# Darvadstrocel for treating complex perianal fistula in Crohn's disease: A Single Technology Appraisal


**Produced by** School of Health and Related Research (ScHARR), The University of Sheffield

**Authors**

Daniel Pollard, Research Associate, ScHARR, University of Sheffield, Sheffield, UK

Abdullah Pandor, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

John W Stevens, Reader in Decision Science, ScHARR, University of Sheffield, Sheffield, UK

Sarah Davis, Senior lecturer, ScHARR, University of Sheffield, Sheffield, UK

Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Chris Carroll, Reader in Systematic Review and Evidence Synthesis, ScHARR, University of Sheffield, Sheffield, UK

Charmian Banks, Consultant Physician and Gastroenterologist, Royal Surrey County Hospital, Guildford, UK

Baljit Singh, Consultant in Colorectal Surgery, Leicester General Hospital, Leicester, UK

Janindra Warusavitarne, Consultant in General Surgery, St Mark's Hospital, London, UK

**Correspondence Author** Daniel Pollard, Research Associate, ScHARR, University of Sheffield, Sheffield, UK



**Date completed** Date completed 21/06/2018

**Source of funding**: This report was commissioned by the NIHR HTA Programme as project number 16/56/20.

**Declared competing interests of the authors**

John W Stevens received grants, through their employer, from Takeda for topics unrelated to the condition and technology under appraisal. Janindra Warusavitarne has received speaker fees from Takeda for a talk unrelated to the technology under appraisal. None of the other authors have any conflicts of interest to declare.

**Acknowledgements**

We would like to thank Paul Tappenden, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

**Rider on responsibility for report**

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

**This report should be referenced as follows:**

Pollard D, Pandor A, Stevens JW, Davis S, Wong R, Carroll C, Banks C, Singh B, Warusavitarne J. Darvadstrocel for treating complex perianal fistula in Crohn's disease: A Single Technology Appraisal. Final report to the National Institute for Health and Care Excellence. School of Health and Related Research (ScHARR), 2018.

**Contributions of authors**

Abdullah Pandor and Chris Carroll summarised and critiqued the clinical effectiveness data reported within the company's submission. John Stevens critiqued the statistical analyses undertaken by the company. Daniel Pollard and Sarah Davis critiqued the health economic analysis submitted by the company. Ruth Wong critiqued the company's search strategy. Charmian Banks, Janindra Warusavitarne and Baljit Singh provided clinical advice to the ERG throughout the project. All authors were involved in drafting and commenting on the final report.

**Standard copyright statement on the front page of the ERG/AG report:**

Copyright belongs to ScHARR

Copyright is retained by Takeda for Figures 3, 4, 5, 6, 7, 8, 14, 19, 20; Tables 10, 12, 23, 24, 25, 26, 27, 39, 40, 41, and; text referenced on pages 18, 33, 34, 36, 40, 41, 42, 43, 50, 51, 55, 65, 68, 69, 88, 94, 95, 97, 99, 101.

**CONTENTS**

**List of tables**

**List of figures**

**Abbreviations**

| | |
|---|---|
| AE | Adverse event |
| AIC | Akaike information criterion |
| BIC | Bayesian information criterion |
| CD | Crohn's disease |
| CI | confidence interval |
| CPC | clinical and patient centric |
| CS | Company's submission |
| CSF | Chronic symptomatic fistula |
| Darv | darvadstrocel |
| ERG | Evidence Review Group |
| eASCs | expanded adipose-derived allogeneic mesenchymal stem cells |
| HR | hazard ratio |
| HRQoL | Health-related quality of life |
| HSUV | health state utility value |
| ICER | Incremental cost-effectiveness ratio |
| ITT | intention-to-treat |
| ln | natural logarithm |
| mITT | modified intention-to-treat |
| MRI | magnetic resonance imaging |
| NICE | National Institute for Health and Care Excellence |
| QALY | Quality-adjusted life year |
| QALYs | quality-adjusted life years |
| RCT | Randomised controlled trial |
| SE | standard error |
| STA | Single Technology Appraisal |
| TEAE | Treatment-emergent adverse event |
| TESAE | Treatment-emergent serious adverse event |

# 1 SUMMARY

## 1.1 Critique of the decision problem in the company's submission

The company's submission (CS) assesses the clinical effectiveness and cost-effectiveness of darvadstrocel (Alofisel®) within its marketing authorisation for the treatment of complex perianal fistulas in adult patients with non-active / mildly active luminal Crohn's disease, when fistulas have shown an inadequate response to at least one conventional or biologic therapy. The company's description of complex perianal fistulae in adults with Crohn's disease is broadly appropriate. The decision problem addressed by the CS is partly in line with the final scope issued by the National Institute for Health and Care Excellence (NICE). The submitted evidence is limited to people with fistulas which have up to two internal openings and up to three external openings. Whilst this restriction is consistent with the information in the summary of product characteristics (SmPC) on the number of internal and external openings that can be treated with a single administration of Darvadstrocel, it is not clear whether patients with more openings can have a subset of their fistula treated (i.e. partial treatment) or if they can have all of their fistula treated by using multiple courses of darvadstrocel. Therefore, it is uncertain whether the population missing from the CS may be able to receive treatment under the marketing authorisation for darvadstrocel. With respect to the population of patients included in the CS, the evidence for darvadstrocel is limited to a single treatment administration; the SmPC advises that "there is currently limited experience with the efficacy or safety of repeat administration."

## 1.2 Summary of clinical effectiveness evidence submitted by the company

The CS includes a systematic review of clinical effectiveness evidence. The ADMIRE-CD study, which forms the main supporting evidence for the intervention, was a Phase III, industry-sponsored, randomised, double-blind, placebo-controlled, multicentre trial (49 sites across eight countries, excluding the UK). ADMIRE-CD was designed to evaluate the efficacy and safety of a single intralesional injection of darvadstrocel (an allogeneic preparation of adipose-tissue-derived mesenchymal stem cells) added on to standard of care in patients (aged ≥18 years) with non-active or mildly active luminal Crohn's disease (defined by a Crohn's Disease Activity Index [CDAI] of ≤220) who had complex perianal fistulas (maximum of 2 internal and 3 external openings that had been draining for at least 6 weeks) that was refractory to conventional therapy. Conventional therapy was defined to consist of at least one of: no therapeutic effect of an antibiotic (recommended treatments were ciprofloxacin and metronidazole) after one month; no response to an immunosuppressant (azathioprine [2-2.5 mg/kg] or 6-mercaptopurine [1-1.5 mg/kg]) after three months, or; no response to an anti-tumour necrosis factor (TNF) inhibitor either 12 weeks after initiation of induction treatment or loss of response after 12 weeks of maintenance treatment under a stable dose.

Prior to randomisation, a pelvic magnetic resonance imaging (MRI) scan was administered (screening visit) and patients' fistula were examined under anaesthesia, curetted and, if indicated, setons were placed during this procedure (preparation visit). If a seton was placed, this was subsequently removed immediately prior to the administration of darvadstrocel. Thereafter, patients were randomly allocated to receive darvadstrocel (24mL containing 120 million expanded allogeneic adipose-derived stem cells) and standard of care (n=107) or placebo sham (saline) and standard of care (n=105) in a 1:1 ratio, with risk stratification based upon previously received therapy (immunomodulators, anti-TNF therapy, both, or neither). After receiving darvadstrocel, patients could be treated with antibiotics for no more than four weeks. Immunomodulators and anti-TNF drugs were maintained at stable doses throughout the study. Initiation or dose increases of these drugs were not allowed. A steroid course was permitted to treat occurrences of luminal disease during the study, with a starting dose of 40mg tapered over a maximum of 12 weeks.

The primary endpoint of the ADMIRE-CD study was combined remission (both clinical and radiologic improvement) at week 24 after study treatment and was defined as the clinical assessment of closure of all treated external openings that were draining at baseline, and the absence of collections >2 cm within the perianal fistula in at least two of three dimensions, confirmed by blinded central MRI. The clinical assessment of closure was defined as the absence of draining despite gentle finger compression. The key secondary endpoints were defined as clinical remission (closure of all treated external openings that were draining at baseline despite gentle finger compression) and response (clinical closure of at least 50% of all treated external openings that were draining at baseline) at week 24. The company also presented two additional *post hoc* analyses - time to clinical and patient centric (CPC) remission and time to relapse from CPC remission. These outcomes were the ones used in the economic model as they were considered by UK clinical experts to be the most relevant way to measure remission and relapse in a population who were refractory to at least one conventional (i.e. antibiotics, immunosuppressants) and/or biological therapy. In addition, following a series of protocol amendments, long-term follow-up was extended to week 52 and then to week 104; however, the efficacy data beyond 52 weeks were limited (only 40/212 [18.9%] patients entered into the 104 week follow-up) as a number of patients had already finished the 52 week trial period. The main efficacy analyses were conducted using the intention-to-treat (ITT) approach (which included all randomly assigned patients, n=212) and the modified ITT (mITT) approach (which included all randomly assigned patients who received study treatment and had at least one efficacy assessment after baseline, n=204). The population used to assess safety was all randomly assigned patients who received study treatment (n= 205).

In the primary ITT population (n=212), a significantly greater proportion of patients in the darvadstrocel group achieved the primary endpoint of combined remission at week 24 compared with the control group (49.5% versus 34.3%, respectively; difference of 15.2%; 97.5% confidence interval [CI] 0.2% to

30.3%; $p$=0.024). With longer follow-up (52 weeks), the beneficial effect of darvadstrocel was maintained in the ITT population with 54.2% of patients achieving combined remission compared with 37.1% in the control group (difference of 17.1%; 97.5% CI: not reported; $p$=0.012). Similar results were observed in the mITT population ($p$=0.021 at week 24 and $p$=0.010 at week 52).

A range of secondary endpoints were evaluated in the ADMIRE-CD study. In general, darvadstrocel demonstrated greater improvements in clinical remission (week 24, $p$=0.064 in ITT population; week 52, $p$=0.013 in mITT population [data not reported for ITT population]) and response (week 24, $p$=0.054 in ITT population; week 52, $p$=0.128 in mITT population [data not reported for ITT population]); however, no significant differences ($p$>0.05 in mITT population) were observed in total Perianal Disease Activity Index, Crohn's Disease Activity Index, Inflammatory Bowel Disease Questionnaire and Van Assche scores ($p$=not reported) at week 24 or week 52.

Adverse events (AEs) were common and were reported by approximately two-thirds of patients receiving darvadstrocel at 24 weeks in the ADMIRE-CD trial. The most common treatment-emergent AEs (TEAEs) were proctalgia (12.6% of patients in the darvadstrocel arm versus 11.8% in the control arm), anal abscess (11.7% versus 12.7%), nasopharyngitis (9.7% versus 4.9%) and diarrhoea (6.8% versus 2.9%). The percentages of patients experiencing the principal TEAEs and severe TEAEs (TESAEs) were generally similar across the darvadstrocel and control arms at 24 weeks. The ERG noted that proctalgia, anal abscess and anal fistulae are symptomatic of the indication in this appraisal and therefore might represent treatment failure, i.e. a lack of efficacy, rather than an AE related to the treatment. Safety data were also available for 52 weeks from the ADMIRE-CD trial. The percentages of TEAEs, treatment-related TEAEs, severe TEAEs (TESAEs), and withdrawals due to treatment-related TEAEs among patients in the darvadstrocel arm were all higher at 52 weeks than at 24 weeks. It was also the case that the percentages of patients experiencing key TEAEs, previously similar between arms at 24 weeks, had by 52 weeks become noticeably higher in the darvadstrocel arm than the control arm: anal abscess (19.4% of patients in the treatment arm versus 13.7% in the control arm, of which 13.6% versus 7.8% were TESAEs); anal fistula (10.7% versus 7.8%) and nasopharyngitis (10.7% versus 4.9%). The ERG also noted that the frequency of treatment-related TEAEs among patients at 24 weeks was higher in the earlier Phase I/II trial than the later ADMIRE-CD trial. This might be explained by the considerably lower dose of darvadstrocel in the earlier trial (<60 million expanded adipose-derived allogeneic mesenchymal stem cells [eASCs] versus 120 million eASC) and the issue that some TEAEs and treatment-related TEAEs might represent a lack of efficacy rather than AEs.

### 1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The systematic review process followed by the company was reasonably comprehensive. Despite minor limitations in the company's search strategy, the ERG is reasonably confident that all relevant published studies of darvadstrocel were included in the CS, including data from ongoing studies. The specified inclusion and exclusion criteria were mostly appropriate and generally reflected the scope. The validity assessment tool used to appraise the ADMIRE-CD was considered appropriate by the ERG.

Although the efficacy (assessed in terms of combined remission and CPC remission) in the ADMIRE-CD study appears favourable, and the safety appears acceptable, there are a number of limitations and uncertainties in the evidence base which warrant caution in its interpretation.

A key limitation of the efficacy and safety data for darvadstrocel relates to the *post hoc* analyses of CPC remission (an outcome used in the economic model) and CPC relapse. These endpoints were not designed or powered to test formal hypotheses. As a result, these results should be treated with caution. It should be noted that the CPC definition of remission was considered by the ERG's and the company's clinical experts to be the most relevant way to measure remission and relapse in a population with complex perianal fistula and non-active / mildly active luminal Crohn's disease who were refractory to at least one conventional and/or biological therapy. Another issue is the lack of a confirmatory study. The effect size in the ADMIRE-CD trial was considered to be modest and less than the 25 percentage difference that it was designed to detect but was considered clinically meaningful given that other treatment options for fistulas had failed. A post-authorisation efficacy and safety trial, ADMIRE-CD-II is expected to help address this concern. However, this study not expected to be complete until October 2021.

The key uncertainties in the clinical evidence for darvadstrocel relate to repeated administration, optimal dosing and long-term efficacy and safety.

### 1.4 Summary of cost effectiveness submitted evidence by the company

The company's *de novo* state transition model assesses the cost-effectiveness of darvadstrocel versus standard care (based on the ADMIRE-CD trial) in adults with complex perianal fistula with non-active/mildly active luminal Crohn's disease, when fistulas have shown an inadequate response to at least one conventional or biologic therapy. Incremental health gains, costs and cost-effectiveness of darvadstrocel are evaluated over a 40-year time horizon from the perspective of the National Health Service (NHS) and Personal Social Services (PSS). The company's model is comprised of eight health states: (1) mild chronic symptomatic complex perianal fistulae (CSF); (2) severe CSF; (3) remission; (4) defunctioning surgery (cycle 1); (5) defunctioning surgery (subsequent cycles); (6) proctectomy (cycle 1); (7) proctectomy (subsequent cycles) and (8) death. The transitions between the remission and

the two CSF health states were generated from analyses of time-to-event data (CPC remission and CPC relapse) from the ADMIRE-CD study. CPC remission and CPC relapse are both modelled using a Gompertz distribution with the differences between the two arms being estimated using a treatment effect covariate (a hazard ratio). A retrospective study at St Marks hospital (a national referral centre for intestinal and colorectal diseases) in the UK was used to determine: the proportion of CSFs which are mild; the proportion of defunctioning surgeries which are successful, and; the proportion of proctectomies which are successful. The annual probability of receiving a defunctioning surgery and the annual probability of receiving a proctectomy were estimated from the literature. Health-related quality of life (HRQoL) is principally determined by the time spent in the different model health states and the incidence of treatment related adverse events; these estimates are informed by a vignette study. Resource use estimates and costs were based on data collected in the ADMIRE-CD trial, clinical expert opinion and routine cost sources. The company states that they believe that Section 6.2.19 of the NICE Methods Guide applies when considering the cost-effectiveness of darvadstrocel and consequently darvadstrocel should be assessed using a discount rate of 1.5% for and quality adjusted life years (QALYs) and 3.5% for costs. The company's rationale is that: (1) as darvadstrocel demonstrates long term healing potential in this population with a significant impact on QoL and the condition often affects young people and has a median age of onset of 15-30 years, and so the benefit of an effective treatment in this young population is likely to provide long term health benefits (>30 years), and; (2) darvadstrocel is unlikely to commit the NHS to significant irrecoverable costs.

Based on the probabilistic version of the model (assuming a 3.5% discount rate for both costs and QALYs), darvadstrocel is expected to generate 1.02 additional QALYs at an additional cost of £21,773 per patient; this corresponds to an incremental cost-effectiveness ratio (ICER) for darvadstrocel versus standard care of £21,417 per QALY gained. The deterministic version of the company's model produces a similar ICER of £20,591 per QALY gained. Assuming a maximum acceptable ICER (MAICER) of £20,000 and £30,000 per QALY gained, the company's model suggests that the probability that darvadstrocel produces more net benefit than standard care is 0.421 and 0.736, respectively.

### 1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG critically appraised the company's economic analysis and double programmed the deterministic version of their model. The ERG's critical appraisal identified eleven issues relating the company's economic analysis and the evidence used to inform it. These include: (1) exclusion of relevant patient groups from the economic analysis; (2) possibility of repeat administrations of darvadstrocel; (3) whether costs and QALYs should be discounted at 1.5% by applying Section 6.2.19 of the NICE Methods Guide, is justified; (4) wastage of darvadstrocel; (5) the company's selection of time to relapse and time to remission time to event functions; (6) the company's expert elicitation exercise to estimate the time to relapse and remission for people on third or later line therapies; (7) the

data used to populate the transitions to the defunctioning and proctectomy health states; (8) missing transitions within the model structure; (9) the company's approach to identifying HRQoL data from the literature; (10) the estimates of utilities from a vignette study; (11) adoption of a 40-year time horizon.

## 1.6 ERG commentary on the robustness of evidence submitted by the company

### *1.6.1 Strengths*

The company undertook a reasonably comprehensive systematic review of darvadstrocel for the treatment of complex perianal fistulae in patients with Crohn's disease. No major limitations were noted with the review. The ADMIRE-CD study was a well-reported and conducted randomised controlled trial (RCT) and measured a range of clinically relevant outcomes.

### *1.6.2 Weaknesses and areas of uncertainty*

Although darvadstrocel offers a novel treatment option with curative intent there are a number of uncertainties in the evidence base: (1) there is no robust supporting data beyond 52 weeks follow-up; (2) there is no evidence on the repeated use of darvadstrocel (licensed dose) when new fistulas open; (3) it is unclear whether patients who have not achieved complete closure with one treatment course would benefit from an additional treatment course, and; (4) whether stem cell therapy would be effective in patients with very complicated perianal fistulising disease who may have more than two internal and/or three external openings.

No evidence was submitted on the cost-effectiveness of darvadstrocel for the treatment of: (1) people who have more than two internal openings or more than three external openings of their complex perianal fistula, or (2) people who receive darvadstrocel as a repeat treatment. The ICER for darvadstrocel versus standard care cannot be estimated in either of these populations and may be substantially different from the ICER for the population considered in the CS. It is unclear whether the ICER would be lower or higher than the base case ICER in these populations.

## 1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook eight sets of exploratory analyses using the deterministic version of the company's model. The ERG's preferred model uses a discount rate of 3.5% for both costs and QALYs and produces a deterministic ICER for darvadstrocel of £23,176 per QALY gained. This model includes: the correction of several minor errors; calibration of the health state occupancy of the defunctioning surgery and proctectomy health states to their data sources; estimating the long term event rates for the salvage therapy arm using the time to event functions for salvage therapy, and; setting the time horizon to 60 years. The ERG undertook a number of further analyses to explore the sensitivity of the ICER to; the inclusion of transitions that were not included in the company's base case model; the impact of under predicting utility values for the CSF mild, successful defunctioning surgery and the successful

proctectomy health states; the impact of using alternative time to event functions; and to assess whether darvadstrocel meets the criteria in Section 6.2.19 of the NICE methods guide. The ERG considers that the exploratory analysis on whether darvadstrocel meets the criteria in Section 6.2.19 of the NICE methods guide indicates that these criteria are not met. Consequently, the ERG considers that both costs and QALYs should be discounted at a rate of 3.5%. The other exploratory analyses suggest that the ICER is sensitive to the time to event functions and any under prediction of the utility values for the CSF mild, successful defunctioning surgery, and/or the successful proctectomy surgery health states. Including additional transitions within the company's model structure has only a minor impact on the ICER.

# 2 BACKGROUND

This report provides a review of the evidence submitted by the company (Takeda) in support of darvadstrocel for treating complex perianal fistulae in people with Crohn's disease. It considers both the company's submission (CS) received on 20th April 2018 and a subsequent response to clarification questions supplied by the company on 24th May 2018.[1, 2]

## 2.1 Critique of company's description of underlying health problem

The CS (pages 14-25) provided a reasonable description of the underlying health problem.[1] The health problem is summarised briefly below.

A perianal fistula is an abnormal passage or tract between the bowel and the perianal region. A complex perianal fistula is difficult to define, but perianal fistulae are usually considered complex if (i) their origin is high enough in the bowel to result in the tract having sphincter involvement or (ii) there are multiple branches with more than one internal or external opening. Information on the aetiology is limited, but in people with Crohn's disease, inflammation of the bowel can lead to repeated abscesses and the development of a perianal fistula and the inflammation of the bowel wall inhibits healing of the fistula.

No direct evidence exists on the incidence of complex perianal fistulae in people with non-active / mildly active Crohn's disease in the UK. In the CS, the company combines evidence on the incidence of Crohn's disease in the UK and data from the Netherlands on the incidence of perianal fistulae to estimate that 7,473 people in the England will have a perianal fistulae and Crohn's disease. The incidence of complex perianal fistulae in people with non-active/mildly active Crohn's disease will be a subset of this population.

Complex perianal fistulae are not associated with mortality, however the available evidence suggests that there is a high morbidity and significant impairment in quality of life (QoL). Symptoms of a perianal fistula include: persistent anal and/or abdominal pain, perianal inflammation, pain during defaecation, continuous malodorous drainage (pus, blood, and faecal material), incontinence, and skin irritation around the anus.[3, 4] As complex perianal fistulae can lead to the development of repeated abscesses, additional effects on QoL include fevers related to an abscess, severe pain, and the abscess itself will require surgical drainage.

## 2.2 Critique of company's overview of current service provision

In general, the CS provides a reasonable overview of current service provision for people with complex perianal fistula and Crohn's disease.[1] The company's description of the treatment pathway is briefly summarised in this section.

First-line treatment for people with Crohn's disease who are diagnosed with a complex perianal fistula consists of examination under anaesthesia (EUA), abscess drainage and loose seton placement. Seton placement involves placing a piece of silicone string into the fistula tract to ensure that the fistula remains open so that it can drain and heal adequately from the middle of the fistula towards the openings. The timing of the seton removal will depend on any other treatments which are given. If the patient has active luminal Crohn's disease, this will be treated in conjunction with the surgical management of the fistula. Immunosuppresants and/or biologics are treatment options to manage any luminal disease that is present in people with complex perianal fistulae who also have mildly luminal Crohn's disease.

Second-line treatment for people who are refractory to first-line treatments is poorly defined and care varies widely across sites even within the UK. Medical decision teams will typically make choices based on their own experience. Data from St Mark's hospital (a UK national referral centre for intestinal and colorectal diseases) indicates that care varies greatly for people in second-line treatment and beyond. Typically, a new seton will be placed and a different medical treatment (from the previous line treatments) will be used. This was also indicated to be the case by the advisors to the ERG. After several lines of failed therapy, patients may go on to receive defunctioning surgery (potentially temporary) or proctectomy (permanent). Defunctioning surgery involves a temporary diversion of the bowel, so that the fistula can heal. A proctectomy involves a permanent removal of the bowel to bypass the perianal fistula.

# 3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

## 3.1 Population

The population defined in the final NICE scope relates to adults with non-active/mildly active luminal Crohn's disease, with complex perianal fistulas which have shown an inadequate response to at least one conventional or biologic therapy.[5]

The population in the CS differs from this population, as it includes only those people with non-active/mildly active luminal Crohn's disease, with complex perianal fistulas which have shown an inadequate response to at least one conventional or biologic therapy who also: (i) have two or less internal openings and three or less external openings of their complex perianal fistula, and; (ii) are naïve to darvadstrocel treatment.

The Summary of Product Characteristics (SmPC) for darvadstrocel, specifies that the full content of four vials must be administered to treat no more than two internal openings or three external openings. It is unclear from the SmPC whether darvadstrocel is licenced to be given more than once; this has two implications.

Firstly, it is unclear from the SmPC whether two procedures could be administered to people who have more than two internal openings or more than thee external openings. The clinical advisors to the ERG stated that care does not currently differ according to the number of external or internal openings of a patient's complex perianal fistula. As such, it is ambiguous whether the population with more than two internal openings or three external openings could be treated with darvadstrocel given the current licence. Therefore, caution may be warranted in interpreting the evidence in this submission for this excluded population, as under the marketing authorisation they may be eligible to receive darvadstrocel.

Secondly, the SmPC does not specify that darvadstrocel can only be administered once per patient, therefore the current licence may allow repeated administration of darvadstrocel. The population included in the final CS does not include any evidence for those people who receive multiple darvadstrocel administrations. As stated in the company's clarification response to question A1, the company have "… *elected to base the submission on single use only….*".[2]

## 3.2 Intervention

The intervention under appraisal is darvadstrocel (24mL dose). Four vials of darvadstrocel are required for a single treatment course. Each vial contains a suspension of 30 million expanded adipose stem cells in a 6mL solution, giving a total dose of 120 million cells per treatment. Darvadstrocel currently holds

an European Union (EU) marketing authorisation for the treatment of complex perianal fistulas in adult patients with non-active/mildly active luminal Crohn's disease, when fistulas have shown an inadequate response to at least one conventional or biologic therapy.[6]

The list price of darvadstrocel stated in the CS (page 12, Table 3) is £13,500 per vial, which corresponds to a total drug cost of £54,000 for one course of treatment. A Patient Access Scheme has been approved by the Department of Health involving a simple price discount. Including the discount, the price of darvadstrocel is [redacted] per vial and a total course of treatment costs [redacted].

Contraindications for darvadstrocel include hypersensitivity to: Dulbecco's Modified Eagle´s Medium, human albumin, or bovine serum.

### 3.3 Comparators

The final NICE scope identified surgical management without darvadstrocel as the only relevant comparator.[5]

The company's review of clinical effectiveness (see Section 0) identified two studies which included direct head-to-head comparisons of darvadstrocel versus surgical management without darvadstrocel. Only the ADMIRE-CD study had a dosing schedule (120 million cells, 4 vials multiplied by 30 million cells per vial) which is in line with the European Medicines Agency (EMA) licence for darvadstrocel.[7] In the other study by de la Portilla *et al*., a dosing schedule of 20 million cells were administered at baseline and in the event of incomplete closure at 12 weeks a further 20 million cells were administered.[8] The clinical evidence which is used to estimate the differences in costs and QALYs between darvadstrocel and surgical management without darvadstrocel in the health economic model is largely based on the ADMIRE-CD study.[7, 9]

### 3.4 Outcomes

The final NICE scope lists the following outcomes[5]:

- Closure of fistula
- Recurrence of fistula
- Continence
- Mortality
- Adverse effects of treatment
- Health-related quality of life (HRQoL)

The CS reports on all of these outcomes, except continence, for patients receiving darvadstrocel or standard care within the ADMIRE-CD study. The ERG's clinical advisors believed that continence was

an important outcome measure, however they thought it was unlikely that incontinence would differ between the darvadstrocel and control arms of ADMIRE-CD.

However, the definition of closure and recurrence of the fistula used in the cost-effectiveness model is defined using a *post hoc* composite outcome, which the company calls clinical and patient-centric (CPC) remission. CPC remission is defined as the closure of external openings as per clinical assessment (not draining when gently compressed with fingers) and the patient does not experience any pain or discharge (defined as a patient scoring 0 in both the pain and discharge sections of the Perianal Disease Activity Index [PDAI] scale). This outcome measure, whilst not pre-specified in the scope, was deemed to be the most relevant outcome by the ERG's clinical advisors for second-line treatment of complex perianal fistulae in people with Crohn's disease.

Fistula recurrence was defined as the lack of continued CPC remission. The ERG's clinical advisors considered that a clinical diagnosis of recurrence of a complex perianal fistula would be made based on clinical factors such as pain, discharge and whether the fistula was adequately draining. However, a successful outcome would not require the fistula to be completely healed.

Mortality was reported in the CS as an adverse event, rather than a primary or secondary outcome in the efficacy analysis.[1] However, this was deemed to be appropriate as there were no deaths in the ADMIRE-CD trial.

HRQoL was captured in the ADMIRE-CD study using the inflammatory bowel disease questionnaire (IBDQ), which is a disease specific measure focused on systemic bowel disease (i.e. luminal Crohn's disease) rather than perianal fistulising disease. The source of utility values for the economic valuation was a separate vignette study (CS,[1] Appendix Q).

### 3.5 Other relevant factors

The CS (page 25) states that there are no equality considerations relevant for the use of darvadstrocel in the treatment of complex perianal fistulae in patients with Crohn's disease.

The company claims that darvadstrocel meets criteria set out in Section 6.2.19 in the NICE Methods Guide (CS, page 64), QALYs should be discounted at 1.5% in the base case.[1, 10] These criteria require that: darvadstrocel restores people to full health for a long period of time (normally at least 30 years); that people receiving standard care have a severely impaired quality of life or would otherwise die, and; that darvadstrocel would not commit the NHS to significant irrecoverable costs. The ERG believes that darvadstrocel does not meet these criteria and, as such, both costs and QALYs should be discounted at 3.5% (see Sections 5.3.4.3 and 5.4).

# 4 CLINICAL EFFECTIVENESS

## 4.1 Critique of the methods of review(s)

### *4.1.1 Searches*

The company performed a single clinical effectiveness search to identify all studies of darvadstrocel and its comparators (broadly called surgical interventions, antibiotics, immunosuppressants, biologics and stem cells) for patients with complex perianal fistula in Crohn's disease.

For the searches, three electronic bibliographic databases including MEDLINE [via Embase.com], MEDLINE in Process [via PubMed], EMBASE [via Embase.com], Cochrane Central Register of Controlled Trials [via Wiley Online Library]) were searched covering the period from inception of the database until January 2018. Several conference proceedings websites (ECC, UEG, AGA/DDW, ESCP, WCG, AIBD, ISPOR) were searched in January 2018 covering the period from 2014 until 2017. The CS did not appear to have searched any clinical trials registers such as clinicaltrials.gov or WHO International Clinical Trials Registry Platform nor did the company report carrying out supplementary searching such as citation searching of included studies. The company's clarification response (question A12) gave details of one ongoing study of darvadstrocel.[2]

In the CS (Appendix D), the company reported the full literature search strategies of the databases searched.[1] The scope of the searches took into account the potential need to make simultaneous comparisons between all interventions (e.g. infliximab, adalimumab, surgical treatment and best supportive care) in the draft NICE scope (the ERG notes that the final scope issued by NICE limited the comparators to surgical management without darvadstrocel only).[5, 11] The search strategy was designed to identify RCTs and systematic reviews of the relevant intervention, darvadstrocel, as well as studies reporting on any comparators relevant to the scope for patients with complex perianal fistula in Crohn's disease (clarification response[2], question A9). Given the broad range of possible comparators, the searches consisted only of terms for 'Crohn disease' or 'fistula' combined/not combined with terms for the comparators and search filters for the relevant study types. However, the strategies did not include all free-text terms for darvadstrocel. At the time of the company searches, darvadstrocel was not indexed in the database (clarification response[2], question A11). The company's amended search, which included keywords for ileostomy, colostomy and new interventions such as stem cells, was provided in the company's clarification response (question A11).[2]

Despite the noted limitations, the ERG considers all the search strategies to be sufficiently comprehensive to retrieve all important and eligible studies of which the ERG and its clinical advisors are aware. However, as no search details/strategies were provided in the CS, it is unclear whether any relevant AE studies have been missed.

*4.1.2 Inclusion criteria*

The CS describes appropriate methods of identifying and screening references for inclusion in the systematic reviews of clinical effectiveness. Two independent reviewers applied pre-specified inclusion and exclusion criteria (via a two-stage sifting process) to citations identified by the searches. Any differences in selection process were resolved through consultation with a third reviewer, if required (CS, Appendix D.1.2). A summary of the inclusion and exclusion criteria is presented in Table 1.

**Table 1: Inclusion/exclusion criteria used select studies of patients with complex perianal fistula and Crohn's disease (adapted from CS,[1] Appendix D, Table 7)**

| Clinical effectiveness | Inclusion criteria | | Exclusion criteria | Rationale |
|---|---|---|---|---|
| **Population** | Patients with perianal fistula in Crohn's disease, irrespective of the age, race, or ethnicity | | Studies which enrolled a mixed population of perianal fistula in Crohn's disease, ulcerative colitis and inflammatory bowel disease of undetermined origin were only included if there was subgroup data for the disease of interest or 80% of the study population met the eligibility criteria of the review | The review is not limited to patients with any particular age group, and does not restrict to any specific gender or race |
| **Intervention** | • 'Cx601' (darvadstrocel) | | | |
| **Comparators** | Surgical interventions<br>• Fibrin glue<br>• advancement flap,<br>• LIFT,<br>• diverting stoma,<br>• proctectomy,<br>• colectomy,<br>• fistula plugs,<br>• fistulotomy,<br>• exam under anaesthesia,<br>• multiple seton placement,<br>• ileostomy,<br>• colostomy,<br>• VAAFT and<br>• Filac | Antibiotics<br>• Ciprofloxacin*<br>• Metronidazole*<br>• Azathioprine*<br><br>Immunosuppressants<br>• Cyclosporine*<br>• Tacrolimus*<br>• Methotrexate*<br>• Thalidomide*<br>• 6-MP*<br><br>Biologics<br>• Infliximab*<br>• Adalimumab*<br>• Certolizumab*<br><br>Other interventions<br>• stem cells* | | Surgical interventions are included in the NICE scope. Antibiotics, immunosuppressants, biologics and other stem cell preparations were included in the search for HRQoL |

| Clinical effectiveness | Inclusion criteria | Exclusion criteria | Rationale |
|---|---|---|---|
| **Outcomes** | • Remission rate<br>• Relapse rate<br>• Definitions of outcomes<br>• No response/failure rate<br>• Fistula closure and partial closure as defined by clinical exam<br>• Fistula internal closure as demonstrated by MRI<br>• Relapse or recurrence rate<br>• Time to remission/relapse<br>• Proportion of patients with draining fistula<br>• Stoma closure<br>• Seton removal time<br>• Mortality<br>• Safety (any adverse events, serious adverse events, specific adverse events) and tolerability (discontinuations due to any reason or due to any adverse event)<br>• HRQoL measures, either disease specific or generic<br>• Perianal Disease Activity Index (PDAI)<br>• Crohn's Disease Activity Index (CDAI)<br>• Inflammatory Bowel Disease Questionnaire (IBDQ)<br>• Short Form 36 Item (SF-36)<br>• EuroQoL-5D (EQ-5D)<br>• Incontinence scores | | |
| **Study design** | • RCT - parallel group<br>• RCT - crossover<br>• Non-randomized controlled clinical trials<br>• Controlled cohort studies (retrospective)<br>• Controlled cohort studies (prospective)<br>• Case-control studies<br>• Cross-sectional studies<br>• Analysis of hospital records/database/chart/claims database<br>• Single arm studies (uncontrolled trials)<br>• For the UK/NICE perspective, only RCTs will be considered for extraction in the clinical review | | |
| **Language restrictions** | English | | |

• CD, Crohn's disease; HRQoL; health related quality of life; MRI, magnetic resonance imaging; RCT, randomised controlled trial

The specified inclusion and exclusion criteria were mostly appropriate and generally reflected the decision problem. It is noteworthy that the CS[1] (page 25) initially considered a wider remit to capture the entire evidence base as part of the inclusion criteria for the review (i.e. all treatments used for the

management of complex perianal fistulae in patients with Crohn's disease) but then restricted the systematic review only to those studies which are directly relevant to the decision problem (i.e. darvadstrocel treatment only [see CS,[1] Section B.2.2]). Despite a request from the ERG to provide separate inclusion and exclusion criteria for two parts of the review, this was not provided by company (clarification response[2], question A9). Ideally, systematic reviews should have clearly focused research questions and inclusion/exclusion criteria at the outset.

The company's systematic review excluded studies which were reported only as abstracts (CS,[1] Appendix D.1.2, Figure 1); however, limited justification for this exclusion was provided. In order to avoid publication bias, a systematic review should aim to include all relevant studies, regardless of publication status. Although differences often occur between data reported in conference abstracts and their corresponding full reports, differences in results are usually not very large.[12] However, the ERG notes that it can be difficult to appraise study quality from limited details provided in an abstract. As a result, sensitivity analyses may be carried out to examine the effect of including data from conference abstracts.[13]

Finally, the reporting of clinical harms is often inadequate in controlled clinical trial publications because they exclude patients at high (or even medium) risk from harms,[12, 14] they may be too short to identify long-term or delayed harms, or they may have insufficient sample sizes to detect rare events.[12, 15] Supplementary sources of evidence may provide additional supporting information concerning safety considerations.[16] The SmPC (pages 11 and 12) suggests that the marketing authorisation was granted with a number of conditions and included the following: periodic safety update reports, adherence to the agreed risk management plan, additional risk minimisation errors (i.e. provide educational material for healthcare professionals on how to give the medicine correctly and on the possibility of passing on an infection to the patient), and conducting a post-authorisation efficacy and safety study - ADMIRE-CD-II (expected to complete in October 2021 [clarification response[2], question A9]).

### *4.1.3 Critique of data extraction*

The data extracted and presented in the clinical section of the CS appear appropriate and comprehensive. Although details of the data extraction process were lacking in the CS, the company's clarification response (question A13) suggests that data extraction was undertaken by two independent reviewers and disagreements were resolved through consultation with a third reviewer.

### *4.1.4 Quality assessment*

The validity assessment tool used to appraise the included studies in the CS[1] (Appendix D.3, p22-23) was based on the minimum criteria for assessment of risk of bias in RCTs, as suggested by the Centre for Reviews and Dissemination.[12] As noted in the company's clarification response (question A13)

methodological quality assessment of included studies was performed by two independent reviewers and disagreements were resolved through consultation with a third reviewer.[2] The ERG acknowledges that the validity assessment tool used in the CS was appropriate.[1]

#### *4.1.5 Evidence synthesis*

The company did not undertake a formal meta-analysis as only one darvadstrocel RCT study was considered relevant to the submission. As a result, the company undertook a narrative synthesis of the evidence for darvadstrocel. However, no explicit details were provided in the CS[1] on how this approach was undertaken. Ideally, a narrative synthesis approach should be justified, rigorous (i.e. describe results without being selective or emphasising some findings over others) and transparent to reduce potential bias.[12, 15] Despite the lack of transparency regarding the methods adopted, the ERG acknowledges that the narrative synthesis approach undertaken by the company was acceptable.

### 4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

#### *4.2.1 Studies included in/excluded from the submission*

The company's Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram relating to the literature searches does not conform exactly to the PRISMA statement flow diagram (http://www.prisma-statement.org/). Despite this, the revised diagram and accompanying narrative provided in the company's clarification response (questions A15 and A16) appear to be a reasonable record of the literature searching and screening process for the systematic literature review of treatments used for the management of complex perianal fistulae in patients with Crohn's disease.[2] Moreover, although the CS initially failed to provide a full and explicit breakdown of the reasons why each citation was rejected (especially after full text papers were retrieved for detailed evaluation), further details were provided by the company in their clarification response (questions A14 and A16).[1, 2]

The company's systematic review of darvadstrocel for the treatment of complex perianal fistulae in patients with Crohn's disease identified two potentially relevant studies (a Phase I/IIa study[8] and a Phase III study).[7, 9] However, as suggested in the CS[1] (p24) and the European Public Assessment Report (EPAR),[17] the design and context of the Phase I /IIa study[8] was not considered to be entirely relevant to the recommended dosing or the licenced indication in the approved product label for darvadstrocel (further details of this study are briefly provided in Section 4.2.4.2). As such, evidence from the Phase III ADMIRE-CD study[7, 9] forms the main pivotal evidence in the CS.[1] Further details of this study are provided in this section.

The company's broader systematic review of all RCTs for complex perianal fistulae in patients with Crohn's disease (which was conducted to assess the feasibility of performing a network meta analysis (NMA) against other treatment options, such as: surgical interventions and medical treatments [i.e. antibiotics, immunosuppressants and biologics; however, these were not included in the final scope issued by NICE])[5] initially identified six potential studies (clarification response,[2] question A15 and A16). Of these, no additional studies to the ADMIRE-CD trial[7, 9] were considered relevant to the decision problem. The company stated that "… *an NMA could not be conducted due to a lack of comparable RCTs and considerable heterogeneity in the studies identified by the systematic review. The assessment found a high level of variability in the comparators, outcomes, patient populations, and sample size across studies.*" (CS,[1] page 53).

- Main evidence (pivotal study: ADMIRE-CD)[7, 9]

The ADMIRE-CD study[7, 9] was a Phase III, company-sponsored, randomised, double-blind, placebo-controlled, multicentre trial designed to assess the efficacy and safety of a single intralesional injection of darvadstrocel (an allogeneic preparation of adipose-tissue-derived mesenchymal stem cells) and standard of care in 212 patients (54.7% male, 92.5% Caucasian) with non-active or mildly active luminal Crohn's disease (defined by a Crohn's Disease Activity Index [CDAI] of ≤220) who had complex perianal fistulae that was refractory to conventional (i.e. antibiotics, immunosuppressants) and/or biological therapy. There is some uncertainty about the repeat use of darvadstrocel in clinical practice, it should be noted that the company states that "*Although some clinicians believe that Alofisel [darvadstrocel] may be beneficial for retreatment in the following patient groups; (i) partial responders; (ii) responders who have relapsed, there is no current evidence to support this treatment approach… therefore elected to base the submission on single use only*" (clarification response[2], question A1). A summary of the study design and population characteristics is provided in Table 2.

The study included patients from 49 hospitals across seven European Union countries (Austria, Belgium, France, Germany, Italy, the Netherlands, and Spain) and Israel. Eligible patients were enrolled between July 2012 to July 2015 and were required to be: (i) ≥18 years old (mean age, 38 years; >65 years, n=7)[17]; (ii) diagnosed with Crohn's disease at least 6 months earlier (in accordance with accepted clinical, endoscopic, histological and/or radiologic criteria); (iii) had complex perianal fistulas with a maximum of 2 internal and 3 external openings (assessed by clinical assessment and MRI) that had been draining for at least 6 weeks (a complex perianal fistula was defined as one or more of the following during its evolution: high intersphincteric, high trans-sphincteric, extra-sphincteric, or supra-sphincteric origin; at least two external openings (tracts); or associated collections); (iv) refractory to antibiotics (ciprofloxacin or metronidazole with lack of response after one month of treatment), immunomodulators (azathioprine, 6-mercaptopurine, or methotrexate with no response after 3 months), or induction or maintenance therapy with anti-TNF therapies. The key exclusion criteria were: (1) a

history of rectovaginal fistulas; (2) rectal and/or anal stenosis and/or active severe proctitis; (3) diverting stomas, an abscess (collection >2 cm) that was not properly drained at the fistula preparation visit; (4) received corticosteroids within the previous 4 weeks; (5) if they had not received previous treatment for perianal fistulising Crohn's disease including antibiotics, and those who underwent previous surgery for the active fistula other than drainage or seton placement.

**Table 2: Characteristics of the ADMIRE-CD study[7, 9]**

| Study | Location (sites) | Design | Population | Intervention | Comparator | Primary outcome measures | Duration |
|---|---|---|---|---|---|---|---|
| ADMIRE-CD (NCT01541579; Cx601-0302)[7, 9]<br><br>Funded by: TiGenix | 49 sites in 8 countries (Austria, Belgium, France, Germany, Italy, the Netherlands, Spain and Israel) | Phase III, randomised, double-blind, parallel group, placebo controlled trial (n=212) | Patients (aged ≥ 18 years) with complex perianal fistulising Crohn´s disease who are refractory to conventional (antibiotics, immunosuppressants) or biological treatment strategies | Darvadstrocel (24 mL containing 120 million expanded allogeneic adipose-derived stem cells) given as a single intralesional injection[a] and standard of care (n=107) | Placebo (24 mL saline solution) given as a single intralesional injection and standard of care (n=105) | Combined remission (clinical and MRI) at 24 weeks[b] | Active treatment consists of one administration of darvadstrocel, follow-up extended from 24 weeks to 52 weeks and then to 104 weeks[c] |

MRI, magnetic resonance imaging

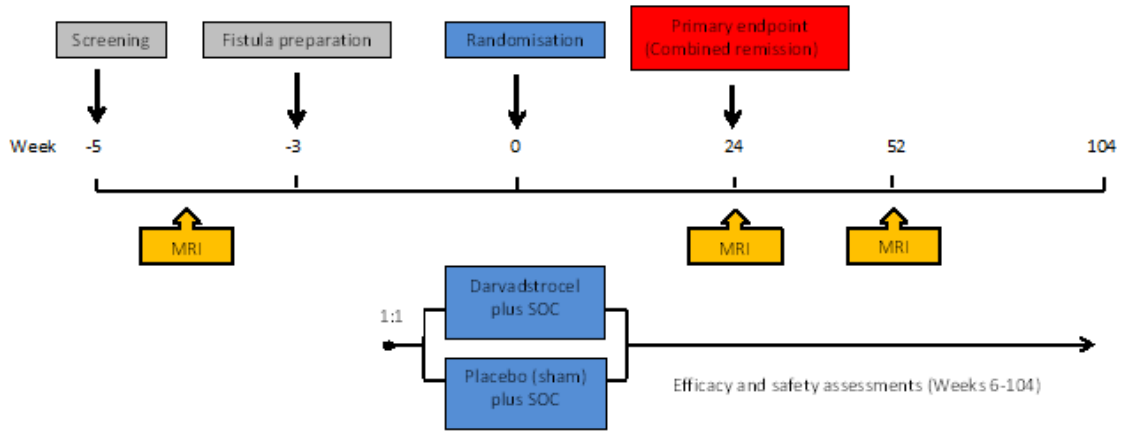
[a] The administration procedure involved the injection of darvadstrocel (or placebo) into the tissues surrounding the tract. Four vials (6mL each) containing approximately 30 million cells were shipped to the hospital for use by the surgeon on the day they were received. The content of two vials (60 million cells) was injected into the fistula walls along the length of the fistula tract and two vials (60 million cells) injected around the internal opening during an Examination Under Anaesthesia. This procedure was done by specialist physicians experienced in the diagnosis and treatment of conditions for which darvadstrocel is indicated.

[b] Defined as the clinical assessment of closure of all treated external openings that were draining at baseline, and the absence of collections > 2 cm of the treated perianal fistula in at least two of three dimensions, confirmed by masked central magnetic resonance imaging. Clinical assessment of closure was defined as the absence of draining despite gentle finger compression.

[c] Following a series of protocol amendments, the follow-up period was extended to 52 weeks (October 2012) and then to 104 weeks (December 2014)[17]

Prior to randomisation, a pelvic MRI was administered (screening visit) and patients' fistulae were examined under anaesthesia, curetted and, if indicated, setons were placed during this procedure (preparation visit). If a seton was placed, this was subsequently removed immediately prior to the administration of darvadstrocel Subsequently, patients were randomly allocated to receive darvadstrocel and standard of care (n=107) or placebo sham (saline) and standard of care (n=105) in a 1:1 ratio, with risk stratification based upon previously received therapy (immunomodulators, anti-TNF therapy, both, or neither). A summary of the ADMIRE-CD trial[7, 9] schema is presented in Figure 1.

**Figure 1: ADMIRE-CD trial schema (adapted from CS,[1] Figure 7)**




MRI, magnetic resonance imaging; SOC, standard of care

After darvadstrocel administration, patients could be treated with antibiotics for no more than four weeks. Immunomodulators and anti-TNF drugs were maintained at stable doses throughout the study. Initiation or dose increases of these drugs were not allowed. A steroid course was permitted to treat occurrences of luminal disease during the study, with a starting dose of 40 mg tapered over a maximum of 12 weeks. Fistula closure was clinically assessed at weeks 6, 12, 18, and 24, 36 and 52; assessing for spontaneous drainage after gentle finger compression was applied to treat external openings. Fistula-associated collections were also radiologically assessed at weeks 24 and 52 by blinded, centrally read pelvic MRI scans. The study protocol was amended five times ( CS,[1] page 31), the ERG considers that the most notable change included extending the trial duration from 24 weeks to 104 weeks to allow assessment of long-term efficacy and clinical and immunological safety of darvadstrocel treatment.

The primary endpoint was combined remission (both clinical and radiologic improvement) at week 24 after study treatment and was defined as the clinical assessment of closure of all treated external openings that were draining at baseline, and the absence of collections >2 cm within the perianal fistula in at least two of three dimensions, confirmed by blinded central MRI. The clinical assessment of closure was defined as the absence of draining despite gentle finger compression. The key secondary

endpoints were defined as clinical remission (closure of all treated external openings that were draining at baseline despite gentle finger compression) and response (clinical closure of at least 50% of all treated external openings that were draining at baseline) at week 24. In addition, long term follow-up was conducted up to week 52 and 104 (CS,[1] page 39). As noted in the CS[1] (page 64) *'...the efficacy data available beyond 52 weeks was limited. This is due to the changes in the protocol whereby the trial duration was extended beyond 104 weeks, which occurred when various patients had already finished the 52 week trial period. This resulted in a low level of patient data, and so generalisation of results beyond 52 weeks is difficult and should be approached with care'*. Other endpoints included safety, time to clinical remission, time to response, relapse, time to relapse and various disease severity measures such as (CS,[1] page 32): Perianal Disease Activity Index (PDAI) and the Van Assche scores (both focus on local perianal fistulising disease activity); Crohn's Disease Activity Index (which focuses on luminal Crohn's disease severity [CDAI]) and the Inflammatory Bowel Disease Questionnaire (a quality of life measure that focuses on systemic bowel disease e.g. luminal Crohn's disease [IBDQ]).

- *Ongoing studies of darvadstrocel for treating complex perianal fistula in non-active or mildly active luminal Crohn's disease*

Although there are no ongoing studies of darvadstrocel that will provide additional evidence in the next 12 months (CS,[1] page 59), the ADMIRE-CD-II study[18] (ClinicalTrials.gov Identifier: NCT03279081; Cx601-0303) is currently recruiting. The company's clarification response to question A12[2] suggests that this is a similar study to the ADMIRE-CD study[7, 9] but is being conducted to include patients from the USA and to satisfy Food and Drug Administration (FDA) requirements (Table 3). This study is expected to complete in October 2021. No other studies are currently planned (see company's clarification response,[2] question A9).

**Table 3: Summary of key ongoing studies**

| Criteria | ADMIRE-CD-II study[18] |
|---|---|
| Title (official) | Phase-III randomised, double-blind, parallel-group, placebo-controlled, multicentre study to assess efficacy and safety of Cx601, allogeneic expanded adipose-derived stem cells for complex perianal fistula(s) in Crohn's disease - ADMIRE-CD-II |
| Study ID number | Clincinaltrials.gov: NCT03279081<br>Other: Cx601-0303; 2017-000725-12 (EudraCT Number) |
| Primary objective | To evaluate the efficacy and safety of darvadstrocel compared to placebo for the treatment of complex perianal fistula(s) in patients with Crohn's disease at week 24 with a follow-up period up to 52 weeks. |
| Study design | Phase III, randomised, double-blind, placebo controlled trial |
| Study location | >120 sites in EU/Israel and Canada/ USA (~60% of all sites) |
| Study population | • Target enrolment: 326 patients to be randomised (>436 to be screened)<br>• Patients (aged 18-75 years) with complex perianal fistulising (maximum of 2 internal openings and a maximum of 3 external openings) Crohn´s disease who are refractory to conventional (antibiotics, immunosuppressants) or biological treatment strategies |
| Intervention/ comparator | • Darvadstrocel (24mL containing 120 million expanded allogeneic adipose-derived stem cells) given as a single intralesional injection and standard of care<br>• Placebo solution given as a single intralesional injection and standard of care |
| Primary endpoint | • Combined remission at week 24 with $\alpha<0.05$ for all treated fistulas |
| Key secondary endpoints at week 24 and relevant at week 52 | • Clinical remission at week 24<br>• Response at week 24<br>• Combined remission, clinical remission/response at week 52<br>• Time to clinical remission / response at week 24, week 52<br>• Safety and tolerability up to week 52<br>• Electronic patient-reported outcomes and quality of life assessments |
| Expected completion date | October 2021 |

### *4.2.2 Details of relevant studies not included in the submission*

The ERG is confident that all relevant studies have been included in the CS[1] and that details of all ongoing trials that are likely to be reporting additional evidence within 12 months were reported.

### *4.2.3 Summary and critique of the company's analysis of validity assessment*

The company provided a formal appraisal of the validity of the included darvadstrocel RCT[7, 9] using standard and appropriate criteria (an adaptation for the Centre for Reviews and Dissemination's guidance for undertaking reviews in healthcare[12]). However, the ERG is unclear as to why the company undertook quality assessments of the potentially relevant studies identified in the broader systematic review of all RCTs for complex perianal fistulae in patients with Crohn's disease (CS,[1] Appendix D.4.5), as none of these studies were included or considered relevant to the decision problem (clarification response[2], question A16 and A25). The completed validity assessment tool for the ADMIRE-CD trial, as reported in the CS,[1] is reproduced (with minor changes) in Table 4.

**Table 4: Quality assessment results for the ADMIRE-CD study, as assessed by the company (adapted from CS,[1] Appendix D3, Table 2)**

| Quality assessment criteria | ADMIRE-CD[7, 9] | |
|---|---|---|
| | Company's assessment | ERG's assessment |
| Was randomisation carried out appropriately? | Yes | Yes |
| Was the concealment of treatment allocation adequate? | Yes | Yes |
| Were the care providers, participants and outcome assessors blind to treatment allocation? | Yes | Yes |
| Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for? | No | No |
| Is there any evidence to suggest that the authors measured more outcomes than they reported? | No | No |
| Did the analysis include an intent-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data? | Yes | Yes |

In general, the ERG considered the ADMIRE-CD trial[7, 9] to be a well-reported and conducted study; however, some further discussion around specific points is required.

In the ADMIRE-CD trial,[7, 9] randomisation was performed using a computer generated randomisation list (stratified according to previously received therapy i.e. immunomodulators, anti-TNF therapy, both, or neither) and allocation concealment was done centrally by a third party. Whilst masking of treatments was not possible due to the visual differences between the darvadstrocel cell suspension and saline solution (i.e. placebo), the double-blind design of the study was maintained by a blinded gastroenterologist and blinded radiologist independently evaluating the clinical and radiological responses, respectively. Unmasked surgeons who administered the treatment were not permitted to share information about the treatment used in the surgical procedure with the gastroenterologist or radiologists, and were also not allowed to participate in any clinical assessment of the fistula during the study. The ERG acknowledges that adequate methods of randomisation, allocation concealment and blinding were used in the conduct of the included trial.

The ADMIRE-CD trial,[7, 9] stratified randomisation according to previously received therapy and did not specify any other relevant prognostic factors. The company's clarification response (question A27) states, "*In a review by Braithwaite et al.(2017), prognostic factors affecting outcomes of perianal disease were examined. This review identified some studies showing significant prognostic factors, yet these were considered insignificant in other identified studies. The heterogeneity observed across the identified studies limits the ability to draw robust conclusions about prognostic markers in this*

*population.*"[2] Prognostic factors should be accounted for in statistical analyses whether or not there is baseline balance between treatments. Nevertheless, the CS[1] (Table 9, pages 34 to 36) suggests that there were slight imbalances in the following key baseline disease characteristics (≥5% difference between the two treatment groups). In the darvadstrocel group 48/107 patients (45%) compared with 31/105 patients (30%) in the control group had more than one fistula tract. The proportion of patients with more than one draining external fistula opening was slightly higher for patients randomised to darvadstrocel (56%, 36%, and 8%, for 1, 2 or >2 draining external openings, respectively [safety population, n=103]) compared with control treatment (72%, 25%, and 4%, respectively [safety population, n=102]). A similar pattern was observed for internal openings, and patients randomised to darvadstrocel were more likely to have two internal openings (0%, 80% and 20% for 0, 1, 2 respectively [safety population, n=103]), compared with patients randomised to control treatment (1%, 88%, 11% respectively [safety population, n=102]). In addition, the majority of patients were receiving concomitant Crohn's disease medication at baseline, although approximately 24 % of patients in the darvadstrocel group and 18 % of the control group did not receive concomitant treatment with either immunosuppressants and/or anti-TNF. The primary endpoint was analysed using a stratified Cochran-Mantel-Haenszel test adjusted for the randomisation strata. PDAI score was analysed using analysis of covariance adjusting for the randomisation strata and baseline response. An imbalance in a variable that is not prognostic is not important. Overall, it is not clear how these baseline differences and ignoring observed variables that may be prognostic may have affected the results.

The CS (Table 11, page 40) showed that during the study period of the ADMIRE-CD study,[7] 19/107 patients (17.8%) in darvadstrocel group and 22/105 patients (21.0%) in the control group did not complete the study protocol due to substantial clinical deterioration, adverse events (AEs) or patient decision/withdrawal of consent.[1] In general, the robustness of an analysis may be threatened if attrition is more than 20%, depending on the method of analysis.[19] In the ADMIRE-CD trial,[7] all patients were accounted for and the key efficacy analyses were conducted using the intention-to-treat (ITT) approach (which included all randomly assigned patients, n=212) or the modified ITT (mITT) approach (which included all randomly assigned patients who received study treatment and had at least one efficacy assessment after baseline, n=204). Therefore, attrition bias should be low in the ADMIRE-CD study.[7, 9]

Although there was no evidence to suggest that the ADMIRE-CD[7, 9] authors measured more outcomes than they reported, based on feedback from clinical experts, the company (CS,[1] page 34 and clarification response[2], question A3) performed and presented two additional *post hoc* analyses - time to clinical and patient-centric [CPC] remission and time to relapse from CPC remission.[1] As noted in the CS (page 10), whilst the key outcomes from the ADMIRE-CD study were combined remission and clinical remission, gastroenterologists and surgeons from the St Mark's Hospital (UK) advised that a more

clinically appropriate outcome of relevance to Crohn's disease patients with perianal fistulae should include a component of pain and discharge in addition to clinical remission. [1] The CS (page 34) considered a patient '*...to achieve CPC remission from complex perianal fistulae when: all the external openings are closed as per clinical assessment, i.e. not draining despite gentle finger compression (i.e. the clinical remission definition of ADMIRE-CD); AND the patient does not experience any pain or discharge, as determined by a score equal to 0 in both the pain and discharge dimensions of the PDAI... The time to CPC remission was the outcome used in the economic model, because expert clinical opinion indicated that this outcome represented more accurately the decision algorithm used in clinical practice*.'[1] The clinical advisors to the ERG also considered CPC remission to be the most clinically relevant outcome to Crohn's disease patients with perianal fistulae. However, the ERG notes that the ADMIRE-CD study was not designed to test hypothesis about these exploratory analyses, as such; the results of these outcomes should be treated with caution.

The ADMIRE-CD trial[7, 9] was performed across several EU countries and Israel; however, no UK sites were included. Based on the findings of a retrospective cohort study of 78 patients, treated by St Mark's Hospital in London (a specialist centre for intestinal and colorectal disorders), the CS[1] (page 63, Appendix Q) and clarification response[2] (questions A21 and A26) suggest that surgical treatments such as an examination under anaesthesia plus/minus seton placement are the most common treatments (approximately 90%) in UK clinical practice for adults with Crohn's disease who have a complex perianal fistula that is refractory to conventional or biologic therapy. In addition, the background therapy received in the trial (antibiotics/immunosuppressants and biologics) was similar to that used in clinical practice. As a result, the CS[1] considered the ADMIRE-CD trial[7, 9] to be reflective of UK practice. Clinical advisors to the ERG agreed with this view.

### *4.2.4 Summary and critique of results*

This section presents the main results from the ADMIRE-CD trial, based on information reported in the CS[1] and trial publications,[7, 9] for the efficacy and safety of darvadstrocel in the treatment of non-active/mildly active luminal Crohn's disease, with complex perianal fistulas which have shown an inadequate response to at least one conventional or biologic therapy. Additional information, not reported in the CS,[1] was provided by the company in the company's clarification response.[2]

#### 4.2.4.1 Efficacy

- Primary outcome (CS,[1] Table 12 and 13, page 42)

In the primary ITT population (n=212), a significantly greater proportion of patients in the darvadstrocel group achieved the primary endpoint of combined remission at week 24 compared with the control group (49.5% versus 34.3%, respectively; difference 15.2%, 97.5% confidence interval [CI]: 0.2 to 30.3; $p$=0.024). Similar results were observed in the mITT (n=204) population (51.5% versus 35.6%;

difference 15.8%, 97.5% CI: 0.5 to 31.2; $p$=0.021) and across all sensitivity analyses ($p$<0.05) used to assess the effects of the imputation conventions for missing data and the impact of use of rescue medication. With longer follow-up (52 weeks), the beneficial effect of darvadstrocel was maintained in the ITT population with 54.2% of patients achieving combined remission compared with 37.1% in the control group (difference 17.1%, 97.5% CI: NR; $p$=0.012).[20] Similar results were observed in the mITT population (56.3% versus 38.6%; 17.7%, 95% CI: 4.2 to 31.2; $p$=0·010). A summary of the key results is presented in Table 5 and Table 6.

**Table 5: Summary of key results from the ADMIRE-CD trial[7, 9] - combined remission, clinical remission and response (adapted from CS[1] Tables 12, 13 and 14)**

| Outcomes | Darvadstrocel n/total N (%) | Control n/total N (%) | Difference (%) | 95 % CI (unless otherwise stated) | *p*-value |
|---|---|---|---|---|---|
| **Analyses at week 24** | | | | | |
| ***Combined remission*** | | | | | |
| ITT population[a] | 53/107 (49.5%) | 36/105 (34.3%) | 15.2% | 97.5% CI: 0.2, 30.3 | 0.024 |
| mITT population | 53/103 (51.5%) | 36/101 (35.6%) | 15.8% | 97.5% CI: 0.5, 31.2 | 0.021 |
| Sensitivity 1[b] | 52/107 (48.6%) | 34/105 (32.4%) | 16.2% | 97.5% CI: 1.3, 31.1 | 0.014 |
| Sensitivity 2[c] | 53/107 (49.5%) | 36/105 (34.3%) | 15.2% | 97.5% CI: 0.2, 30.3 | 0.024 |
| Sensitivity 3[d] | 53/107 (49.5%) | 36/105 (34.3%) | NA | NA | 0.017 |
| ***Clinical remission*** | | | | | |
| ITT population | 57/107 (53.3%) | 43/105 (41.0%) | 12.3% | -1.0, 25.7 | 0.064 |
| ***Response*** | | | | | |
| ITT population | 71/107 (66.4%) | 56/105 (53.3%) | 13.0% | -0.1, 26.1 | 0.054 |
| **Analyses at week 52** | | | | | |
| ***Combined remission*** | | | | | |
| ITT population[20] | 58/107 (54.2%) | 39/105 (37.1%) | 17.1% | NR | 0.012 |
| mITT population | 58/103 (56.3%) | 39/101 (38.6%) | 17.7% | 4.2, 31.2 | 0.010 |
| ***Clinical remission*** | | | | | |
| mITT population | 61/103 (59.2%) | 42/101 (41.6%) | 17.6% | 4.1, 31.1 | 0.013 |
| ***Response*** | | | | | |
| mITT population | 68/103 (66.0%) | 56/101 (55.4%) | 10.6% | -2.8, 23.9 | 0.128 |

CI, Confidence interval; LOCF, Last observation carried forward; (m)ITT, (modified) intention-to-treat; NA, not applicable; NR, not reported

[a] Primary analysis of the ADMIRE-CD trial[7]

[b] Sensitivity analysis 1: ITT, non-response/non-remission imputed for all missing data and after rescue therapy (no LOCF) (rescue therapy was defined as corticosteroids at 40 mg prednisone equivalent for ≥12 weeks; new anti-TNF compared with baseline therapy for ≥8 weeks; new immunosuppressant compared with baseline therapy for ≥12 weeks; or surgical intervention for the treated fistula)

[c] Sensitivity analysis 2: ITT, missing = non-response/non-remission after LOCF applied. Rescue medication not considered as failure

[d] Sensitivity analysis 3: ITT, missing = non-response/non-remission after LOCF applied. Logistic analysis including stratification factor and number of baseline external openings as factors

**Table 6: Time to combined remission, clinical remission and response of perianal fistula by week 24, ITT Population (adapted from CS,[1] Table 15)**

| | Darvadstrocel (N=107) | Control (N=105) | Hazard ratio (95% CI) |
|---|---|---|---|
| ***Combined remission*** | | | |
| Combined remission, n (%) [a] | 53 (49.5%) | 36 (34.3%) | |
| Censored cases, n (%) | [redacted] | [redacted] | |
| Kaplan-Meier estimates, Median (95% CI), weeks | 25.0 (24.7, 26.1) | 28.1 (24.7, 36.0) | 0.74 (0.48, 1.14) |
| ***Clinical remission*** | | | |
| Clinical remission, n (%) [a] | [redacted] | [redacted] | |
| Censored cases, n (%) | [redacted] | [redacted] | |
| Kaplan-Meier estimates, Median (95% CI), weeks | 6.7 (6.4, 11.9) | 14.6 (11.9, 22.9) | 0.57 (0.41, 0.79) |
| ***Response*** | | | |
| Response, n (%) [a] | [redacted] | [redacted] | |
| Censored cases, n (%) | 18 (16.8%) | 30 (28.6%) | |
| Kaplan-Meier estimates, Median (95% CI), weeks | 6.3 (6.0, 6.6) | 11.7 (6.7, 12.9) | 0.59 (0.43, 0.81) |

CI, Confidence interval; ITT, Intention-to-treat
[a] Achieved at least once during the 24-week follow-up

- Secondary and other outcomes (CS,[1] p42-49)

A range of secondary endpoints were evaluated in the ADMIRE-CD study. A summary of the results is presented in Table 5 and Table 6. The key secondary endpoints were clinical remission and clinical response at week 24.

In the ITT population, 53.3% of the patients treated with darvadstrocel achieved clinical remission compared with 41.0% of the control patients (difference 12.3%; *p*=0.064) at week 24. Similar results were observed in the mITT population (55% and 43%, respectively; difference 12.8%; *p*=0.057).[9] The time to achieve clinical remission was significantly faster by 7.9 weeks for the darvadstrocel group compared with the control group (6.7 versus 14.6 weeks, respectively; hazard ratio [HR]: 0.57, 95% CI: 0.41 to 0.79; *p* = not reported [NR]). With longer follow-up (52 weeks), clinical remission in the mITT population (data not reported for ITT population) was 59.2% in the darvadstrocel group and 41.6% in the control group with a difference of 17.6% (*p*=0.013).

In the ITT population, response was achieved in 66.4% of the patients treated with darvadstrocel compared with 53.3% of the control patients (difference 13.0%; *p*=0.054) at week 24. Similar results were observed in the mITT population (69% and 55%, respectively; difference 13.5%; *p*=0.045).[9] The time to response was significantly faster by 5.4 weeks with darvadstrocel compared with the control group (6.3 versus. 11.7 weeks, respectively; HR: 0.59, 95% CI: 0.43 to 0.81; *p* = NR;). At week 52, response in the mITT population (data not reported for ITT population]) was achieved in 66.0% in the darvadstrocel group and 55.4% in the control group with a difference of 10.6% (*p*=0.128).

The ERG notes that the times to clinical remission and clinical response are interval censored such that events could have occurred at any time between assessments; this may result in exaggerated estimates of treatment effect. According to the CS, time to clinical remission and response were analysed using Cox regressions adjusted for the randomisation stratum, although HRs from this model are not presented in the CS. The company's clarification response[2] (question A28) states that the results of the Cox regression could be found in Tables 14.1.4.3.1, 14.1.4.4.1 and 14.1.4.5.1 of the week 24 CSR, although the ERG could not find these. Furthermore, the company clarification response stated, "*As there is no evidence of non-homogeneity in the treatment effect across strata and the trial was not powered to detect differences in treatment effect between these randomisation strata, these analyses were not included in the CS*." (clarification response, [2] A28)

Various other disease severity outcomes measures (PDAI, CDAI and Van Assche score) and quality of life (IBDQ) were assessed in the ADMIRE-CD trial. Detailed results for these outcomes are presented in the CS (pages 45-49) and in Panes *et al*.[9] Briefly, total PDAI scores in the mITT population decreased in both treatment groups at all visits (week 6, 12 and 18) and at week 24 (treatment difference, -0.8; 95% CI: -1.8 to 0.2; *p*=0.101) and week 52 (treatment difference, -0.7; 95% CI: -1.7 to 0.3; *p*=0.186),[9] with the improvement (i.e. decrease) being greater in the darvadstrocel group compared with the control group. However, the differences between treatments did not reach statistical significance ($p > 0.05$). Similarly, in the mITT population, there were no significant differences ($p>0.05$ for all) between the groups at weeks 24 or 52 for total and subdomain IBDQ, CDAI and Van Assche scores. The CS (page 48) stated that '*...darvadstrocel did not have an effect on instruments designed primarily to assess the impact of luminal CD, such as the CDAI or IBDQ…. Since patients with active luminal disease were excluded from the study, CDAI scores were low and IBDQ scores were high throughout as expected'*.[1] A summary of these results is presented in

Table 7.

**Table 7: Other secondary outcomes - PDAI, CDAI, IBDQ and Van Assche score, mITT population (adapted from CS[1] Table 17)**

| Outcome | Darvadstrocel (N=103) | Control (N=101) | Treatment difference (95% CI) | *p*-value |
|---|---|---|---|---|
| ***PDAI, mean (SD)*** [a] | | | | |
| Baseline | 6.7 (2.5) | 6.5 (2.8) | NR | |
| Week 24 | 4.4 (3.6) | 5.1 (3.9) | NR | |
| Change from baseline | -2.3 (3.8) | -1.3 (3.5) | -0.8 (-1.8 to 0.2) | 0.101 |
| Week 52 | 4.4 (3.8) | 5.0 (4.0) | NR | |
| Change from baseline | -2.3 (4.1) | -1.4 (3.7) | -0.7 (-1.7 to 0.3) | 0.186 |
| ***IBDQ*** [b] ***mean (SD)*** | | | | |
| Baseline | 173.5 (31.6) | 169.4 (36.1) | NR | NR |
| Week 24 | 178.3 (34.6) | 174.7 (36.2) | NR | NR |
| Change from baseline | 3.8 (25.5) | 4.0 (25.6) | 0.3 (-6.6, 7.3) | 0.923 |
| Week 52 | 176.1 (38.1) | 172.7 (40.6) | NR | NR |
| Change from baseline | 2.1 (27.4) | 1.7 (25.0) | 0.7 (-6.7, 8.2) | 0.849 |
| ***CDAI*** [c] ***mean (SD)*** | | | | |
| Baseline | 87.8 (48.3) | 93.3 (55.0) | NR | NR |
| Week 24 | 92.5 (66.5) | 94.1 (76.1) | NR | NR |
| Change from baseline | 5.7 (62.2) | 2.2 (65.5) | 1.8 (-16.0, 19.7) | 0.839 |
| Week 52 | 97.4 (82.7) | 99.2 (77.8) | NR | NR |
| Change from baseline | 11.1 (80.5) | 7.6 (67.3) | -1.3 (-19.6, 22.1) | 0.906 |
| ***Van Assche Score*** [d] | | | | |
| Baseline | 9.0 | 9.4 | NR | NR |
| Week 24 | 8.6 | 9.0 | 0.004 (-0.686, 0.694) | NR |
| Change from baseline | NR | NR | NR | NR |
| Week 52 | [redacted] | [redacted] | [redacted] | NR |
| Change from baseline | NR | NR | NR | NR |

CDAI, Crohn's Disease Activity Index; CI, Confidence interval; IBDQ, Inflammatory Bowel Disease Questionnaire; mITT, Modified intention-to-treat; PDAI, Perianal Disease Activity Index; SD, Standard deviation

[a] Data from Panes *et al.*[9]
[b] IBDQ score ranges from 32 to 224, whereby a higher score indicates a better quality of life
[c] CDAI score ranges from 0 to 600, whereby a higher score indicates that the disease is more active / severe
[d] Van Assche score ranges from 0-22, whereby a higher score indicates more severe disease

In the mITT population, a subgroup analysis across four randomisation strata (i.e. Crohn's disease treatment being received at the time of randomisation) found that the effect of darvadstrocel on combined remission was proportionally greater than control with the difference between groups being greatest in patients receiving neither (difference 33.1%, 95% CI: 6.0 to 60.2; *p*=NR) or both anti-TNF and immunosuppressant treatments (20.0%, 95% CI: –5.2 to 45.2; *p*=NR) at week 24; however, the difference in the treatment effect between the four stratification groups was not significant (*p*=0.47).

The CS[1] (page 52) notes that *'…The trial was not powered for the subgroup analyses due to the small patient numbers in these subgroups… Due to low patient numbers during the 52 week follow up, it is not possible to analyse the relapse rates within these subgroups'.*

- *Post hoc* analyses (CS,[1] pages 49-51)

The company presented two additional *post hoc* analyses: (i) time to CPC remission (used in the economic model, because expert clinical opinion to the company indicated that this outcome represented more accurately the decision algorithm used in UK clinical practice), and (ii) time to relapse from CPC remission (as these outcomes were considered by clinical experts to the most relevant outcome). The time to CPC remission in the ITT population, improved by 14.1% in the darvadstrocel group as compared with control treatment (55.1% versus 41.0%, respectively), and the median time to CPC remission was 6.5 weeks faster (28.7 versus 35.2 weeks, HR 0.61; 95% CI: 0.42 to 0.91; *p*=NR). The CS[1] (page 50) noted that *'…this analysis yields very similar results to the combined remission results…'*. Moreover, fewer patients relapsed with darvadstrocel as compared with control treatment (50.8% versus 59.6%, respectively). The time to loss of CPC remission was extended with darvadstrocel compared with control (48.7 versus. 12.9 weeks; HR: 1.38; 95% CI: 0.89 to 2.12; *p*=NR). A summary of these data, adapted by the ERG, is presented in Table 8. However, the company's clarification response[2] (question A20) stated that the HR and 95% confidence interval for CPC relapse is from a Gompertz model. The HR under this model suggests that the effect of darvadstrocel on CPC relapse is worse than control, although the sample data suggest otherwise. It is unclear whether a Gompertz model was also used to estimate the HR for CPC remission.

**Table 8: *Post hoc* analyses - time to CPC remission and time to relapse from CPC remission (adapted from CS[1] Tables 18 and 19)**

| | **Darvadstrocel** | **Control** | **Hazard ratio (95% CI)** |
|---|---|---|---|
| ***CPC remission*** | | | |
| Patients at risk | N=107 | N=105 | |
| CPC remission, n (%) | 59 (55.1%) | 43 (41.0%) | |
| Kaplan-Meier estimates, Median (95% CI), weeks[a] | 28.7 (17.7, 37.0) | 35.2 (24.4, NA) | 0.61 (0.42, 0.91) |
| Log-rank test | | | $X_1^2$=6.0, p=0.014 |
| ***CPC relapse*** | | | |
| Patients at risk | N=59 | N=47 | |
| CPC relapse, n (%) | 30 (50.8%) | 28 (59.6%) | |
| Kaplan-Meier estimates, Median (95% CI), weeks | 48.7 (18.9, NA) | 12.9 (12.0, 33.0) | 1.38 (0.89, 2.12) |
| Log-rank test | | | $X_1^2$=4.9, p=0.0262 |

CI, Confidence interval; CPC, Clinical and patient-centric
[a] Restricted mean with upper limit of 52 weeks

4.2.4.2 Safety and tolerability

This section provides the main safety evidence, as reported by the company, for all patients who received study treatment within the ADMIRE-CD trial (safety population). Additional safety data were also reported from a Phase I/IIa study.[8]

The CS[1] (page 54) states that darvadstrocel is well tolerated, with an AE profile similar to control treatment (CS,[1] Tables 20, 21 and 22), although no test was performed to determine whether there was a statistically significant difference between trial arms for any specific AE. The majority of the data were for AEs up to week 24 of the ADMIRE-CD trial,[7] although some longer-term, 52-week safety data were also provided.[9] The CS,[1] published papers[7, 9] and clinical study reports[21, 22] reported treatment-emergent AEs (TEAEs), defined as any AE reported during the trial; treatment-related TEAEs, defined as '*events with relationship certain, probable or possible with the study treatment …*';[21] serious AEs (TESAEs), defined as '*events that threaten patient life or functions*';[21] and severe TEAEs, defined as an event that '*causes a significant interference with function*'.[21]

In the ADMIRE-CD trial at 24 weeks, TEAEs were common (66.0% of patients in the darvadstrocel arm compared with 64.7% in the control arm, see Table 9). The most common treatment-related TEAEs were proctalgia (12.6% of patients in the darvadstrocel arm versus 11.8% in the control arm), anal abscess (11.7% versus 12.7%) and nasopharyngitis (9.7% versus 4.9%), respectively. Diarrhoea was also more frequent in the darvadstrocel arm (6.8%) compared with the control arm (2.9%). The reported frequency of most TEAEs in patients was similar between the darvadstrocel and control arms of the ADMIRE-CD trial at 24 weeks. In many instances, the reported frequency of AEs was higher in the placebo arm than the treatment arm. This is because, as acknowledged in the CS[1] and reported in the EPAR,[17] some AEs, including anal abscess and proctalgia, are associated with the indication and might represent treatment failure, i.e. a lack of efficacy, rather than an AE related to the treatment (CS,[1] page 54). This explains why, for example, after 24 weeks, fewer patients treated with darvadstrocel compared with control experienced treatment-related TEAEs (17.5% of patients receiving darvadstrocel versus 29.4% receiving control, see Table 9), and why the reported frequency of withdrawal from the trial to due TEAEs was similar between arms (4.9% of patients receiving darvadstrocel versus 5.9% receiving control). Clinical advice received by the ERG indicated that such outcomes should have been treated as efficacy outcomes rather than AEs.

The CS[1] (Table 23, page 58) also reported so-called procedure-emergent, non-treatment emergent events (PENTE) for ≥2 patients up to week 24 in the ADMIRE-CD trial. These events are defined as AEs 'starting prior to administration of study treatment, but after [the] curettage procedure'.[21] None of these specific events were reported to affect more than [redacted] in any treatment arm, and only [redacted]

[redacted] were more frequent in the darvadstrocel arm compared with the control arm.

**Table 9: TEAEs, treatment-related TEAEs (≥10 patients), severe TEAEs and TESAEs up to week 24 in ≥2 patients in either treatment group, of ADMIRE-CD, safety population (adapted from CS,[1] Table 21)**

| Number patients (%) | TEAE | | Treatment related TEAE | | Severe TEAE | | TESAE | |
|---|---|---|---|---|---|---|---|---|
| | Darvadstrocel N=103 | Control N=102 | Darvadstrocel N=103 | Control N=102 | Darvadstrocel N=103 | Control N=102 | Darvadstrocel N=103 | Control N=102 |
| Number of patients | 68 (66%) | 66 (64.7%) | 18 (17.5%) | 30 (29.4%) | [redacted] | [redacted] | 18 (17.5%) | 14 (13.7%) |
| Withdrawals due to AE | 5 (4.9%) | 6 (5.9%) | | | | | [redacted] | [redacted] |
| Gastrointestinal disorders | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Proctalgia | 13 (12.6%) | 11 (10.8%) | 5 (4.9%) | 9 (8.8%) | [redacted] | [redacted] | | |
| Anal fistula | 3 (3%)[c] | 6 (6%)[c] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Infections and Infestations | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Anal abscess | 12 (11.7%) | 13 (12.7%) | 6 (5.8%) | 9 (8.8%) | [redacted] | [redacted] | 9 (8.7%)[d] | 7 (6.9%)[d] |
| Nasopharyngitis | 10 (9.7%) | 5 (4.9%) | | | | | | |
| General disorders and administration site conditions | [redacted] | [redacted] | [redacted] | [redacted] | | | [redacted] | [redacted] |
| Musculoskeletal and connective tissue disorders | [redacted] | [redacted] | [redacted] | [redacted] | | | | |

CSR, Clinical study report; TEAE, Treatment-emergent adverse event; TESAE, Treatment-emergent serious adverse event

[redacted] b-[redacted] d-Treatment-related TESAEs: 5% in both arms (Panes *et al.*)[7] and [redacted] ‡

[redacted] (see Table 10). The latter increased to [redacted] arm at 52 weeks.[22] There were no deaths recorded due to AEs or during the trial. The outcomes from TEAEs and TESAEs were similar across trial arms, although there was [redacted], where there was recovery, there was a [redacted] (see Table 10).

**Table 10: Summary of treatment-emergent adverse events and treatment-emergent serious adverse events up to week 24 in ADMIRE-CD, safety population[7, 21] (reproduced from CS,[1] Table 20)**

| | TEAE | | TESAEs | |
|---|---|---|---|---|
| | **Darvadstrocel N=103** | **Control N=102** | **Darvadstrocel N=103** | **Control N=102** |
| | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| **TEAEs/TESAEs** | 68 (66.0%) | 66 (64.7%) | 18 (17.5%) | 14 (13.7%) |
| **Intensity of TEAEs** | | | | |
| Mild | [redacted] | [redacted] | [redacted] | [redacted] |
| Moderate | [redacted] | [redacted] | [redacted] | [redacted] |
| Severe | [redacted] | [redacted] | [redacted] | [redacted] |
| Missing | [redacted] | [redacted] | [redacted] | [redacted] |
| **Outcome of TEAEs/TESAEs** | | | | |
| Death | [redacted] | [redacted] | [redacted] | [redacted] |
| Not recovered | [redacted] | [redacted] | [redacted] | [redacted] |
| Recovered with sequelae | [redacted] | [redacted] | [redacted] | [redacted] |
| Recovered without sequelae | [redacted] | [redacted] | [redacted] | [redacted] |
| Changed intensity | [redacted] | [redacted] | [redacted] | [redacted] |
| Unknown | [redacted] | [redacted] | [redacted] | [redacted] |

TEAE, Treatment-emergent adverse event; TESAE, Treatment-emergent serious adverse event

Safety data for 52 weeks follow-up have been published[9] and are reported in the CS[1] (Table 22, page 58). As with the 24-week data, the frequency of patients with some TEAEs was similar across the treatment and control arms, but the trend changed at 52 weeks for key AEs such as anal abscess (19.4% of patients in the treatment arm versus 13.7% in the control arm) and anal fistula (10.7% versus 7.8%) and nasopharyngitis (10.7% versus 4.9%), with higher frequencies of patients affected in the darvadstrocel arm than the control arm (see Table 11). This trend was the same for the TESAEs of anal abscess (13.6% of patients in the treatment arm versus 7.8% in the control arm) and anal fistula (3.9% versus <1.0%).

For the 52-week data, compared with the 24-week data, there were higher frequencies of patients with TEAEs, treatment-related TEAEs and TESAEs. For example, for darvadstrocel 76.7% of patients experienced a TEAE by week 52 compared to 66.0% of people experiencing a TEAE by week 24. Equivalently for standard care, 72.5% of patients experienced a TEAE by week 52 compared to 64.7% for week 24. The withdrawals due to TEAEs also increased over time (darvadstrocel. 8.7% of patients for week 52 versus 4.9% for week 24; standard care 8.8% versus 5.9%). The ERG noted that there was a sizeable increase in the frequency of patients with TESAEs at 52 weeks across arms compared with week 24 (24.3% at week 52 versus 17.5% at week 24 for darvadstrocel, and 20.6% at week 52 versus 13.7% at week 24 for control). The CS[1] (page 59) also reported that there were no immune reactions or TEAEs associated with the development of donor-specific antibodies, and no association between positivity for donor-specific antibodies and therapeutic response.

**Table 11: Longer-term safety from ADMIRE-CD, safety population, ≥4 patients, (adapted from CS,[1] Table 22, with data from Panes et al,[7, 9] and TiGenix Clinical Study Reports[21, 22])**

| Number patients (%) | Week 24 | | Week 52[22] | |
|---|---|---|---|---|
| | Darvadstrocel N=103 | Control N=102 | Darvadstrocel N=103 | Control N=102 |
| **TEAEs** | 68 (66.0%) | 66 (64.7%) | 79 (76.7%) | 74 (72.5%) |
| Proctalgia | 13 (12.6%) | 11 (10.8%) | 15 (14.6%) | 12 (11.8%) |
| Anal abscess | 12 (11.7%) | 13 (12.7%) | 20 (19.4%)[a] | 14 (13.7%)[a] |
| Anal fistula | 3 (3%) | 6 (6%) | 11 (10.7%)[a] | 8 (7.8%)[a] |
| Nasopharyngitis | 10 (9.7%) | 5 (4.9%) | 11 (10.7%) | 5 (4.9%) |
| **Treatment-related TEAEs** | 18 (17.5%) | 30 (29.4%) | 21 (20.4%) | 27 (26.5%) |
| Withdrawn due to AEs | 5 (4.9%) | 6 (5.9%) | 9 (8.7%) | 9 (8.8%) |
| **Treatment-related AEs in ≥5% of patients** | | | | |
| Anal abscess | 6 (5.8%) | 9 (8.8%) | [redacted][b] | [redacted][b] |
| Anal fistula | [redacted] | [redacted] | [redacted][b] | [redacted][b] |
| Proctalgia | 5 (4.9%) | 9 (8.8%) | 5 (4.9%) | 8 (7.8%) |
| **Serious TEAEs** | 18 (17.5%) | 14 (13.7%)c | 25 (24.3%) | 21 (20.6%) |
| Anal abscess | 9 (8.7%) | 7 (6.9%) | 14 (13.6%)[a] | 8 (7.8%)[a] |
| Anal fistula | [redacted] | [redacted] | 4 (3.9%)[a] | 1 (<1.0%)[a] |
| **Treatment-related TESAEs in ≥2% of patients** | | | | |
| **TESAEs** | 5 (5%)[d] | 7 (7%)[d] | 7 (6.8%) | 7 (6.9%) |
| Anal abscess/fistula | 5 (5%)[d] | 5 (5%)[d] | 7 (6.8%)[e] | 5 (4.9%)[e] |

AE, Adverse event; TEAE, Treatment-emergent adverse event

a - TiGenix Clinical Study Report[22] and EPAR[17]
b - Unpublished data TiGenix Clinical Study Report[22]
c -Erroneously reported in CS, Table 22 as n=6 (5.9%).
d -Panes *et al*.[7]
e - Unpublished data TiGenix Clinical Study Report[22] on anal abscess only [redacted]

The frequency of AEs in the ADMIRE-CD trial was generally similar to that reported for an earlier Phase I/II trial,[8] which also had 24-week follow-up (CS, Appendix F). However, the frequency was lower for 24-week data for the ADMIRE-CD trial for some events (see Table 11). For example, the frequency of patients with treatment-related TEAEs at 24 weeks in ADMIRE-CD was 4.9% (5/103) compared with 21% (5/24) in the Phase I/II trial; the frequency of patients with treatment-related anal abscess was 5.8% (6/103) in the ADMIRE-CD trial compared with 12.5% (3/24) in the Phase I/II trial; and the frequency of TESAEs was 5% (5/103) in ADMIRE-CD compared with 8% (2/24) in the Phase I/II trial (CS,[1] Appendix F). It is not clear why the reported frequency of patients with treatment-related TEAEs in particular was lower in the ADMIRE-CD trial compared with the earlier Phase I/II trial. However, this might be explained by the lower dose of darvadstrocel in the Phase I/II trial, i.e. up to a maximum of 60 million expanded adipose-derived allogeneic mesenchymal stem cells (eASCs)[8] compared with 120 million cells in the ADMIRE-CD trial[7] and the potential for these key AEs to be considered as efficacy rather than safety outcomes.

In summary, TEAEs were common and were reported by approximately two-thirds of patients receiving darvadstrocel. The most common TEAEs were proctalgia, anal abscess, nasopharyngitis and diarrhoea.

The frequency of the principal TEAEs was generally similar across the treatment and control arms, but the ERG noted that proctalgia, anal abscess and anal fistulae are symptomatic of the indication in this appraisal and are therefore indicative of treatment failure rather than being treatment-related AEs. The ERG also noted that the percentages of TEAEs, treatment-related TEAEs, TESAEs, and withdrawals due to treatment-related TEAEs among patients in the darvadstrocel arm were all higher at 52 weeks than at 24 weeks. It was also the case that the percentages of patients experiencing key TEAEs, previously similar between arms at 24 weeks, had become noticeably higher in the darvadstrocel arm than the control arm of the trial at 52 weeks. The ERG also noted that the frequency of treatment-related TEAEs at 24 weeks was higher in the earlier phase I/II trial[8] than the later ADMIRE-CD trial,[7, 9] which might be explained by the much lower dose of darvadstrocel in the earlier trial (≤60 million eASCs versus 120 million eASC) and the issue that some AEs might represent a lack of efficacy rather than AEs.

### 4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

No indirect comparison was undertaken by the company to supplement the direct evidence as there is only one trial that has evaluated the use of darvadstrocel in the treatment of non-active/mildly active luminal Crohn's disease, with complex perianal fistulas which have shown an inadequate response to at least one conventional or biologic therapy (CS,[1] Section B.29, pages 53-54). The ERG agrees with this position, which is in line with the final scope issued by NICE.

### 4.4 Critique of the indirect comparison and/or multiple treatment comparison

No indirect comparison was undertaken by the company (see Section 4.3).

### 4.5 Additional work on clinical effectiveness undertaken by the ERG

As the company undertook a reasonably comprehensive systematic review (no major limitations were noted) of darvadstrocel for treating complex perianal fistula in Crohn's disease, no additional work was undertaken by the ERG.

### 4.6 Conclusions of the clinical effectiveness section

#### *4.6.1 Completeness of the CS with regard to relevant clinical studies and relevant data within those studies*

The clinical evidence in the CS[1] is based on a systematic review of darvadstrocel for the treatment of complex perianal fistulae in patients with Crohn's disease. The ERG is confident that all relevant controlled trials (published and unpublished) were included in the CS,[1] including data from ongoing/planned studies.

*4.6.2 Interpretation of treatment effects reported in the CS in relation to relevant population, interventions, comparator and outcomes*

A key limitation of the efficacy and safety data for darvadstrocel reported in the CS[1] relates to the *post hoc* analyses of CPC-remission (an outcome used in the economic model) and CPC relapse. These endpoints were not designed or powered to test formal hypotheses. As such, these results should be treated with caution. Another issue is the lack of a confirmatory study. As noted in the EPAR,[17] the effect size in the ADMIRE-CD trial was considered to be modest and less than the 25 percentage difference that it was designed to detect, yet this was considered clinically meaningful given that other treatment options for fistulas had failed. A post-authorisation efficacy and safety trial, ADMIRE-CD-II[18] is expected to help address this concern. This study is similar to the ADMIRE-CD study in that it is a Phase-III, randomised, double-blind, parallel-group, placebo-controlled, multicentre study evaluating the efficacy and safety of darvadstrocel compared to placebo for the treatment of complex perianal fistula(s) in patients with Crohn's disease at week 24 with a follow-up period up to 52 weeks. This study is being conducted to include patients from the USA and to satisfy FDA licensing requirements. However, the study is expected to complete in October 2021 and the final clinical study report to the EMA is expected in 2022.[17]

*4.6.3 Uncertainties surrounding the reliability of the clinical effectiveness*

The key uncertainties in the clinical evidence for darvadstrocel relate to repeated administration, optimal dosing and long-term efficacy and safety. Further details are provided below.

- *Repeated administration*

The EPAR[17] states that '*While treatment with Alofisel [darvadstrocel] is proposed for single dose administration, the need for repeated treatment in the clinical setting seems foreseeable in the targeted patient population*'. The company's clarification response to question A1[2] suggest that '*Although some clinicians believe that Alofisel [darvadstrocel] may be beneficial for retreatment in the following patient groups; (i) partial responders; (ii) responders who have relapsed, there is no current evidence to support this treatment approach… therefore elected to base the submission on single use only. Some patients who have responded to Alofisel treatment and achieved healing over a significant period of time may develop a new fistula tract (recurrence). We believe this should be considered as a new fistula and should therefore be treated as such*.' The ERG notes that although darvadstrocel offers a novel treatment option with curative intent, there are no robust supporting data beyond 52 weeks follow-up; there is no evidence on the repeated use of darvadstrocel (licensed dose) when new fistulas open and it is unclear whether patients who have not achieved complete closure with one injection would benefit from an additional injection.

- *Optimal dosing*

In the ADMIRE-CD study,[7, 9] patients with complex perianal fistulising (maximum of 2 internal openings and a maximum of 3 external openings) Crohn´s disease who were refractory to conventional (antibiotics, immunosuppressants) or biological treatment strategies received a single intralesional injection containing 120 million darvadstrocel cells. Although no formal dose finding studies have been conducted (see clarification response,[2] question A6), it remains unclear whether alternative dosage regimens may have been clinically effective with fewer AEs or whether stem cell therapy would be effective in patients with very complicated perianal fistulising disease who may have more than two internal and three external openings (see clarification response,[2] question A4).

- *Long-term efficacy and safety*

In the ADMIRE-CD,[7, 9] the follow-up was extended from 24 weeks to 104 weeks to allow for the assessment of long-term efficacy and clinical and immunological safety of darvadstrocel treatment. However, as noted in the CS[1] (page 64), the available efficacy data beyond 52 weeks were limited because the protocol change occurred when various patients had already finished the 52 week trial period. The CS states *'...This resulted in a low level of patient data, and so generalisation of results beyond 52 weeks is difficult and should be approached with care*'. As a result, there is uncertainty regarding the long-term efficacy and safety of darvadstrocel. The SmPC[6] and EPAR[17] for darvadstrocel also advise for monitoring and reporting of any suspected adverse reactions after authorisation for signs of infection after administration and immunogenicity/ all-immunoreactions.

# 5 COST EFFECTIVENESS

## 5.1 ERG's comment on company's review of cost-effectiveness evidence

### *5.1.1 Objective of cost effectiveness review*

The company performed two broad searches. The first search was undertaken to identify economic evaluations, resource use and costing studies in Crohn's disease and people with perianal fistulas. Terms for Crohn's disease were combined with a cost-effectiveness filter (CS,[1] Appendix G). The following sources were searched: MEDLINE [via Embase.com], MEDLINE In-Process [via PubMed], Embase [via Embase.com] NHS EED [via Wiley Online Library] and EconLit [via AEAweb.org]. Supplementary searches in Research papers in Economics (RePEC) and the cost-effectiveness analysis (CEA) Registry were carried out by the company to identify further resource use and cost data studies in people with perianal fistulas and Crohn's disease (CS,[1] Appendix I). The search covered the period from 2000 up to 22 January 2018.

The second search was undertaken to identify HRQoL studies in Crohn's disease where terms for the disease were combined with a QoL filter. Full details of the searches carried out in MEDLINE [via Embase.com], MEDLINE In-Process [via PubMed], Embase [via Ovid] and NHS EED [via Wiley Online Library] are presented in the CS (Appendix H).[1] Supplementary searches included searching several online websites: Tufts CEA Registry database, NICE and School of Health and Related Research Health Utilities Database (ScHARR HUD). The search covered the period from 2000 up to 22 January 2018.

The ERG considers that the searches were fully reported in the CS (Appendices G, H and I) that they were sufficiently comprehensive.[1] There were no studies that the ERG or their clinical advisors were aware of that were missed.

### *5.1.2 The inclusion and exclusion criteria used in the study selection*

The inclusion criteria for the systematic review of the cost-effectiveness evidence is briefly summarised

Table 12. It is unclear why the company applied intervention criterion in the inclusion criteria, as the objective of the review was to identify relevant cost-effectiveness studies in the same disease area. However, as the inclusion criteria cover most relevant interventions for people with Crohn's disease and complex perianal fistulae, it is unlikely that any relevant studies will have been missed.

**Table 12: Inclusion criteria used in the company's review of cost-effectiveness evidence (reproduced from CS,[1] Appendix G, Table 18)**

| Studies to include | |
|---|---|
| Study Design | • Cost studies/surveys/analyses<br>• Cost/economic burden of illness<br>• Resource use studies<br>• Cost-effectiveness analyses<br>• Cost-utility analyses<br>• Cost-benefit analyses<br>• Cost-minimization analyses<br>• All economic evaluation studies based on models<br>• Budget impact models<br>• Database analyses with cost |
| Population | • Patients with perianal fistula in Crohn's disease<br>• No age, gender or race restriction |
| Intervention/ Comparator | • Cx601/darvadstrocel<br>• Ciprofloxacin<br>• Infliximab<br>• Adalimumab<br>• Certolizumab<br>• Fibrin glue<br>• Metronidazole<br>• Azathioprine<br>• 6-MP<br>• Cyclosporine<br>• Tacrolimus<br>• Methotrexate<br>• Thalidomide<br>• Surgery (fibrin glue, advancement flap, LIFT, diverting stoma, proctectomy, colectomy, fistula plugs, fistulotomy, exam under anaesthesia, multiple seton placement, ileostomy, colostomy, stem cells, VAAFT and Filac) |
| Language | English only |
| Country | No restriction |
| Publication timeframe | 2000-2018 |

LIFT - Ligation of the inter-sphincteric fistula tract; VAAFT - Video assisted anal fistula treatment

### *5.1.3 Findings of the cost-effectiveness review*

Following de-duplication, the company's searches found 335 publications. Two hundred and fifty six publications were excluded at the abstract review and a further 72 publications were excluded at the full text stage. This left seven remaining publications. A further two publications were identified through searching of conference records and bibliographic searching. In total, nine publications (reporting on seven studies) were identified; two of these studies reported cost-utility analyses. Only one study, by Lindsay *et al*, related to a UK health care setting.[23] Lindsay *et al*. assessed the cost-effectiveness of infliximab versus standard care for luminal and fistulising Crohn's disease patients in England and

Wales. Whilst a useful source of information, this study was not directly relevant to the cost-effectiveness of darvadstrocel compared with standard care.

#### *5.1.4 Conclusions of the cost-effectiveness review*

The CS concludes that the existing evidence is insufficient to determine the cost-effectiveness of darvadstrocel as a specific treatment for complex perianal fistula in Crohn's disease patients, as the previous model examined a patient population treated for both luminal and fistulising Crohn's disease.[1] As such, it was necessary to develop a *de novo* model for this appraisal. The ERG agrees with this conclusion.

### 5.2 Summary of the company's submitted health economic analysis

#### *5.2.1 Population*

The population included in the company's health economic analysis reflects people with complex perianal fistulae and Crohn's disease who have two or less internal openings and three or less external openings of their complex perianal fistula; are naïve to darvadstrocel treatment, and; are refractory to conventional first-line therapy. Failure of conventional first-line therapy was defined to consist of at least one of: no therapeutic effect of an antibiotic (recommended treatments were ciprofloxacin and metronidazole) after one month; no response to an immunosuppressant (azathioprine [2-2.5 mg/kg] or 6-mercaptopurine [1-1.5 mg/kg]) after three months, or; no response to an anti-TNF either 12 weeks after initiation of induction treatment or loss of response after 12 weeks of maintenance treatment under a stable dose.

#### *5.2.2 Interventions and comparators*

In the ADMIRE-CD study, four vials of darvadstrocel (total dose = 120 million cells) were administered as an intralesional injection during an EUA after the fistula had been conditioned. Conditioning of the fistula consisted of: an EUA; curetting (scraping anything out of) the fistula tract; and if indicated, setons (surgical cords used to open the fistula so that it drains) were placed during the EUA. If setons were placed whilst conditioning the fistula, they were removed immediately prior to the administration of darvadstrocel. Darvadstrocel injections were given in addition to standard care therapies for people who were already refractory to first-line treatment.

In the UK, standard care for people who are refractory to conventional therapy consists of at least one of the following options: surgically managing the fistula; antibiotics; immunosuppressants and/or, biologics. Whilst surgical treatments are similar between first- and second-line treatments, different antibiotics, immunosuppressants and/or biologics than the treatment which failed at first-line will typically be used.

If a patient does not respond to their initial treatment (either darvadstrocel or standard care) within one year or if the patient relapses after achieving remission of their fistula, they subsequently receive salvage therapy. Salvage therapy is similar to standard care in that one of the following treatments will be used: surgically managing the fistula; antibiotics; immunosuppressants and/or biologics. Typically, different medical management of the fistula will be undertaken (antibiotics, immunosuppressants and biologics) and possibly different surgical procedures will be considered. After several failed lines of salvage therapy, last resort surgeries are considered. These consist of defunctioning surgery, in which the fistula is temporarily bypassed to allow healing, and proctectomy, in which a proportion of the bowel is permanently bypassed.

### *5.2.3 Perspective, time horizon and discounting*

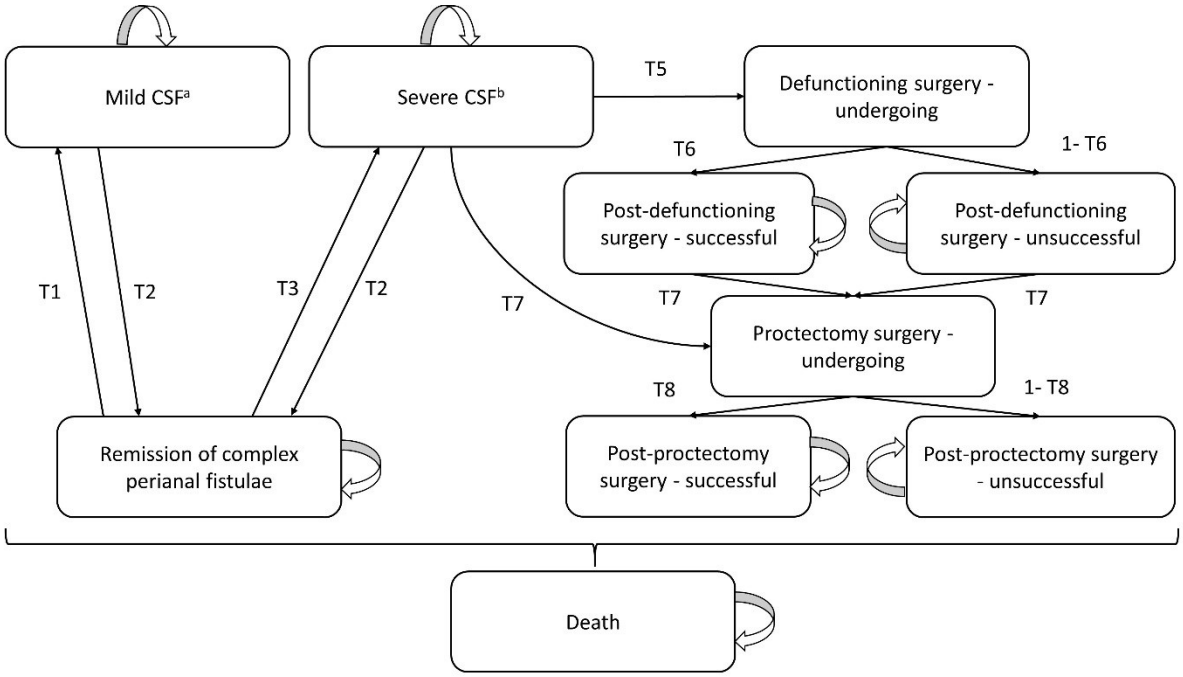
The base case model adopts an NHS and Personal Social Services (PSS) perspective. The time horizon of the base case model was 40 years from the model start. Costs and QALYs were discounted at 3.5% and at 1.5% respectively, as the company states that "*It was considered that a non-reference discount rate of 1.5% per annum for health outcomes was applicable, as darvadstrocel demonstrates long term healing potential in this population with a significant impact on QoL … as per the NICE methods guide.*"(CS,[1] page 74) The ERG notes that section 6.2.19 of the NICE methods guide states that "*In cases when treatment restores people who would otherwise die or have a very severely impaired life to full or near full health, and when this is sustained over a very long period (normally at least 30 years), cost-effectiveness analyses are very sensitive to the discount rate used. In this circumstance, analyses that use a non-reference-case discount rate for costs and outcomes may be considered. A discount rate of 1.5% for costs and benefits may be considered by the Appraisal Committee if it is highly likely that, on the basis of the evidence presented, the long-term health benefits are likely to be achieved. Further, the Appraisal Committee will need to be satisfied that the introduction of the technology does not commit the NHS to significant irrecoverable costs.*"(page 66 -67).[10] This means that the originally presented analyses in the CS are out of scope, as the NICE methods guide does not advocate differential discounting of costs and QALYs, even if these criteria are met.[1, 10] Two sets of in scope analyses, one using discount rates of 1.5% for both costs and QALYs and the other 3.5% for both costs and QALYs were provided in the company's clarification response (question B7).[2] The in scope analyses are the focus of the ERG's summary of the company's submitted analyses (see Sections 5.2.7 and 5.2.8).

### *5.2.4 Model structure*

The company's model adopts a state transition approach and is constructed in Microsoft Excel® (see Figure 2). The model includes eight main health states: (1) mild chronic symptomatic complex perianal fistulae (CSF); (2) severe CSF; (3) remission; (4) defunctioning surgery (cycle 1); (5) defunctioning surgery (subsequent cycles); (6) proctectomy (cycle 1); (7) proctectomy (subsequent cycles) and (8) death. Patients with mild or severe CSF (model states 1 or 2) experience AEs (abscesses and proctalgia)

which are dependent on treatment and the severity of their CSF. The defunctioning surgery (subsequent cycles) and the proctectomy surgery (subsequent cycles) were both split into successful and unsuccessful surgeries. Transitions to the proctectomy and defunctioning surgery health states are assumed to not be possible from either the remission or the mild CSF health states. Patients who have had defunctioning surgery are not able to have this reversed in the model; as such, the only possible transitions from the defunctioning surgery states are to proctectomy or death.

**Figure 2: Model diagram (adapted from CS,[1] Figure 17)**




a – 40.1% of people start the model in this health state; b – 59.9% of people start the model in this health state; T1 – time to relapse * probability that a CSF is mild; T2– time to remission; T3 - time to relapse * (1- probability that a CSF is mild); T5 – time to defunctioning surgery; T6 – probability that a defunctioning surgery is successful; T7 – time to proctectomy; T8 – probability that a proctectomy is successful

Patients enter the model in either one of the two CSF health states (40.1% mild, 59.9% severe) at a mean age of 38.27 years. Health state transitions are estimated over 520 4-weekly cycles (approximately 40 years); at this time point, only 31.7% of patients in each treatment group have died. The treatment-specific transitions from the CSF mild (state 1) and CSF severe health states (state 2) to the remission health state (state 3) are based on the same parametric model (Gompertz distribution) fitted to the CPC remission outcome from the ADMIRE-CD.[1] The treatment-specific transitions to the CSF mild (state 1) and CSF severe health states (state 2) from the remission health state (state 3) are based on the same parametric model (Gompertz distribution) fitted to the CPC relapse outcome from the ADMIRE-CD trial and the probability that a CSF is mild.[1] The Gompertz distributions are different in the darvadstrocel and standard care groups, as a treatment effect covariate (HR) is estimated for both the

time to relapse and time to remission Gompertz distributions. After one completed line of either darvadstrocel or standard care (defined as achieving remission or remaining in the CSF health state for more than 13 model cycles), patients go on to receive salvage therapy in both arms. To estimate the time to remission and relapse for people who have received salvage therapy, these transitions are estimated by applying a HR based on an expert elicitation exercise to the respective time to event function for patients receiving standard care. The probability that a CSF is mild is estimated from the ADMIRE-CD trial data.[1] Transitions to the defunctioning surgery state were based on a parametric model (exponential distribution) fitted to digitised individual-level patient data (IPD) from a subgroup of people with a complex perianal fistulae in a prospective cohort study on surgical outcomes in people with perianal fistulae and Crohn's disease by Mueller *et al*.[24] The digitised IPD were reconstructed from the Kaplan-Meier time-to-event function using the Guyot *et al*[25] method. Transitions to the proctectomy state were based on an analysis of the St Mark's retrospective dataset.[1] The St Mark's retrospective data set is retrospective cohort study of 78 consecutive patients who presented at St Mark's hospital, London, with complex perianal fistulae in Crohn's disease between January 1st 2008 and July 1st 2017. Transitions to the death state from all states are based on general population life tables.[26]

5.2.4.1 Modelling HRQoL impacts

The model assumes that HRQoL is principally determined by time spent in each health state and therefore the patient's HRQoL is driven by time to remission, time to relapse and the timing of defunctioning surgery or proctectomy. Whilst patients were receiving darvadstrocel, standard care or salvage therapy, utility decrements for the incidence of TEAEs were applied, resulting in different HRQoL in the mild CSF (state 1) and severe CSF (state 2) health states across the three treatment groups. The HRQoL effects associated with each health state are not age-adjusted.

5.2.4.2 Modelled treatment pathway and associated costs

The company's model includes the following cost components: (1) drug acquisition; (2) drug administration; (3) TEAEs, and (4) health state resource use (hospital visits and tests). The only differences in the model pathways between the darvadstrocel and standard care arms are that in the initial CSF health states (either mild or severe). Patients in the darvadstrocel arm receive a single course of darvadstrocel in addition to the standard care treatments; therefore patients in the darvadstrocel arm receive different time to remission and time to relapse functions which influence the transitions to and from the remission health state (state 3). Consequently, this leads to a differences in the amount of time spent at risk of receiving defunctioning (state 5) or protectomy surgery (state 7) between the treatment groups. Patients experience different rates of TEAEs in the two treatment groups. Upon the first relapse (transition from remission (state 3) to a CSF health state (state 1 or state 2)), patients in both the standard care and darvadstrocel arms are assumed to receive salvage therapy.

Within the standard care group, the model assumes the following treatment pathway:

- The average patient receives surgical (EUA and/or seton placement) and medical management for their complex perianal fistula. The exact treatments used for surgical and medical management are based on data from the ADMIRE-CD study.

Within the darvadstrocel group, the model assumes the following treatment pathway:

- All patients receive a single course of four vials (120 million cells) of darvadstrocel within the first cycle (four weeks).
- Darvadstrocel is administered using one additional EUAs compared to standard care (two EUAs in total). The first EUA is used to condition the fistula and the second to administer darvadstrocel.
- Patients receiving darvadstrocel also receive the same medical management of their fistula as people in the standard care group
- Upon relapse, no further administrations of darvadstrocel are given.

Upon relapse, all patients in both groups receive salvage therapy. This consists of surgical and medical management. The exact treatments used for surgical and medical management are different from the standard care group and are based on expert clinical opinion.

### *5.2.5 Key structural assumptions employed within the company's model*

The company's model employs the following structural assumptions:

- All patients enter the model in either the mild active complex perianal fistula health state or the severe complex perianal fistula health state.
- HRQoL is principally determined by time spent in remission (state 3) and CSF (state 1 and state 2), post-defunctioning surgery (state 5) and post-proctectomy (state 6) health states.
- All darvadstrocel administration is completed within the first model time cycle (4 weeks).
- The hazard rate for time to remission is assumed to follow a Gompertz distribution in the darvadstrocel, standard care and salvage therapy groups.
- The hazard rate for time to relapse is assumed to follow a Gompertz distribution in the darvadstrocel, standard care and salvage therapy groups.
- Patients are only eligible to receive one line of treatment (i.e. darvadstrocel or standard care); following relapse, patients are assumed to receive salvage therapy.
- Patients who do not achieve remission within one year of treatment with either darvadstrocel or standard care are assumed to receive salvage therapy.
- The probabilities of undergoing proctectomy and defunctioning surgery are assumed to be constant with respect to time.

- It is only possible to enter the defunctioning surgery health state from the severe CSF health state.
- It is only possible to enter the proctectomy surgery health state from either the severe CSF health state or either of the post-defunctioning surgery health states.
- It is not possible for a proctectomy or a defunctioning surgery to be reversed.
- It is not possible for a successful proctectomy to become unsuccessful or *vice versa.*
- It is not possible for a successful defunctioning surgery to become unsuccessful or *vice versa.*

The structural assumptions in the company's model are commented on by the ERG in the critical appraisal section (see Section 5.3.4.8)

### *5.2.6 Evidence used to inform the company's model parameters*

The evidence sources used to inform the model parameters are summarised in

Table 13. These are discussed in further detail in the subsequent sections.

**Table 13: Evidence sources used to inform the company's model parameters**

| Parameter type | Parameter | Source(s) |
|---|---|---|
| Time-to-event parameters | Remission – darvadstrocel | CPC definition of remission in the ADMIRE-CD trial[1] |
| | Remission – standard care | |
| | Relapse – darvadstrocel | CPC definition of relapse in the ADMIRE-CD trial[1] |
| | Relapse – standard care | |
| | Remission – HR of salvage therapy versus standard care | Company's expert elicitation exercise[1] |
| | Relapse – HR of salvage therapy versus standard care | |
| | Time to defunctioning surgery | Mueller *et a l*prospective cohort study[24] |
| | Receiving a proctectomy surgery | Bell *et al.*prospective study[27] |
| Time independent probabilities | Probability complex perianal is mild | ADMIRE-CD trial[1] |
| | Probability proctectomy is successful | St Mark's retrospective study[1] |
| | Probability defunctioning surgery is successful | St Mark's retrospective study[1] |
| Mortality | Age-dependent probability of death | ONS[26] |
| HRQoL | Health utility – all model health states | Vignette study[1] |
| | Disutility associated with abscesses | Vignette study[1] |
| | Disutility associated with proctalgia | Assumption[1] |
| Resource use and costs | Health state related inpatient, outpatient resource use and associated costs | Expert opinion,[1] NHS Reference Costs 2016-17,[28] PSSRU,[29] NICE TA 329,[30] NICE DG11[31] |
| | Darvadstrocel acquisition cost (including PAS) | Company[1] |
| | Frequency of use for different surgical and drug treatments for complex perianal fistulae | ADMIRE-CD trial,[1] expert opinion[1] |
| | Unit costs of surgical procedures used to treat complex perianal fistulae | NICE MIB 102,[32] NICE MIB 105,[33] NHS Reference Costs 2016-17[28] |
| | Unit costs and dosing related to drug treatments | BNF,[34] SmPC,[6] NICE TA187[35] |

5.2.6.1 Time-to-event analyses

*CPC definition of remission*

The company fitted parametric survival functions to time-to- remission data from the ADMIRE-CD trial. In the company's base case analysis, remission was defined as the interval from the date of treatment completion for darvadstrocel (four weeks post-randomisation) to the time of remission of the fistula, which was defined as the fistulae not draining when gently compressed and the patient reporting a PDAI score of 0 in the pain and discharge dimensions (CPC remission). Relapse was defined as the interval from achieving CPC remission to either the fistulae re-opening (determined by gentle finger compression) or the patient reporting a PDAI score of ≥ 1 in the pain or discharge dimensions.

The company fitted a range of standard parametric time to event distributions (exponential, Weibull, log normal, log logistic, generalised gamma, and Gompertz) to the data. The goodness-of-fit of each model was assessed using the methods detailed in NICE Decision Support Unit Technical Support

Document 14 (comparing Akaike information criterion [AIC] and Bayesian information criterion [BIC], and by visual assessment).[36] An assessment of the proportional hazards assumption was carried out only for the time to relapse functions, because the remission time-to-event functions for the darvadstrocel and standard care groups were not extrapolated beyond the 1-year follow-up data (CS,[1] page 79).It should be noted that when patients received salvage therapy, the time to remission function was extrapolated. An assessment of other plausible assumptions (e.g. accelerated failure time) were not conducted. In all analyses a treatment effect covariate (either a constant HR or constant acceleration factor, depending on the model type) was included in the statistical models to estimate the treatment effect parameter (the difference between the time-to-event for patients receiving darvadstrocel versus those receiving standard care). Piecewise exponential models were also fitted to the data, however the ERG notes that, it is unclear how these functions were fitted and which goodness-of-fit tests, if any, were conducted in these cases. The Gompertz distributions for time to remission and time to relapse were presented to the company's clinical experts to assess the clinical plausibility of the extrapolation (CS,[1] page 79).



Table 14 presents the AIC and BIC statistics for each of the fitted parametric time-to-event functions. These indicate that when the CPC definition of remission is used, the generalised gamma distribution provides the best fit to the observed time to remission data and the Gompertz distribution provides the best fit to the observed time to relapse data (although there is very little to distinguish between the Gompertz and the log normal models).

**Table 14: AIC and BIC statistics for time-to-event functions fitted to data on time to remission and relapse using the CPC definition of remission, excluding the piecewise exponential model (adapted from CS,[1] Tables 32 and 38)**

| | **Remission** | | **Relapse** | |
|---|---|---|---|---|
| | **AIC** | **BIC** | **AIC** | **BIC** |
| Exponential | 980.8393 | 987.4459 | 539.436 | 544.606 |
| Weibull | 965.6205 | 975.5305 | 528.702 | 536.457 |
| Gompertz | 946.2664 | 956.1763 | ***517.572*** | ***525.327*** |
| Log normal | 954.7821 | 964.6920 | 518.216 | 525.971 |
| Log logistic | 954.7821 | 964.6920 | 521.644 | 529.399 |
| Generalised gamma | ***931.1734*** | ***944.3866*** | 522.156[a] | 532.496[a] |

AIC – Akaike information criterion; BIC – Bayesian information criterion; a - the stacy parametrisation used for the generalised gamma rather than the default prentice parameterisation
Text in ***bold and italics*** indicates the lowest value out of the converged time-to-event functions in each column

The appropriateness of the proportional hazards assumption was assessed by examining the log cumulative hazard plot. The log cumulative hazard plot for CPC remission is presented in

Figure 3; this plot shows that the lines are approximately parallel and do not cross, thereby indicating that the proportional hazards assumption is not violated. The plot of the empirical hazard function and fitted hazard function is given in

Figure 4. In Figure 4 the solid black lines represent the empirical hazard, the solid coloured line represents the central estimate of the fitted hazard for each treatment group, and the dotted coloured lines represent the 95% CI around the fitted hazard. This plot shows that the empirical hazards stay within the confidence interval of the predicted hazard for the Gompertz curve, but not for the other parametric time to event functions.

**Figure 3: The log cumulative hazard plot for CPC remission data (reproduced from clarification response,[2] question B3)**

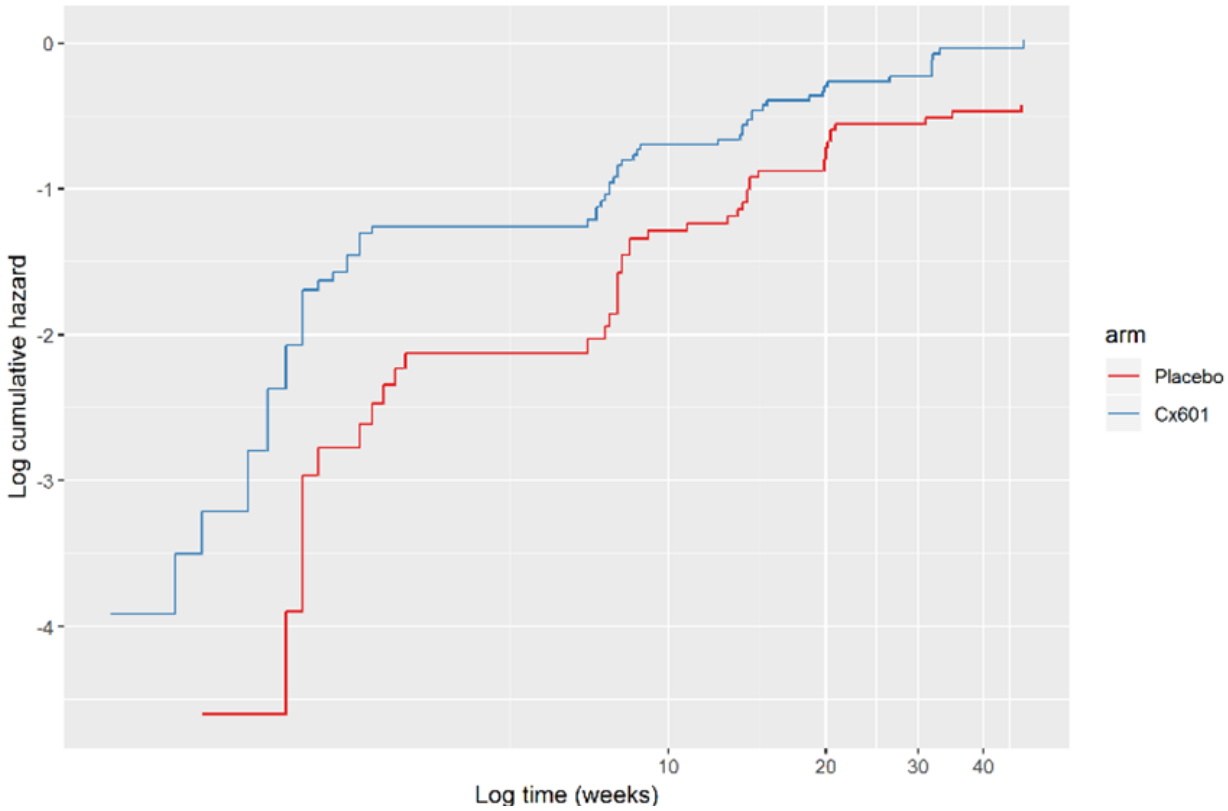


**Figure 4: Empirical versus predicted hazards for CPC remission (reproduced from clarification response,[2] question B3)**

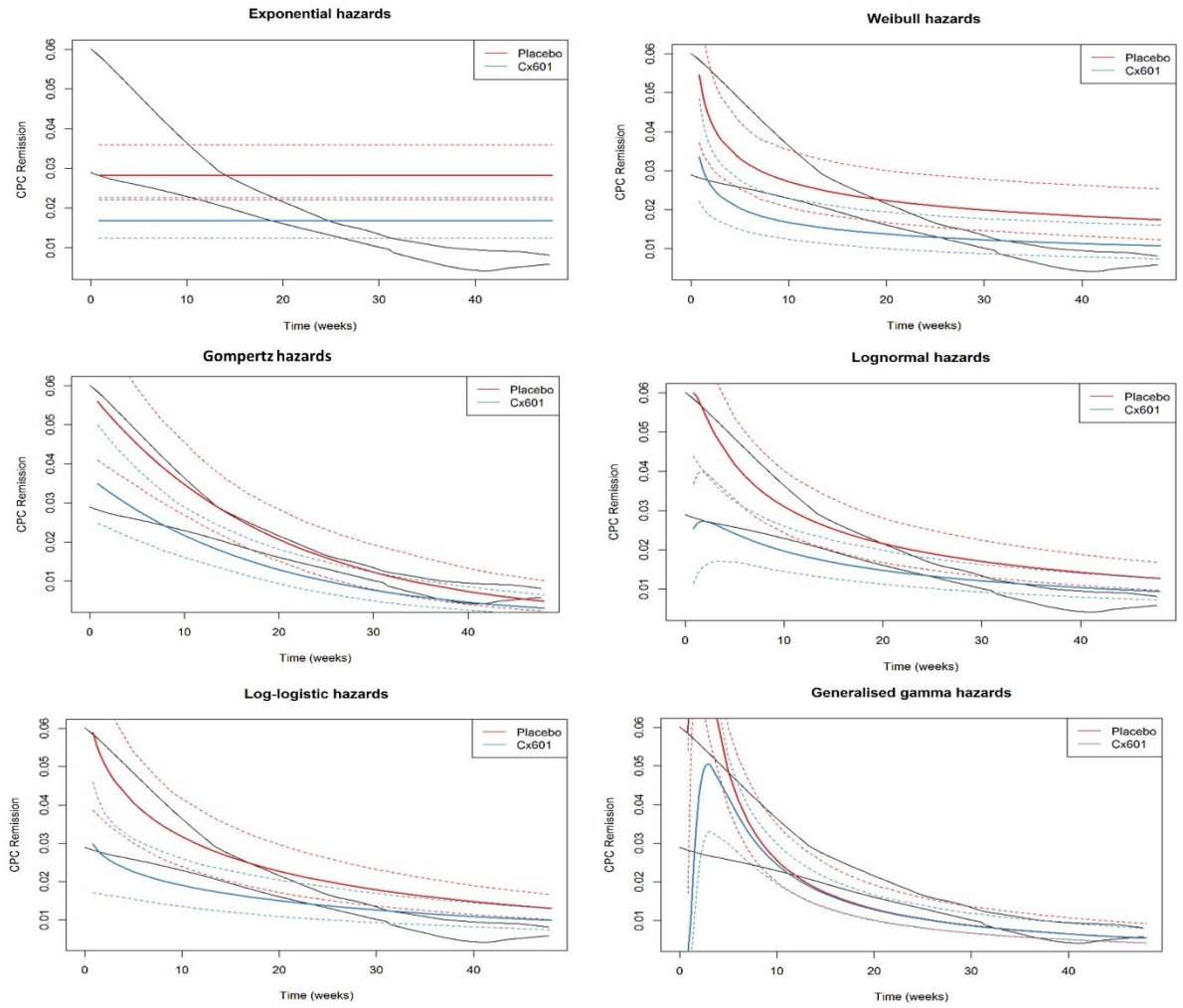



Figure 5 shows that for CPC relapse, the curves cross early and then separate at a later time point. The company states that the curves are approximately parallel in the medium to long-term. This indicates that the proportional hazards assumption for CPC relapse is likely to be inappropriate. A plot of the empirical and predicted hazard functions is given in

Figure 6; this shows that the shape of the empirical hazard function is not consistent with any of the fitted parametric curves across the full time period plotted, and that the Gompertz curve provides a reasonable fit up to around 40 weeks. The other curves fitted tend to over-predict the hazard in the darvadstrocel arm prior to 40 weeks.

**Figure 5: Log cumulative hazard plot for CPC remission relapse data (reproduced from clarification response,[2] question B3)**

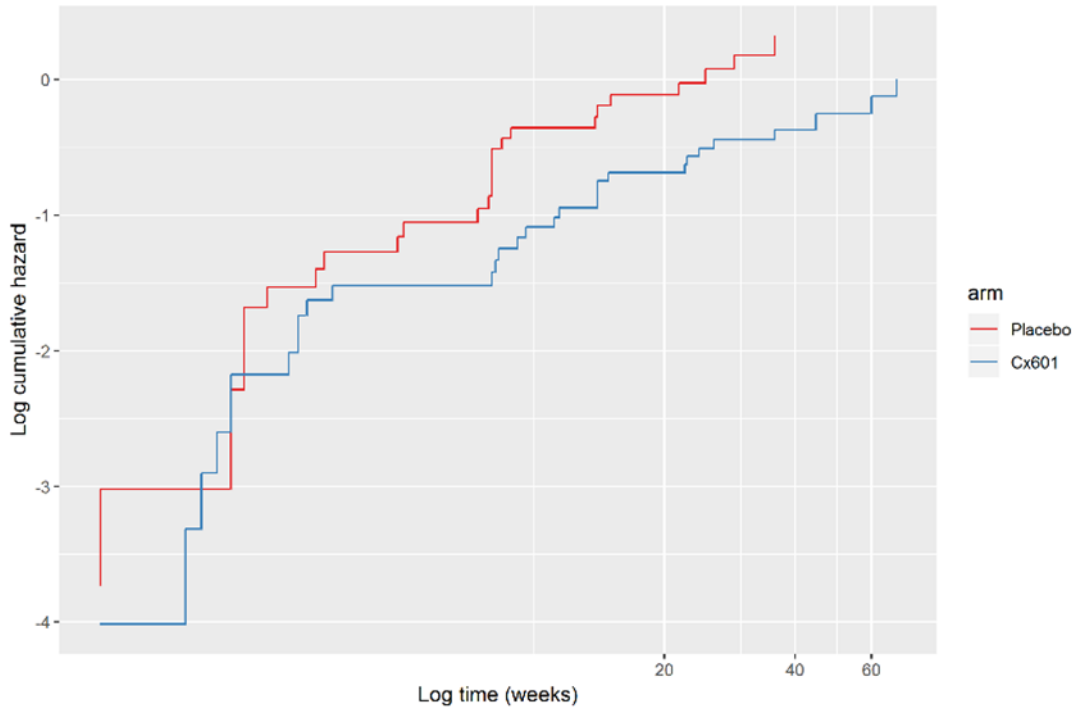
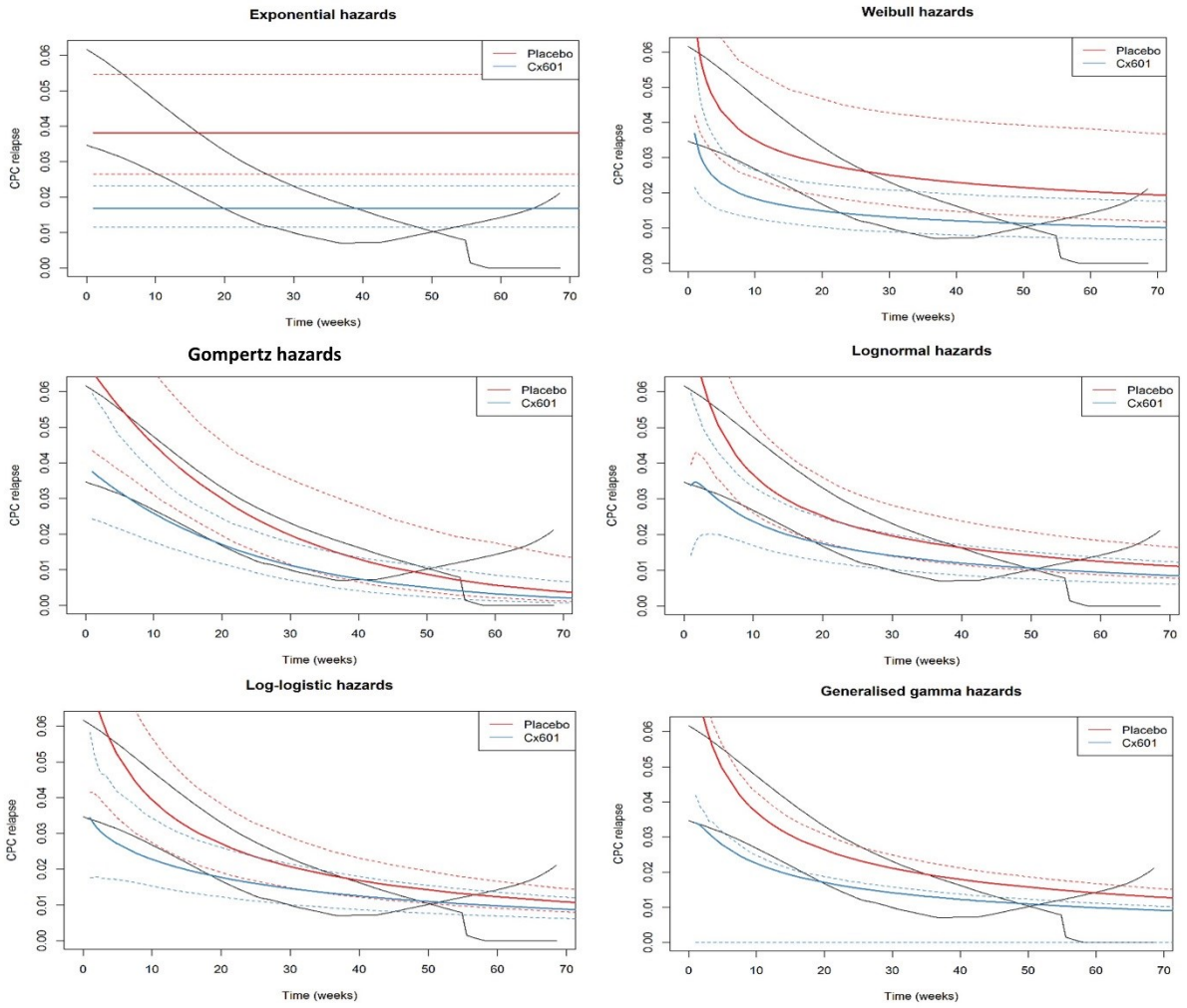


**Figure 6: Empirical hazard versus predicted hazard plots for CPC remission (reproduced from clarification response,[2] question B3)**




The company obtained expert opinion in two phases on the plausibility of the long-term extrapolations of the time-to-event functions. In the first phase, general opinions around the expected time-to-event function were sought from 10 experts (see clarification response,[2] question B4). This opinion indicated that "*... the risk of relapse for patients who have been in remission a long time would decrease…"*. This rationale was used to support the company's selection of the Gompertz time-to-event function in the base case. Visual comparison of the different parametric time-to-event models against the Kaplan-Meier time-to-event function are presented for CPC remission and CPC relapse in Figure 7 and

Figure 8, respectively. In both cases, the Gompertz time-to-event function was selected for use in the company's health economic model and was presented to a panel of seven clinical experts to assess the plausibility of that curve alone (clarification response,[2] question B4). Based on the information on the model diagnostics, clinical opinion around the long-term event hazards, and the fact that the company's elicitation exercise produced a HR, the company selected a Gompertz distribution for both the CPC remission and CPC relapse time-to-event functions.

**Figure 7: Parametric time-to-event functions compared to the Kaplan-Meier time-to-event function for CPC remission (reproduced from clarification response,[2] question B13)**

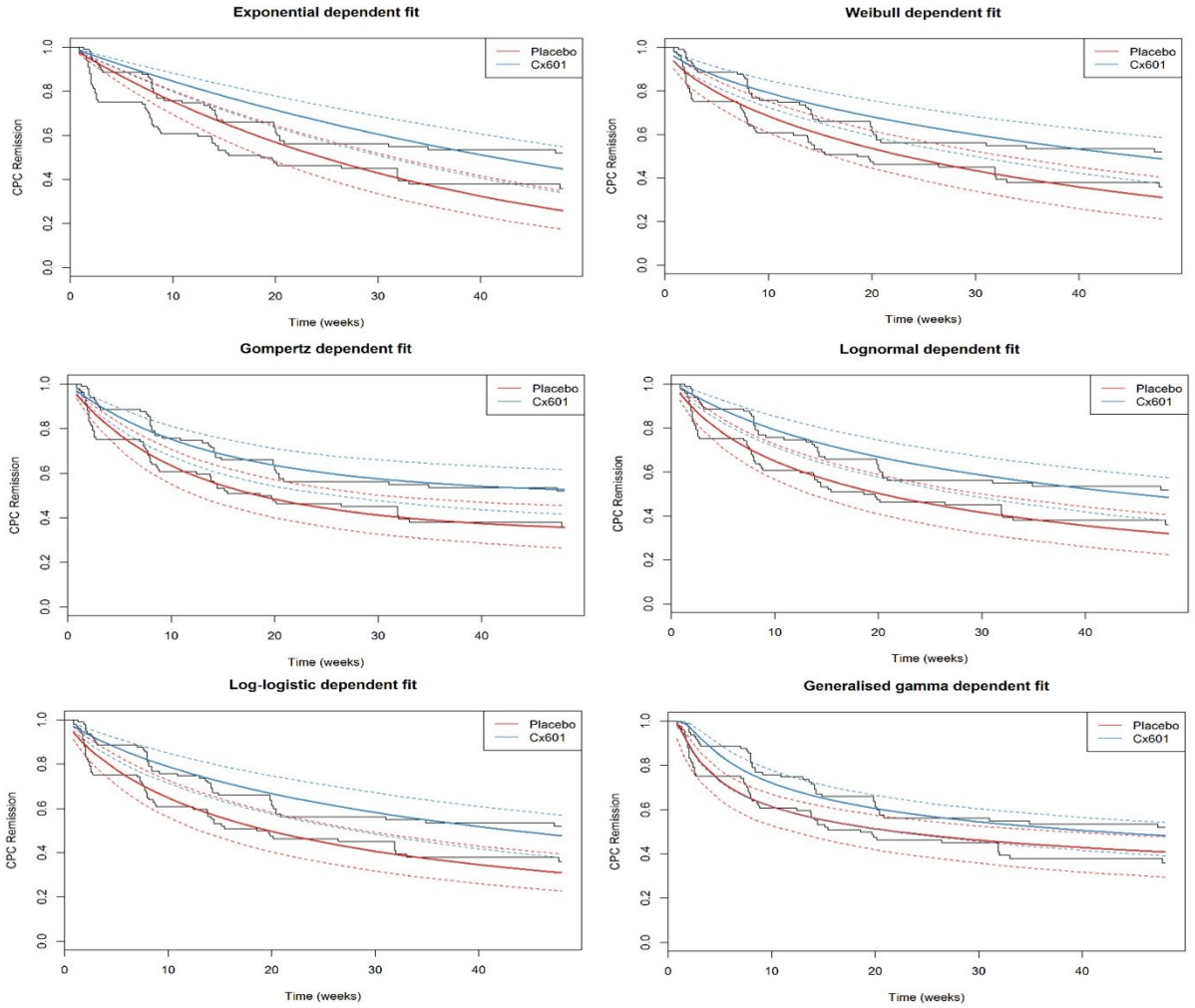
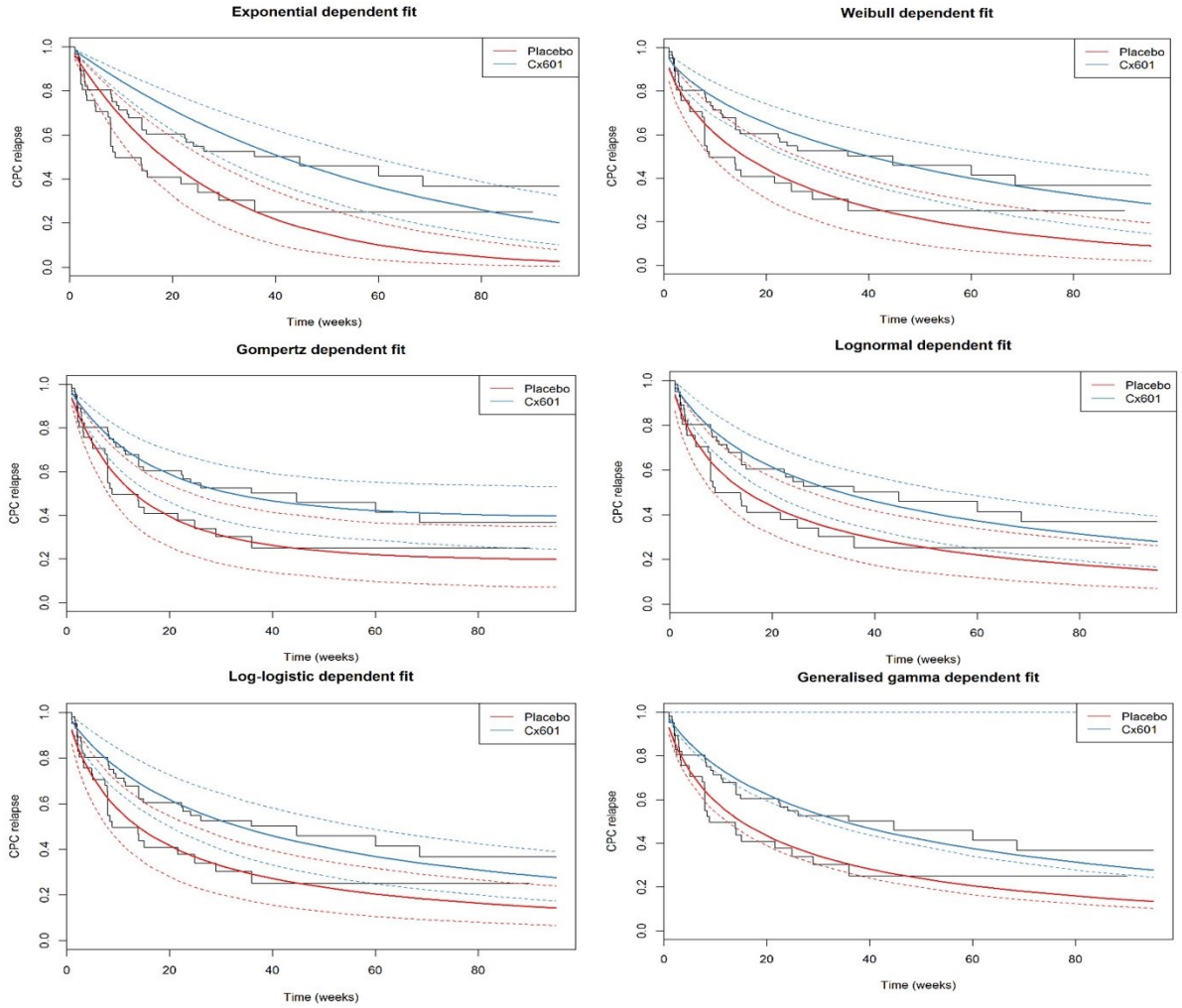


**Figure 8: Parametric time-to-event functions compared to the Kaplan-Meier time-to-event function for CPC relapse (reproduced from clarification response,[2] question B13)**




With respect to the long-term extrapolation, the company used the statistical time-to-event functions to estimate the probability of relapse (for darvadstrocel, standard care, and salvage therapy) or remission (for salvage therapy only) functions up to the 24th model cycle (approximately two years). After this point the company, estimated a time-to-event function specific constant probability of relapse or remission. This probability was assumed to be constant and calculated using two points of each fitted time-to-event function: (1) the cumulative probability of relapse or remission at 100 weeks post-randomisation, and (2) the cumulative probability of relapse or remission at 160 weeks post-randomisation. The company considered this approach to be appropriate, as their clinical advisors stated that they would expect there to be a higher risk of relapse in the long-term than was predicted by the Gompertz time-to-event functions. The company presented the resulting curve to seven clinical experts, which they deemed to be clinically plausible. (see clarification response[2], question B4) The ERG's critique of this approach is provided in Section 5.3.4.5. It should be noted that the time to remission functions were not extrapolated beyond 1 year for the darvadstrocel or standard care groups.

*Time to remission and relapse for salvage therapy*

In the modelled population, the time to remission and time to relapse need to be estimated for patients receiving salvage therapy. The company estimated the time to remission and time to relapse for people receiving salvage therapy by estimating treatments of salvage therapy compared to standard care in an expert elicitation exercise. The company estimated these treatment effects as HRs of 0.6 for time to remission and 1.0 for time to relapse.[1]

In response to a request for clarification from the ERG (question B16), the company provided the following additional details regarding the expert elicitation exercise.[2] The expert elicitation exercise followed no formal protocol. Six experts from the EU (three of whom were from the UK), were asked to identify the scenario regarding the effectiveness of salvage therapy compared with control that they believed best represented the effectiveness of future lines of therapy compared with standard care. The expert elicitation exercise was only designed to elicit a HR and other plausible treatment effect assumptions were not elicited from the six experts. The company's justification for this was that this assumption was "… *validated by clinical experts in Europe and the UK*…" and that "… *both control and salvage therapy broadly consisted of the same interventions, those being EUA +/- seton placement with background therapy consisting of antibiotics, immunosuppressants and biologic therapy* …" (clarification response,[2] question B16). Other details on the company's elicitation process and information presented to the experts during this process are unclear. The ERG's concerns regarding the expert elicitation exercise and the implementation of the estimated HRs are presented in Section 5.3.4.6.

The logic used to implement the time to remission and time to relapse functions for patients on salvage therapy is that: before 24 model cycles (approximately 2 years), the time-to-event functions for the salvage therapy group (standard care Gompertz distribution with a HR applied) is used; after 24 model cycles the constant probability of remission or relapse from the standard care arm is used (i.e. no HR is applied). This issue is further discussed in Section 5.3.3.

*Probability of receiving defunctioning surgery*

Mueller *et al.* was a prospective cohort study of 102 consecutive patients with Crohn's disease who presented with their first manifestation of perianal fistula or perianal abscess in a German outpatient ward between 1992 and 1995.[24] Out of the 102 patients recruited, 46 subjects had a complex perianal fistula. A Kaplan-Meier time-to-event function for the time to permanent faecal diversion from the year since each patient first presented with Crohn's disease was produced for the subgroup of study participants with a complex perianal fistula. The company calculated the time to defunctioning surgery, by digitising the Kaplan-Meier time-to-event function from Mueller *et al*, using the Guyot *et al.* method for reconstructing time-to-event data.[24, 25] The company fitted only an exponential distribution time-to-

event to the data as "… *an assumption required for simplification of the model structure*…" (clarification response,[2] question B6, page 63), i.e. they chose a distribution with a constant hazard rate to avoid the need for time dependent probabilities for transitions out of this health state which simplified the implementation of the model.

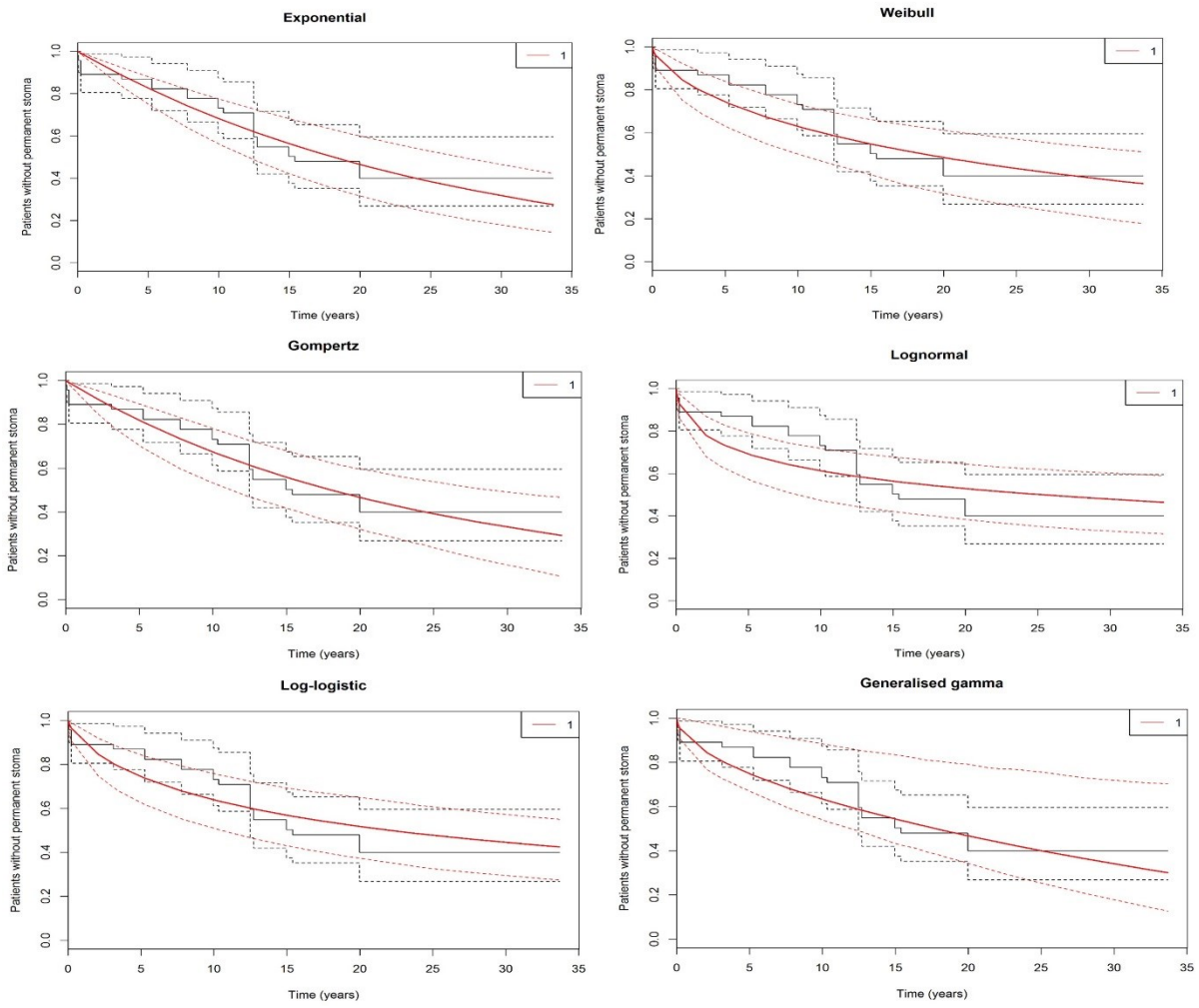
In response to a request for clarification from the ERG (question B6), the company provided AIC and BIC statistics and visual comparisons of the parametric time-to-event functions to the Kaplan-Meier curve. The AIC and BIC statistics are presented in Table 15; these show that the generalised gamma function provides the best fit to the underlying data based on the AIC criterion and that the Weibull function provides the best fit based on the BIC criterion. The visual plot of the parametric time-to-event models and the Kaplan-Meier curves are given in Figure 9. In this plot, the black line indicates the Kaplan-Meier curve, the red line indicates the fitted parametric time-to-event function, and the dotted lines indicate the 95% confidence intervals around these estimates. It is clear in Figure 9 that none of the curves provide a particularly good fit to the observed data, but the central estimates of the exponential and Gompertz functions provide the best approximation of the shape of the Kaplan-Meier curve.

**Table 15: AIC and BIC statistics for the fitted parametric curves to the time to permanent stoma (adapted from clarification response,[2] question B6)**

| | **AIC** | **BIC** |
|---|---|---|
| Exponential | 215.1842 | 217.0129 |
| Weibull | 210.6724 | **214.3297** |
| Gompertz | 217.1266 | 220.7839 |
| Log normal | 222.7323 | 226.3896 |
| Log logistic | 214.5136 | 218.1709 |
| Generalised Gamma | **209.5852** | 215.071 |

AIC –Akaike information criterion; BIC – Bayesian information criterion
Text in ***bold and italics*** indicates the lowest value in each column (best fitting to the data)

**Figure 9: Observed and predicted time-to-event curves for permanent stoma (from clarification response,[2] question B6)**




*Probability of receiving proctectomy*

The probability of receiving a proctectomy surgery was estimated by the company from Bell *et al*.[27] This prospective study collected data on the clinical course of 87 patients with Crohn's disease related fistulae between January 1993 and December 1994. Approximately 21%(18/87) of people with an active fistula and Crohn's disease subsequently received a proctectomy at a median time of 6 years (range 0.23 to 28.2 years) after their first presentation of a fistula. The company calculated the annual probability of undergoing a proctectomy to be 0.0385.

5.2.6.2 Time-independent probabilities

*Probability that a CSF is mild*

Data on the probability that a CSF is mild was obtained from the ADMIRE-CD trial.[1] The company defined mild CSF to be any person with a complex perianal fistulae and non-active / mildly active luminal Crohn's disease who had a score of 1 on either the pain or discharge dimensions of the PDAI and a score of ≤1 on the other dimension. Severe CSF was defined as any complex perianal fistulae

for people who had non-active / mildly active luminal Crohn's disease that were not either mild or in remission. The company estimated the proportion of cases that were mild and severe by taking an average of the PDAI score of people with CSF. Patients with missing data or in remission were excluded from these calculations. It was assumed that these probabilities were constant with respect to time.

*Probabilities that a proctectomy or defunctioning surgery are successful*

The probability that a proctectomy was successful and the probability that a defunctioning surgery was successful were obtained from the St Mark's retrospective cohort study (CS,[1] Appendix Q). In this prospective study, data was collected from 78 consecutive patients who presented with a complex perianal fistula and Crohn's disease at St Marks hospital between from 1st January 2008 to July 1st 2017. Data were collected at baseline, routine visits and study termination (lost to follow up, transferred to another hospital, or patient death). In this data source, the probability that a proctectomy was successful was 0.62 and the probability that a defunctioning surgery was successful was 0.80.



*Mortality*

The age-dependent probability of death was taken from general population life tables for England and Wales in 2013-15.[26]

*HRQoL*

The ADMIRE-CD trial[1] did not include a preference-based measure of HRQoL. The CS states that there are no disease-specific measures of HRQoL available for patients with perianal fistula.[1] The only patient reported outcome measure included in ADMIRE-CD was the IBDQ. The company considered whether it was possible to map from the PDAI, CDAI or IBDQ scores obtained in the trial to the EQ-5D. The CS states that there is insufficient conceptual overlap between the content of the PDAI and CDAI, which are considered to be measures of disease activity, and the relevant components of HRQoL.[1] The company cites a mapping study by Buxton *et al.*(2007)[37] which they claim supports the poor performance of CDAI as a predictor of utility. The ERG notes that the mapping algorithms reported by Buxton *et al*[37] were derived and validated in studies that included patients with moderately to severely active Crohn's disease. The company does not consider mapping from IBDQ to be appropriate because IBDQ is focused on luminal disease and not complex perianal fistulae. The ERGs clinical advisors agreed that IBDQ was a Crohn's disease specific measure of health. The company conducted a systematic review of HRQoL studies, but concluded that none of the studies identified were suitable for informing utility values in the model.

The health state utility values (HSUVs) used in the company's model were taken from a vignette study reported by Fountain *et al*.[38] which was funded by Takeda (the full study report is provided in

CS,[1] Appendix R). Vignettes were developed describing eight health states that were relevant to the model structure: (1) remission, (2) CSF with mild symptoms, (3) CSF with moderate symptoms, (4) abscess, (5) defunctioning surgery with positive outcome, (6) defunctioning surgery with negative outcome, (7) proctectomy with positive outcome and (8) proctectomy with negative outcome. The health state descriptions were derived with the input of both patients and clinicians. These were valued used a time-trade off (TTO) methodology by both a representative sample of the general public (n=835) and by a sample of patients with Crohn's disease, but not specifically CSF (n=162). The values generated by the general public sample were used in the company's base case analysis; the values from Crohn's disease patients were explored in a sensitivity analysis. The CS also reported in detail the validation of the utility values by EU and UK clinical experts (CS,[1] Appendix P).

The utility values applied in the company's base case analysis are summarised in Table 16. Whilst the vignette study measured the utility of CSF with abscess as a separate health state, the company incorporated a utility decrement associated with abscess into the model by calculating the difference between the utility values for CSF with abscess and CSF with severe symptoms. This resulted in a mean disutility of 0.16 (SE 0.026, 95% CI 0.11 to 0.21). The company's model assumes that there is no additional decrement associated with proctalgia as this event may be experienced by patients having CSF and was therefore already accounted for within the HSUVs for CSF.

**Table 16: Vignette study results, general population sample (adapted from CS,[1] Table 46)**

| Health state | | Observations | Mean utility | Standard deviation | Standard error | 95% confidence interval |
|---|---|---|---|---|---|---|
| **Remission** | | 835 | 0.865 | 0.24 | 0.008 | [0.85; 0.88] |
| **Chronic symptomatic fistulae** | **Mild symptoms** | 835 | 0.578 | 0.44 | 0.015 | [0.55; 0.61] |
| | **Severe symptoms** | 835 | 0.383 | 0.50 | 0.017 | [0.35; 0.42] |
| **Abscess** | | 835 | 0.223 | 0.55 | 0.019[a] | [0.19;0.26] |
| **Defunctioning** | **Undergoing** | Assumed equal to CSF with severe symptoms | | | | |
| | **Successful** | 835 | 0.567 | 0.46 | 0.016 | [0.54; 0.60] |
| | **Unsuccessful** | 835 | 0.193 | 0.56 | 0.019 | [0.15; 0.23] |
| **Proctectomy** | **Undergoing** | Assumed equal to CSF with severe symptoms | | | | |
| | **Successful** | 835 | 0.564 | 0.50 | 0.017 | [0.53; 0.60] |
| | **Unsuccessful** | 835 | 0.202 | 0.57 | 0.020 | [0.16; 0.24] |

Abbreviations: CSF, chronic symptomatic fistulae. Notes: **, assumed equal to chronic symptomatic fistulae with severe symptoms.
Source: Takeda, data on file.
a calculated by ERG

*Resource use and costs*

*Health state related inpatient and outpatient resource use*

The health state resource use per 4-weekly model cycle, the unit cost of each resource use type and the total cost associated with each type of resource use for each health state are summarised in Table 17.

The unit costs of each resource use item were obtained from a variety of sources (NHS Reference Costs 2016-17[28], the Personal Social Services Research Unit (PSSRU),[29] NICE TA329[30] and NICE DG11[31]). For each health care resource use item, the number of items used in each 4-weekly cycle were obtained from the ADMIRE-CD trial and/or clinical expert opinion.

**Table 17: Health state resource use and associated costs used in the company's model (adapted from CS,[1] Tables 52 and 53)**

| | Unit cost | | Resource use (number of visits / tests) per 4 weekly cycle | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Resource item** | **Cost per item of resource use (£)** | **Source** | **Remission** | **CSF** | | **Defunctioning** | | | **Proctectomy** | | |
| | | | | **Mild** | **Severe** | **Undergoing** | **S** | **U** | **Undergoing** | **S** | **U** |
| **Healthcare professional resource use** | | | | | | | | | | | |
| GP visits | 37.00 | PSSRU[29] | 0.06 | 0.12 | 0.14 | 1.38 | 0.10 | 0.21 | 1.38 | 0.10 | 0.25 |
| Gastroenterologist visits | 149.76 | NHS Reference costs[28] | 0.13 | 0.17 | 0.31 | 2.00 | 0.10 | 0.31 | 2.00 | 0.12 | 0.31 |
| Surgeon visits | 127.09 | NHS Reference costs[28] | 0.04 | 0.10 | 0.22 | 2.25 | 0.10 | 0.29 | 3.25 | 0.12 | 0.48 |
| Nurse appointments | 51.15 | NHS Reference costs[28] | 0.06 | 0.16 | 0.27 | 1.75 | 0.12 | 0.35 | 2.75 | 0.15 | 0.56 |
| Nutritionist visits | 81.33 | NHS Reference costs[28] | 0.02 | 0.02 | 0.08 | 0.25 | 0.04 | 0.12 | 0.25 | 0.06 | 0.12 |
| ***Total cost of health care professional visits per four weekly cycle*** | | | ***£31.70*** | ***£52.04*** | ***£99.35*** | ***£746.38*** | ***£39.21*** | ***£117.66*** | ***£924.62*** | ***£48.06*** | ***£154.34*** |
| Monitoring resource use | | | | | | | | | | | |
| Rectal MRI | 162.23 | NHS Reference costs[28] | 0.01 | 0.06 | 0.13 | 1.00 | 0.02 | 0.10 | 1.25 | 0.04 | 0.13 |
| Endoscopy | 182.10 | NHS Reference costs[28] | 0.06 | 0.06 | 0.13 | 1.00 | 0.06 | 0.13 | 1.25 | 0.00 | 0.06 |
| Stoma care* | 1,961.00 | NICE TA 329[30] | 0.00 | 0.00 | 0.00 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 |
| Computerised tomography | 85.56 | NHS Reference costs[28] | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Colonoscopy | 334.76 | NHS Reference costs[28] | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ***Total cost of monitoring patients per four weekly cycle*** | | | ***£12.07*** | ***£19.87*** | ***£44.60*** | ***£495.18*** | ***£164.47*** | ***£190.83*** | ***£581.26*** | ***£157.09*** | ***£183.19*** |
| Laboratory resource use | | | | | | | | | | | |
| Blood count | 1.69 | NHS Reference costs[28] | 0.15 | 0.12 | 0.23 | 2.25 | 0.15 | 0.28 | 2.50 | 0.15 | 0.35 |
| C-reactive protein | 1.13 | NHS Reference costs[28] | 0.17 | 0.13 | 0.27 | 2.25 | 0.15 | 0.31 | 2.50 | 0.15 | 0.37 |
| Haemoglobin | 3.06 | NHS Reference costs[28] | 0.17 | 0.12 | 0.23 | 2.25 | 0.15 | 0.28 | 2.50 | 0.15 | 0.35 |
| Faecal calprotectin | 22.79 | NICE DG11[31] | 0.13 | 0.13 | 0.27 | 1.50 | 0.10 | 0.15 | 1.75 | 0.12 | 0.15 |
| ***Total cost of laboratory tests per four weekly cycle*** | | | ***£3.77*** | ***£7.54*** | ***£4.05*** | ***£47.42*** | ***£3.10*** | ***£5.19*** | ***£54.58*** | ***£3.53*** | ***£5.56*** |

| | Unit cost | Resource use (number of visits / tests) per 4 weekly cycle | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Total health state resource use costs per four weekly cycle** | | **£47.82** | **£75.67** | **£151.49** | **£1288.97** | **£206.78** | **£313.68** | **£1560.46** | **£208.68** | **£343.09** |

CSF – chronic symptomatic fistula; S – successful; U – unsuccessful; GP – general practitioner; PSSRU - Personal Social Services Research Unit; NHS – National Health Service; MRI – magnetic resonance imaging; NICE – National Institute for Health and Care Excellence; TA – technology appraisal; DG – diagnostics guidance; * - the unit cost applied is an annual cost

*Darvadstrocel acquisition cost (including PAS)*

Drug acquisition costs for darvadstrocel were provided by the company. The company has a Patient Access Scheme in place for darvadstrocel which takes the form of a simple price discount. Including the PAS, the price per vial of darvadstrocel is [redacted], giving a total cost of [redacted] per course of treatment. The model assumes that four vials are used in the EUA procedure in the darvadstrocel arm, which occurs in cycle 1 of the model. This is in line with the marketing authorisation for darvadstrocel.[6]

*Frequency of different surgical and drug treatments for complex perianal fistulae*

The proportion of patients who receive the different types of surgical and medical treatments are given by health state and treatment line in Table 18. These proportions were estimated from the ADMIRE-CD trial data for the darvadstrocel and standard groups when they were in the CSF mild or CSF severe health state. For the other health states and the people receiving salvage therapy group in the CSF mild or CSF severe health states, the proportions were estimated using UK expert clinical opinion. The exact number of experts used is unclear.

*Costs of use for different surgical and medical treatments for complex perianal fistulae*

The costs of the different surgical and medical treatments depends upon the health state in which the they are used. When patients received their initial treatment in the CSF health states (either darvadstrocel or standard care), the procedures and associated costs were split into those that would be delivered in the first cycle only and those that were delivered in all cycles. These costs were then applied at appropriate times within the state transition structure of the health economic model. For all other model health states, the mean cost of treatment over 13 model cycles was used to calculate the costs of treatment regardless of how many cycles patients spent in that particular health state. The cost of surgical and medical treatments are given by health state for cycle 1, subsequent cycles and the average over all cycles is given in

Table 19. In addition to the treatment costs, additional costs were applied relating to the administration of these treatments, these administration costs are provided in Table 20.

**Table 18: Percentage of patients receiving each treatment by health state and treatment group (adapted from CS,[1] Table 54)**

| Treatment mix | Mild CSF | | | Severe CSF | | | Rem | Defunctioning | | Proctectomy | | Sources and assumptions |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DARV | Control | Salvage | DARV | Control | Salvage | | S | U | S | U | |
| **Darvadstrocel** | | | | | | | | | | | | |
| Darvadstrocel | 100 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **Antibiotics** | | | | | | | | | | | | |
| Ciprofloxacin | 29.76 | 29.76 | 11.25 | 29.78 | 29.78 | 57.50 | 0 | 0 | 0 | 0 | 0 | ADMIRE CD trial data |
| Metronidazole | 38.05 | 38.05 | 55.28 | 38.05 | 38.05 | 58.75 | 11.20 | 18.56 | 57.81 | 1.09 | 32.66 | |
| **Immunosuppressants** | | | | | | | | | | | | |
| Azathioprine | 46.23 | 46.23 | 46.37 | 46.23 | 46.23 | 47.50 | 51.32 | 58.99 | 46.88 | 45.01 | 52.50 | ADMIRE CD trial data, clinical expert opinion |
| Methotrexate | 0 | 0 | 9.05 | 0 | 0 | 0.5 | 7.29 | 0.00 | 5.84 | 11.66 | 0 | |
| 6-MP | 0 | 0 | 7.50 | 0 | 0 | 26.75 | 10.00 | 11.88 | 11.88 | 0 | 0 | |
| **Biologics** | | | | | | | | | | | | |
| Adalimumab | 33.59 | 33.59 | 30.65 | 33.59 | 33.59 | 19.17 | 31.76 | 21.32 | 27.03 | 12.86 | 25.47 | ADMIRE CD trial data, clinical expert opinion |
| Infliximab | 27.26 | 27.26 | 30.65 | 27.26 | 27.26 | 35.83 | 32.39 | 21.32 | 27.03 | 12.86 | 25.47 | |
| Adalimumab dose escalation | 0 | 0 | 5.94 | 0 | 0 | 7.5 | 4.92 | 3.38 | 10.21 | 0.75 | 8.75 | |
| Infliximab dose escalation | 0 | 0 | 5.94 | 0 | 0 | 7.5 | 4.92 | 3.38 | 10.21 | 0.75 | 8.75 | |
| Vedolizumab | 0 | 0 | 8.67 | 0 | 0 | 0 | 8.24 | 5.08 | 7.69 | 3.36 | 7.36 | |
| **Surgery** | | | | | | | | | | | | |
| Seton | 95 | 95 | 20.56 | 95 | 95 | 48.5 | 5.21 | 11.54 | 11.96 | 0 | 2.50 | ADMIRE CD trial data, clinical expert opinion |
| Fistulotomy | 0 | 0 | 1.51 | 0 | 0 | 16.5 | 0 | 0 | 5.84 | 0 | 0 | |
| Anal plug | 0 | 0 | 12.50 | 0 | 0 | 11.25 | 0 | 0 | 0 | 0 | 0 | |
| Fibrin glue | 0 | 0 | 0 | 0 | 0 | 6.25 | 0 | 0 | 0 | 0 | 0 | |
| Rectal flap | 0 | 0 | 0 | 0 | 0 | 12.5 | 0 | 0 | 0 | 0 | 0 | |
| EUA alone | 0 | 0 | 43.09 | 0 | 0 | 0 | 11.12 | 6.59 | 37.38 | 0 | 26.43 | |
| VAAFT | 0 | 0 | 4.52 | 0 | 0 | 0 | 0 | 6.73 | 0 | 0 | 0 | |

CSF – chronic symptomatic fistulae; Rem – remission; DARV – darvadstrocel; Control – standard care; S – successful; U - unsuccessful; EUA, examination under anaesthesia; 6-MP, 6-mercaptopurine.

**Table 19: Cost of pharmacological and surgical treatments given to each patient (adapted from CS,[1] Table 54)**

| Treatment | Unit cost | Doses per item | Source | Doses given in cycle 1 | Doses given in subsequent cycles | Cost in cycle 1 | Cost in subsequent cycles | Average Cycle cost across 13 model cycles |
|---|---|---|---|---|---|---|---|---|
| Darvadstrocel | | | | | | | | |
| Darvadstrocel | [redacted] | 1 unit | Takeda | 4 units | 0 units | [redacted] | £0 | Not applicable |
| Antibiotics | | | | | | | | |
| Ciprofloxacin | £0.089 | 500mg | BNF | 56 | 56 | £4.94 | £4.94 | £4.94 |
| Metronidazole | £0.195 | 400mg | BNF | 76.20 | 76.20 | £14.88 | £14.88 | £14.88 |
| Immunosuppressants | | | | | | | | |
| Azathioprine | £0.039 | 50mg | BNF | 91.44 | 91.44 | £3.56 | £3.56 | £3.56 |
| Methotrexate | £0.054 | 2.5mg | BNF | 28 | 28 | £1.51 | £1.51 | £1.51 |
| 6-MP | £1.966 | 50mg | BNF | 50.80 | 50.80 | £99.88 | £99.88 | £99.88 |
| Biologics | | | | | | | | |
| Adalimumab | £352.14 | 40mg | BNF | 2 | 2 | £704.28 | £704.28 | £704.28 |
| Infliximab | £377.00 | 100mg | BNF | 1.81 | 1.81 | £684.01 | £684.01 | £684.01 |
| Adalimumab dose escalation | £352.14 | 40mg | BNF | 4 | 4 | £1368.02 | £1368.02 | £1368.02 |
| Infliximab dose escalation | £377.00 | 100mg | BNF | 3.63 | 3.63 | £1408.56 | £1408.56 | £1408.56 |
| Vedolizumab | £2050 | 300mg | BNF | 1.00 | 0 | £2050 | 0 | £78.85 |
| Surgical procedures | | | | | | | | |
| Seton | £0 | 1 set | Assumption | 1 | 0 | £0 | £0 | £0 |
| Fistulotomy | £1,170.21 | 1 operation | NICE MIB 102 | 1 | 0 | £1,170.21 | £0 | £78.85 |
| Anal plug | £1,170.21 | 1 operation | Assumed equal to fisulotomy | 1 | 0 | £1,170.21 | £0 | £78.85 |
| Fibrin glue | £724.19 | 1 set | NICE MIB 105 | 1 | 0 | £724.19 | £0 | £55.71 |
| Rectal flap | £1,170.21 | 1 operation | Assumed equal to fisulotomy | 1 | 0 | £1,170.21 | £0 | £78.85 |
| EUA | £1,170.21 | 1 operation | NHS reference costs[28] | 1 | 0 | £1,170.21 | 0 | £90.02 |
| VAAFT | £1,195.40 | 1 operation | NICE MIB 102 | 1 | 0 | £1,195.40 | 0 | £91.95 |

BNF – British National Formulary; 6–MP - 6-mercaptopurine; NICE – national institute for health and care excellence; MIB – Medtech Innovation Briefing; EUA – examination under anaesthesia; VAAFT - video-assisted anal fistula treatment

**Table 20: Cost of treatment administration methods (adapted from CS,[1] Table 55)**

| Administration method | Unit Cost | Source | Treatments delivered[a] |
|---|---|---|---|
| EUA | See Table 19 | NHS reference costs[28] | Darvadstrocel, seton, fibrin glue |
| IV infusion | £284.49 | NHS reference costs[28] | Infliximab, dose escalated infliximab, vedolizumab |
| SC injection | £0 | Assumed to be self-administered | Adalimumab, dose-escalated adalimumab |
| Oral | £0 | Assumption to be self-administered | Ciprofloxacin, Metronidazole, Azathioprine, Methotrexate, 6-MP |

a - any treatment not included in this table did not have an administration cost
EUA – examination under anaesthesia; IV – intravenous; SC - subcutaneous

### *5.2.7 Cost effectiveness results*

In the CS, the company discounts costs at a rate of 3.5% and QALYs at a rate of 1.5%.[1] The ERG considers this to be inappropriate, as differential discounting of costs and QALYs is not supported in the NICE Methods guide.[10] A further consideration is that the company states that they believe that: darvadstrocel restores people with complex perianal fistulae and non-active / mildly active luminal Crohn's disease to full health over a long period of time; people receiving standard care have a severely impaired quality of life, and; and that darvadstrocel would not commit the NHS to irrecoverable costs.[1] Consequently, section 6.2.19 of the methods guide may apply.[10] The company believes that "*…darvadstrocel demonstrates long term healing potential in this population with a significant impact on QoL…*" (CS,[1] page 74). The ERG has concerns about whether darvadstrocel meets the criteria in Section 6.2.19 of the Method Guide (see Section 5.3.4).

The ERG considers that the analyses presented in the original CS are out of scope, as differential discounting of costs and QALYs are used.[1] In the company's clarification response to question B7, two sets of in scope analyses were provided, in the first both costs and QALYs are discounted at a rate of 3.5% and in the second both costs and QALYs are discounted at a rate of 1.5%.[1] For completeness, the ERG presents the company's base case analysis both when using the company's preferred differential discounting and when using 3.5% discounting for both costs and QALYs as per the NICE Reference Case. The company's results using discount rates of 1.5% for both costs and QALYs are presented in Appendix 2.

Table 21 shows the results of the company's base case analysis in both the deterministic analysis and the PSA analysis when discount rates of 1.5% and 3.5% are used for QALYs and costs respectively.

Based on the probabilistic version of the company's model, darvadstrocel is expected to generate an additional 1.35 QALYs at an additional cost of £21,774, compared with standard care. The corresponding incremental cost-effectiveness ratio is £16,121 per QALY gained. The deterministic version of the company's model produces a similar ICER of £15,471 per QALY gained. These results are based on differential discounting of costs and QALYs, so the ERG urges caution in using these values.

**Table 21: Company's base case results, including the patient access scheme for darvadstrocel, assuming 1.5% discount rate for QALYs and a 3.5% discount rate for costs (adapted from CS,[1] Table 66 and Table 67)**

| Treatment | Total QALYs | Total costs | ICER (£ per QALY gained) | Probability that the intervention is the most cost-effective at a maximum acceptable ICER of: | |
|---|---|---|---|---|---|
| | | | | £20,000 per QALY gained | £30,000 per QALY gained |
| Probabilistic Sensitivity Analysis – based on rerun by the ERG | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | 0.650 | 0.870 |
| Standard care | [redacted] | [redacted] | - | 0.350 | 0.130 |
| Incremental | 1.35 | £21,773 | £16,102 | - | - |
| Deterministic | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | - | - |
| Standard care | [redacted] | [redacted] | - | - | - |
| Incremental | 1.40 | £21,639 | £15,471 | - | - |

QALYs – quality adjusted life years; PAS – Patient Access Scheme; ICER – incremental cost-effectiveness ratio
As there was no differential mortality between the darvadstrocel and standard care group, both arms accrued 36.65 undiscounted life years gained over the 40-year time horizon.

Table 22 shows the results of the company's revised analysis using a discount rate of 3.5% for both costs and QALYs in both the deterministic analysis and a rerun of the PSA analysis by the ERG. Based on the probabilistic version of the company's model, darvadstrocel is expected to generate an additional 1.02 QALYs at an additional cost of £21,773, compared with standard care. The corresponding incremental cost-effectiveness ratio is £21,417 per QALY gained. The deterministic version of the company's model produces a similar ICER of £20,591 per QALY gained. As shown in

Table 22, increasing the discount rate to 3.5% for both costs and QALYs increases the ICER, compared to the company's original base case presented in the CS.[1]

**Table 22: Company's revised base case results, including the patient access scheme for darvadstrocel, assuming 3.5% discount rate for both costs and QALYs (adapted from clarification response,[2] question B7, Table 23 and Table 24)**

| Treatment | Total QALYs | Total costs | ICER (£ per QALY gained) | Probability that the intervention is the most cost-effective at a maximum acceptable ICER of: | |
|---|---|---|---|---|---|
| | | | | £20,000 per QALY gained | £30,000 per QALY gained |
| Probabilistic Sensitivity Analysis – based on rerun by the ERG | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | 0.421 | 0.736 |
| Standard care | [redacted] | [redacted] | - | 0.579 | 0.264 |
| Incremental | 1.02 | £21,773 | £21,417 | - | - |
| Deterministic | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | - | - |
| Standard care | [redacted] | [redacted] | - | - | - |
| Incremental | 1.05 | £21,639 | £20,591 | - | - |

QALYs – quality adjusted life years; PAS – Patient Access Scheme; ICER – incremental cost-effectiveness ratio

Figure 10 and Figure 11 present the results of the company's PSA in the form of a cost-effectiveness plane and CEACs, based on a re-run of the company's original submitted model (based on discount rates of 3.5% for costs and 1.5% for QALYs). Assuming a maximum acceptable ICER (MAICER) of £20,000 per QALY gained, the company's model suggests that the probability that darvadstrocel produces more net benefit than standard care is 0.650. Assuming a MAICER of £30,000 per QALY gained, the company's model suggests that the probability that darvadstrocel produces more net benefit than standard care is 0.870.

**Figure 10: Cost-effectiveness plane, including the patient access scheme for darvadstrocel, comparing darvadstrocel to standard care, using a discount rate of 3.5% for costs and 1.5% for QALYs**

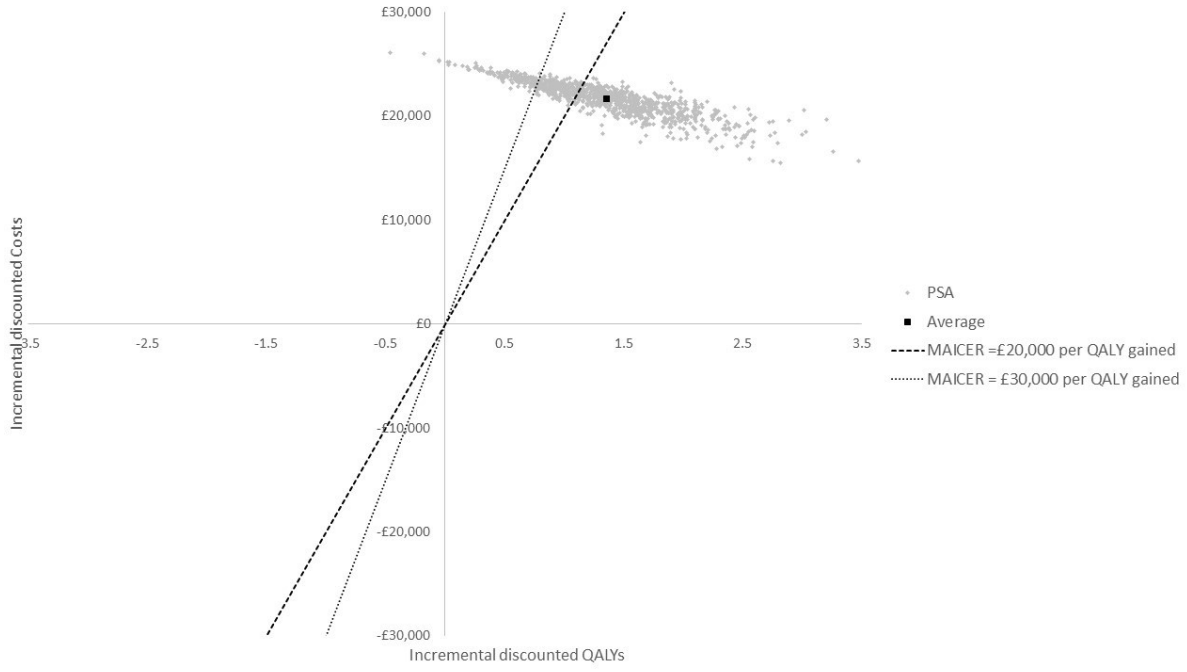


**Figure 11: Cost effectiveness acceptability curve, including the patient access scheme for darvadstrocel, using a discount rate of 3.5% for costs and 1.5% for QALYs**

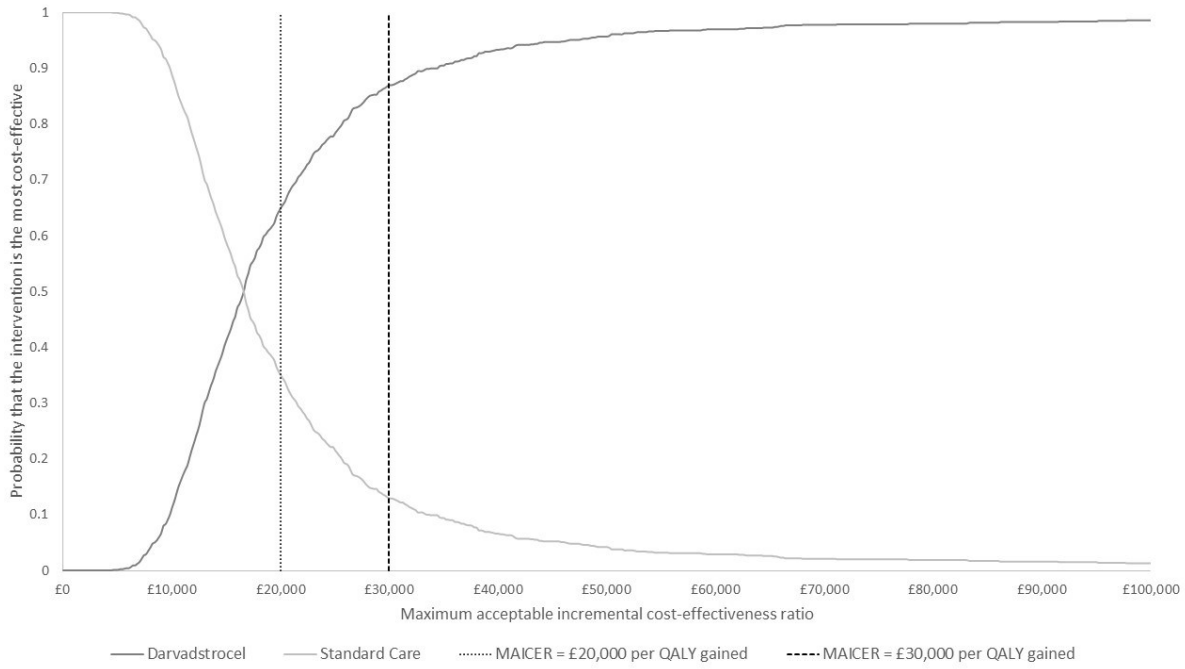


Figure 12 and
Figure 13 present the results of the company's was used in the form of a cost-effectiveness plane and a CEAC, based on a re-run of the company's model (using a discount rate of 3.5% for both costs and QALYs). Assuming a MAICER of £20,000 per QALY gained, the company's model suggests that the probability that darvadstrocel produces more net benefit than standard care is 0.421. Assuming a MAICER of £30,000 per QALY gained, the company's model suggests that the probability that darvadstrocel produces more net benefit than standard care is 0.736.

**Figure 12: Cost-effectiveness plane, comparing darvadstrocel to standard care, including the patient access scheme for darvadstrocel, using a discount rate of 3.5% for both costs and QALYs**

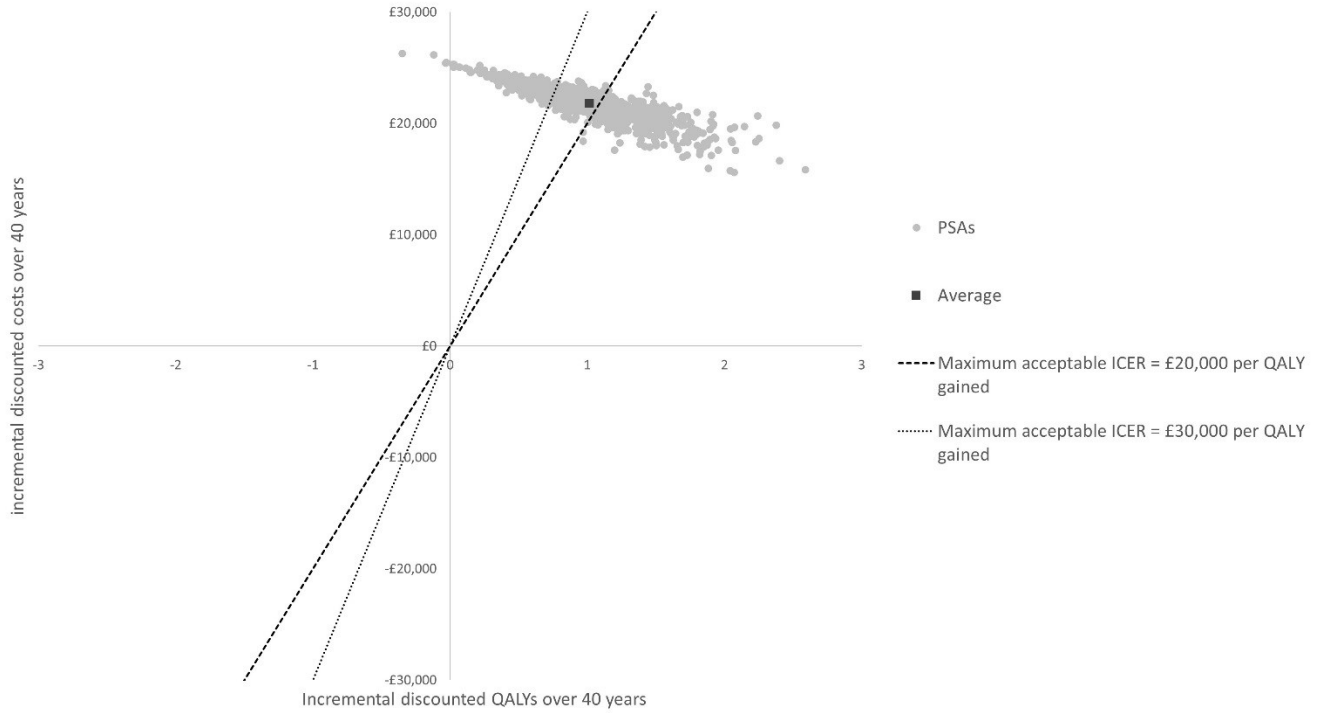
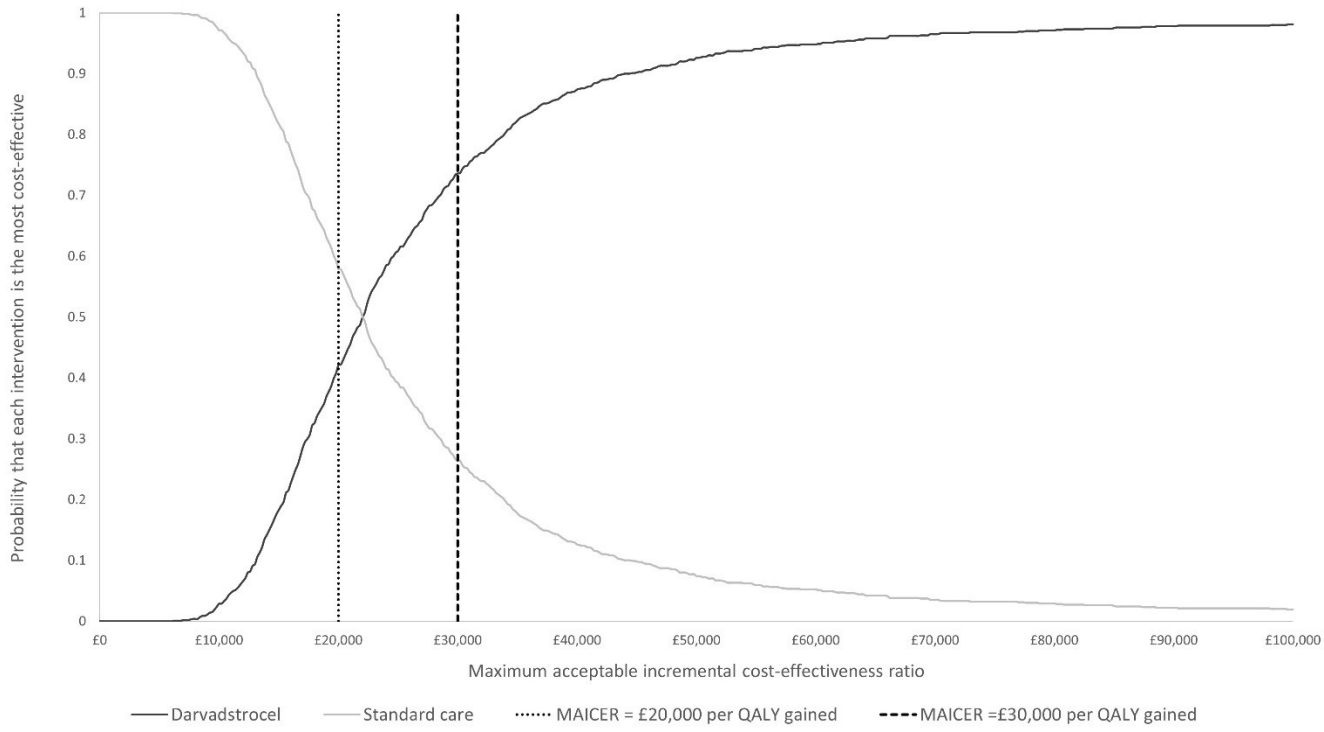



**Figure 13: Cost-effectiveness acceptability curve including the patient access scheme for darvadstrocel, using a discount rate of 3.5% for both costs and QALYs**



*5.2.8 Sensitivity analyses*

For the sensitivity analyses, the results corresponding to the company's original sensitivity analyses are presented when a using a discount rate of 3.5% for cost and QALYs. Sensitivity analyses using a discount rate of 1.5% for both costs and QALYs are presented in Appendix 2. The company's original sensitivity analyses using a discount rate of 1.5% for QALYs and 3.5% for costs are presented in the CS[1]; as the ERG considers these to be inappropriate, for brevity, these results are not reproduced here.

The company conducted a wide range of sensitivity analyses, which included: (i) a tornado diagram to show the influence of uncertainty in individual model parameters on the ICER; (ii) assessing the impact of using alternative data and/or assumptions on the ICER; (iii) assessing the impact of using alternative parametric time-to-event functions on the ICER; (iv) assessing the impact of using different definitions of remission and relapse on the ICER, and; (v) assessing the impact of directly using data collected in the St Mark's retrospective cohort study.

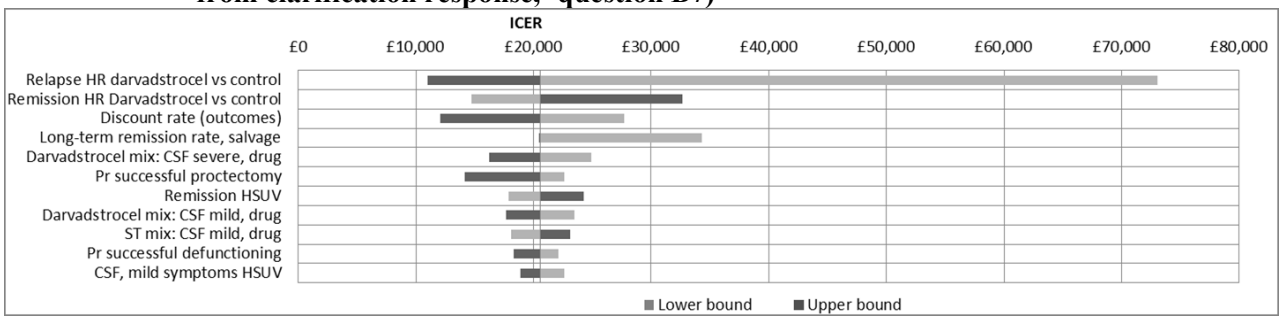
5.2.8.1 Tornado diagram

The tornado diagram presented in

Figure 14 shows the ten most influential parameters in the company's base case model, assuming a discount rate for costs and QALYs of 3.5% per annum. Within this sensitivity analysis, parameters which were included in the PSA were assessed at the upper and lower limits of their 95% CIs; parameters which were not included in the PSA were assessed at 70% of their mean value for the lower bound and 130% of their mean value for the upper bound. This analysis indicates that the company's model is particularly sensitive to:

- The HR of darvadstrocel compared to standard care for remission
- The HR of darvadstrocel compared to standard care for relapse
- The estimated remission rate for salvage therapies in year two onwards
- The probability that a proctectomy is successful
- The HSUV for remission
- The overall cost of treatments for people receiving darvadstrocel (including the fixed cost per vial of darvadstrocel) in the mild CSF health state
- The overall cost of treatments for people receiving standard care in the mild symptomatic CSF health state
- The probability that a defunctioning surgery is successful
- The HSUV for mild symptomatic complex perianal fistulae.

**Figure 14: Company's tornado diagram showing the one way sensitivity analyses conducted by the company using 3.5% discounting for both costs and QALYs (reproduced from clarification response,[2] question B7)**




5.2.8.2 Impact of alternative data sources and assumptions

The company undertook several additional sensitivity analyses (see Table 23). In these analyses, the ICER ranges from £11,380 per QALY gained to £28,438 per QALY gained. Across the range of analyses presented, the lowest ICER was generated from the sensitivity analysis in which costs and QALYs were undiscounted, the highest ICER was generated from the scenario in which the discount rates for costs and health outcomes were set equal to 6%.

**Table 23: Sensitivity analyses conducted by the company (reproduced from clarification response, question B7,[2] Table 25)**

| Scenario description | Total costs | | | Total QALYs | | | ICER (£ per QALY gained) |
|---|---|---|---|---|---|---|---|
| | Darv | Standard care | Difference | Darv | Standard care | Difference | |
| Base case, 3.5% discount for costs and QALYs | [redacted] | [redacted] | £21,639 | [redacted] | [redacted] | 1.05 | £20,591 |
| 0% discount rate for costs and QALYs | [redacted] | [redacted] | £20,400 | [redacted] | [redacted] | 1.79 | £11,380 |
| 6% discount rate for costs and QALYs | [redacted] | [redacted] | £22,233 | [redacted] | [redacted] | 0.78 | £28,438 |
| 10% annual proctectomy probability post defunctioning | [redacted] | [redacted] | £22,024 | [redacted] | [redacted] | 1.04 | £21,124 |
| 50% annual stoma reversal probability from successful defunctioning state | [redacted] | [redacted] | £21,186 | [redacted] | [redacted] | 1.04 | £20,312 |
| Upper bound of annual stoma care costs (£2,682 per year) | [redacted] | [redacted] | £20,944 | [redacted] | [redacted] | 1.05 | £19,930 |
| Infusion costs halved (£142.25) | [redacted] | [redacted] | £21,514 | [redacted] | [redacted] | 1.05 | £20,472 |
| HSUVs based on CD patients vignette study set | [redacted] | [redacted] | £21,639 | [redacted] | [redacted] | 0.98 | £22,095 |
| Relapse HR for salvage therapy vs. control equal to 1.20 | [redacted] | [redacted] | £21,566 | [redacted] | [redacted] | 1.07 | £20,131 |
| Time horizon: 20 years | [redacted] | [redacted] | £21,846 | [redacted] | [redacted] | 0.78 | £28,181 |
| Time horizon: 60 years | [redacted] | [redacted] | £21,706 | [redacted] | [redacted] | 1.10 | £19,719 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| No inclusion of Biologic usage within salvage therapy (all other assumptions as per base case) | [redacted] | [redacted] | £17,557 | [redacted] | [redacted] | 1.05 | £16,707 |
| Wastage assumed to result in 5% additional cost for darvadstrocel | [redacted] | [redacted] | £22,889 | [redacted] | [redacted] | 1.05 | £21,781 |

QALYs - quality-adjusted life years; ICER - incremental cost-effectiveness ratio; Darv – darvadstrocel; HSUV - health state utility value; CD - Crohn's disease; HR - hazard ratio

5.2.8.3 The use of alternative parametric time-to-event functions

Table 22 shows the sensitivity of the company's model to the choice of the two best fitting models (in terms of AIC and BIC) for both the time to remission and time to relapse outcomes. This shows that the model is highly sensitive to the choice of the parametric function used to model these data. The lowest ICER of £20,591 per QALY gained is produced when a Gompertz distribution is used to model both the remission and relapse time-to-event functions. The highest ICER of £133,311 per QALY gained is produced when the generalised gamma distribution is used to model the time to remission and the log normal distribution is used to model the time to relapse. These limited results also appear to indicate that the model is more sensitive to the time to relapse function than it is the time to remission function.

**Table 24: Impact of different parametric time-to-event functions on the company's base case using a discount rate of 3.5% for both costs and QALYs (reproduced from clarification response,[2] question B7, Table 26)**

| Time to remission function | Time to relapse function | Total costs | | | Total QALYs | | | ICER |
|---|---|---|---|---|---|---|---|---|
| | | Darv | SC | Incr | Darv | SC | Incr | |
| Gompertz (base case) | Gompertz (base case) | [redacted] | [redacted] | £21,639 | [redacted] | [redacted] | 1.05 | £20,591 |
| Generalised gamma | Gompertz (base case) | [redacted] | [redacted] | £22,653 | [redacted] | [redacted] | 0.75 | £30,064 |
| Gompertz (base case) | Log-normal | [redacted] | [redacted] | £24,740 | [redacted] | [redacted] | 0.24 | £104,398 |
| Generalised gamma | Log-normal | [redacted] | [redacted] | £24,754 | [redacted] | [redacted] | 0.19 | £133,311 |

Darv – darvadstrocel; SC – standard care; Incr – incremental difference between darvadstrocel and standard care; ICER - incremental cost-effectiveness ratio; QALYs – quality adjusted life years

5.2.8.4 Using different definitions of remission and relapse

Three other definitions of remission were evaluated within the CS. These were: (1) clinical remission alone, (2) CPC + MRI remission, and (3) combined remission (clinical +MRI remission). Clinical remission was defined as "…*closure of all treated external openings that were draining at baseline despite gentle finger compression*…" (CS,[1] page 30). Combined remission (ADMIRE-CD primary outcome measure) was defined as the closure of all treated external openings that were draining at baseline despite gentle finger compression, and the absence of collections larger than 2 cm of the treated perianal fistula in at least two of three dimensions, confirmed by masked central MRI (CS,[1] page 30). CPC + MRI remission was defined as CPC remission and the absence of collections larger than 2 cm of the treated perianal fistula in at least two of three dimensions, confirmed by masked central MRI.

For clinical remission, Kaplan-Meier curves were produced and parametric time-to-event functions were fitted to the underlying data. Details of the goodness-of-fit of these parametric time-to-event functions to the clinical remission data are provided in Appendix 1. In summary, AIC and BIC statistics both indicate that the log normal distribution provides the best fit to the time to remission and time to relapse, when the clinical definition of relapse is used (see Appendix 1,

Table 37). It is unclear from the CS whether an assessment of the clinical plausibility of the curves fitted to the clinical remission definition was conducted. The company again adopted the Gompertz distribution as the preferred model for the clinical definition of relapse and remission (CS,[1] page 81, page 86)

For combined remission and CPC +MRI remission, Kaplan-Meier time-to-event functions were not produced by the company, *"... due to the limited time points that combined remission was reported in the ADMIRE-CD trial.*"(clarification response,[2] question B3). Instead, HRs were estimated for the effect of MRI on the time to relapse and time to remission for both definitions of remission (CPC or clinical). This was done by comparing the number of events including an MRI definition of remission (at 24 and 52 weeks post-darvadstrocel administration) with the number of events without including the MRI criterion. The number of events at 24 and 52 weeks post-darvadstrocel administration were pooled to estimate a HR between using MRI in the definition of remission and not using MRI in the definition of remission. This process was conducted separately for the two definitions of remission (clinical remission and CPC). The HRs estimated from this process are presented in Table 25. The exact statistical process used by the company to estimate these HRs is unclear.

**Table 25: Hazard ratios applied for the calibration of the remission time-to-event functions to incorporate MRI criterion in the definition of achievement of remission (reproduced from CS,[1] page 83, Table 37)**

| Definition comparison | HR | SE [ln(HR)] | 95% CI |
|---|---|---|---|
| CPC vs. CPC + MRI | 0.922 | 0.135 | [0.708, 1.200] |
| Clinical vs. Combined (Clinical + MRI) | 0.896 | 0.111 | [0.721; 1.113] |

HR - hazard ratio; SE - standard error; ln – natural logarithm; CI - confidence interval; CPC – clinical and patient centric; MRI – magnetic resonance imaging

Table 26 shows the sensitivity of the model results to the different definitions of remission in the ADMIRE-CD study. It should be noted that the choice of parametric model did not differ in the different scenarios on the underlying remission survivor function. In response to clarification question B3, the company clarified that they believed that the Gompertz parametric model provided the best fit to the clinical and CPC definition of relapse and remission.[2]

**Table 26: Results of the scenario analyses surrounding the definition of relapse in the company's submitted economic model (adapted from company's clarification response,[2] question B7, Table 27)**

| **Scenario** | Definition of remission, parametric function | Total Costs | | | Total QALYs | | | ICER |
|---|---|---|---|---|---|---|---|---|
| | | Darv | SC | Incr | Darv | SC | Incr | |
| **Base case** | CPC, Gompertz | [redacted] | [redacted] | £21,639 | [redacted] | [redacted] | 1.05 | £20,591 |
| **1** | Clinical, Gompertz | [redacted] | [redacted] | £23,343 | [redacted] | [redacted] | 0.68 | £34,177 |
| **2** | CPC+ MRI, Gompertz | [redacted] | [redacted] | £21,755 | [redacted] | [redacted] | 1.01 | £21,446 |
| **3** | Clinical + MRI, Gompertz | [redacted] | [redacted] | £23,367 | [redacted] | [redacted] | 0.68 | £34,295 |

QALYs- quality-adjusted life years; ICER – incremental cost-effectiveness ratio Darv – darvadstrocel; SC –standard care; CPC – clinical and patient centric; MRI – magnetic resonance imaging

5.2.8.5 Using the St Mark's retrospective study data directly in the company's model

Table 27 shows the impact of using data from the St Mark's retrospective study instead of the model base case parameters to inform: (i) the transition probabilities related to salvage therapy, proctectomy and defunctioning surgery health states; (ii) the salvage therapy treatment mix; (iii) maintenance and post-surgery treatment mixes and (iv) health care resource utilisation.

**Table 27: Effect of using data from the St Mark's retrospective cohort study (reproduced from clarification response,[2] question B7, Table 28)**

| Scenario | Total costs | | | Total QALYs | | | *ICER* |
|---|---|---|---|---|---|---|---|
| | Darv | Control | Incremental | Darv | Control | Incremental | |
| **Base case** | [redacted] | [redacted] | £21,639 | [redacted] | [redacted] | 1.05 | £20,591 |
| **St Mark's retrospective data set** | [redacted] | [redacted] | £26,201 | [redacted] | [redacted] | 1.11 | £23,524 |

Darv – darvadstrocel; ICER - incremental cost-effectiveness ratio.

## 5.3 Critical appraisal of the company's submitted evidence

This section presents a critical appraisal of the health economic analysis presented in the CS.

### *5.3.1 Methods for reviewing the company's economic evaluation and health economic model*

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic evaluation and the underlying health economic model upon which this was based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists[39] to critically appraise the company's model and analysis.
- Scrutiny of the company's model by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the model structure, to draw out any unwritten assumptions and to identify any apparent errors in the implementation of the model.
- Examination of the correspondence between the description of the model reported within the CS and the company's executable model.
- Where possible, checking of parameter values used in the company's model against their original data sources.
- The use of expert clinical input to judge the credibility of the company's economic evaluation.

### *5.3.2 Adherence of the company's model to the NICE reference case*

The company's economic model is generally in line with the NICE Reference Case.[10] The ERG notes that the model excludes relevant patient subgroups, which are included in the scope and may be covered by the marketing authorisation. In addition, there is a lack of evidence on repeated administration of darvadstrocel, but the licence does not indicate that darvadstrocel should be a single use treatment. The ERG also notes that analyses presented in the original CS, were out of scope as they discounted at a rate of 1.5% for QALYs and 3.5% for costs.[1] The NICE Methods Guide does not

support differential discounting.[10] In scope analyses using discount rates of 3.5% for both costs and QALYs and 1.5% for both costs and QALYs were provided by the company at clarification.[2]

**Table 28: Adherence of the company's model to the NICE Reference case**

| Element | Reference case | ERG comments |
|---|---|---|
| Defining the decision problem | The scope developed by NICE | The model reflects people with non-active / mildly active luminal Crohn's disease and complex perianal fistulae. However, a subgroup of the patient population whose complex perianal fistulae have more than two internal openings or more than three external openings are not considered within the company's analysis of the available evidence or the company's submitted model. It is unclear whether this missing population is included within the licence population for darvadstrocel (see Section 3.1) |
| Comparator(s) | As listed in the scope developed by NICE | The company's model compares darvadstrocel against standard care surgical interventions combined with associated medical management. |
| Perspective on outcomes | All direct health effects, whether for patients or, when relevant, carers | Health gains accrued by patients are modelled in terms of QALYs gained. |
| Perspective on costs | NHS and PSS | The model takes an NHS and PSS perspective |
| Type of economic evaluation | Cost-utility analysis with fully incremental analysis | The company's economic evaluation takes the form of a cost-utility analysis. The results of the analysis are presented in terms of the incremental cost per QALY gained for darvadstrocel versus standard care |
| Time horizon | Long enough to reflect all important differences in costs or outcomes between the technologies being compared | The model adopts a 40-year time horizon. By this time point, only 38.1% of people have died in each group. |
| Synthesis of evidence on health effects | Based on systematic review | Based on the ADMIRE-CD study, which is the only study of the effectiveness of darvadstrocel in this population at the dose stated in the marketing authorisation. |
| Measuring and valuing health effects | Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults. | Health effects are expressed in QALYs. A vignette study, using time-trade off (TTO) valuations by members of the general public was used to inform HRQoL parameters in the model.<br>EQ-5D data were not available from the ADMIRE-CD trial and mapping from the trial outcomes to the EQ-5D was not considered appropriate by the company. |
| Source of data for measurement of health-related quality of life | Reported directly by patients and/or carers | No. The utility values used in the model were based on vignettes, not a description of HRQoL provided directly by patients. Patients did have input into the health state descriptions. |
| Source of preference data for | Representative sample of the UK population | Yes. The vignette study used a representative sample of the UK population to value the health states using the time |

| | | |
|---|---|---|
| valuation of changes in HRQoL | | trade off method. Patient valuations of the vignettes using TTO methodology were considered in a scenario analysis |
| Equity considerations | An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit | No additional equity rating is applied to estimate QALY gains |
| Evidence on resource use and costs | Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS | Resource components include those relevant to the NHS and PSS. Whilst not explicitly stated in the CS, unit costs are valued in 2016/17 prices |
| Discount rate | The same annual rate for both costs and health effects (currently 3.5%) | The base case in the CS used 3.5% discounting for costs and 1.5% discounting for benefits, as the company claims that Section 6.2.19 of the NICE Methods Guide applies (see Section 5.2.3).[10]<br><br>In response to clarification question B7, the company provided analyses where both health effects and costs are discounted at 3.5% and analyses where both the health effects and costs are discounted at 1.5%. |



*5.3.3 Model validation and face validity check*

The ERG rebuilt the deterministic version of the company's base case model in order to verify its implementation.

Table 29 shows that the ERG's rebuilt model produces very similar estimates of undiscounted life years gained, health gains, costs and cost-effectiveness. This double-programming exercise led to the identification of three minor implementation errors:

i. When estimating the average risk of relapse and the average risk of remission across weeks 104 to 164, to inform the long-term relapse and remission rates, the company divides by 16 instead of 15 cycles.
ii. The per-cycle probability of all-cause mortality was subject to a minor error which led to a small over-prediction of the number of deaths throughout the model time horizon.
iii. The long-term remission rates in the salvage therapy arm were specific to the standard care arm time-to-event function, not the salvage therapy time-to-even function.

**Table 29: Comparison of the company's base case model and the ERG's rebuilt model including PAS and using 3.5% discounting for both cost and QALYs**

| Treatment | Total Life years gained (undiscounted) | Total QALYs | Total costs (with PAS) | ICER (£ per QALY gained) |
|---|---|---|---|---|
| *The company's deterministic base case model* | | | | |
| Darvadstrocel | 36.65 | [redacted] | [redacted] | - |
| Standard care | 36.65 | [redacted] | [redacted] | - |
| Incremental | 0 | 1.05 | £21,639 | £20,591 |
| *The ERG's rebuild of the company's deterministic base case model* | | | | |
| Darvadstrocel | 36.85 | [redacted] | [redacted] | - |
| Standard care | 36.85 | [redacted] | [redacted] | - |
| Incremental | 0 | 1.05 | £21,657 | £20,639 |

QALYs – quality-adjusted life years; PAS – patient access scheme; ICER – incremental cost-effectiveness ratio

Given the results of the rebuild of the company's base case economic model, the ERG is satisfied that the company's model has been implemented without any significant errors.

### *5.3.4 Main issues identified within the critical appraisal*

The main issues identified by the ERG within the ERG's critical appraisal of the company's economic analysis are given in Box 1.

**Box 1: Summary of the issues raised by the ERG in the critical appraisal of the company's cost-effectiveness evidence**

1. Exclusion of relevant patient groups from the economic analysis
2. Possibility of **Error! Reference source not found.**
3. **Error! Reference source not found.**, is justified
4. Wastage of darvadstrocel
5. **Error! Reference source not found.**
6. Concerns regarding the company's expert elicitation exercise to
7. **Error! Reference source not found.**
8. Missing transitions within the model structure
9. The company's approach to identifying HRQoL data from the literature
10. The estimates of utilities from the vignette study
11. Adoption of a 40-year time horizon

5.3.4.1 Exclusion of relevant patient groups from the economic analysis

The EPAR and the final NICE scope relate to the use of darvadstrocel for "… *the treatment of complex perianal fistulas in adult patients with non-active/mildly active luminal Crohn's disease, when fistulas have shown an inadequate response to at least one conventional or biologic therapy. Alofisel [darvadstrocel] should be used after conditioning of fistula*" (EPAR, page 82).[17] The clinical effectiveness evidence used to populate the comparative effectiveness of darvadstrocel compared with standard care in the CS is based solely on the ADMIRE-CD trial (CS,[1] page 77). In this trial, only people whose fistula had two or less internal openings and three or less external openings were eligible for inclusion in the study (CS,[1] page 28). In response to a request for clarification from the ERG (question A4), the company stated "… *SPC [SmPC] for darvadstrocel specifies that 4 vials must be administered for the treatment of up to two internal openings and up to three external openings*" and "*Without further data we cannot be certain that 120 million cells is sufficient to adequately treat disease that is characterised by a greater number of internal and external openings*" (Company's clarification response,[2] question A4, pages 10-11). Whilst the use of darvadstrocel within the clinical trial was consistent with the posology and method of administration described within the SmPC, the ERG is unclear as to whether people with more than two internal openings or more than three external openings would be ineligible for any treatment with darvadstrocel. It is possible that these patients may have some but not all of their fistula treated with a single course of darvadstrocel or they may have multiple courses of treatment over multiple procedures. Neither of these scenarios have been modelled by the company. As such, no evidence is provided on the effectiveness or cost-effectiveness of darvadstrocel in this population.

On the basis of the evidence submitted in the CS[1], the ERG believes that it is not possible to produce a reliable estimate of the cost-effectiveness of darvadstrocel in this excluded population group.

5.3.4.2 Possibility of repeat administrations of darvadstrocel

The ADMIRE-CD trial only tested a single use of darvadstrocel. The company's model is consistent with the single use of darvadstrocel observed in the ADMIRE-CD trial In response to a request for clarification from the ERG (question A1), the company stated *"Although some clinicians believe that Alofisel [darvadstrocel] may be beneficial for retreatment in the following patient groups; (i) partial responders; (ii) responders who have relapsed, there is no current evidence to support this treatment approach.*"(Company's clarification response,[2] question A1). The company's clarification response also states "*Some patients who have responded to Alofisel [darvadstrocel] treatment and achieved healing over a significant period of time may develop a new fistula tract (recurrence). We believe this should be considered as a new fistula and should therefore be treated as such."*(Clarification response[2], question A1)

The ERG notes that there are two key uncertainties with this statement. Firstly, the clinical effectiveness, and consequently the cost-effectiveness, of darvadstrocel upon a repeat administration is unknown. Secondly, it is unclear what is meant by a "significant period of time". Two of the clinical advisors to the ERG believed that darvadstrocel may be reused if the time to relapse was more than two years. The ERG believes that it is not possible to make a reliable estimate of the cost-effectiveness of darvadstrocel for use in treating a fistula which has relapsed following prior darvadstrocel administration, regardless of time since relapse, as it is unknown how effective darvadstrocel will be upon repeat administration. In addition, the ERG notes that the cost-effectiveness of using darvadstrocel for the first time is likely to be affected by the costs of downstream therapies used to treat patients who have relapsed. Therefore, any future use of darvadstrocel may increase the ICER compared to the company's analyses which assume no repeated use.

5.3.4.3 Whether the discounting of costs and QALY at 1.5%, in accordance with Section 6.2.19 of the NICE Methods Guide, is justified

The company claims that darvadstrocel meets the criteria in Section 6.2.19 of the NICE Methods Guide (see Section 5.2.3).[10] These criteria require that: (1) standard care would result in death or a severely impaired quality of life for the population being considered; (2) darvadstrocel would restore this population to near full health over a very long period (usually 30 years), and (3) that the Appraisal Committee is satisfied that the introduction of darvadstrocel does not commit the NHS to significant irrecoverable costs. No quantitative analyses were provided by the company to demonstrate that these criteria had been met. The ERG considers that exploratory analyses should have been conducted by the company in which undiscounted QALYs were presented and compared to undiscounted life years gained so that it can be assessed whether darvadstrocel meets the first and second of these criteria (see Section 5.4).

5.3.4.4 Wastage of darvadstrocel

The EPAR states that darvadstrocel has a shelf life of 48 hours (EPAR,[17] page 75). The ERG has concerns that in clinical practice, some doses of darvadstrocel could be wasted and that this was not accounted for in the company's model. In their clarification response[2] (question B18), the company stated that "… *no wastage was observed for the 107 patients assigned to darvadstrocel*…". As part of their clarification response on this issue, the company presented an additional sensitivity analysis in which 5% wastage for darvadstrocel was assumed; this resulted in an ICER of £15,911 per QALY gained when a 1.5% discount rate was used for both costs and QALYs are used. This compares to a deterministic base case ICER of £15,017 per QALY gained when a 1.5% discount rate for both costs and QALYs are used in the company's model. This is a modest increase in the ICER. One of the clinical advisors to the ERG believed that this represented a high estimate of wastage and that they would expect

that 5% would likely be an overestimate of wastage in clinical practice. Consequently, 5% wastage is likely to represent an upper limit of the impact of wastage in clinical practice on the company's ICER.

5.3.4.5 The company's selection of time to relapse and time to remission time-to-event functions

a) Ignoring the interval censored nature of the data

In the ADMIRE-CD trial, remission and relapse were effectively assessed at 6 week intervals at which the PDAI survey was administered to patients. This raises a concern about interval censoring, as people who report a remission/relapse on the PDAI score may have experienced the remission/relapse at any time between the six-weekly data collection points. As the PDAI score is a key component of CPC remission, this means that interval censoring is potentially a consideration. Interval censoring is a minor issue when the interval between assessments is short compared to the average time to relapse.[40] The ERG is concerned that the CPC time to relapse analysis may not meet this criterion, as the median time to relapse was 12.9 weeks in the standard care arm (see Table 8). Not accounting for the interval censoring is likely to bias the fitted parametric time to event functions. However, it is unclear whether this bias is favourable to darvadstrocel when it is compared to standard care. Consequently, the direction and magnitude of any changes in the ICER due to not adjusting the time to event analyses for interval censoring is unknown. It should be noted that the company's analyses demonstrated that the ICER is highly sensitive to the curve selection for time to relapse for people on darvadstrocel (see Table 24). The ERG considers the parametric time to event functions should have been fitted using interval censoring techniques, as detailed in Chapter 9 of Collett.[40]

b) Method used to extrapolate the time-to-event functions

In the company's model, the fitted statistical models are not used to extrapolate the time-to-event functions beyond two years (see Section 5.2.6). Instead, a time-invariant probability was calculated based on the follow up data at 104 weeks post-baseline to 164 weeks post-baseline (note this includes the 4 week period in which the time-to-event functions were not estimated due to the structural absence of events). The rationale for this is unclear and does not appear to be supported by data or clinical opinion that the hazard rate would change at 104 weeks. Furthermore, it is unclear why the time invariant event hazard used in the extrapolated period should be based on the points of the time-to-event function at 104 and 164 weeks. As such, the ERG does not consider the company's approach to be a reliable estimate of the time-to-event function over the long-term. The ERG notes that mixture cure models may have provided a more plausible long-term fit, given the company's clinical expert advice. However, the company's submitted model would require significant adaptation to use parametric functions over the full model time horizon due to the current model structure having only a limited number of tunnel states (24 tunnel states per health state).

5.3.4.6 Concerns regarding the company's expert elicitation exercise to The company's expert elicitation exercise to estimate the time to relapse and remission for people on third or later line therapies

The ERG notes that there are three key issues when considering the robustness of the evidence generated by the company to estimate the effectiveness of salvage therapy compared to standard care which are: (i) the methodological rigour of the exercise; (ii) the design of the expert elicitation exercise, and (iii) the estimation of uncertainty in the exercise.

The expert elicitation exercise conducted by the company did not follow a formal elicitation protocol (clarification response,[2] question B16). Despite additional information provided during the clarification process, it was unclear what information was presented to the experts at the elicitation exercise. This is a source of uncertainty which is not captured in the economic model. This means that the ERG cannot adequately asses if the estimate of the treatment effect of salvage therapy compared with standard care (both time to relapse and the time to remission) generated from the elicitation process is likely to be robust or unbiased.

The ERG notes that the effectiveness of salvage therapy compared to standard care was only elicited as a HR. The rationale for only eliciting a HR was that the proportional hazards assumption was "…*validated by clinical experts in Europe and the UK. This assumption was originally based on the fact that both control and salvage therapy broadly consisted of the same interventions, those being EUA +/- seton placement with background therapy consisting of antibiotics, immunosuppressants and biologic therapy*."(Company's clarification response,[2] question B16). Given this justification, it is unclear how the assumption of proportional hazards was validated with clinicians and the relevance of the justification provided by company does not appear to support eliciting only a HR. The ERG considers it possible that the most appropriate treatment effect was not elicited within the company's exercise. Consequently, the ICER may not be a robust estimate of the cost-effectiveness of darvadstrocel compared to standard care.

Finally, the ERG notes that uncertainty was not elicited from the company's clinical experts and instead it was assumed that the variance of the HR was equal to 15% of the mean. Several formal elicitation procedures include methods for formally eliciting uncertainty from experts, which capture the magnitude and distribution of the experts' uncertainty.[41] Consequently, the uncertainty in the ICER may have been overestimated or underestimated within the CS. The company provided several exploratory analyses exploring the uncertainty in the HR in response to a request for clarification by the ERG (question B16).[2] The assumed variance in the HR was changed to 30% and 60% of the mean HR and the PSA was rerun, a 1.5% discount rate was assumed for both costs and QALYs. A summary of these results is provided in Table 30. The ICER increased slightly when the variance in the HR was increased,

with the ICER being £15,017 per QALY gained when the variance was 15% of the mean HR increasing to £15,666 per QALY gained when the variance was 60% of the mean HR. Even though the analyses indicate that the ICER is relatively robust to increases in the assumed coefficient of variance in the HRs for salvage therapy versus control, it may be the case that the experts were more uncertain than the scenarios presented by company and their distributions could be different to the one's assumed by the company. It is unclear what direction directly eliciting the uncertainty and the associated probability distribution would move the ICER. However, these sensitivity analyses indicate that any changes in the uncertainty due to following an elicitation process which can capture uncertainty is likely to have only a modest effect on the ICER.

**Table 30: Sensitivity analysis on the assumed hazard ratio for the effectiveness of salvage therapy compared to standard care using a discount rate of 1.5% for both costs and QALYs and including the PAS for darvadstrocel (adapted from clarification response,[2] questions B7, Table 30 and Table 35)**

| Scenario | Incremental costs[a] | Incremental QALYs[a] | ICER (£ per QALY gained) | Probability that darvadstrocel provides the most net benefit at: | |
|---|---|---|---|---|---|
| | | | | £20,000 per QALY gained | £30,000 per QALY gained |
| Base case – variance is equal to 15% of the mean | £21,161 | 1.35 | £15,017 | 0.66 | 0.87 |
| Variance is equal to 30% of the mean | £21,011 | 1.35 | £15,311 | 0.67 | 0.88 |
| Variance is equal to 60% of the mean[a] | £21,140 | 1.35[b] | £15,666 | 0.67 | 0.87 |

a – incremental differences were calculated as the mean value for darvadstrocel – the mean value for standard care; b – recalculated by the ERG, as the reported incremental QALYs were inconsistent with; the difference between the mean QALYs for darvadstrocel and standard care, and; the reported ICER

5.3.4.7 The data used to populate the transitions to the defunctioning and proctectomy health states

The ERG noted that the model outputs for defunctioning surgery and proctectomy do not match the data used to populate the model. These two issues are dealt with separately in the subsequent sections.

*Defunctioning surgery*

The CS, suggests that there is an annual probability of 3.75% for people with complex fistulising Crohn's disease receiving a defunctioning surgery over a median time of 16 years after the person's fistulae first presented.[1] This estimated is based on the exponential curve fitted to the Mueller *et al*. data (see Section 5.2.6). Between year 0 and year 1, the company's model predicts that 1.48% of people in the darvadstrocel arm receive a defunctioning surgery and 1.75% of people in the standard care arm

receive a defunctioning surgery. The reason for the discrepancy is that the data used to populate the model relate to all people with complex fistulising Crohn's disease, whereas this transition probability is only applied to a subset of the population (those patients in the model who are in the severe CSF health state). Consequently, the ERG considers that the company's model underestimates the risk of receiving a defunctioning surgery for those people in the severe CSF health state. This will have an impact on the ICER, as increasing this probability will reduce the time spent in the severe CSF health state and increase the time spent in the post defunctioning health state. This is associated with the potential for patients to have lower or higher utility than severe CSF (see Table 16) and higher health state resource use (see Table 17). The impact of this factor on the ICER is addressed in the ERG's exploratory analyses (see Section 5.5).

*Proctectomy*

The data used in the CS suggests that approximately 20.7% (18/87) people with complex perianal fistulae and Crohn's disease would have a proctectomy after a median time of 6 years since the first presentation of fistulising disease.[27] This data was obtained from the Bell *et al.* prospective study (see Section 5.2.6). The company's model suggests that by year 6 of the company's base case model: 8.5% of people in the darvadstrocel group have received a proctectomy, and; 10.2% of people in the standard care group have received a proctectomy. Similar to the lack of fit to the defunctioning surgery data, the reason for this discrepancy is that the company's model structure only allows patients in the severe CSF and defunctioning surgery health states to transition to proctectomy. The reason for the discrepancy, is that the data used to populate the model relate to all people with complex fistulising Crohn's disease, whereas this transition probability is only applied to a subset of the population (those patients in the model who are in the severe CSF or post-defunctioning health states).

The ERG also notes that some of the assumptions regarding the equal probability of transitioning to the proctectomy health state from the severe CSF and post-defunctioning health states may not be clinically plausible. The company's clinical advisors noted that "… *at least 9 out of 10 defunctioned patients would eventually go on to receive proctectomy; therefore, the rate of proctectomy events derived from Bell et al. (2003) is likely to underestimate the transition probability from the post-defunctioning surgery health state*…". The clinical advisors to the ERG agree that proctectomy is more likely for a patient who has had a defunctioning surgery than a patient who has not. However, the company's model assumes that the probability of transitioning to the proctectomy state is the same for people in the severe CSF and post-defunctioning health states. Consequently, the ERG considers that the model's assumptions do not reflect clinical reality. The impact of all three points on the ICER are explored in the ERG exploratory analysis (see Section 5.5).

5.3.4.8 Missing transitions within the model structure

The ERG noted that the company conducted an analysis of data of 78 patients who presented at St Mark's hospital from 1st January 2008 to July 1st 2017 (CS,[1] Appendix Q). Data were collected at baseline, routine visits and study termination (lost to follow up, transferred to another hospital, or patient death). Transition probabilities to each of the company's health economic model health states were estimated from the data using a statistical Markov multi state model for panel data.[42] The observed data in the CS (Appendix Q, Table 29) suggest that it was possible for people with: a successful defunctioning surgery to transition to an unsuccessful defunctioning surgery state; a successful proctectomy to transition to a unsuccessful proctectomy state; and an unsuccessful proctectomy to a successful proctectomy state.[1] All other transitions in this fitted model are possible either directly (from one health state to another) or indirectly (the patient has to move from one health state, to a second health state, to a third health state) within the company's submitted health economic model. Despite the small sample size of the St Mark's retrospective cohort study (n=78), the ERG considers it inappropriate to assume that these transitions cannot occur (directly or indirectly) in the company's submitted model. The impact on the ICER of using these specific transitions from the St Mark's data set is explored in the ERG's exploratory analyses (see Section 5.5).

5.3.4.9 The company's approach to identifying HRQoL data from the literature

In general, the ERG was satisfied with the company's rationale for not mapping from the ADMIRE-CD outcomes (PDAI, CDAI or IBDQ) to EQ-5D. It was therefore reasonable for the company to look for alternative estimates of HSUVs from published or *de novo* studies. The ERG agrees that none of the studies identified in the company's review of HRQoL studies provide relevant and methodologically robust utility values for inclusion within the company's model. However, the CS does not provide sufficient information to determine whether any relevant studies were discarded from the company's HRQoL review. Specifically, 35 of the 37 included studies appear to have been discarded based on their relevance to the model; without more information, it was not possible for the ERG to determine whether these decisions were reasonable.

5.3.4.10 The estimates of utilities from the vignette study

The ERG notes that the use of utility values obtained from direct valuation of health states vignettes is not consistent with the NICE Reference Case.[10] The ERG considers that the valuations of the vignettes by the general population were closer to the Reference Case requirements than those obtained from the sample of patients with Crohn's disease.

The ERG has some concerns regarding the face validity of some of the estimates obtained from the vignette study. The ERG notes that the clinical experts at the EU Advisory Board felt that the utility values for CSF with severe symptoms were slightly higher than expected (CS,[1]Appendix P) and that

three of seven experts at the UK Advisory Board felt that the utility values for the CSF with mild symptoms state were underestimated (CS,[1] Appendix P ); this issue was also noted by one of the ERG's clinical advisors. In addition, one of the clinical advisors to the ERG believed that the utility values for a successful outcome following surgery were underestimated; this would underestimate the benefits to patients of a successful surgical procedure.

The report by Fountain *et al*. (2017)[38] (which is provided in the CS,[1] Appendix R) assessed the external validity of the estimates derived from the vignettes by comparing them to values reported in the literature from 21 studies. Seventeen of these studies focussed on Crohn's disease and four studies focussed on IBD or UC but reported surgical states which are similar to the surgical states described in this study.[38] Seven of these studies reported values obtained from the EQ-5D (Richards 2001[43], Kuruvilla 2012[44], Casellas 2005[45], Stark 2010,[46] Benedini 2012[47], Casellas 2000[48], Casellas 2007[49]). Fountain *et al*. (2017)[38] conclude that *"all health states valued in [the vignette] study had lower utility estimates than other studies reporting utilities in Crohn's disease; however it is not possible to make direct comparisons due to the lack of data for many of the specific states and conditions included in [the vignette] study"*. The ERG noted in particular, that many of the studies estimating the utility values in patients following surgical intervention gave higher utility estimates than the utilities for those patients with positive surgical outcomes estimated in the Fountain *et al*. vignette study. In particular, in the study by Casellas *et al.*(2000)[48], the EQ-5D estimates for patients in remission following surgery were much closer to those for patients in medically induced remission (median values of 0.87 vs 0.86, respectively in Casellas 2000). This suggests that the benefits to patients of defunctioning or proctectomy surgery may be underestimated in the company's model. However, the ERG accepts that any differences between the utility values obtained in the vignette study and those identified from the literature may be due to differences in the population studied, as few of the studies were specific to patients with mildly or inactive Crohn's disease and complex perianal fistulea. Fountain *et al*.[38] also state, "*Lower utility estimates could have been generated because of use of condition specific vignettes (as opposed to generic measure) that may cause a focussing effect, whereby attention is drawn to health problems that may not be considered as so severe when placed in the context of a broader description of health (Brazier and Tsuchiya, 2010).*[50]" This supports the ERG's concern regarding the use of a non-Reference Case method of measuring utility. The potential impact of this on the ICER is explored in the ERG's exploratory analyses (see Section 5.5)



5.3.4.11 Adoption of a 40-year time horizon

The ERG noted that in the company's submitted model only 38.1% of people in the model are in the death health state at the end of the model's 40-year time horizon. The ERG considers that it is possible that the company's base case model may not capture all important differences in costs and QALYs between darvadstrocel and standard care. The company did submit a scenario analysis in

which, the time horizon was set to 60 years (CS[1] and clarification response[2]). Changing the time horizon to 60 years decreases the ICER from £20,591 per QALY gained in their base case to £19,719 per QALY gained. The ERG considers this to be a more appropriate time horizon, as at this time point 97.0% of people have died in both treatment groups.

### 5.4 Exploratory and sensitivity analyses undertaken by the ERG

*Exploratory analysis 1: Correction of errors.*

Within this analysis, the three programming errors identified during the ERG's rebuild of the company's deterministic base case model (see Section 5.3.3) were rectified

*Exploratory analysis 2: Probability of proctectomy and defunctioning surgery.*

The ERG had concerns about how well the model fitted the data used to populate the transitions to the defunctioning surgery and proctectomy health states (see Section 5.3.4). Two general approaches were taken in this exploratory analysis. In the first approach, the company's model was calibrated using the Solver Excel add-in so that the company's model matched the data sources for the probability of proctectomy (18/87 people received a proctectomy after 6 years) and the probability of defunctioning surgery (average 0.0375 annual probability of receiving a defunctioning surgery after 16 years). This was done for defunctioning surgery (analysis 2a) and proctectomy (analysis 2b) separately and then again for both surgical treatments together (analysis 2c). When both surgical treatments were calibrated, the ERG selected the combination of the two annual probabilities of defunctioning surgery and proctectomy that minimised the company's ICER.

In the second approach (analysis 2d) data presented in the CS (Appendix Q, Table 28) on the yearly probability of transitioning between the model health states observed in the St Mark's retrospective cohort study was used. The data were from 78 consecutive patients with Crohn's disease and complex perianal fistulae from St Mark's Hospital. These transition probabilities were derived by fitting a statistical model called a Markov multi-state model (for panel data) to the data. Further details on this statistical model are given in the CS.[1] The results of this exploratory analysis should be interpreted with caution as: the goodness of fit of the company's statistical model and the follow up duration are unclear. However, the values produced from this analysis of the St Mark's data has a higher risk of receiving proctectomy for someone who has received a defunctioning surgery compared to someone who is the CSF severe health state. This is consistent with advice from the ERG's clinical advisors, who consider that people who have previously had a defunctioning surgery are more likely to have a proctectomy than someone who has not previously has a defunctioning surgery.

A comparison of the company's annual probabilities of proctectomy and defunctioning surgery, to the ones used by the ERG in this exploratory analysis are given in Table 31.

**Table 31: Comparison of three different annual transition probabilities used in the company's base case analysis and those used this exploratory analysis**

| Transition | | Annual probabilities | | |
|---|---|---|---|---|
| **From health state** | **To health state** | **Values used in the company's base case model** | **ERG calibrated values[a]** | **St Mark's retrospective data** |
| CSF severe | Defunctioning surgery | 0.0375 | 0.2929 | 0.1975 |
| CSF severe | Proctectomy | 0.0385 | 0.0797 | 0.1555 |
| Defunctioning surgery | Proctectomy | 0.0385 | 0.0797 | 0.1706 |

ERG –evidence review group; CSF – chronic symptomatic fistulae
a – these values are from the calibration of the company's model to both the proctectomy and defunctioning surgery data

*Exploratory analysis 3: Long-term remission rate for salvage therapy*

The ERG had concerns that the long term rate used to extrapolate the company's curves had a treatment effect applied between the darvadstrocel and standard care groups but did not have a treatment effect applied between the standard care and salvage therapy groups (see Section 5.3.4). This resulted in the long term extrapolation rates being the same for the standard care and salvage therapy groups, whilst the rates differed for the darvadstrocel group. In this sensitivity analysis the ERG amended the long term rates so that the long term rates were based on the salvage therapy time to event functions and not on the standard care time to event functions.



*Exploratory analysis 4: Setting the model time-horizon to 60 years*

As the ERG believes that a longer-term (60 year) time-horizon is more appropriate than the shorter term time horizon applied in the company's base case (40 years). This analysis by the ERG replicates the company's analysis of the model time horizon presented in Table 23.

*The ERG's preferred base case model*

The ERG's preferred base case model combines ERG analyses 1, 2c, 3 and 4. Unless otherwise stated, all subsequent analyses start from the ERG preferred base case analysis and include discounting of 3.5% for both costs and QALYs.

*Exploratory analysis 5: Exploration of the extent to which darvadstrocel restores people with complex perianal fistulae and Crohn's disease to near full health*

The ERG has concerns about whether darvadstrocel meets two of the criteria set out in the NICE Methods Guide for the Committee to consider using discount rates of 1.5%. These are that over a long period of time (usually 30 years): (1) currently people will die or have a very severely impaired quality of life; and (2) the treatment restores these people to full or near full health.

The ERG explored the extent to which darvadstrocel meets these two criteria. In order to do this, the discount rate was set to equal to 0% and the time horizon of the model was set to 30 years. The mean utility value accrued in each treatment group per year was then calculated by dividing the undiscounted QALYs by the undiscounted life years gained. These average utility values accrued per year, were then compared to the highest utility value used in the model (0.865 for the remission health state). As the model utilities were not adjusted for age, a simple division of the mean utility accrued each year by the highest utility value used in the model was conducted to calculate the proportion of the maximum available health gain in each treatment group. This exploratory analysis was conducted with both the ERG's preferred base case and the company's base case model.

*Exploratory analysis 6: Inclusion of missing transitions*

The ERG had concerns that the St Mark's retrospective study indicated that some transitions were possible, yet these were not permitted to occur within the company's submitted model structure (see Section 5.3.4.8). In this sensitivity analysis, three additional transitions were added to the company's model structure based on the four weekly transitions probabilities estimated from the St Mark's retrospective study (CS,[1] Appendix Q, Table 29). These were: successful defunctioning surgery to unsuccessful defunctioning surgery (4-weekly probability 0.03); successful proctectomy to unsuccessful proctectomy (4-weekly probability 0.02), and; unsuccessful proctectomy to successful proctectomy (4-weekly probability 0.05).

*ERG exploratory analysis 7: CSF mild, successful defunctioning surgery and successful proctectomy health states have the same utility value as the remission health state*

The ERG is concerned that the vignette study may have underestimated the utility of people in the CSF mild, successful defunctioning surgery and successful proctectomy health states as the differences between these health states and the remission health states are larger than those observed in other literature (see Section 5.3.4). To provide an upper limit on the effect of under predicting the utility in these health states, the ERG set the utility for these health states equal to those of remission (0.865). This scenario should be interpreted with caution, as it is intended only to inform the direction and maximum magnitude of any changes in the ICER due to the possible under prediction of utility in these three health states. For this reason, it is not incorporated in the ERG's preferred base.

*ERG exploratory analysis 8: Use of different parametric distributions for the time to relapse and time to relapse.*

The ERG has concerns that the company may not have fitted the most appropriate parametric model. In order to explore the impact of alternative functions on the ICER, this analysis replicates the company's sensitivity analysis on the parametric time-to-event functions in the ERG's preferred base case model.

### 5.5 Impact on the ICER of Additional Clinical and Economic Analyses Undertaken by the ERG

The results of each set of exploratory analyses are addressed below. In these analyses, costs and QALYs are discounted at 3.5%, unless otherwise specified. The results of the ERG exploratory analyses using a 1.5% discount rate for costs and QALYs are given in Appendix 4.

*Exploratory analyses 1 to 4*

Table 32 shows the results of the ERG exploratory analyses 1 to 4. Each analysis was conducted individually on the company's base case model. When combined these four exploratory analyses form the ERG's preferred base case, also provided in

Table 32.

Table 32 shows that in the ERG preferred base darvadstrocel is expected to generate an additional 1.01 QALYs at an additional cost of £23,978. The corresponding ICER is £23,176 per QALY gained. This compares to an ICER of £20,591 per QALY gained in the company's base case. The results of each individual change suggest that the key driver of the differences between the ERG's preferred base case and the company's base case is the calibration of the probabilities of proctectomy and defunctioning surgery (i.e. analysis 2c). The other three factors have a modest impact on the ICER.

**Table 32: The results of the ERG exploratory analyses for analysis sets 1 to 4, including the PAS for darvadstrocel**

| Treatment | Total QALYs | Total costs (with PAS) | ICER (£ per QALY gained) |
|---|---|---|---|
| Company's base case | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.05 | £21,639 | £20,591 |
| 1) ERG exploratory analysis – correction of implementation errors | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.05 | £21,666 | £20,700 |
| 2a) ERG exploratory analysis – only proctectomy calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.01 | £23,127 | £22,887 |
| 2b) ERG exploratory analysis – only defunctioning surgery calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.01 | £22,024 | £21,824 |
| 2c) ERG exploratory analysis – proctectomy and defunctioning surgery calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 0.96 | £23,241 | £24,115 |
| 2d) ERG exploratory analysis – proctectomy and defunctioning surgery probabilities were obtained from the St Mark's retrospective cohort study | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 0.95 | £24,530 | £25,530 |
| 3) ERG exploratory analysis – long term remission and relapse rates for salvage therapy are obtained from the salvage therapy arm | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.05 | £21,628 | £20,540 |
| 4) Time horizon is set to 60 years (replication of the company's scenario analysis) | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard Care | [redacted] | [redacted] | - |
| Incremental | 1.10 | £21,706 | £19,719 |
| ERG base case: 1 + 2c + 3 + 4 | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.01 | £23,978 | £23,176 |

QALYs – quality-adjusted life years; PAS – patient access scheme; ICER – incremental cost-effectiveness ratio; ERG – Evidence Review Group

*Exploratory analysis 5: Analysis of the extent that darvadstrocel restores people with complex perianal fistulae and Crohn's disease to near full health*

Table 33 shows that in the ERG's preferred model over a 30-year time horizon; patients in both treatment groups accrue 28.82 life years; patients in the standard care group accrue [redacted] undiscounted QALYs, and; patients in the darvadstrocel group accrue [redacted] undiscounted QALYs. This results in darvadstrocel accruing an average utility of [redacted] per year and standard care accruing an average utility of [redacted] per year. These two values correspond to [redacted] and [redacted] of the utility value for the remission health state, respectively.

The equivalent values using the company's base case model show that over a 30-year time horizon; patients in both treatment groups accrue 28.78 life years; patients in the standard care group accrue [redacted] undiscounted QALYs, and; patients in the darvadstrocel group accrue [redacted] undiscounted QALYs. This results in darvadstrocel accruing an average utility of [redacted] per year and standard care accruing an average utility of [redacted] per year. These two values correspond to [redacted] and [redacted] of the utility value for the remission health state, respectively.



**Table 33: Assessment of the proportion of health achieved in each model arm using the company's and the ERG's base case model over a 30-year time horizon and a 0% discount rate**

| Treatment | Undiscounted life years | Undiscounted QALYs | Mean utility accrued per year | Highest health state utility value | Percentage of maximum health achieved |
|---|---|---|---|---|---|
| Company's base case model | | | | | |
| Standard Care | 28.78 | [redacted] | [redacted] | 0.865 | [redacted] |
| Darvadstrocel | 28.78 | [redacted] | [redacted] | 0.865 | [redacted] |
| ERG's base case model | | | | | |
| Standard Care | 28.82 | [redacted] | [redacted] | 0.865 | [redacted] |
| Darvadstrocel | 28.82 | [redacted] | [redacted] | 0.865 | [redacted] |

QALYs – quality-adjusted life years

On the basis of these results the ERG believes that: (1) the average patient with complex perianal fistulae and Crohn's disease does not have a very severely impaired quality of life when treated with standard care and (2) that darvadstrocel does not restore the average patient with complex perianal fistulae and Crohn's disease to full or near full health. As such, the ERG considers that darvadstrocel does not meet the criteria described in Section 6.2.19 of the guide to the NICE Methods Guide.[10] Consequently, the ERG believes that costs and QALYs should be discounted at a rate of 3.5% for both costs and QALYs.

*Exploratory analysis 6: Inclusion of missing transitions*

Table 34 shows the impact of adding transitions between: (1) the successful and unsuccessful defunctioning surgery health states; (2) the successful and unsuccessful proctectomy health states, and; (3) the unsuccessful and successful proctectomy health states. This suggests that adding these transitions will decrease the ICER to £19,452 per QALY gained from the ERG's base case ICER of £23,176 per QALY gained.

**Table 34: Impact of three additional transitions on the ICER the ERG's base case model, including the PAS for darvadstrocel**

| Treatment | Total QALYs | Total costs (with PAS) | ICER |
|---|---|---|---|
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.11 | £21,655 | £19,452 |

QALYs – quality-adjusted life years; PAS – patient access scheme; ICER – incremental cost-effectiveness ratio

*Exploratory analysis 7: CSF mild, successful defunctioning surgery and successful proctectomy health states have the same utility value as the remission health state*

The results in Table 35 indicate that the ICER for the ERG's preferred base case scenario would increase from £23,176 per QALY gained to £63,721 per QALY gained, if the utilities in the CSF mild, successful defunctioning surgery and successful proctectomy health states were the same as the utilities in the remission health sate. This indicates that applying lower utility values to these three health states produces a more favourable ICER for darvadstrocel, and also that, the ICER is sensitive to changes in the utility values for these health states. Consequently, if the utility values for these health states have been significantly under-predicted, then the ICER may have also been significantly underestimated.

**Table 35: The effect of setting the utility for patients in the CSF mild, successful defunctioning surgery and successful proctectomy health states to the same value as patients in the remission health state, including the PAS for darvadstrocel**

| Treatment | Total QALYs | Total costs (with PAS) | ICER (£ per QALY gained) |
|---|---|---|---|
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 0.37 | £23,738 | £63,721 |

QALYs – quality-adjusted life years; PAS – patient access scheme; ICER – incremental cost-effectiveness ratio

*Exploratory analysis 8*: *The use of different parametric distributions for the time to relapse and time to relapse*

The results of the ERG's exploratory analysis on the base case curve selection is presented in Table 36. These analyses show that the ICER is particularly sensitive to the time to relapse function (Gompertz

distribution). As the ERG is concerned that the time-relapse-function may have been biased due to informative censoring (see Section 5.3.4.5), this analysis indicates that the ICER may be significantly higher or lower than those presented by the ERG and company. The direction of bias will depend on whether the impact of informative censoring is favourable or unfavourable to darvadstrocel.

**Table 36: The effect of changing the time-to-event functions on the ICER in the ERG's base case model, including the PAS for darvadstrocel**

| Time to remission function | Time to relapse function | Total costs | | | Total QALYs | | | ICER |
|---|---|---|---|---|---|---|---|---|
| | | Darv | SC | Incr | Darv | SC | Incr | |
| Gompertz (base case) | Gompertz (base case) | [redacted] | [redacted] | £23,378 | [redacted] | [redacted] | 1.01 | £23,176 |
| Generalised gamma | Gompertz (base case) | [redacted] | [redacted] | £24,033 | [redacted] | [redacted] | 0.82 | £29,200 |
| Gompertz (base case) | Log-normal | [redacted] | [redacted] | £25,,084 | [redacted] | [redacted] | 0.21 | £119,514 |
| Generalised gamma | Log-normal | [redacted] | [redacted] | £25,146 | [redacted] | [redacted] | 0.18 | £143,131 |

Darv – darvadstrocel; SC – standard care; Incr – incremental difference between darvadstrocel and standard care; ICER - incremental cost-effectiveness ratio; QALYs – quality adjusted life years

## 5.6 Conclusions of the cost effectiveness section

The ERG were satisfied that the company's review of published economic evaluations did not exclude any cost-effectiveness studies which were relevant to the scope of this appraisal.

The CS argues darvadstrocel should be assessed using a discount rate of 1.5% for QALYs and 3.5% for costs.[1] The ERG notes that the NICE methods guide specifies that in the Reference Case a discounting rate of 3.5% should be used for both costs and QALYs and that a rate of 1.5% for both costs and benefits may be considered by the Appraisal Committee under specific circumstances.[10] The ERG therefore notes that the use of differential discounting is not supported within the NICE methods guide.

Based on the probabilistic version of the model in the CS (using a discount rate of 1.5% for QALYs and 3.5% for costs) darvadstrocel is expected to generate an additional 1.35 QALYs at an additional cost of £21,773, compared with standard care: the corresponding incremental cost-effectiveness ratio is £16,102 per QALY gained.[1] The deterministic version of the company's model produces a similar ICER of £15,471. At clarification the company's presented additional analyses using: (1) a discount rate of 3.5% for both costs and QALYs and (2) a discount rate of 1.5% for both costs and QALYs.[2] When a discount rate of 3.5% was used for both costs and QALYs, the updated model suggested that darvadstrocel is expected to generate an additional 1.02 QALYs at an additional cost of £21,773, compared with standard care, giving an ICER of £21,417 per QALY gained. The results of the analysis using 1.5% discount rates are presented in Appendix 2.

The ERG critically appraised the company's economic analysis and double programmed the deterministic version of their model. The ERG's critical appraisal identified eleven issues relating to the company's economic analysis and the evidence used to inform it, each of these are addressed in turn.

The ERG believes that a longer-term (60 year) time-horizon is more appropriate than the shorter term time horizon applied in the company's base case (40 years). The ERG considered that the model submitted by the company did not adequately predict the data used in the model for the receipt of defunctioning surgery or proctectomy. The ERG considered that the long term event rates for the salvage therapy arm should have been estimated using the time to event functions for salvage therapy. The ERG's preferred base case analysis addressed these issues, and corrected several minor errors in the company's model. This resulted in a moderate increase in the deterministic ICER from £20,591 per QALY gained in the company's deterministic base case to £23,176 per QALY gained in the ERG's preferred base case

The CS did not include any data on the cost-effectiveness of darvadstrocel for people with complex perianal fistulae and Crohn's disease whose fistulae has more than two internal openings or more than three external openings, however the marketing authorisation does not specifically exclude this population. The ERG considers that an ICER for darvadstrocel cannot be reliably estimated for this population. The marketing authorisation for darvadstrocel does not preclude people who have previously been treated with darvadstrocel receiving another treatment course, however the submitted evidence only relates to a single use of darvadstrocel. The ERG considers that the ICER for darvadstrocel may increase if repeated use were to be included compared to the company's analyses which assume no repeated use.

The ERG considers that a discounting rate of 3.5% should be applied to both costs and QALYs, as per the NICE Reference Case, because the company has not demonstrated that: (1) standard care would result in death or a severely impaired quality of life for the population being considered; and (2) darvadstrocel would restore this population to near full health over a very long period (usually 30 years). The ERG considers that the exploratory analysis on whether darvadstrocel meets the criteria in Section 6.2.19 of the NICE methods guide indicates that these criteria are not met.[10]

The ERG were concerned that in clinical practice doses of darvadstrocel would be wasted, as it has a shelf life of 48 hours. An analysis by conducted by the company in response to a clarification question suggested that wastage would have a minor impact on the ICER. The ERG's advisors indicated that the assumed wastage in the company's analysis was likely to be an upper estimate of what would be observed in clinical practice.

The ERG had concerns that the company's estimated time to event functions did not control for interval censoring, which may bias these functions, and the long term extrapolations were not reliable. The direction and magnitude of any changes in the ICER is unknown, however the company's sensitivity analyses and the ERG's exploratory analyses indicate that the ICER is highly sensitive to the assumed time to event function.

The ERG had two concerns that the expert elicitation exercise: firstly, the exercise was not adequately reported, so the ERG could not assess whether the estimated produced from the exercise were robust or unbiased; and secondly, the exercise did not capture the uncertainty that the experts had in their elicitation. Analyses conducted by the company in response to clarification suggested that different assumed uncertainty in the elicited values had a modest impact on the ICER, but the effect on the ICER of any bias in the elicited values is unknown.

The ERG noted that the company's analysis of the St Mark's dataset suggests that some transition probabilities which were not possible within their model structure, were possible in clinical practice. The ERG explored the effect of adding these transition probabilities to the company's model in an exploratory analysis. This exploratory analysis suggested that adding these transitions would moderately decrease the ICER

The ERG notes that the method used to estimate the utility values incorporated in the economic analysis was not consistent with the NICE reference case and that in general the method used to estimate utilities may influence the values obtained. The ERG were concerned that the utility values applied to some model states may have been underestimated, based both on comparisons made with published estimates and the opinion of clinical experts. The ERG's exploratory analyses suggest that applying higher utility values for those model states that may have been underestimated would tend to increase the ICER, but the ERG was unable to identify a more plausible estimate of utilities than those used by the company.

The ERG considers the following to represent the key uncertainties within the company's health economic analysis:

- The absence of comparative clinical evidence for darvadstrocel versus standard care within people with complex perianal fistulae and Crohn's disease whose fistula has more than two internal openings or more than three external openings.
- The absence of clinical evidence regarding the repeat administration of darvadstrocel.
- The potential introduction of bias in the estimation of the time to event functions, as interval censoring techniques were not applied.

# 6 END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when both the criteria referred to below are satisfied:

1) The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
2) There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

The company does not claim that darvadstrocel meets NICE's end of life criteria. The ERG concurs with this view.

# 7 OVERALL CONCLUSIONS

*Clinical effectiveness*

The efficacy (in terms of combined remission and CPC remission) and safety of a single intralesional injection of darvadstrocel added on to standard of care (compared with placebo sham and standard of care) was positively demonstrated in the ADMIRE-CD study. However, there are a number of limitations and uncertainties in the evidence base which warrant caution in its interpretation. Whilst the study was generally well reported and conducted, a key limitation of the efficacy and safety data for darvadstrocel reported in the CS relates to the post hoc analyses of CPC-remission (an outcome used in the economic model) and CPC relapse.[1] These endpoints were not designed or powered to test formal hypotheses. Another issue is the lack of a confirmatory study. As noted in the EPAR, the effect size in the ADMIRE-CD trial was considered to be modest and less than the 25 percentage difference that it was designed to detect, yet this was considered clinically meaningful given that other treatment options for fistulas had failed.[17] A post-authorisation efficacy and safety trial, ADMIRE-CD-II is expected to help address this concern. However, this study is not expected to be complete until October 2021. The key uncertainties in the clinical evidence for darvadstrocel relate to repeated administration, optimal dosing and long-term efficacy and safety

*Cost-effectiveness*

Notwithstanding uncertainties regarding the statistical analysis of the time to event data and the utilities for the CSF mild, successful defunctioning and successful proctectomy health states, the ERG's preferred base case increases the ICER for darvadstrocel versus standard care from £20,591 per QALY gained to £23,176 per QALY gained. On the basis of an exploratory analysis conducted by the ERG, the ERG does not consider that darvadstrocel meets the criteria in Section 6.2.19 of the NICE Methods Guide.[10] Consequently, the ERG believes that costs and QALYs should both be discounted at a rate of 3.5%. Additional exploratory analyses indicate that including additional transitions in the company's model structure only has a minor impact on the ICER for darvadstrocel versus standard care. Conversely, the selected time to event distributions for time to relapse and time to remission and the utility values for the CSF mild, successful defunctioning surgery and successful proctectomy health states have a significant impact on the ICER for darvadstrocel versus standard care. The ERG notes that no comparative clinical or economic evidence is available for the comparison of darvadstrocel versus standard care in patients with complex perianal fistula and Crohn's disease whose fistula has more than two internal openings and/or more than three external openings. Furthermore, no comparative clinical or economic evidence is available in which repeated administration of darvadstrocel is compared to either the single use of darvadstrocel or standard care.

### 7.1 Implications for research

The ERG considers that future research should be undertaken in the four key areas. Firstly, a confirmatory study that is statistically powered to detect a difference in remission and relapse using the CPC definition should be conducted. Secondly, a study is required to evaluate the optimal dose and treatment duration of darvadstrocel. Thirdly, a study to investigate efficacy and safety of repeat administration of darvadstrocel and administration of darvadstrocel to people with more than two internal openings and/or more than three external openings of their complex perianal fistula is required. Finally, longer term epidemiological studies and clinical experience are required to estimate the long term remission and relapse rates and fully assess the risk of AEs associated with darvadstrocel.

# REFERENCES


1. Takeda UK Ltd. Darvadstrocel for treating complex perianal fistula in Crohn's disease [ID960]. Company's evidence submission. 2018.
2. Takeda UK Ltd. Darvadstrocel for treating complex perianal fistula in Crohn's disease [ID960] Response to Clarification Letter. 2018.
3. Marzo M, Felice C, Pugliese D, Andrisani G, Mocci G, Armuzzi A*, et al.* Management of perianal fistulas in Crohn's disease: an up-to-date review. *World J Gastroenterol* 2015;21:1394-403.
4. Gionchetti P, Dignass A, Danese S, Magro Dias FJ, Rogler G, Lakatos PL*, et al.* 3rd European Evidence-based Consensus on the Diagnosis and Management of Crohn's Disease 2016: Part 2: Surgical Management and Special Situations. *J Crohns Colitis* 2017;11:135-49.
5. National Institute for Health and Care Excellence. Darvadstrocel for treating complex perianal fistula in Crohn's disease: Final scope. London: National Institute for Health and Care Excellence; 2018.
6. Takeda Pharma A/S. Alofisel® (Darvadstrocel) 5 million cells/mL suspension for injection: Summary of Product Characteristics [updated 4 April 2018]. 2018.
7. Panes J, Garcia-Olmo D, Van Assche G, Colombel JF, Reinisch W, Baumgart DC*, et al.* Expanded allogeneic adipose-derived mesenchymal stem cells (Cx601) for complex perianal fistulas in Crohn's disease: a phase 3 randomised, double-blind controlled trial. *Lancet (London, England)* 2016;388:1281-90.
8. de la Portilla F, Alba F, Garcia-Olmo D, Herrerias JM, Gonzalez FX, Galindo A. Expanded allogeneic adipose-derived stem cells (eASCs) for the treatment of complex perianal fistula in Crohn's disease: results from a multicenter phase I/IIa clinical trial. *International journal of colorectal disease* 2013;28:313-23.
9. Panes J, Garcia-Olmo D, Van Assche G, Colombel JF, Reinisch W, Baumgart DC*, et al.* Long-term Efficacy and Safety of Stem Cell Therapy (Cx601) for Complex Perianal Fistulas in Patients With Crohn's Disease. *Gastroenterology* 2018;154:1334-42.e4.
10. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013; 2013.
11. National Institute for Health and Care Excellence. Darvadstrocel for treating complex perianal fistula in Crohn's disease: Draft Scope (pre-referral). London: National Institute for Health and Care Excellence; 2017.
12. Centre for Reviews and Dissemination. Systematic review: CRD's guidance for undertaking reviews in health care.: CRD, University of York; 2009.
13. Dundar Y, Dodd S, Williamson P, Dickson R, Walley T. Case study of the comparison of data from conference abstracts and full-text articles in health technology assessment of rapidly evolving technologies: does it make a difference? *International journal of technology assessment in health care* 2006;22:288-94.
14. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet (London, England)* 2005;365:82-93.
15. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. *The Cochrane Collaboration,* 2011.
16. Loke YK, Golder SP, Vandenbroucke JP. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. *Therapeutic advances in drug safety* 2011;2:59-68.
17. European Medicines Agency (EMA). Assessment Report for Alofisel® (darvadstrocel). Procedure No EMEA/H/C/004258//0000. London: EMA; 2017.
18. TiGenix S.A.U. Adult Allogeneic Expanded Adipose-derived Stem Cells (eASC) for the Treatment of Complex Perianal Fistula(s) in Patients With Crohn's Disease (ADMIRE-CD-II). 2018. https://clinicaltrials.gov/ct2/show/study/NCT03279081 (Accessed 24 May 2018).
19. Dumville JC, Torgerson DJ, Hewitt EE. Reporting attrition in randomised controlled trials. *BMJ (Clinical research ed)* 2006;332:969-71.
20. Brussels: Tigenix. Press release: TiGenix announces positive 52-week Phase III results of Cx601 in complex perianal fistulas in Crohn's disease patients. Press

release.http://tigenix.com/wp-content/uploads/2017/05/f56ddc5245fdfa7.96276707_PR-ADMIRE-CD-52-week-data-070316-ENG-FINAL.pdf
21. TIGENIX S.A.U. ADMIRE-CD Clinical Study Report - Week 24. In. Spain: TIGENIX; 2016.
22. TiGenix S.A.U. Clinical Study Report - Week 52. In. Spain: TIGENIX; 2016.
23. Lindsay J, Punekar YS, Morris J, Chung-Faye G. Health-economic analysis: cost-effectiveness of scheduled maintenance treatment with infliximab for Crohn's disease--modelling outcomes in active luminal and fistulizing disease in adults. *Aliment Pharmacol Ther* 2008;28:76-87.
24. Mueller MH, Geis M, Glatzle J, Kasparek M, Meile T, Jehle EC*, et al.* Risk of fecal diversion in complicated perianal Crohn's disease. *J Gastrointest Surg* 2007;11:529-37.
25. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12:9.
26. Office for National Statistics. National life tables: England and Wales 2013-15. *Online Source* 2017;Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandandwalesreferencetables. .
27. Bell SJ, Williams AB, Wiesel P, Wilkinson K, Cohen RC, Kamm MA. The clinical course of fistulating Crohn's disease. *Aliment Pharmacol Ther* 2003;17:1145-51.
28. Department of Health. NHS Reference costs 2016/17. 2017. https://improvement.nhs.uk/resources/reference-costs/ (Accessed 17/05/2018).
29. Personal Social Services Research Unit. Unit Costs of Health and Social Care 2017. 2017, http://www.pssru.ac.uk/project-pages/unit-costs/.
30. National Institute for Health and Care Excellence. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. London: National Institute for Health and Care Excellence; 2015.
31. National Institute for Health and Care Excellence. Faecal calprotectin diagnostic tests for inflammatory diseases of the bowel. London: National Institute for Health and Care Excellence; 2013.
32. National Institute for Health and Care Excellence. VAAFT for treating anal fistulae. London: National Institute for Health and Care Excellence; 2017.
33. National Institute for Health and Care Excellence. Permacol for treating anal fistulae. London: National Institute for Health and Care Excellence; 2017.
34. British National Formulary. BNF March 2018. 2018. www.bnf.org/ (Accessed 4 June 2018).
35. National Institute for Health and Care Excellence. Infliximab and adalimumab for the treatment of Crohn's disease. London: National Institute for Health and Care Excellence; 2010.
36. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.; 2011.
37. Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. *Value Health* 2007;10:214-20.
38. Fountain DL, L; Singh, J; Cowper, T. Development of health-related utility data in complex perianal fistula in Crohn's disease for Cx601. London, UK: PMHR.
39. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D*, et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013;346:f1049.
40. Collett D. Modelling Survival Data in Medical Research. 3rd Edition edn. Boca Raton, Florida: Taylor & Francis Group; 2015.
41. Oakely JE, O'Hagan A. SHELF: the Sheffield Elicitation Framework (version 3.0). 2016. (http://tonyohagan.co.uk/shelf) (Accessed 7th June 2018).
42. Jackson JH. Multi-State Models for Panel Data: The msm Package for R. *Journal of Statistical Software* 2011;38:1 - 28.

43. Richards DM, Hughes SA, Irving MH, Scott NA. Patient quality of life after successful restorative proctocolectomy is normal. *Colorectal Dis* 2001;3:223-6.
44. Kuruvilla K, Osler T, Hyman NH. A comparison of the quality of life of ulcerative colitis patients after IPAA vs ileostomy. *Dis Colon Rectum* 2012;55:1131-7.
45. Casellas F, Arenas JI, Baudet JS, Fabregas S, Garcia N, Gelabert J*, et al.* Impairment of health-related quality of life in patients with inflammatory bowel disease: a Spanish multicenter study. *Inflamm Bowel Dis* 2005;11:488-96.
46. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflamm Bowel Dis* 2010;16:42-51.
47. Benedini V, Caporaso N, Corazza GR, Rossi Z, Fornaciari G, Cottone M*, et al.* Burden of Crohn's disease: economics and quality of life aspects in Italy. *Clinicoecon Outcomes Res* 2012;4:209-18.
48. Casellas F, Lopez-Vivancos J, Badia X, Vilaseca J, Malagelada JR. Impact of surgery for Crohn's disease on health-related quality of life. *Am J Gastroenterol* 2000;95:177-82.
49. Casellas F, Rodrigo L, Nino P, Pantiga C, Riestra S, Malagelada JR. Sustained improvement of health-related quality of life in Crohn's disease patients treated with infliximab and azathioprine for 4 years. *Inflamm Bowel Dis* 2007;13:1395-400.
50. Brazier J, Tsuchiya A. Preference-based condition-specific measures of health: what happens to cross programme comparability? *Health Econ* 2010;19:125-9.

# APPENDICES

**Appendix 1: The goodness of fit of the company's parametric models to relapse and remission data when remission is defined using the clinical remission criterion**

**Table 37: The AIC and BIC statistics for the different fitted parametric time-to-event functions to the time to remission and relapse using the clincal definition of remission, excluding the piecewise exponential model (adapted from CS Table 35 and Table 41)**

| | **Remission** | | **Relapse** | |
|---|---|---|---|---|
| | **AIC** | **BIC** | **AIC** | **BIC** |
| Exponential | 1156.866 | 1163.463 | 791.794 | 797.774 |
| Weibull | 1127.301 | 1137.196 | 763.665 | 772.636 |
| Gompertz | 1089.373 | 1099.268 | 757.079 | 766.050 |
| Log normal | ***1017.138*** | ***1030.331*** | ***749.776*** | ***758.747*** |
| Log logistic | 1091.477 | 1101.372 | 756.516 | 765.487 |
| Generalised Gamma | Not converged | Not converged | 754.526 | 766.488 |

AIC –Akaike information criterion; BIC – Bayesian information criterion;
Text in ***bold and italics*** indicates the lowest value out of the converged time-to-event functions in each column

**Figure 15: Log cumulative hazard plot for clinical remission data (from clarification response,[2] question B3, Figure 8)**

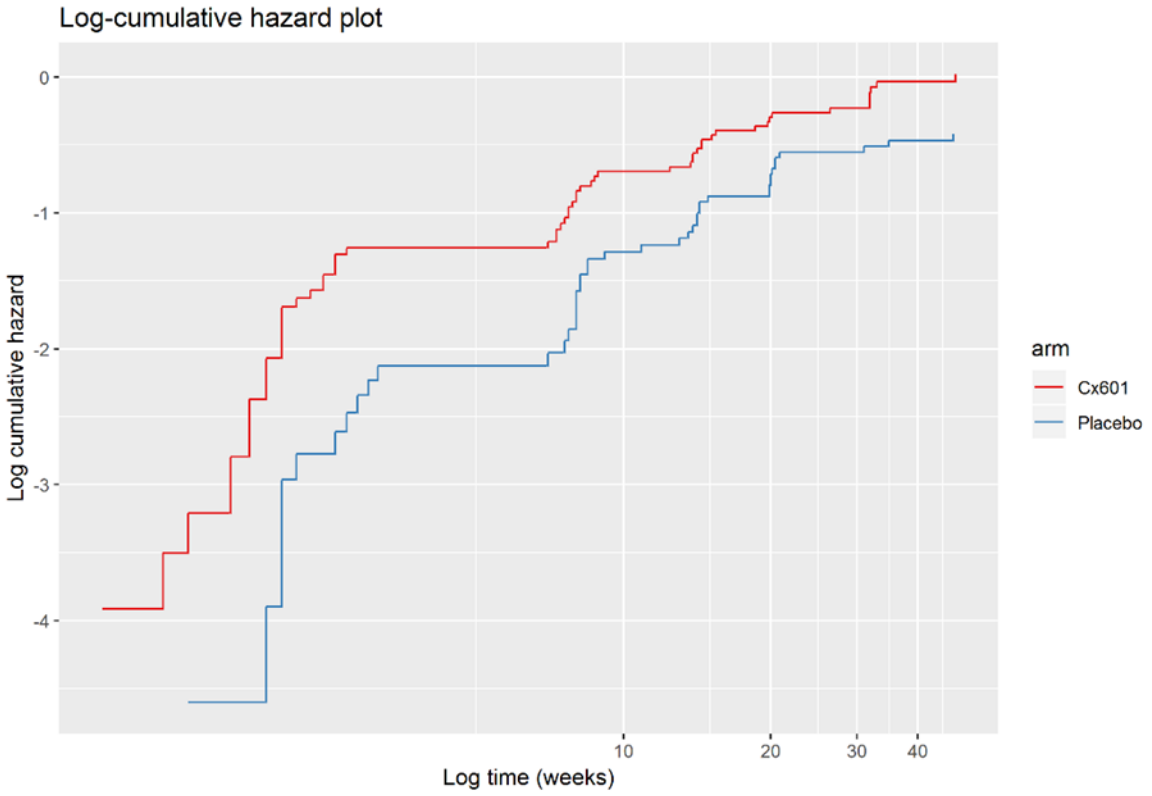


**Figure 16: Empirical hazard plot vs. predicted Gompertz hazards for clinical remission outcome (from clarification response,[2] question B3, Figure 10)**

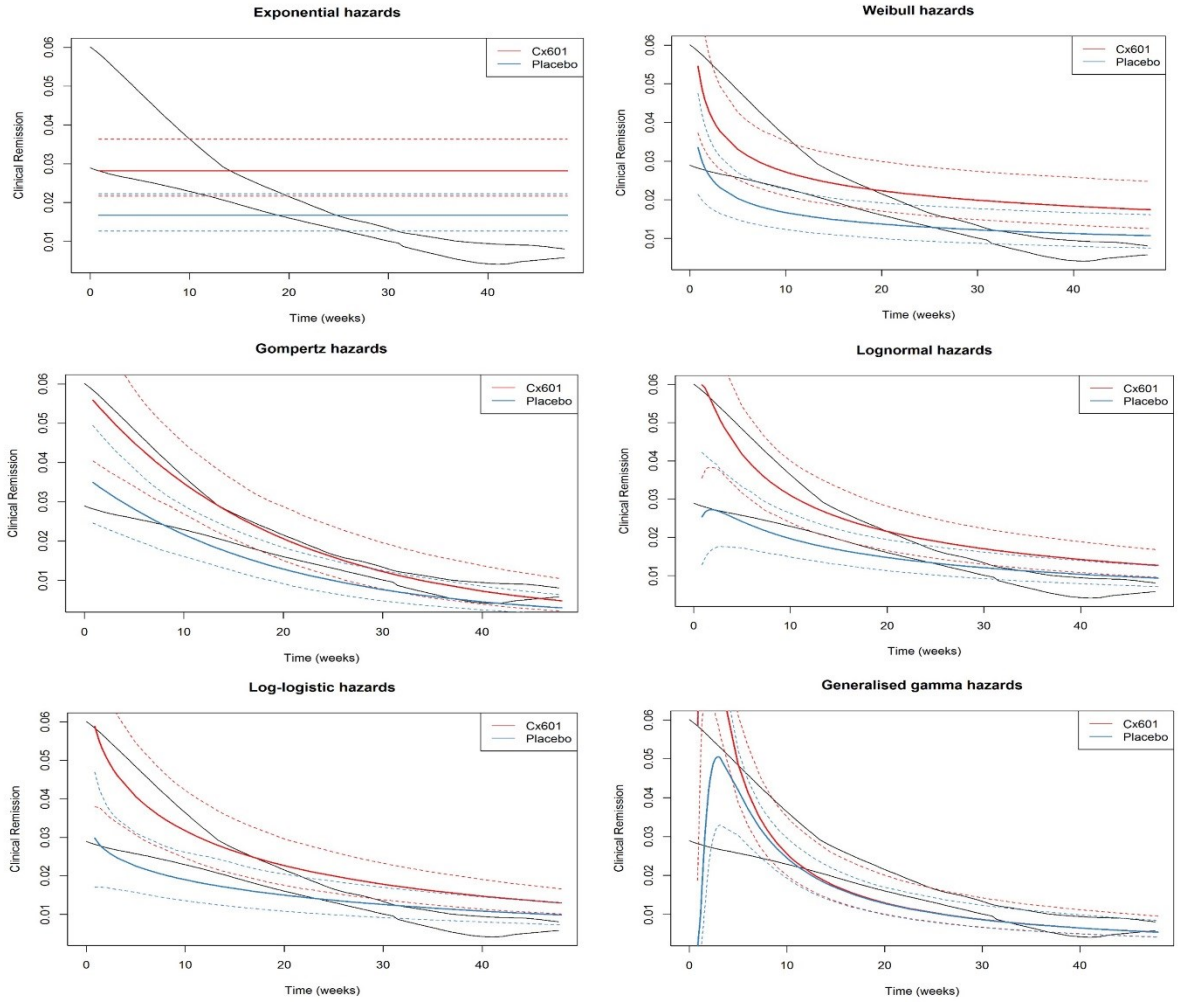



**Figure 17: Log cumulative hazard plot for clinical relapse data (from Clarification response[2], question B3, Figure 15)**

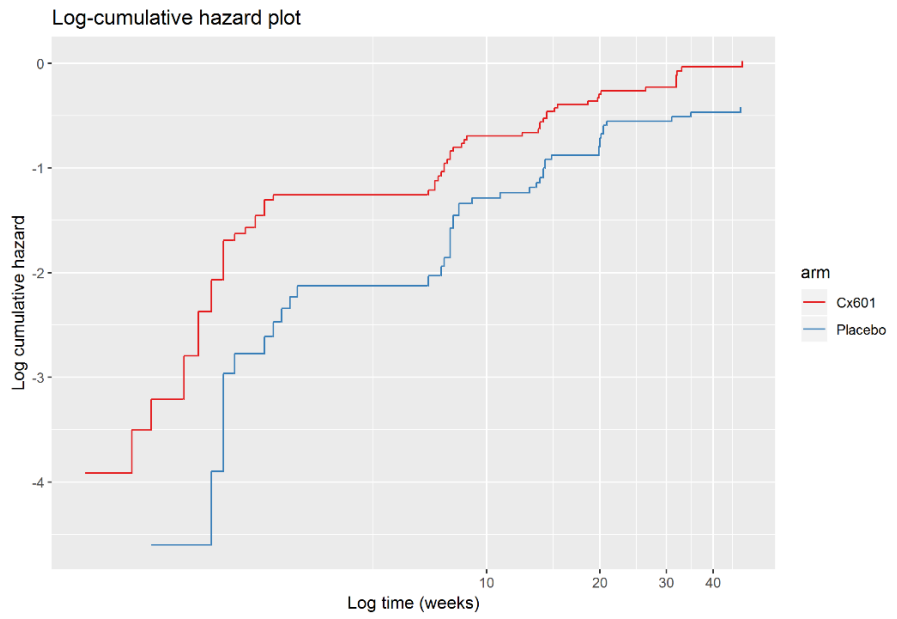
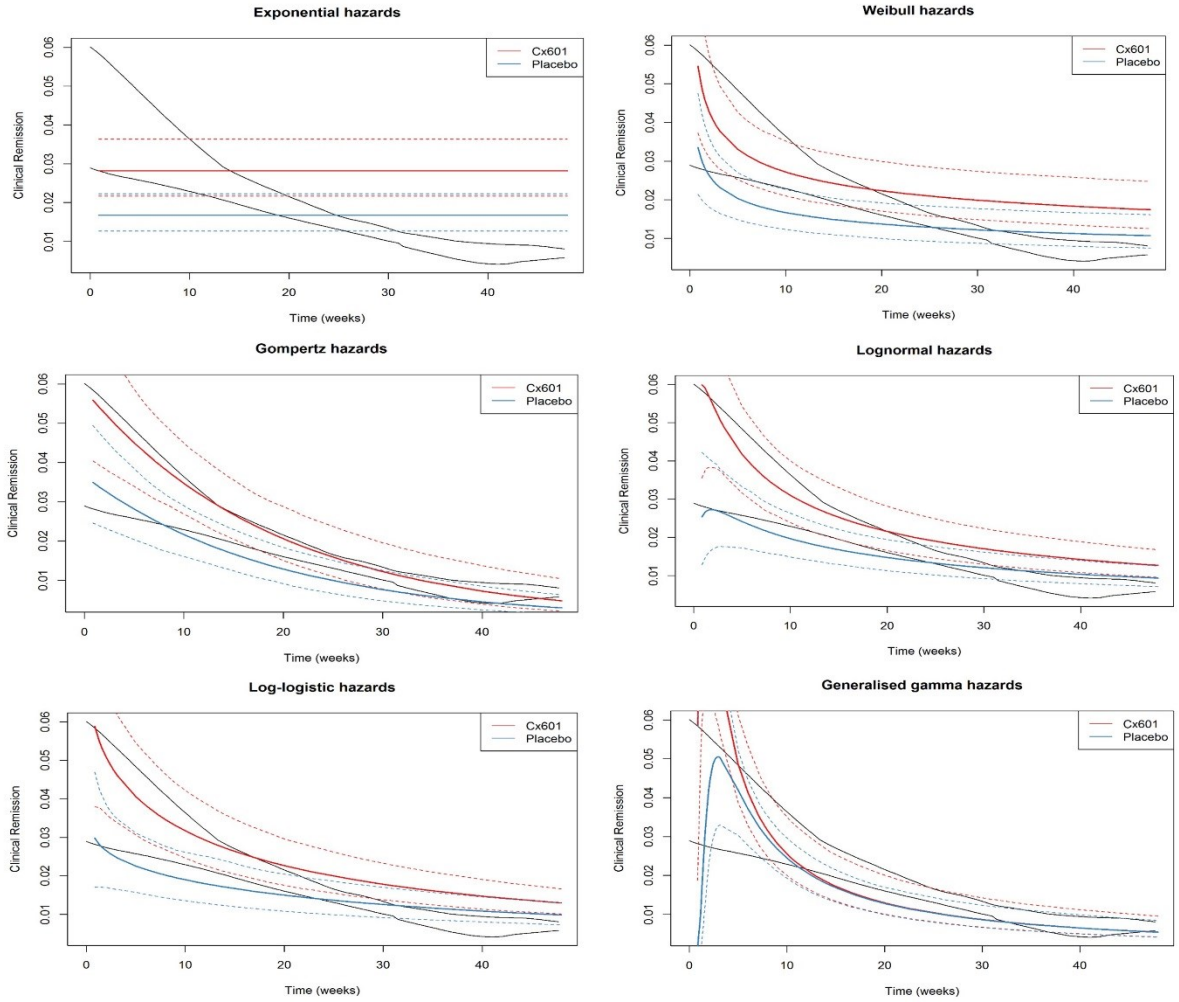


**Figure 18: Empirical hazard plot vs. predicted Gompertz hazards for clinical remission outcome (from clarification response,[2] question B3, Figure 17)**



## Appendix 2: Technical Appendix - The company's results, when a discount rate of 1.5% for both costs and QALYs are used

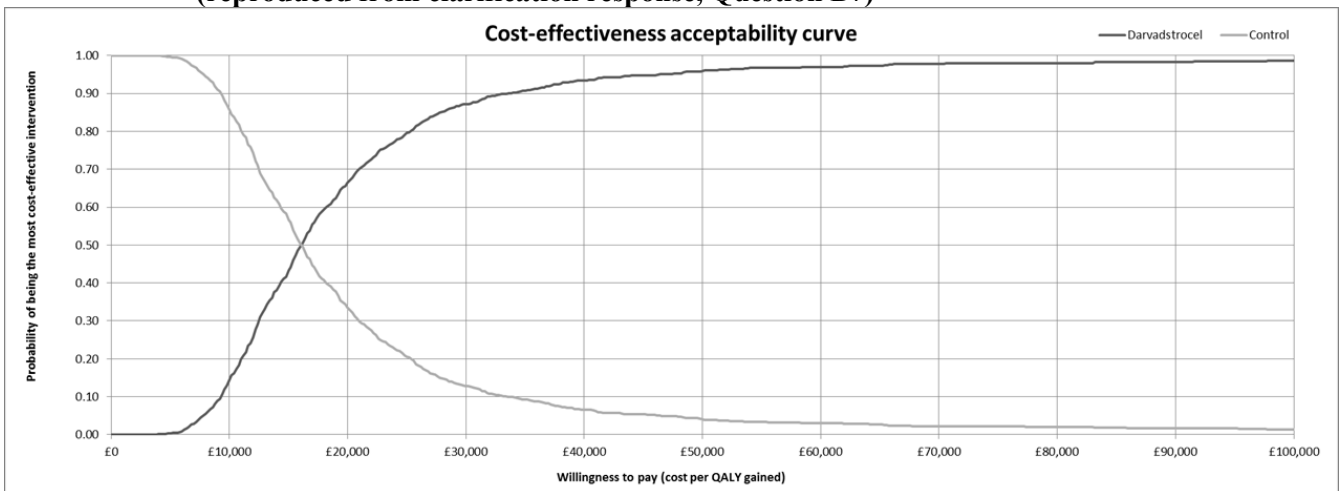
Table 38 shows the results of the company's base case analysis in both the deterministic analysis and the PSA analysis. Based on the PSA version of the company's model, darvadstrocel is expected to generate an additional 1.40 QALYs at an additional cost of £21,004, compared with standard care. The corresponding incremental cost-effectiveness ratio is £15,017 per QALY gained. The deterministic version of the company's model produces a similar ICER of £15,649 per QALY gained.

**Table 38: Company's base case results, including the patient access scheme for darvadstrocel, assuming 1.5% discount rate for both costs and QALYs (adapted from clarification response, question B7)**

| Treatment | Total QALYs | Total costs | ICER (£ per QALY gained) | Probability that the intervention is the most cost-effective at a maximum acceptable ICER of: | |
|---|---|---|---|---|---|
| | | | | £20,000 per QALY gained | £30,000 per QALY gained |
| Probabilistic Sensitivity Analysis | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | 0.66 | 0.87 |
| Standard care | [redacted] | [redacted] | - | 0.34 | 0.13 |
| Incremental | 1.35 | £21,161 | £15,649 | - | - |
| Deterministic | | | | | |
| Darvadstrocel | [redacted] | [redacted] | - | - | - |
| Standard care | [redacted] | [redacted] | - | - | - |
| Incremental | 1.40 | £21,004 | £15,071 | - | - |

QALYs – quality adjusted life years; PAS – Patient Access Scheme; ICER – incremental cost-effectiveness ratio

**Figure 19: Cost effectiveness acceptability curve, including the patient access scheme for darvadstrocel, using a discount rate of 1.5% for both costs and QALYs (reproduced from clarification response, Question B7)**




QALY – quality-adjusted life year

**Figure 20: Company's tornado diagram showing the one way sensitivity analyses conducted by the company using a discount rate of 1.5% for both costs and QALYs (reproduced from clarification response, question B7)**

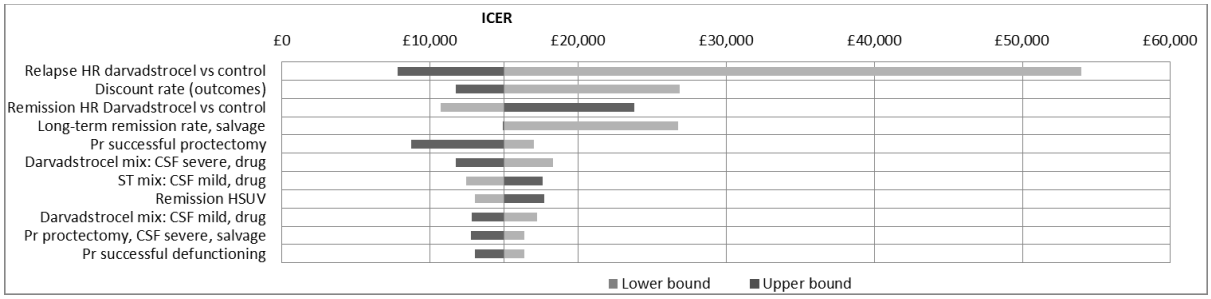



**Table 39: Sensitivity analyses conducted by the company using a discount rate of 1.5% for both costs and QALYs (reproduced from clarification response, question B7, Table 31)**

| Scenario description | Total costs | | | Total QALYs | | | ICER (£ per QALY gained) |
|---|---|---|---|---|---|---|---|
| | Darv | Standard care | Difference | Darv | Standard care | Difference | |
| Base case, 3.5% discount for costs and QALYs | [redacted] | [redacted] | 21,004 | [redacted] | [redacted] | 1.40 | 15,017 |
| 0% discount rate for costs and QALYs | [redacted] | [redacted] | 21,625 | [redacted] | [redacted] | 1.39 | 15,603 |
| 6% discount rate for costs and QALYs | [redacted] | [redacted] | 20,313 | [redacted] | [redacted] | 1.39 | 14,651 |
| 10% annual proctectomy probability post defunctioning | [redacted] | [redacted] | 19,972 | [redacted] | [redacted] | 1.40 | 14,280 |
| 50% annual stoma reversal probability from successful defunctioning state | [redacted] | [redacted] | 20,809 | [redacted] | [redacted] | 1.40 | 14,878 |
| Upper bound of annual stoma care costs (£2,682 per year) | [redacted] | [redacted] | 21,004 | [redacted] | [redacted] | 1.31 | 16,057 |
| Infusion costs halved (£142.25) | [redacted] | [redacted] | 20,922 | [redacted] | [redacted] | 1.43 | 14,676 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| HSUVs based on CD patients vignette study set | [redacted] | [redacted] | 21,323 | [redacted] | [redacted] | 0.92 | 23,191 |
| Relapse HR for salvage therapy vs. control equal to 1.20 | [redacted] | [redacted] | 21,172 | [redacted] | [redacted] | 1.52 | 13,926 |
| Time horizon: 20 years | [redacted] | [redacted] | 15,297 | [redacted] | [redacted] | 1.40 | 10,937 |
| Time horizon: 60 years | [redacted] | [redacted] | 22,254 | [redacted] | [redacted] | 1.40 | 15,911 |
| No inclusion of Biologic usage within salvage therapy (all other assumptions as per base case) | [redacted] | [redacted] | 21,004 | [redacted] | [redacted] | 1.40 | 15,017 |
| Wastage assumed to result in 5% additional cost for darvadstrocel | [redacted] | [redacted] | 21,625 | [redacted] | [redacted] | 1.39 | 15,603 |

QALYs - quality-adjusted life years; ICER - incremental cost-effectiveness ratio; Darv – darvadstrocel; HSUV - health state utility value; CD - Crohn's disease; HR - hazard ratio

**Table 40: Impact of different parametric time-to-event functions on the company's base case ICER using a discount rate of 1.5% for both costs and QALYs (reproduced from clarification response, question B7, Table 33)**

| Time to remission function | Time to relapse function | Total costs | | | Total QALYs | | | ICER |
|---|---|---|---|---|---|---|---|---|
| | | Darv | SC | Incr | Darv | SC | Incr | |
| Gompertz (base case) | Gompertz (base case) | [redacted] | [redacted] | 21,004 | [redacted] | [redacted] | 1.40 | 15,017 |
| Generalised gamma | Gompertz (base case) | [redacted] | [redacted] | 22,316 | [redacted] | [redacted] | 0.99 | 22,432 |
| Gompertz (base case) | Log-normal | [redacted] | [redacted] | 24,952 | [redacted] | [redacted] | 0.25 | 99,339 |
| Generalised gamma | Log-normal | [redacted] | [redacted] | 24,924 | [redacted] | [redacted] | 0.20 | 123,732 |

Darv – darvadstrocel; SC – standard care; Incr – incremental difference between darvadstrocel and standard care; ICER - incremental cost-effectiveness ratio; QALYs – quality adjusted life years

**Table 41: Results of the scenario analyses surrounding the definition of relapse in the company's submitted economic model (reproduced from clarification response, question B7, Table 34)**

| Scenario | Total costs | | | Total QALYs | | | |
|---|---|---|---|---|---|---|---|
| | Darv | Control | Incremental | Darv | Control | Incremental | |
| **Base case** | [redacted] | [redacted] | 21,004 | [redacted] | [redacted] | 1.40 | 15,017 |
| **St Mark's retrospective data set** | [redacted] | [redacted] | 27,893 | [redacted] | [redacted] | 1.51 | 18,529 |

Darv – darvadstrocel; ICER - incremental cost-effectiveness ratio.

## Appendix 3: Technical appendix detailing methods for applying the ERG's exploratory analyses within the company's model

Note when using the company's model, the discount rates for costs and QALYs should be changed to either 3.5% for both or 1.5% for both. To do this change the discount rates in Sheet "Settings", cells E29 and E30.

Exploratory analysis 1

1) Start with the Company's model
2) Go to the sheet "TimeToRemission", cell AC65
3) Change the formula to "=-(LN(1-(AC48-AC63)/AC48)/(COUNT(AC48:AC63)-1))"
4) Paste the formula to cells AF6, AI65, AL65, AO65, AR65
5) Go to the Sheet "TimeToRelapse", cell AA65
6) Change the formula to "=-(LN(1-(AA48-AA63)/AA48)/(COUNT(AA48:AA63)-1))"
7) Drag the formula across to cell AR65
8) Go to the sheet "Patient flow – Darvadstrocel", cells E7:GE7
9) Change the array formula to "=MMULT(E6:GE6,'Transition matrices'!$D$6:$GD$188)*(1-VLOOKUP(ROUNDDOWN(D6,0),Mortality!$R$13:$Y$116,8,FALSE))"
10) Go to cell GF7
11) Change the formula to: "=GF6+(SUM(E6:GE6)*VLOOKUP(ROUNDDOWN(D6,0),Mortality!$R$13:$Y$78,8,FALSE))"
12) Select cells E7:GF7
13) Copy the formulae down to the row 786
14) Go to the sheet "Patient flow-Control", cells E7:GE7
15) Change the array formula to:"=MMULT(E6:GE6,'Transition matrices'!$D$193:$GD$375)*(1-VLOOKUP(ROUNDDOWN(D6,0),Mortality!$R$13:$Y$78,8,FALSE))"
16) Select cell GF7, ant type the formula "=GF6+SUM(E6:GE6)*(VLOOKUP(ROUNDDOWN(D6,0),Mortality!$R$16:$Y$78,8,FALSE))"
17) Select cells E7:GF7
18) Copy the formulae down.

Exploratory analysis 2

*1)* For all parts of exploratory analysis 2, enable the solver add in to Excel, if you have not already done so.

*2a) Proctectomy*

1) Start with the Company's model
2) Go to Sheet "Clinical inputs" cell E128, change the formula to "='Patient flow-Control'!$E$2"
3) Go to Sheet "Clinical inputs" cell E127, change the formula to "='Patient flow-Control'!$E$2"
4) Go to Sheet "Clinical inputs" cell E125, change the formula to "='Patient flow-Control'!$E$2"
5) Go to the sheet Patient flow-Control'
6) Open solver and use the following settings:
    a. Set objective HL$84
    b. To: value of 0.2068965517 (18/87 to 10 dp)
    c. By changing variable cells: $E$2
    d. No constraints
    e. Solving method: GRG Nonlinear



*2b) Defunctioning*

1) Start with the Company's model
2) Go to Sheet "Clinical inputs" cell E111, change the formula to ='Patient flow-Control'!$F$2
3) Go to Sheet "Clinical inputs" cell E113, change the formula to ='Patient flow-Control'!$F$2
4) Go to Sheet "Patient flow-Control"
5) Go to cell G2 and input the following formula "=HK214"
6) Go to cell H2 and input the following formula: "=-(LN(1-G2))/16"
7) Go to cell I2 and input the following formula "=1-EXP(-H2*1)"
8) Set up solver with the following settings
    a. Set objective I2
    b. To: value of 0.03752771 (value given elsewhere in the model for the annual probability of undergoing a defunctioning surgery)
    c. By changing variable cells: $F$2
    d. Constraints: $F$2 ≤ 1
    e. Solving method: GRG Nonlinear

2c)

1) Start with the Company's model
2) Do 2a, steps 1 to 3
3) Do 2b, steps 1 to 6

4)

5) Run solver with the following settings
    a. Set objective I2
    b. To: value of 0.03752771 (value given elsewhere in the model for the annual probability of undergoing a defunctioning surgery)
    c. By changing variable cells: $E$2:$F$2
    d. Constraints: HL$84 = 0.2068965517; $F$2 ≤ 1; $F$2 ≥ 0; $E$2 ≤ 1; $E$2 ≥ 0
    e. Solving method: GRG Nonlinear
6) Put the following formula in cell J2 "=dICER"
7) Run a new solver with the following settings
    a. Set objective J2
    b. To: Min
    c. By changing variable cells: $E$2:$F$2
    d. Constraints: I2 = 0.03752771; HL$84 = 0.2068965517; $F$2 ≤ 1; $F$2 ≥ 0; $E$2 ≤ 1; $E$2 ≥ 0
    e. Solving method: Evolutionary

2d)

1) Start with the Company's model
2) Go to Sheet "Clinical Inputs", cell E111 & cell E113, set the formula to "=0.111609+0.085896"
3) Go to Sheet "Clinical Inputs", cell E125 & cell E127, set the formula to "=0.118666+0.036848""
4) Go to Sheet "Clinical Inputs", cell E128. Set the formula to "=E116*(0.041258+0.022391)+E117*(0.228007+0.117161)"

ERG exploratory analysis 3

1) Start with the Company's model
2) Go to Sheet "TimeToRemission", insert new columns AD, AH, AL, AP, AT, AX
3) In cell AD23 type the formula "=AC23^'Clinical inputs'!$E$68"
4) Copy the formula down to row 63
5) Copy the formula in AD23 and paste into the cells AH23, AL23, AP23, AT23, AX23
6) Copy these new formulae down to row 63
7) In cell AC65 change the formula to "=-(LN(1-(AD48-AD63)/AD48)/(COUNT(AD48:AD63)))"
8) Copy the formula in cell AC65 and paste into cells AG65, AK65, AO65, AS65, AW65
9) Go to Sheet "TimeToRelapse", insert new columns AD, AH, AL, AP, AT, AX

10) In cell AD23 type the formula "=AB23^'Clinical inputs'!$E$95"
11) Drag the formula down to row 63
12) Copy the formula in cell AD23 and paste it to cells AH23, AL23, AP23, AT23, AX23
13) Copy the formulae down to row 63
14) Go to cell AC65 and change the formula to "=-(LN(1-(AD48-AD63)/AD48)/(COUNT(AD48:AD63)))"
15) Copy the formula in cell AC65 and paste to cells AG65, AK65, AO65, AS65, AW65

ERG exploratory analysis 4

1) Start with the Company's model
2) Go to Sheet "Settings", cell E18 and change the value to 60

ERG preferred base case

1) Follow the steps in ERG exploratory analysis 1
2) Follow the steps in ERG exploratory analysis 2c
3) Follow steps 1 to 5 in ERG exploratory analysis 3
4) In Sheet "TimeToRemission", cell AC65 change the formula to "=-(LN(1-(AD48-AD63)/AD48)/(COUNT(AD48:AD63)-1))"
5) Follow steps 7 to 12 in ERG exploratory analysis 3
6) In Sheet "TimeToRelapse" change the formula to "=-(LN(1-(AD48-AD63)/AD48)/(COUNT(AD48:AD63)-1))"
7) Follow step 14 in ERG exploratory analysis 3
8) Follow the steps in ERG exploratory analysis 4

ERG exploratory analysis 5

1) Start with the ERG preferred base case or the Company's base case model (as appropriate)
2) Go to Sheet "Settings", go to cells E29 and E30 and set the value to 0
3) Go to cell E18 and set the value to 30
4) Go to Sheet "Results", go to cell E49 and input the formula "=E42/E48"
5) Go to cell E50 and input the formula "=E49/'Clinical inputs'!$E$185"
6) Copy cells E49:E50, paste the formulae into cells G49:G50

ERG exploratory analysis 6

1) Start with the ERG preferred base case
2) Go to Sheet "transition matrices", cell GB 185, input the value 0.031640929
3) Go to cell GD187, input the value 0.016770373
4) Go to cell GC188, input the value 0.048945715

5) Go to cell GB372, input the value 0.031640929
6) Go to cell GD374, input the value 0.016770373
7) Go to cell GC375, input the value 0.048945715

ERG exploratory analysis 7

1) Start with the ERG preferred base case
2) Go to Sheet “Clinical inputs”, go to cell E186 and input the formula “=$E$185”
3) Copy cell E186
4) Paste the formula into cells E190 and E192

**Appendix 4: Technical appendix detailing the results of the ERG exploratory analyses when a discount rate of 1.5% for both costs and QALYs are used**

**Table 42: The results of the ERG exploratory analyses for analysis sets 1 to 4, including the PAS for darvadstrocel when a discount rate of 1.5% for both costs and QALYs is used**

| Treatment | Total QALYs | Total costs (with PAS) | ICER (£ per QALY gained) |
|---|---|---|---|
| Company's base case | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.40 | £21,004 | £15,017 |
| 1) ERG exploratory analysis – correction of implementation errors | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.39 | £21,046 | £15,117 |
| 2a) ERG exploratory analysis – only proctectomy calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.34 | £23,155 | £17,231 |
| 2b) ERG exploratory analysis – only defunctioning surgery calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.35 | £21,548 | £16,015 |
| 2c) ERG exploratory analysis – proctectomy and defunctioning surgery calibrated | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.28 | £23,315 | £18,152 |
| 2d) ERG exploratory analysis – proctectomy and defunctioning surgery probabilities were obtained from the St Mark's retrospective cohort study | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.27 | £24,665 | £19,465 |
| 3) ERG exploratory analysis – long term remission and relapse rates for salvage therapy are obtained from the salvage therapy arm | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.40 | £20,988 | £14,973 |
| 4) Time horizon is set to 60 years (replication of the company's scenario analysis) | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard Care | [redacted] | [redacted] | - |
| Incremental | 1.52 | £21,172 | £13,926 |
| ERG base case: 1 + 2c + 3 + 4 | | | |
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.40 | £23,639 | £16,198 |

QALYs – quality-adjusted life years; PAS – patient access scheme; ICER – incremental cost-effectiveness ratio; ERG – evidence review group

ERG exploratory analysis 5

No change, as this exploratory analysis is based on undiscounted costs and QALYs

ERG exploratory analysis 6

**Table 43: Impact of three additional transitions on the ICER the ERG's base case model, including the PAS for darvadstrocel**

| Treatment | Total QALYs | Total costs (with PAS) | ICER |
|---|---|---|---|
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 1.53 | £31,352 | £13,922 |

ERG exploratory analysis 7

**Table 44: The effect of setting the utility for patients in the CSF mild, successful defunctioning surgery and successful proctectomy health states to the same value as patients in the remission health state, including the PAS for darvadstrocel**

| Treatment | Total QALYs | Total costs (with PAS) | ICER (£ per QALY gained) |
|---|---|---|---|
| Darvadstrocel | [redacted] | [redacted] | - |
| Standard care | [redacted] | [redacted] | - |
| Incremental | 0.48 | £23,639 | £49,610 |

ERG exploratory analysis 8

**Table 45: The effect of changing the time-to-event functions on the ICER in the ERG's base case model, including the PAS for darvadstrocel**

| **Time to remission function** | **Time to relapse function** | **Total costs** | | | **Total QALYs** | | | **ICER** |
|---|---|---|---|---|---|---|---|---|
| | | **Darv** | **SC** | **Incr** | **Darv** | **SC** | **Incr** | |
| Gompertz (base case) | Gompertz (base case) | [redacted] | [redacted] | £23,639 | [redacted] | [redacted] | 1.40 | £16,918 |
| Generalised gamma | Gompertz (base case) | [redacted] | [redacted] | £34,627 | [redacted] | [redacted] | 1.15 | £21,487 |
| Gompertz (base case) | Log-normal | [redacted] | [redacted] | £25,342 | [redacted] | [redacted] | 0.22 | £113,960 |
| Generalised gamma | Log-normal | [redacted] | [redacted] | £25,470 | [redacted] | [redacted] | 0.19 | £134,063 |

Darv – darvadstrocel; SC – standard care; Incr – incremental difference between darvadstrocel and standard care; ICER - incremental cost-effectiveness ratio; QALYs – quality adjusted life years