNFα failure population due to the age of the main trial and hence this relevant comparator was not included in the analysis.

The ERG considered the design of the maintenance trial IM-UNITI trial appropriate; however, the use of a withdrawal design, in which responders to treatment were re-randomised to either active treatment or placebo, means this trial provides only limited evidence on the long term effectiveness of the ustekinumab relative to conventional therapy. Those re-randomised to placebo were still likely to be receiving some benefit from their induction treatment for the first few weeks; therefore remission may have been achieved as a result of the biologically persistent ustekinumab, rather than on placebo alone. Furthermore, the randomised population included all participants who had achieved a clinical response to ustekinumab during the induction trials UNITI-1 and UNITI-2. As a result, the IM-UNITI population contained patients who had responded to an unlicensed dose of ustekinumab (130mg), yet no separate analyses were provided for those patients receiving their treatment as per UK practice.

Like the IM-UNITI trial, the four maintenance trials of the comparators were ‘withdrawal trials’ As such, they were not an ideal source of evidence for evaluating the relative efficacy of the various biologics, as there was no common comparator group against which they could be assessed in the form of a population remaining continuously on placebo throughout both induction and maintenance.

### **Data synthesis and interpretation**

#### Short-term outcomes

The network meta-analyses presented in the company’s submission were of mixed quality, likewise was their relevance to the decision problem. The company’s analyses were inconsistent in their selection of trial evidence and relevance to the decision problem. The submission presented no sound justification for the exclusion of the CERTIFI trial from the NMA, included the irrelevant comparator vedolizumab in the conventional care failure NMA, and in all analyses aggregated results from the 130 mg unlicensed ustekinumab dose with the ~6 mg/kg dose. Significant differences in assessment timings between the ustekinumab and comparator trials may have caused the overestimation of the relative efficacy of ustekinumab. In addition, the aforementioned lack of comparability between the trial and NHS populations introduced further uncertainty in the validity and applicability of the syntheses performed, though the ERG acknowledged these analyses were likely appropriate given the quality of available data.

#### Long-term outcomes

Analysis of long-term treatment efficacy was subject to a number of methodological challenges; the maintenance trial designs shared the comparability issues seen in the induction NMA, and lacked a true common comparator, while the data as a whole did not strictly address the decision problem. The issue stems primarily from the design of the maintenance trials which enrolled responders to biologic treatment and did not follow up those patients who received placebo therapy at induction. The resulting maintenance phase ‘placebo arm’ could more accurately be described as a ‘treatment withdrawal arm’, as all individuals randomised to placebo had previously been treated with and responded to biologic therapy in the induction phase. To overcome these shortcomings in the trial data the company used the treatment sequence methodology to attempt to account for the fact that patients enrolled in the maintenance trials had responded to induction treatment. The ERG however, had a number of concerns regarding the interpretation and application of the treatment sequence analysis results. Firstly, the treatment sequence analysis does not represent realistic long-term treatment outcomes, as maintenance trial outcomes were conditional upon induction response; therefore, relative long-term efficacy is exaggerated due to the omission of data on those patients not exhibiting an initial response to therapy. Furthermore, as the analysis does not include long-term data on non-responders, it cannot represent the relative effectiveness of the compared treatments over a typical year of clinical use, i.e. where patients are randomised to either active therapy or placebo with outcomes assessed over one year, regardless of their response to biologic induction. The analysis as presented gives only a sense of relative efficacy of ustekinumab and its comparators, and cannot provide true efficacy of the treatments in a typical patient population over one year as stated in the CS. Secondly, the treatment sequence analysis relied on imputed data from the UNITI trials to provide a common control arm for all biologics, whereby response rates of comparator treatments were adjusted according to each trial’s placebo response rates, resulting in an exaggeration of ustekinumab’s relative efficacy. This analysis also omits all randomised placebo data from the maintenance trials of infliximab, adalimumab and vedolizumab, meaning the results are no longer based on randomised comparisons and lack the associated scientific rigour. This introduces the risk of confounding due to differences in setting, therapies received, and severity of disease between the comparator trial treatment arms and the UNITI placebo control.

In addition to the above, the ERG also raised a number of concerns with respect to the comparability of the designs and outcomes of trials included in the treatment sequence analysis, and identified a number of instances in which the data inputs used by the company were incorrect.

## **Cost-effectiveness evidence**

The CS included a systematic review for relevant cost-effectiveness studies; however, it did not identify any studies relating directly to the decision problem. The company therefore presented a *de novo* economic analysis evaluating the cost-effectiveness of ustekinumab for Crohn’s disease.

The *de novo* model presented by the company assessed the cost-effectiveness of ustekinumab in two population groups, conventional care failure patients (patients who have failed conventional therapies, but are naive to anti-TNFα therapies) and anti-TNFα failure patients (patients who have either failed to respond to, or lost response to anti-TNFα therapy). The comparator therapies considered in each analysis were dependent on the population under consideration. In the conventional care failure population ustekinumab was compared with conventional care and adalimumab. A scenario analysis was also presented which included infliximab as a comparator (Infliximab was not considered in the base-case analysis due to limitation in the efficacy data for infliximab). In the anti-TNFα failure population ustekinumab was compared with conventional care and vedolizumab.

The model structure adopted by the company was based on a previous economic model developed by Bodger *et al*.[31], a variant of which was used in the NICE appraisal of the comparator therapy vedolizumab (TA352). The *de novo* analysis presented by the company consists of two parts: a short-term induction phase, represented by a decision tree (Figure 1), and a long-term maintenance phase, represented by a Markov model (Figure 2). The model includes four alive health states plus death. The four alive health states are: moderate to severe (CDAI >220), mild (CDAI 150 to 220), remission (CDAI <150) and surgery. The time horizon considered was 60 years, which was stated to represent a lifetime horizon. Costs and benefits in the model were discounted at an annual rate of 3.5% as per the NICE reference case.

**Figure 1 Schematic of the company’s model structure: Induction phase (decision tree; adapted from Figure 40 CS)**



**Figure 2 Schematic of the company’s model structure: Maintenance phase (Markov model; adapted from Figure 41 CS)**

 **Key:** CDAI,Crohn’s Disease Activity Index.

Patients entered the model in the moderate to severe health state at the start of the treatment induction period. The length of the induction period was allowed to vary by treatment to reflect the market authorisations of the respective treatments. Patients receiving ustekinumab, vedolizumab or adalimumab who failed to respond to the initial induction dose(s) were assumed to receive a further induction dose to allow for a delayed response to treatment. This is in line with the marketing authorisations for the respective drugs. At the end of the induction phase all patients move into the Markov model. Patients who had achieved a 100 point improvement in CDAI (CDAI-100) during anti-TNFα induction were classed as responders, and continued to receive maintenance therapy for a maximum of one year. After this point, biologic therapy was assumed to stop, with all patients going on to receive conventional care for the remaining duration of the model. Patients who failed to respond in the induction period were assumed to move directly to conventional care. Patients who initiated on conventional care were assumed to continue to receive conventional care regardless of their induction phase response.

At any point during the maintenance phase of the model, patients in the moderate to severe health state can receive surgery. Patients in the moderate to severe health state were assumed to be at constant risk of surgery in the model and could receive multiple surgeries throughout their lifetime. Surgery therefore did not impact on the likelihood of future surgeries or affect future prognosis.

Clinical effectiveness inputs used in the induction phase of the model was based on UNITI-1, UNITI-2, and the NMA, which included data from induction trials for all relevant comparators. Clinical effectiveness data used in the maintenance phase of the model was based on the IM-UNITI trial and a treatment sequence analysis (see Section 3.2 above for details) which included data from the induction and maintenance trials for all comparators of interest.

Utilities for all health states were the same regardless of whether a patient is receiving biologic therapy or conventional care, based on EQ-5D utility values mapped from Inflammatory Bowel Disease Questionnaire data collected during the UNITI-1 and UNITI-2 pivotal trials.

The model included costs associated with drug acquisition, drug administration, monitoring and management of CD and costs of adverse events. Drug acquisition costs were sourced from the BNF and MIMS with doses based on the relevant marketing authorisation. Drug administration costs were also included for treatments delivered by intravenous drip (infliximab, vedolizumab and ustekinumab induction treatment. No drug administration costs were included for those delivered by subcutaneous injection (adalimumab and infliximab maintenance treatment). The health state costs associated with Crohn’s disease included in the model were estimated based on an elicitation exercise involving 12 clinicians. This included costs associated with outpatient visits, radiological tests, endoscopies and hospitalisations. Adverse event (AE) costs were included for five AEs: serious infection (defined as septicaemia, pneumonia, urinary tract infections, respiratory infections and bronchitis), tuberculosis, hypersensitivity, injection site reactions, and lymphoma. The costs for all AEs, except for lymphoma, were sourced from NHS Reference Costs and were based on values used in TA352 [13].

The results of the company’s base-case analysis estimated ustekinumab to be more effective and at a lower cost compared with adalimumab and conventional care. In the anti-TNFα failure subgroup, the company’s model estimated ustekinumab to be more effective and lower cost compared to vedolizumab and conventional care. Note these results include a confidential price agreement with the commercial medicines unit regarding the price of induction doses of ustekinumab.

## **Critique of the Cost-effectiveness Evidence**

The ERG highlighted a number of issues with the company’s cost-effectiveness evidence.

### **Model Structure**

While the model structure was based on a previous economic model developed by Bodger *et al.*[31]*,* the ERG identified a number of weaknesses with the model structure. Firstly, the model failed to capture the progressive and chronic nature of CD and did not account for the relapsing-remitting nature of the condition. Secondly, the model structure fails to distinguish between different types of surgery, and does not recognise that patients who receive surgery are likely to have a quite different prognosis and treatment pathway to patients receiving drug therapy. The model also fails to recognise the significant impact of surgery on HRQoL and chance of future surgery. Thirdly, the model makes a number of structural assumptions that are inconsistent with UK clinical practice. These included the assumption that all non-responders have moderate to severe disease; that responders with moderate to severe CD and non-responders are assumed to have equal costs and HRQoL; that patients cannot re-initiate biologic treatment upon future relapse, and the use of a drop of ≥100 points in the CDAI score to define response to induction treatment.

### **Long-term effectiveness of treatment**

The ERG had a number of substantial concerns regarding the data, assumptions and methods used to calculate the transition probabilities used in the maintenance phase of the model. The focus of these criticisms was twofold:

*Assumption that non-responders remain in the moderate to severe health state for the entire maintenance period:* The treatment sequence analysis used to populate the economic model does not include data on patients who failed to respond to treatment in the induction period. There was therefore an implicit model assumption that patients who did not respond to induction therapy would remain in the moderate to severe health state with no possibility of spontaneous improvement. This is inconsistent with evidence from the maintenance trials of biologic therapies which show some patients receiving conventional care do respond to therapy and enter remission during maintenance treatment [12, 27, 28, 32].

*Calculation of transition probabilities:* The company makes use of a calibration technique to estimate the transition probabilities of patients in the maintenance phase of the model. This method relies on imposing a series of constraints and selecting a series of starting values. The Excel solver function is then used to estimate transition probabilities that fit with the limited clinical data available. The constraints implied in this process are however, only partially justified, and while based on those used in TA352, are largely arbitrary. The ERG in particular noted that the transition probabilities predicted using this method contradicted data from the IM-UNITI trial on the proportion of patients maintaining remission, and noted that the transition probabilities implied that conventional care is a more effective than adalimumab. This is inconsistent with evidence from the CHARM trial [27] which shows adalimumab to be more effective than placebo during the maintenance phase.

### **Health state costs**

The ERG noted that the health state costs used in the model were particularly high, and far higher than those used in the appraisal of vedolizumab (TA352). Furthermore, they are also substantially higher than was estimated in recent costing studies [33].The ERG also noted that the health state costs included additional surgical costs above and beyond those included through the surgery health state.

### **Duration of treatment**

In the base-case analysis presented by the company, the maximum duration of treatment with biologic therapy is assumed to be one year. This one year stopping rule is in alignment with the NICE recommendation for currently available biologics [34]. The ERG noted that in practice it was typical for patients to continue on therapy for much longer than a year, indeed, the annual IBD audit [35] suggested that ~90% of UK patients receiving biologic therapies had been on treatment for over a year.

### **Additional issues**

In addition to the principal issues outlined above, the ERG had the following concerns:

* The economic model in line with the SPC’s for ustekinumab, adalimumab and vedolizumab allows for a delayed response to induction therapy. However, the maintenance phase transitions for these patients are based on data from initial responders. This may overestimate the effectiveness of these biologic therapies in the maintenance phase if secondary responders are also less responsive to maintenance treatment.
* In the CS model, the effectiveness of conventional care reflects the effectiveness of the concomitant therapies used in the placebo arm of the UNITI trials. These are made up of a combination of therapies including corticosteroids and immunomodulators. A significant proportion of trial participants (18.7% to 30.1%) also received no concomitant therapies. This mix of therapies is not reflective of current practice in the UK, as a significant proportion of placebo patients were left untreated.
* Treatment effectiveness in scenarios where biologic therapy was assumed to continue beyond one year assumed the same effectiveness as in the first year; due to lack of long-term effectiveness of biologic therapies. This is likely to overestimate the effectiveness of biologic therapy as it is common for patients to lose response to therapy over time.
* Adverse event costs: The value used for injection site reactions was £5,000, far in excess of the £1,363 value used in TA352, and was considered to overestimate the costs associate with treating infection site reactions.

## **Conclusions of the ERG review**

Based on the short-term clinical effectiveness results, ustekinumab appears to be more effective than placebo in terms of clinical response and remission in both the conventional care failure and anti-TNFα failure populations. The results of IM-UNITI indicate that around half of patients who respond to ustekinumab are in clinical remission at Week 44. A higher proportion of patients randomised to the two ustekinumab dosages (UST 90 mg every 8 weeks or 12 weeks) retained their responder status and a higher proportion were in remission at Week 44 than those randomised to placebo

Since there were no head-to-head comparative analyses to allow a direct comparison of ustekinumab with its comparators in CD, an NMA was conducted. The analysis of effectiveness in the maintenance phase was, however, not straightforward, and required a ‘treatment sequence’ analysis. This found ustekinumab to be comparable in terms of induction of clinical response and remission with the other biologics for both populations. However, the results of this highly complex analysis based on uncertain data and methods are likely to be unreliable. No evaluation of the relative efficacies of the biologics beyond one year was possible due to lack of data.

The ERG considers the economic analysis presented by the company to be inadequate in fully addressing the decision problem specified in the NICE scope. The analysis contains serious problems pertaining to the model structure and the way in which clinical data was incorporated. The ERG was unable to fully rectify all the identified issues with the company’s model, but was able to carry out a number of analyses using assumptions and data inputs it believed to be more plausible than those used in the company’s base-case analysis.

The focus of ERG’s exploratory analysis was on alternative assumptions regarding the long-term effectiveness of ustekinumab and other biologic therapies, and exploring the impact of alternative assumptions regarding the maximum duration of treatment. In the ERG’s base-case analysis, which assumed a maximum duration of 1 year, the ICER for ustekinumab compared with conventional care was £109,279 per QALY in the conventional care failure subpopulation and £110,967 per QALY in the anti-TNFα failure subpopulation. In additional exploratory analysis exploring the impact of alternative assumptions regarding maximum duration of therapy, the ICER for ustekinumab compared with conventional care ranged between to be £131,811 and £160,165 per QALY in the conventional care failure subpopulation and between £111,122 and £116,268 per QALY in the anti-TNFα failure subpopulation. In all scenarios and in both populations ustekinumab yielded similar or greater QALY gains than other biologics at lower total costs. The ERG’s analysis demonstrated that all biologic therapies currently recommended by NICE - infliximab, adalimumab, vedolizumab and ustekinumab, are unlikely to be cost-effective relative to conventional care. Relative to existing biologic therapies, however, ustekinumab is likely to offer similar benefits at lower total costs to the NHS. Note the results assume that MIMS prices are indicative of local prices paid for biologic therapies; however, it was indicated at the committee stage that this is unlikely to be the case.

# **Final NICE guidance**

The Appraisal Committee reviewed the presented evidence on the clinical and cost effectiveness of ustekinumab following the consideration of information and testimony provided to them by patients with the condition, those representing them, and clinical experts, who emphasised the value placed on the benefits of an additional treatment option. It also took into account the effective use of NHS resources. Following this, the Committee issued the following final guidance:

*“Ustekinumab is recommended, within its marketing authorisation, as an option for treating moderately to severely active Crohn’s disease, that is, for adults who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a TNF-alpha inhibitor or have medical contraindications to such therapies.

The choice of treatment between ustekinumab and other biological therapy should be made on an individual basis after discussion between the patient and their clinician about the advantages and disadvantages of the treatments available. If more than 1 treatment is suitable, the least expensive should be chosen (taking into account administration costs, dosage and price per dose).”*

Although the Committee acknowledged and agreed with the reservations and concerns of the ERG regarding the reliability of the data and analyses, the general weakness of the model structure, and cost effectiveness analysis, they considered the evidence adequate for the purposes of decision making in light of a cost minimisation approach suggested by the company, and deemed ustekinumab a cost-effective option for use in the NHS.

# **ERG Conclusion**

This STA highlighted a number of issues relating to the modelling of CD that are likely to impact on future appraisals in both CD and other indications in which biologic therapies are used. Particularly apparent were the difficulties of modelling CD; specifically, the complexities of modelling the cyclical (& deteriorating) nature of relapses; the multiple lines of treatment patients typically receive and the role surgery plays in managing the condition. Related to these model structure issues, this STA also highlights the limitations of the clinical data available in this area which has tended to examine the effectiveness in induction and maintenance periods separately. This generates problems for comparisons of biologic therapies and introduces complexities in parameterising the clinical effectiveness data in the economic model.

In additional to these modelling issues this STA also raises questions about the wider cost-effectiveness of biologic therapies in the treatment of CD. The ERG base-case analysis demonstrated that none of the currently available biologic therapies are likely to be cost-effective relative to conventional care. This has a significant bearing on the cost-effectiveness of ustekinumab, as while ustekinumab is likely cost-effective relative to other biologic therapies, its cost-effectiveness relative to conventional care is uncertain. This is primarily a consequence of restrictive assumptions made in previous appraisals (to ensure biologics were cost-effective) that biologic therapy should not be continued beyond one year except in particular circumstances. In practice these recommendations have been largely ignored by clinicians, who will often keep their patients on treatment until loss of response to improve their outcomes. This raises questions about how meaningful such restrictions are, and whether greater care is needed in optimising the availability of therapies to account for the pragmatic nature of clinicians and their natural desire to pursue optimal treatment options, even when that option is not cost-effective.

This STA also raises a number of issues with NICE’s current TA processes. Firstly, this STA highlights potential issues with the new cost-comparison fast track appraisal process recently introduced by NICE. The cost-comparison fast-track appraisal (FTA) will allow new technologies to be recommended if they can demonstrate similar health benefits with similar overall costs and resource use to technologies already recommended by NICE for the same indication. The intention being that scrutiny of cost-effectiveness of technologies such as ustekinumab which have similar effectiveness and costs to existing technologies is over-burdensome, and that a lighter-touch assessment is sufficient in such circumstances. This approach, however, assumes that the cost-effectiveness of all comparators is certain, and the FTA process does not allow the committee to fully explore the implications of recommending a technology in the context of previous decisions. Secondly, this STA raises broader issues of what should be done where the cost-effectiveness of previously recommended technologies is uncertain. Currently, there is limited scope for the re-evaluation of health technologies which have already been recommended, and clear practical and political difficulties associated with the reversal of NICE recommendations where new evidence suggests a previously recommended technology may not be cost-effective. This issue also has implications for NICE processes which currently do not formally incorporate the costs of uncertainty. As outlined in Claxton *et al.*[36], formal consideration of the costs associated with uncertainty and the costs of incorrect decisions is possible and within the NICE decision making framework and should be potentially considered by NICE as when they look to revise their processes.

**Acknowledgements** The authors thank Melissa Harden who provided a critical review of the company’s search strategy and Professor Alan Lobo, Consultant Gastroenterologist and Honorary Professor of Gastroenterology at the University of Sheffield for clinical advice throughout the project.

**Author contributions:** Robert Hodgson, Mousumi Biswas, Teumzghi Mebrahtu, Matthew Walton, and Nerys Woolacott all formed part of the ERG that produced the ERG report that this paper describes. Robert Hodgson and Matthew Walton wrote the first draft of the manuscript. All authors commented on the manuscript and approved the final version. This summary has not been externally reviewed by PharmacoEconomics.

**Funding** This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (project number 16/10/12). This summary of the ERG report was compiled after the Appraisal Committee’s review and incorporates additional information and comment from the authors on the STA process and iterations of the NICE guidance not covered by the HTA report. The views and opinions expressed herein are those of the authors and do not necessarily reflect those of NICE or the Department of Health. This work is Crown copyright (UK).

**Conflict of interest** Robert Hodgson, Matthew Walton, Mousumi Biswas, Teumzghi Mebrahtu, and Nerys Woolacott have no conflicts of interest.

**References**

1. National Institute for Health and Care Excellence (NICE). Guide to the processes of technology appraisal. London: NICE; 2014. https://www.nice.org.uk/article/pmg19/chapter/foreword. Accessed 7 May 2017.

2. Janssen. Ustekinumab for previously treated moderate to severe active Crohn’s disease; submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA).

3. Hodgson R, Biswas M, Mebrahtu T, Walton M, Harden M, Woolacott N. Evidence Review Group’s Report: Ustekinumab for treating moderately to severely active CD after prior therapy [TA456]. York: University of York; 2017.

4. National Institute for Health and Care Excellence (NICE). Ustekinumab for previously treated moderate to severe active Crohn’s disease. Technology appraisal guidance [TA456] London: NICE; 2017. https://www.nice.org.uk/guidance/ta456. Accessed 21 Jun 2017.

5. National Institute for Health and Care Excellence (NICE). Crohn's disease overview. NICE Pathways. London: NICE; 2016. http://pathways.nice.org.uk/pathways/crohns-disease. Accessed 15 May 2017.

6. Henriksen M, Jahnsen J, Lygren I, Aadland E, Schulz T, Vatn MH, et al. Clinical course in Crohn's disease: results of a five-year population-based follow-up study (the IBSEN study). Scandinavian journal of gastroenterology. 2007;42(5):602-10.

7. Munkholm P, Langholz E, Davidsen M, Binder V. Disease activity courses in a regional cohort of Crohn's disease patients. Scandinavian journal of gastroenterology. 1995;30(7):699-706.

8. Solberg IC, Vatn MH, Hoie O, Stray N, Sauar J, Jahnsen J, et al. Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study. Clinical gastroenterology and hepatology: the official clinical practice journal of the American Gastroenterological Association. 2007;5(12):1430-8.

9. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology. 2012;142(1):46-54.

10. Ananthakrishnan AN. Epidemiology and risk factors for IBD. Nature reviews Gastroenterology & hepatology. 2015;12(4):205-17.

11. The IBD Standards Group. IBD Standards: Standards for the Healthcare of People who have Inflammatory Bowel Disease (IBD). British Society of Gastroenterology; 2013. https://www.crohnsandcolitis.org.uk/improving-care-services/health-services/ibd-standards. Accessed 23 Jun 2017.

12. Sandborn WJ, Feagan BG, Rutgeerts P, Hanauer S, Colombel JF, Sands BE, et al. Vedolizumab as induction and maintenance therapy for Crohn's disease. N Engl J Med. 2013;369(8):711-21.

13. National Institute for Health and Care Excellence (NICE). Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy. Technology appraisal guidance [TA352]. London: NICE; 2015. https://www.nice.org.uk/guidance/ta352. Accessed 12 Dec 2015.

14. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med. 2012;367(16):1519-28.

15. Committee for Medicinal Products for Human Use. Summary of opinion (Stelara®). London: European Medicines Agency; 2016. http://www.ema.europa.eu/docs/en\_GB/document\_library/Summary\_of\_opinion/human/000958/WC500146594.pdf. Accessed 1 Jul 2017.

16. National Institute for Health and Care Excellence (NICE). Ustekinumab for the treatment of adults with moderate to severe psoriasis. Technology appraisal guidance [TA180]. London: NICE; 2009. https://www.nice.org.uk/guidance/ta180. Accessed 10 May 2017.

17. National Institute for Health and Care Excellence (NICE). Ustekinumab for treating active psoriatic arthritis. Technology appraisal guidance [TA340]. London: NICE; 2015. https://www.nice.org.uk/guidance/ta340. Accessed 10 May 2017.

18. Sandborn W, Gasink C, Blank M, Lang Y, Johanns J, Gao LL, et al. O-001 A multicenter, double-blind, placebo-controlled phase 3 study of ustekinumab, a human IL-12/23P40 mAB, in moderate-servere Crohn's disease refractory to anti-TNFalpha: UNITI-1. Inflamm Bowel Dis. 2016;22 Suppl 1:S1.

19. Janssen Research and Development. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction therapy in subjects with moderately to severely active Crohn’s disease (UNITI 2); 2015 13 October.

20. Sandborn WJ, Gasink C, Gao LL, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med. 2012;367(16):1519-28.

21. Feagan BG, Sandborn WJ, Gasink C, al e. Ustekinumab as Induction and Maintenance Therapy for Crohn's Disease. N Engl J Med. 2016;375(20):1946-60.

22. Targan SR, Hanauer SB, van Deventer SJ, al e. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. N Engl J Med. 1997;337(15):1029-35.

23. Hanauer SB, Sandborn WJ, Rutgeerts P, al e. Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial. Gastroenterology. 2006;130(2):323-33.

24. Watanabe M, Hibi T, Lomax KG, Paulson SK, Chao J, Alam MS, et al. Adalimumab for the induction and maintenance of clinical remission in Japanese patients with Crohn's disease. J Crohns Colitis. 2012;6(2):160-73.

25. Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Colombel JF, Panaccione R, et al. Adalimumab induction therapy for Crohn disease previously treated with infliximab: a randomized trial. Ann Intern Med. 2007;146(12):829-38.

26. Sands BE, Feagan BG, Rutgeerts P, al e. Effects of vedolizumab induction therapy for patients with Crohn's disease in whom tumor necrosis factor antagonist treatment failed. Gastroenterology. 2014;147(3):627-.

27. Colombel JF, Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Panaccione R, et al. Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. Gastroenterology. 2007;132(1):52-65.

28. Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, et al. Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. Lancet (London, England). 2002;359(9317):1541-9.

29. Abrams K. Evaluating complex interventions using multi-phase nested trial designs-Management of urinary incontinence across the primary/secondary care interface. Controlled Clinical Trials. New York: Elsevier Science Inc; 2003.

30. Thorlund K, Druyts E, Toor K, Mills EJ. Comparative efficacy of golimumab, infliximab, and adalimumab for moderately to severely active ulcerative colitis: a network meta-analysis accounting for differences in trial designs. Expert review of gastroenterology & hepatology. 2015;9(5):693-700.

31. Bodger K, Kikuchi T, Hughes D. Cost-effectiveness of biological therapy for Crohn's disease: Markov cohort analyses incorporating United Kingdom patient-level cost data. Aliment Pharmacol Ther. 2009;30(3):265-74.

32. Janssen Research and Development. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab maintenance therapy in subjects with moderately to severely active Crohn's disease (44-week Clinical Study Report). (EDMS-ERI-97847519, 1.0) Janssen Research and Development; 2015.

33. National Institute for Health and Care Excellence (NICE). TA352. Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy. Costing report. 2015 26 August 2015 [cited 2016 15 October]; Available from: https://www.nice.org.uk/guidance/ta352/costing

34. National Institute for Health and Care Excellence (NICE). Vedolizumab for treating moderately to severely active Crohn’s disease after prior therapy. Costing report. London: NICE; 2015. https://www.nice.org.uk/guidance/ta352/costing. Accessed 20 Oct 2016.

35. Royal College of Physicians. National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit. Royal College of Physicians; 2015. https://www.rcplondon.ac.uk/projects/outputs/national-clinical-audit-report-biological-therapies-adult-report-2015. Accessed 15 Jun 2017.

36. Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, Soares M, et al. A Comprehensive Algorithm for Approval of Health Technologies With, Without, or Only in Research: The Key Principles for Informing Coverage Decisions. Value Health. 2016;Sep-Oct; 19(6):885-91.