Addendum to ASTIClite Statistical Analysis Plan Version 1

Study Title	Autologous Stem cell Transplantation In refractory Crohn's disease – Low Intensity Therapy Evaluation (ASTIClite)
Funding body	**NIHR EME** (project number 15/178/09)

1.1 Background

The ASTIClite trial was paused on 30th December 2019 whilst a number of serious and unexpected adverse reactions were investigated. Two further serious and unexpected adverse reactions were reported in May 2020. In June 2020, the DMEC and TSC held a joint meeting to discuss the unexplained adverse events, the outcomes of the trial team's investigations and the impact of the coronavirus pandemic. The DMEC and TSC were in agreement that;

- recruitment to the trial should cease,
- patients that were in screening, or randomised but had not yet received treatment should be withdrawn,
- patients that had completed the intervention or were randomised to usual care should continue to be followed up as normal.

In addition, the coronavirus pandemic has impacted the ability of trial sites to conduct all aspects of patient visits within the pre-specified time point. Some visits have been delayed and some trial procedures that require attendance in hospital have been omitted.

The last patient last visit is due in November 2020 and the database is due to be finalised in December 2020.

1.2 The Addendum

In light of the early termination and the reduced sample size, together with the possible missing visit data due to COVID-19, the original analyses have largely been scaled back to descriptive summaries. The changes are described in more detail below but two general changes are as follows:

- Additional summaries which present outcome data in relation to treatment received and incidence of SUSAR, in addition to randomised treatment group. This will be done alongside the usual ITT analysis. The rationale for this is the importance of understanding the mechanism for treatment response and for harm attributable to the treatment itself
- A wider visit window will be considered, based on i) the extent of missing data within the existing time window and ii) the availability within a reasonable time window beyond this . The decision will be determined following a blinded data review prior to database lock. Wider visit windows will be allowed for longer term (24 and 48 week) clinical, radiological and endoscopic outcomes. In cases where follow-up is outside and after the visit window but the participant has shown a positive response, the participant will be assumed to be a responder at that timepoint.

The following are proposed changes to the intended statistical analysis as specified in ASTIClite Statistical Analysis Plan (SAP) version 1, which was approved in April 2019. For any SAP section that is not listed, the analysis outlined in that section will remain unchanged.

SAP Section	Change
6.2 Definition of the analysis population	A third analysis population will be included in addition to the ITT and PP populations pre-specified in the SAP. This is an extended ITT population which includes all randomised patients *including* those found to be ineligible post randomisation. All safety summaries and mechanistic analyses will be presented on the extended ITT population in order to retain as much useful information as possible, particularly in relation to harm. Reporting of efficacy outcomes specified in section 7.6 will be presented on the ITT set, with presentation of primary outcome results repeated for the PP set.
7.6.1. Analysis of the primary outcome	The primary outcome will be analysed as specified in the SAP (using a mixed effects logistic regression model) assuming the model will converge etc with the low number of observations. This will be used to provide an estimated odds ratio for absence of ulceration in HSCTlite compared to control with a corresponding 95% CI. P-values will not be reported. **Sensitivity analyses** The following two sensitivity analysis will be conducted 1. Removing participants that had surgery for Crohn's in the usual care group. 2. Using 24 week colonoscopy data to impute missing 48 week primary outcome. The other sensitivity analyses will not be conducted (CACE, removing unrelated deaths, adjusting for baseline missingness predictors, multiple imputation, worst case imputation). The primary outcome will be summarised for the PP group alongside the ITT analysis but a model will not be fit.
7.6.2 Secondary outcomes	No statistical models will be fit on any secondary outcomes, and as such, no treatment differences and confidence intervals will be reported for any secondary outcomes. The secondary outcome data will be summarised by group using tables and spaghetti plots.
7.8 Re-introduction of anti-TNF	Due to low numbers of patients receiving re-introduction of anti-TNF, any details regarding the re-introduction of anti-TNF and subsequent disease activity will be reported descriptively in the text.
7.9 Late effects of HSCT	Late affects will be reported as part of adverse event summaries.
7.10 Subgroup and moderator analysis	Subgroup analysis will not include any modelling or interaction statistical tests. Instead, the primary outcome will be summarised separately by treatment and subgroup for the four pre-specified subgroups. For the moderator analysis, graphical displays of the pre-specified covariate against outcome will be produced but effect size and 95% confidence intervals will not be reported.
7.11 Mechanistic evaluation	No formal mediation analyses will be undertaken. The summaries will comprise descriptive and graphical displays of the markers among three subgroups: -participants that received HSCT and experienced a SUSAR -participants that received HSCT and did not experience a SUSAR -participants that did not receive HSCT (either by randomisation or by treatment switch/withdrawal) This may include assessing differences in mechanistic data between patients undergoing HSCT who met the primary endpoint and patients undergoing HSCT who did not meet the primary endpoint.

ASTIClite Statistical Analysis Plan Version 1

Study Title	Autologous Stem cell Transplantation In refractory Crohn's disease – Low Intensity Therapy Evaluation (ASTIClite)
Funding body	**NIHR EME** (project number 15/178/09)

Table of contents

List of abbreviations used

AE	Adverse Event
ATG	Anti-thymocyte globulin
CD	Crohn's Disease
CDAI	Crohn's Disease Activity Index
CACE	Complier average causal effect
CI	Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
CTRU	Clinical Trials Research Unit
DMEC	Data Monitoring and Ethics Committee
EME	Efficacy and Mechanism Evaluation
EQ-5D	EuroQol Five Dimensions Questionnaire
G-CSF	Granulocyte-colony stimulating factor
HIV	Human Immunodeficiency Virus
HSCT	Haematopoietic stem cell transplantation
IBD	Inflammatory Bowel Disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICC	Intraclass correlation coefficient
ITT	Intention to Treat
MaRIA	Magnetic resonance index of activity
MRI	Magnetic resonance imaging
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NHS	National Health Service
NIHR	National Institute for Health Research
NIMP	Non-Investigational Medicinal Product
Qol	Quality of life
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SES CD	Simple Endoscopic Score for Crohn's Disease
SD	Standard Deviation
SOP	Standard Operating Procedure
TMG	Trial Management Group
TNF	Tumour Necrosis Factor
TSC	Trial Steering Committee
WPAI	Work Productivity and Activity Impairment

Summary Table

Trial title	Autologous Stem cell Transplantation In refractory Crohn's disease – Low Intensity Therapy Evaluation (ASTIClite)
Trial design	Open label, superiority, multicentre, parallel group, randomised controlled trial with internal pilot
Trial participants	Patients aged 18-60 with a diagnosis of CD for at least six months, refractory to biological therapy
Sample size	99
Follow-up	48 weeks past day 0 (day 0: stem cell infusion for HSCT-lite group; 49 days post randomisation for usual care group)
Internal pilot	Ability to recruit to target will be assessed at month 10 of recruitment with STOP/GO criteria set at 60% of the anticipated recruitment rate and to ensure that HSCTlite achieves adequate mobilisation without causing a flare up of CD activity prior to conditioning (reported in other conditions)
Primary analysis	The proportion of patients with absence of ulceration at 48 weeks will be compared between allocated groups using mixed effects logistic regression controlled for study centre as a random effect.
Secondary analyses	Secondary outcomes (clinical disease activity and quality of life) will be compared between groups using parametric regression models as appropriate to the distribution of the outcome.

2 Introduction, study design and key trial objectives

This statistical analysis plan (SAP) is written in conjunction with the International Conference on Harmonisation topic E9 (Conference et al., 1999), guidance for the content of SAPs in clinical trials (Gamble et al., 2017), applicable statistical standard operating procedures from the University of Sheffield Clinical Trials Research Unit (CTRU) and trial documents (Protocol v6 and Data Validation Specification). The trial will be conducted in accordance with Good Clinical Practice in Clinical Trials (ICH Harmonised Tripartite Guideline, 1996) and Medicine for Human Use (Clinical Trials) Regulations (UK Statutory Instruments, 2004).

This SAP will guide the Trial Statistician during the statistical analysis of all quantitative outcomes in order to answer the objectives of the study. It excludes the health economic evaluation (which will be described elsewhere if undertaken).

All analysis will be performed in a validated statistical software package such as STATA version 15 (StataCorp, 2017).

2.1 Study outline

The ASTIClite study is a multicentre, parallel group, randomised controlled trial to evaluate the efficacy of HSCTlite compared with standard care at inducing regression of ileo-colonic ulceration in patients with refractory CD. Participants will be recruited from eight sites that have tertiary referral IBD clinics

and randomised to receive either autologous stem cell transplantation using the HSCTlite regimen, or standard care, in the ratio 2:1, and the primary endpoint will be assessed at week 48.

An internal pilot will be incorporated to confirm whether the HSCTlite mobilisation regimen delivers effective stem cell harvest without a flare up of CD activity. The DMEC will assess efficacy and safety of the HSCTlite mobilisation regimen after 10 patients and subsequently at each DMEC meeting. Should the protocol fail to mobilise 2 x 10_6/kg CD34+ cells (haematopoietic stem and progenitor cells) in more than 10% patients, or if greater than 10% patients experience a disease flare up (increase in Harvey Bradshaw Index of >30% from baseline associated with a rise in CRP) during mobilisation, a protocol amendment will be submitted to modify the mobilisation regimen for subsequent patients. Ability to recruit to target will be assessed at month 10 of recruitment with STOP/GO criteria set at 60% of the anticipated recruitment at that time.

The study is funded by the National Institute for Health Research (NIHR) Efficacy and Mechanism Evaluation (EME) programme. Barts Health NHS Trust will act as the sponsor for this study.

2.2 Study objectives

Internal pilot objectives

1. To assess the feasibility of trial recruitment.
2. To assess efficacy and safety of HSCTlite.

Primary objective

To assess the efficacy of HSCTlite compared to standard care at inducing regression of ileo-colonic ulceration in patients with refractory CD at week 48.

Secondary objectives

1. To assess whether low dose cyclophosphamide and G-CSF is a safe and effective mobilisation regimen for patients with refractory CD.
2. To assess the impact of HSCTlite on clinical disease activity, quality of life and adverse events compared to standard care.
3. To assess the safety and efficacy of anti-TNF therapy in patients who demonstrate endoscopic disease recurrence at week 24 after HSCTlite.

Mechanistic objectives

1. Intestinal MRI will be performed to determine the early impact on mucosal disease at week 4.
2. Immune profiling of peripheral blood and mucosal biopsies will:

a. Characterise immune re-constitution after HSCT, and assess impact of HSCT on disease activity
b. Assess immunological events that precede disease recurrence post HSCT
c. Assess the mechanism of restoration of responsiveness to anti-TNF therapies
d. Serum will be stored for future assessment of response to vaccination post HSCT

3 Outcome measures

Internal pilot endpoint

The endpoint for the pilot phase is a recruitment STOP/GO criterion

1. 60% of the anticipated recruitment rate met at 10 months (15 participants across 8 centres).

Adequate mobilisation achieved by HSCTlite will be assessed by the DMEC at month 10 to ensure:

a. Mobilisation of 2 x 10^6/kg CD34+ cells (haematopoietic stem and progenitor cells) in more than 10% patients;
b. or if less than 10% patients experience a disease flare up (increase in Harvey Bradshaw Index of >30% from baseline associated with a rise in CRP) during mobilisation

A protocol amendment will be submitted to modify the mobilisation regimen for subsequent patients if mobilisation does not achieve these criteria.

Primary outcome

1. Proportion of patients with absence of ulceration on endoscopic assessment at the 48 week assessment. Full definition of treatment success can be found in Section 10.1.

Secondary outcomes

Clinical endpoints

1. Clinical remission (CDAI <150)
2. Steroid free clinical remission (CDAI <150)
3. Clinical remission (Harvey Bradshaw Index ≤4)
4. Clinical remission (PRO2 – mean scores taken from 7 days of data – abdominal pain ≤1, stool frequency ≤1.5)
5. Absolute CDAI at week 48
6. Absolute SES CD at week 48
7. Change in CDAI and SES CD between baseline and week 48
8. Proportion of patients in complete endoscopic remission (SES CD score of 0)
9. Absolute MaRIA score at week 48

Safety endpoints

1. Toxicity of chemotherapy using NCI CTCAE (National Cancer Institute Common Terminology Criteria for Adverse Events) criteria version 4.03
2. Adverse events (AEs) and Serious Adverse Events (SAEs), including mortality

Patient-reported endpoints

1. Disease specific quality of life using the IBDQ
2. Disease specific quality of life using the IBD Control
3. Quality of life using the EQ-5D-5L
4. Health care resource utilisation questionnaire (this endpoint will form the basis of a potential future economic analysis, if undertaken)

Exploratory secondary endpoints

1. Efficacy of re-introduction of anti-TNF therapy in patients with disease recurrence post-HSCT (change in CDAI at 6 weeks and change in SES CD at 22 weeks after initiation)
2. Safety of re-introduction of anti-TNF therapy in patients with disease recurrence post-HSCT
3. Presence of any of the late side effects of HSCT.

4 Sample Size

The assumptions in the calculations are based upon the endoscopic assessment post HSCT reported in the ASTIC trial program (Lindsay JO, Allez M, Clark M, Labopin M, Ricart E, Satsangi J, Rogler G, Rovira M, Farge D, Hawkey C, 2017). For the primary outcome, to detect a significant difference in the proportion of patients with absence of ulceration on endoscopic assessment of 35%, based on 50% in the HSCT group and no more than 15% in the control group, with 90% power at 5% significance level, with a 2:1 (intervention:control) allocation ratio, requires 62 patients in the HSCT group, and 31 in the control group. Therefore, 93 patients will be recruited at baseline and allocated to intervention or control group, using 2:1 randomisation. Due to the nature of the condition, the design of the intervention and control group, the definition of the primary endpoint and our experience in the ASTIC trial, we anticipate a 6% drop out rate and will therefore recruit 99 patients (66 in the intervention group and 33 in the control group).

Based on experience in ASTIC, recruitment is anticipated to take 36 months. Patients will be recruited at 8 UK NHS Trusts, at an anticipated rate of 2.75 per month across all sites, or approximately 4 patients per site per year.

5 Randomisation & Blinding

Once eligibility has been confirmed and baseline data recorded, participants will be centrally randomised using the CTRU online randomisation system (SCRAM) hosted by epiGenesys, a wholly owned subsidiary of the University of Sheffield. Participants will be randomly allocated to either the HSCTlite arm or usual care, in the ratio 2:1. The randomisation schedule will be generated by the blinded trial statistician prior to the start of the study. The trial statistician will generate the schedule via SCRAM but will remain blinded to the allocation as they will not be able to access the schedule. The doctor or nurse will access the web-based randomisation system, patient details (ID, date of birth) will be entered and the treatment allocation will be returned. Randomisation will be stratified by centre, using permuted blocks of variable size which were chosen to ensure enough participants are allocated in the correct ratio each arm of the trial within each centre. Test sequences will be generated to test the choice of block sizes before the final randomisation schedule is generated.

In view of the nature of HSCT, neither patients nor their treating physicians will be blinded to the treatment allocation. However, an adjudication panel blind to both the timing of procedure and treatment allocation will assess videos of all endoscopic procedures used to determine the primary endpoint. Likewise, expert physicians unaware of the timing of investigation or prior treatment will perform central MRI review for completed scans, and calculate the MaRIA score using anonymised electronic copies of the appropriate images. All completed MRI scans will have the MaRIA score calculated centrally to ensure consistency across centres.

The trial statistician(s) will remain blinded throughout the study, but will be unblinded at database freeze, for analysis. The Senior Statistician within the CTRU will be unblinded to the treatment allocation throughout the trial, but will review and approve the statistical analysis plan (SAP) version 1 before seeing any outcome data. Any subsequent changes to the SAP will be documented in detail. Reports to the Data Monitoring and Ethics Committee (DMEC) will be prepared by the CTU Data Management team, who are unblinded. The Senior Statistician will be available to prepare unblinded statistical reports and provide advice to Data Management and the DMEC where required.

6 Interim analyses, data monitoring committees etc.

The following committees will be established:

1. **Data Management and Ethics Committee (DMEC)** – established with an independent chair, statistician, gastroenterologist and haematologist that will adhere to the Standard Operating Procedure of the CTRU and the DMEC charter. The DMEC will meet 6-monthly with meetings comprising an open session to which members of the study team may attend, followed by a closed session to which unblinded data will be available.

2. **Trial Steering Committee (TSC)** - consist of an independent chair, gastroenterologist and haematologist, statistician and two patient representatives. The committee will meet approximately every 6 months from the start of the trial and will see blinded data summaries from the trial.
3. **Study Management Group (TMG)** – oversee the day-to-day management of the trial and will comprise the core members of the team (Chief Investigator, Project Manager and direct research staff).

This trial has been designed with a fixed sample size and one formal statistical analysis at the scheduled end; no formal interim analyses and efficacy/futility stopping guidelines are set in advance. However, the DMEC will review efficacy and safety of the HSCTlite mobilisation regimen after 10 patients and subsequently at each DMEC meeting, in particular any data related to the safety of low dose cyclophosphamide/GCSF mobilisation, in accordance with the DMEC charter. The DMEC may recommend the trial is stopped or modified on the basis of the data, in writing, to the chair of the TSC.

7 Data Sources, data evaluability and analysis populations

7.1 Data sources

The randomisation list will be held on the CTRU's randomisation system. Trial data will be extracted from source documents and entered onto the CTRUs in house data management system (PROSPECT). The data management team in the Sheffield CTRU will validate and query electronic data for inconsistencies during the course of the trial (as stipulated in SOP DM005). The trial statistician will conduct any additional validation checks where appropriate before the data lock and sign off (as guided by DM005 and DM012). Details of data collected at each timepoint are given in Table 1.

Table 1: Details of data collected at each timepoint

Assessments	Screening[1]	Baseline	HSCT procedure (intervention) or continuation on current treatment (control)	Week 4 HSCT only	Week 8	Week 14	Week 24	Week 32	Week 40	Week 48
Eligibility assessment	✓	✓								
Consent	✓									
Standard Pre-HSCT work (including chest x-ray and MUGA scan)	✓									
Serology for HBV, HCV, HIV	✓									
Demographics	✓									
Medication history	✓	✓		✓	✓	✓	✓	✓	✓	✓
Concomitant medications	✓	✓		✓	✓	✓	✓	✓	✓	✓
Adverse events				✓	✓	✓	✓	✓	✓	✓
General Medical History	✓									
History of CD	✓									
General Physical Examination	✓	✓			✓	✓	✓	✓	✓	✓
Urinalysis	✓	✓								
Pregnancy test	✓	✓								
Smoking History	✓									
Crohn's Disease Activity Index (CDAI)	✓	✓			✓	✓	✓	✓	✓	✓
Harvey Bradshaw Index	✓	✓			✓	✓	✓	✓	✓	✓
Patient Reported Outcome 2 questionnaire (PRO2)	✓				✓	✓	✓	✓	✓	✓
Ileo-colonoscopy (Simple Endoscopic Score for Crohn's Disease (SES CD))	✓						✓			✓
Biopsies[2]	✓						✓			✓
MRI Intestine	✓			✓			✓			✓
MRI Pelvis	✓									
Routine Clinical Care blood test		✓			✓	✓	✓	✓	✓	✓
Serum[3]		✓			✓	✓	✓	✓		✓
Whole Blood[3]		✓			✓	✓	✓	✓		✓
Peripheral Blood mononuclear cells [3]		✓			✓	✓	✓	✓		✓
Stool sample[3]		✓			✓	✓	✓	✓		✓
Inflammatory Bowel Disease Questionnaire (IBDQ)	✓				✓	✓	✓	✓	✓	✓
Inflammatory Bowel Disease Control Questionnaire (IBD-Control)	✓				✓	✓	✓	✓	✓	✓
100 day safety (collection of Adverse Events for transplant endpoint)						✓				
EQ-5D-5L	✓				✓	✓	✓	✓	✓	✓
Work Productivity and Activity Impairment questionnaire (WPAI)	✓				✓	✓	✓	✓	✓	✓
Health care resource use questionnaire	✓				✓	✓	✓	✓	✓	✓
Patient Global Impression of Change (PGIC)										✓
For participants in HSCTlite arm only:										
JACIE and HTA recommended routine tests		✓								
Anti-TNF therapy initiated (if required)							✓			
Adherence to re-vaccination policy						✓	✓	✓	✓	✓

7.2 Definition of the analysis populations

The following analysis populations will be studied in the analyses:

Name	Participants included	Treatment group
Intention to treat (ITT)	All randomised participants according to the randomised treatment assignment (including those who do not complete therapy) with the following exclusions: • • No recorded consent information • Found to be ineligible after randomisation	As randomised
Per protocol	A subset of the ITT analysis population. Exclusions will apply to: • Participants in the treatment group that did not receive HSCTlite as planned (defined further below) • Participants in the usual care group that received stem cell transplantation. The per protocol analysis will be used as the basis of complier average causal effect (CACE) analysis as described below.	As received

Per protocol

A participant in the HSCTlite group is defined as per protocol if they have received the stem cell transplant as intended. Patients that do not undergo mobilisation, stem cell harvest, conditioning and transplant will be defined as non-compliant. For the remainder, a case review (blind to outcome data) will be undertaken in which the chief investigator and other members of the study team review

- Mobilisation: therapies used (including dosage and duration), days taken to achieve a CD34+ count of $10x10^6$ L (or maximum count if target not reached)
- Harvest: number of CD34+ cells harvested
- Conditioning: therapies used (including dosage and duration), days taken to achieve a neutrophil count of $1x10^9$ L for two consecutive days (or maximum count if target not reached)
- Any reasons or comments recorded during the transplantation process (e.g. in cases where therapy is changed, where patients experience serious reactions and other observations)

A maximum of two attempts is allowed for mobilisation. If the first attempt is unsuccessful, the review will consider data collected from the second attempt.

The review will be completed before the end of study, following which patients will be defined “protocol compliant” or “non-protocol compliant”.

In the unlikely event of an error in the dose calculation or administration of study products, this will be reported as a protocol non-compliance. Any non-compliance in the dose or administration of study products will also be presented to the DMEC for discussion and they will decide whether this patient should be included in the per-protocol set at the end of the study.

A participant in the usual care group is defined as per-protocol if they did not receive stem cell transplantation, but analyses will retain all data collected up to the point of transplant.

7.2.1 The role of the analysis populations

Analysis will be conducted on the primary outcome (absence of ulceration at 48 weeks) using both analysis sets (ITT, CACE). To minimise potential selection bias, preserve randomisation and provide an unbiased estimate of the treatment effect, the primary analysis population will be based on Intention-to-treat, and analysis of the CACE population will be considered as secondary. For all secondary outcomes, the analysis will be reported on the ITT population only.

By contrast to ITT, CACE analysis estimates the effect of HSCTlite compared to usual care in the (hypothetical) scenario where it is always used as intended. CACE analysis is preferable to per-protocol analyses since it aims to separate the effect of non-uptake from the bias emanating from selective uptake. Specifically, the characteristics of “non-compliant” participants may be different to those that “comply” (for instance in terms of their age or disease severity), and ignoring this means a simple protocol-compliant analysis does not compare like-for-like (Little, Long, & Lin, 2009). In brief, CACE attempts to compare participants randomised to and receiving HSCTlite against a comparator group who are “likely” to have complied to the intervention had they been randomised to receive it. Further details are given in section 9.2.

8 Outline of analyses

Data will be reported according to the Consolidated Standards Of Reporting Trials (CONSORT) statement for individually randomised parallel group trials (Schulz, Altman, & Moher, 2010).

8.1 General considerations

Summaries of continuous variables will comprise the number of observations used, mean, median, standard deviation, inter-quartile range, minimum and maximum as appropriate for the distributional form of the data. Summaries of categorical variables will comprise the number of observations used, and the number and percentage of observations in each category.

Tables containing the results of the statistical modelling will present the overall difference between treatment groups with two-sided 95% confidence intervals (CI) and p-values. Hypothesis tests will use a two sided 5% significance level.

Complete details of data derivations and methods of handling missing data are covered in sections 9.1 and 10.

8.2 Internal Pilot Analysis

The TSC and DMEC will assess the feasibility outcomes at the end of the internal pilot phase. They will consider whether the trial should continue in light of the feasibility results against the STOP/GO criteria listed in section 1. The results and recommendations will be communicated to the funder (NIHR EME). If the result of the internal pilot analysis is to stop the trial, the trial will be written up and reported according to the updated CONSORT statement for pilot studies (Eldridge et al., 2016). Summary statistics and confidence intervals will be presented; however, no hypotheses will be tested (i.e. no comparative hypothesis testing will be undertaken or p-values reported) at the end of the internal pilot phase.

8.3 Data Completeness

A CONSORT style flow diagram will be presented to summarise the flow of participants through the trial, from screening, during follow up and inclusion in the primary analysis. An example CONSORT diagram is shown in section 2.2 Figure 1. For the purpose of the CONSORT diagram, data completeness will be based on the primary outcome as defined in section 9.1. This information will be made available to the TMG, TSC and DMEC on their request during the course of the trial. A table giving detailed reasons for withdrawal by group and stage will be presented alongside the CONSORT flow diagram.

8.4 Baseline characteristics

Summaries of screening/baseline variables will be presented by treatment group and overall (as in section 2.2; Table 3). If information is collected at both screening and baseline (e.g. lab tests, CDAI), only baseline findings will be presented. The baseline data will be assessed for comparability between groups, any noted differences will be described and considered for adjustment in sensitivity analyses on the primary outcome. No statistical testing will be undertaken on baseline data and no confidence intervals for differences between randomised groups on baseline variables will be presented. The following summaries will be presented:

Demographics	Categorical variables - Centre - Sex - Ethnicity

	- Smoking status (never smoker, current smoker, previous smoker (stopped last year, stopped 1-5years, stopped 5+years)) - Current tobacco intake and cumulative tobacco intake among current smokers (calculations defined in section 10.1.2) Continuous variables - Age - BMI
Crohn's disease characteristics	Categorical variables - Family history of inflammatory bowel disease (yes/no, plus each that apply of Crohn's, Ulcerative colitis, or other unknown) - Montreal age at onset classification (A1 below 16 years, A2 between 17 and 40 years, A3 above 40 years (Satsangi, Silverberg, Vermeire, & Colombel, 2006)) - Behaviour of CD (B1 non-stricturing, non-penetrating, B2 stricturing, B3 penetrating (Satsangi et al., 2006)) - Perianal CD (yes/no) - Stoma (no/yes (ileostomy, colostomy, end stoma, loop stoma)) - Disease location (L1 ileal, L2 colonic, L3 ileocolonic, L4 isolated upper disease (Satsangi et al., 2006) - Extraintestinal involvement (yes/no, plus each that apply of joints, skin, eyes, other) - Extraintestinal manifestations (yes/no, plus each that apply of erythema nodosum, pyoderma gangrenosum, …, other) Continuous variables - Age at CD onset - Duration of CD Number of previous operations for CD - Intestinal surgery - Perianal surgery - Other surgery for CD
Medical history	Co-morbidities (hypertension, established cirrhosis, respiratory disease, established renal disease, psychiatric disease, other)
Drug history	Drugs used - Immunosuppressant (Azathioprine / 6-Mercaptopurine, Methotrexate, Anti-TNF agents, other immune suppressants) - Steroid use (Budesonide, Methylprednisolone, Prednisolone, other steroids) - Biological therapy (Adalimumab, Infliximab, Ustekinumab, Vedolizumab, other) Number of drugs used (immunosuppresants only)

	Months used (immunosuppresants only)
Lab tests	Haemoglobin (Hb) Platelet count Albumin, C-reactive protein (CRP)
Disease activity & colonoscopy	CDAI PRO2 Harvey Bradshaw Activity Index Segments examined in colonoscopy SES-CD score MaRIA score
Qol and patient reported outcomes	IBDQ IBD-Control EQ-5D-5L, EQ-VAS WPAI

8.5 Treatment Summaries

The following treatment summaries will be presented for the HSCTlite group:

Treatment summaries	Mobilisation • Mobilisation successful (yes/no, and if no mobilisation repeated yes/no) • Number of cycles of mobilisation (NB: If patient has had more than one cycle, the following will be presented for final cycle only) • Dose of cyclophosphamide • Number of days GCSF required • Days between cyclophosphamide and harvest of stem cells • Number of cells harvested (Total no of nucleated cells, CD 34^{+}) Disease activity after mobilisation • Karnofsky performance status • Harvey Bradshaw Index • Lab tests (haemoglobin, platelet count, albumin, CRP) Conditioning and Transplantation • Days between harvest and stem cell reinfusion • Number of stem cells reinfused

	• Time to engraftment (see section 10.1.4) • Proportion achieving engraftment (see Section 10.1.4 for a definition) • Number of blood products transfused (red cells and platelets)
Early termination of treatment	N(%) patients that withdrew from study and from treatment Reasons for termination of treatment

Summaries for the usual care group will provided descriptively within the manuscript text.

8.6 Efficacy

The usual care arm will be the reference group for the analysis.

8.6.1 Analysis of the primary outcome

The primary hypothesis of a between group difference in the proportion of patients with absence of ulceration (ulcer subscore of 0 in all segments examined, see Section 10.1.1 for a full definition) will be tested by estimating the proportions for each group. Mixed effects logistic regression will be used to estimate the odds ratio for absence of ulceration in HSCTlite in comparison to conventional therapy, controlling for baseline SES-CD ulcer subscale score as a fixed effect and study centre as a random effect. The intraclass correlation coefficient (ICC) for study centre will be reported. No other variables will be controlled for in the primary analysis.

Sensitivity analysis on primary outcome

Sensitivity analyses on the primary outcome will be undertaken based on the following populations:

- CACE analysis
- Removing participants that died (if not a direct result of Crohn's Disease – this will be assessed by the Chief Investigator based on SAE data, SAE details will be blinded by arm where possible (i.e. unless the detailed description of the SAEs and death reveals the treatment arm))
- Removing participants that had surgery for Crohn's in the usual care group. A pre-requisite of inclusion into the trial is 'surgery is considered not appropriate or has been declined'. Participants that undergo surgery for Crohn's during the follow up period are considered a treatment failure by the definition of the primary outcome, which may overestimate the observed treatment benefit in the (potential) scenario where patients decline pre-trial surgery in hope of receiving HSCTlite and then go on to have surgery if they have been randomised to the usual care group. Removing participants who have had surgery post randomisation in the usual care group will lead to a more conservative treatment difference, thereby assessing robustness of the results of the primary analysis.
- Adjusting for baseline predictors of missingness (see section 8.1)

- Imputing missing primary outcome data
    - Imputation of “worst case” treatment failure for all patients with missing primary outcome data
    - Using 24 week colonoscopy data to impute for 48 week missing colonoscopy data
    - If there is more than 6% missing primary outcome data, further imputation methods will be used including multiple imputation (see section 9.1 for more details).

The primary analysis will be repeated for each of these data sets and displayed alongside the ITT analysis results (see Table 7). They will also be reported using a forest plot, as illustrated in Figure 2.

8.6.2 Secondary outcomes

Secondary continuous outcomes (CDAI at week 48, SES CD at week 48, CDAI change from baseline at week 48, SES CD change from baseline at week 48, IBDQ, IBD Control, EQ-5D-5L, MaRIA score) will be compared between treatment groups using mixed effects linear regression with study centre as a random effect and the corresponding baseline assessment of the variable (or screening if measure only taken at screening). The mean difference, 95% CI and p-value will be presented for all secondary continuous outcomes.

Secondary categorical outcomes (CDAI<150, CDAI<150 + no steroid use, Harvey Bradshaw Index ≤4, PRO2 abdominal pain mean score≤1, stool frequency mean score≤1.5, SES CD score of 0) will be compared between groups using mixed effects logistic regression adjusted for study centre as a random effect. The odds ratio, 95% CI and p-value will be presented for all secondary categorical outcomes. Adjustment for baseline assessment of these secondary outcomes will be made for each categorical outcome as follows

- CDAI<150 and CDAI<150 + no steroid use will be adjusted for CDAI baseline score
- Harvey Bradshaw index ≤4 will be adjusted for Harvey Bradshaw Index baseline score
- PRO2 (abdominal pain mean score≤1 and stool frequency mean score≤1.5) will be adjusted for baseline PRO2 score
- SES-CD score of 0 will be adjusted for SES-CD at screening

Spaghetti plots (where outcome over time is plotted for each patient on one graph) will be presented for SES-CD and CDAI across all recorded timepoints (see Figure 3).

8.7 Safety and Harms

Any untoward medical occurrence affecting the patient to whom a medicinal product has been administered (irrespective of relationship) will be recorded.

An AE will be recorded as a serious adverse event (SAE) if it:

- Results in death

- Is life-threatening
- Requires hospitalisation or prolongation of existing inpatients hospitalisation
- Results in persistent or significant disability or incapacity

Descriptive statistics of Adverse Events and Serious Adverse Events will be presented. Safety data will be presented on the intention to treat population, but an AE or SAE that occurred after a treatment switch will be highlighted.

The following summaries of the safety data will be presented overall, by treatment group and by the following time periods:

- Mobilisation phase (period from start of mobilisation to start of conditioning). Randomisation to day 0 for usual care participants;
- Transplant phase (period from start of conditioning up to 100 days from day 0 i.e. day of autologous transplant/infusion). Day 0 to day 100 for usual care participants;
- Follow up phase (from +100 days post-transplant phase to one year assessment; see Table 11 for an example). Day 100 to one year assessment in usual care participants.

  The time periods used for the intervention and usual care arms are not exactly equivalent but have been included so safety data can be compared across arms by the key stages of the intervention.

AEs	Number (%) participants experiencing ≥1 AE Number (%) participants experiencing ≥1 AE by relationship to intervention - (Cyclophosphamide, Filgrastim, Fludarabine, Rabbit ATG, MENSA, Methyl-prednisolone, another drug) Number of all AEs including repeat events Number of AEs by category (see below for details of categorisation) Number of all AEs by relationship to intervention - (Cyclophosphamide, Filgrastim, Fludarabine, Rabbit ATG, MENSA, Methyl-prednisolone, another drug)
Serious AEs (SAEs)	Number (%) participants experiencing ≥1 SAE Number of all SAEs including repeat events Number of SAEs by Seriousness (Death, Life threatening, Inpatient hospitalisation, Prolongs hospitalisation, Persistent or significant disability/incapacity, congenital abnormality/birth defect) Number of SAEs by Outcome (Recovered, Improved, Unchanged, Deterioration, Persisted, Death)

	SAE by Action taken (None, treatment withdrawn, NIMP dose alteration, conmed dose change, specific treatment, other)

The following summaries will be presented by treatment group and by CTCAE grade (grade 1-2, grade 3, grade 4)

AEs	Number (%) participants experiencing ≥1 AE Number of all AEs including repeat events Number of all AEs by category
SAEs and AEs by CTCAE grade	Number (%) participants experiencing ≥1 SAE Number of all SAEs including repeat events Number of all SAEs by category

The following by-patient line listings will be presented

All AEs	A listing of all AEs including i. Treatment group (if the patient switches treatment, details will be included) ii. Timing of AE (days – anything pre day 0 will be recorded as a negative number i.e. -15 is 15 days prior to day 0) - Relationship (Cyclophosphamide, Filgrastim, Fludarabine, Rabbit ATG, MENSA, Methyl-prednisolone, another drug) - CTCAE category and grading iii. Outcome (if SAE) iv. Seriousness (if SAE)
All SAEs	A listing of all SAEs (as "all AEs" with the omission of "seriousness")
All treatment-related AEs	A listing of all treatment-related AEs ordered by intervention (as "all AEs" with the omission of "relationship")

AE and SAE category will be defined based on NCI CTCAE categories and the free text entered within the "Adverse event term" section of the CRF. They will be approved by the Chief investigator. Example categories are Infectious (Viral, Sepsis, Localised), GI (disease flare, non-flare symptoms), Haematological (anaemia, neutropenia, pancytopenia). Categories with low prevalence or not of particular importance to the disease and treatment under investigation will be categorised as other.

8.8 Re-introduction of anti-TNF

Patients in the intervention arm with evidence of disease activity at scheduled week 24 colonoscopy or MRI will receive induction and maintenance anti-TNF therapy. The efficacy of this re-introduction will be assessed by calculating:

- The proportion achieving CDIA<150 at 6 weeks after initiation (32 weeks)
- The proportion achieving CDIA<150 at 24 weeks after initiation (48 weeks)
- The change in CDAI at 6 weeks after initiation (the difference between 24 and 32 week CDIA)
- The change in CDAI at 22 weeks after initiation (the difference between 24 and 48 week CDAI)

The proportions and mean change scores will be presented with 95% confidence intervals and will be calculated for the participants in the HSCTlite arm who have evidence of disease activity on either the colposcopy or MRI and have antiTNF medication listed on the concomitant medication form. A spaghetti plot (where CDAI over time from 24 to 48 weeks is plotted per participant on one graph) will be presented for this group to display change in disease activity post re-introduction of anti-TNF.

The safety of re-introduction of anti-TNF therapy in patients with disease recurrence post-HSCT will be investigated by tabulating AEs and SAEs that occurred after initiation for the same population. AE and SAE data for patients in the HSCTlite group who did not receive anti-TNF therapy will be tabulated alongside those that received anti-TNF in order to act as a reference population. The following will be tabulated by time period (i) within 6 weeks of initiation and (ii) between 6 and 24 weeks after initiation:

AEs	Number (%) participants experiencing ≥1 AE Number of all AEs including repeat events Number of all AEs by category Number of all AEs related to anti-TNF therapy
SAEs and AEs by CTCAE grade	Number (%) participants experiencing ≥1 SAE Number of all SAEs including repeat events Number of all SAEs by category

8.9 Late effects of HSCT

The long term screening for late effects of HSCT is listed in 2.1. Screening for late effects of HSCT is assessed as part of standard care following a stem cell transplant. The presence of late effects will be extracted from adverse event data entered on or after week 14, week 24 or week 48 depending on the late effect assessed (see Section 2.1 Table 2). The adverse events for participants in the HSCTlite arm recorded on or after week 14 will be reviewed and agreed by two clinical members of the TMG and categorised as a late effect where appropriate. A table of number and percentage for each late effect will be presented (see Table 13).

8.10 Subgroup and moderator analysis

The objective of an exploratory subgroup analysis is to explore heterogeneity in the intervention effects across pre-defined subgroups. An exploratory subgroup analysis will be performed using mixed effects

logistic regression with the primary outcome, mucosal healing without surgery or death at 48 weeks, as the response. An interaction statistical test between the randomised treatment group and pre-defined subgroup will be used to directly examine the strength of evidence for the difference between treatment group (HSCTlite vs usual care) varying between subgroups. Four subgroups of interest have been pre-specified based on factors that were found to predict benefit or harm in the ASTIC trial (Lindsay JO, Allez M, Clark M, Labopin M, Ricart E, Satsangi J, Rogler G, Rovira M, Farge D, Hawkey C, 2017) and on recommendation by the TSC:

- Disease location (ileal, colonic, ileocolonic)
- Perianal Crohn's Disease (yes/no)
- Current smoking status (yes/no)
- Current CD treatment at screening (on treatment/not on treatment)

Subgroup analysis will be performed regardless of the results of the primary analysis. The odds ratio (and 95% CI) for absence of ulceration in HSCTlite compared to conventional care will be computed for each subgroup category and visually displayed using a forest plot (Cuzick, 2005). The regression coefficient for the interaction between treatment group and subgroup will be presented with the associated confidence interval and p-value. We will not calculate separate p-values within each subgroup category (Assmann et al., 2000; Pocock et al., 2002; Wang et al., 2007). Results will be presented as shown in Table 14.

The relationship between treatment and the following continuous covariates found to predict benefit or harm in the ASTIC trial will also be investigated:

- CDAI at baseline
- SES-CD at screening
- Disease duration

The association will be depicted visually using a lowess smoother, with the probability of response plotted against the covariate as shown in Figure 4. The effect size together with its 95% CI will be visualised using a non-parametric kernel density estimator as shown in Figure 5 (Cattaneo & Jansson, 2018).

Influential covariates identified in the above analyses will be incorporated into a predictive model for the mechanistic evaluation as described below.

8.11 Mechanistic evaluation

A secondary mediation analysis will investigate putative mediational factors using modern causal inference methods. This involves using parametric regression models to test for mediation of HSCT on

treatment success through biomarkers. Analyses will adjust for baseline measures of the marker, and possible measured confounders.

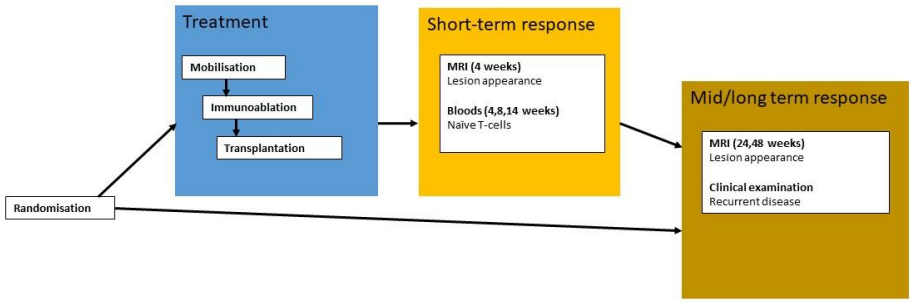
The biological rationale of stem-cell transplant is that the T-cells will be “reset”. In order to test this mechanism, biomarkers will be analysed to assess the following:

1) the change in markers from baseline to 4 weeks, both within each group and between groups (the hypothesis is that inflammatory markers will reduce in the HSCT group but not the control)

2) the trajectory of markers with time (i.e. whether markers of inflammatory activity stabilise or return to pre-baseline levels) and

For the purpose of this mechanistic analysis, patients in the HSCT arm that do not undergo transplant will not be included in the analysis.

In the first stage, the surrogacy of biomarkers and short-term measures will be evaluated.



Mechanistic immunology

The complex datasets will be integrated, analysed and interpreted using established artificial neural network (ANN) and computational intelligence-based approaches. The expertise is available in the John van Geest Cancer Research Centre and has been used by them previously (26,27). We will use adaptions of existing neuro-fuzzy computational intelligence models (28) to answer the questions posed. Importantly, these approaches will provide mechanistic insight into underlying responsiveness to anti-TNF and events that are associated with patients becoming refractory to it after HSCT.

9 Detailed statistical methods and calculations

9.1 Missing data

We anticipate minimal drop out/attrition (up to 6%) based on the original ASTIC trial. Baseline variables of participants with complete primary outcome data and participants without complete primary

outcome data will be compared between treatment group and overall in order to check for differential predictors of missing outcomes (see Section 2.2 Table 4). The baseline variables that will be compared will include:

- Demographics: age and sex
- Crohn's disease characteristics: disease location, perianal, age at onset
- Disease activity: CDAI, SES-CD, Harvey Bradshaw Index
- Qol and patient reported outcomes: IBDQ, IBD-Control

Any predictors that qualitatively appear imbalanced will be included as a covariate in a sensitivity analysis model for the primary outcome. This accounts for missing outcome data under a missing at random assumption, conditional on the covariates included in the model (Hippel, 2015). Given the constraints of sample size, these covariates will be added one by one rather than in combination.

Three further methods will be used to impute missing primary outcome data regardless of the proportion of missing data observed:

- Imputation of "worst case" treatment failure for all patients with missing primary outcome data
- Using 24 week SES-CD ulcer subscale score to impute missing 48 week SES-CD ulcer sub-scale score. This will only be done in those who are missing the primary outcome, for example those that are lost to follow up or withdrawn, not participants that cannot complete colonoscopy due to surgery or extent of disease.
- Using 48 week MRI data to impute missing 48 week SES-CD ulcer sub scale score in participants who's segment is not explored but has not been resected (see section 10.5).

If more than 6% of missing data is observed, the following imputation sensitivity analyses will also be performed:

- Multiple imputation: One hundred multiple imputation data sets will be created using chained equations (including the baseline variables found to be predictive of missing data status).

Further sensitivity analysis using a weighting approach as proposed by Carpenter et al (Carpenter, Kenward, & White, 2007) will be used to investigate the sensitivity of MI to the missing at random assumption.

9.2 CACE analysis

CACE analysis is an attempt to compare the 'protocol compliers' in the HSCTlite group to those in the usual care group who are 'likely' to have undergone treatment had they been randomised to HSCTlite intervention. CACE analysis will be performed in the following steps (Peng, Little, & Raghunathan, 2004):

1. Using participants in the HSCTlite group, derive a logistic regression model to predict the probability of being a non-complier (i.e. not receiving the stem cell transplant). Possible predictor covariates will include baseline demographics, CD characteristics and measures of disease activity.
2. Apply these predictions to the Usual Care group, so that each participant is given a probability of receiving the HSCTlite intervention as planned (if they had been randomised to receive it) which is based on their covariates.
3. For each participant in the Usual Care group calculate a re-weighted outcome defined as the original outcome multiplied by the predicted probability of receiving as planned.
4. Compare the subset of participants in the HSCTlite group that are deemed to have complied with intervention with the re-weighted outcomes amongst participants in the Usual Care group.

CACE analysis will be conducted by a two stage regression, the first will use mixed effects logistic regression including site as a random effect to predict non-compliance. The second model used in step 4 will be the mixed effects model as used in the primary analysis.

9.3 Mixed effects model checks

Model goodness of fit of any mixed effects logistic regression will be investigated by plotting predicted proportion as predicted by the model against observed proportion and using the Pearson statistic utilizing the scaled chi-square distribution (Evans & Hosmer, 2004).

Model assumptions for mixed effects regression will be checked via graphical methods (e.g. histograms of residuals and scatterplots of residuals vs. covariates). Influential observations and outliers will also be investigated and sensitivity analyses at the discretion of the trial statistician will be undertaken and reported.

In the event that the goodness-of-fit is poor or the model assumptions have not been met, the following alternative models will be fitted and assessed (in order):

i. non-linear transforms of baseline SES-CD ulcer score using fractional polynomials (Royston & Altman, 1994)
ii. interactions between treatment and SES-CD baseline score
iii. alternative covariates as identified in the mechanistic analysis

Any changes to the model specification and their justification will be described in the final report.

Irrespective of the model fit, the consistency of response across the treatment centres will be assessed by qualitatively assessment of the random effect variance term; if there are no between-site differences

this term should be close to 0. If this is not the case, further exploratory analysis will be undertaken to assess the characteristics of patients between centres, in order to assess whether differences reflect case mix confounding or are true site heterogeneity.

10 Data manipulation and definitions

10.1 Definitions

10.1.1 Primary outcome

The primary outcome (treatment success at week 48) will be defined as mucosal healing (no endoscopic ulceration (SES CD ulcer sub score = 0 (SES-CD score of 0 on the sub score for size of ulcer in all bowel segments), assessed by adjudication panel blind to allocation and time of assessment)) without surgery for CD or death. If a patient has surgery and a SES CD score can still be assessed (e.g. an abscess requiring draining then surgery) treatment success or failure will be defined based on SES-CD score. Patients who do not complete the week 48 colonoscopy due to surgery for CD, or worsening disease will be categorised as treatment failures. Any surgery post screening is documented in the 'Further Interventions' section of the CRF; the CI will review and confirm whether the surgery was performed for CD. Patients who do not complete the week 48 assessment because they are lost to follow up, or have withdrawn from follow up will be excluded from the primary analysis but included in a sensitivity analysis of the primary outcome (see section 8.6.1) – patient withdrawal and loss to follow up is recorded on the 'Study completion/discontinuation' form.

10.1.2 Current tobacco intake

The calculation of current tobacco intake and cumulative tobacco intake among current smokers is made on the basis of 50 g tobacco per week = 2 oz tobacco per week = 100 cigarettes per week = 50 cigars per week (NHS, n.d.):

$$Current\ tobacco\ intake\ in\ cigarettes\ per\ day\ equivalents\ =$$
$$Number\ of\ cigarettes\ per\ day\ +\ 2 \times Number\ of\ cigars\ per\ day$$
$$+\ (2/7) \times tobacco\ per\ week\ (g)$$

$$Current\ intake\ of\ tobacco\ in\ tobacco\ per\ week\ equivalents\ (g)\ =$$
$$(7/2) \times current\ intake\ in\ cigarettes\ per\ day$$

Pack years are defined as the cumulative number of years exposure to the equivalent of a pack of cigarettes per day, where a pack is defined as 20 cigarettes. To illustrate this, a person who smokes 10

cigarettes (0.5 pack) per day for 12 yeas has a pack-years of 12×0.5 = 6 pack years. Extending this to incorporate cigars and tobacco gives

$$Cumulative\ tobacco\ intake\ in\ pack\ year\ equivalents\ =$$
$$Number\ of\ cigarettes\ per\ day\ \times\ Number\ of\ years\ smoking\ cigarettes$$
$$+\ 2 \times Number\ of\ cigars\ per\ day\ \times\ Number\ of\ years\ smoking\ cigars$$
$$+\ (2/7) \times tobacco\ per\ week\ (g)\ \times\ Number\ of\ years\ smoking\ tobacco$$

10.1.3 Disease location

Disease location will be defined using the Montreal classification (Satsangi et al., 2006) and will be categorised using information from the Crohn's Disease history, Colonoscopy and MRI (all taken at screening). The categories are as follows:

Category	**Description**
L1 Ileal	Ileum
L2 Colonic	Any/all of • Caecum • Ascending colon/hepatic flexure • Transverse colon • Descending colon • Sigmoid colon • Splenic fixture • Rectum
L3 Ilealcolonic	If both L1 and L2 are present
L4 isolated upper disease	Disease proximal to the Ileum, any or all of • Jejunum • Duodenum • Stomach • Oesophagus L4 displayed on its own is isolated upper disease L4 is a modifier that can be added to L1-L3 when concomitant upper gastrointestinal disease is present.

10.1.4 Engraftment

A participant is defined as achieving engraftment if they have three consecutive days with a persistent blood cell count above a pre-defined level in the period immediately after stem cell reinfusion (Labopin & Iacobelli, 2008). The limit for each cell type is described below:

- White cell engraftment: White Cell count >1 x 10^9/L
- Platelet engraftment: Platelet count >20 x 10^9/L
- Neutrophil engraftment: Neutrophils >0.5 x 10^9/L

A second definition of platelet engraftment using *platelet count of >50 x 10^9/L* will also be used and presented alongside the definition described above as part of the treatment summaries (Section 8.5). Each cell type will be summarised separately.

10.2 CDAI

The Crohn's Disease Activity Index (CDAI) will be calculated using the standard scoring criteria (Best, Becktel, Singleton, & Kern, 1976). A cap of minus 10 will be used in relation to minus scores for those participants exceeding standard weight ranges. If there is data available for 5 out of the 7 days, the missing days of data will be imputed as the average of the available days and the CDAI score will be calculated. The CDAI score ranges from 0 to over 600: a score of less than 150 is considered to be remission, score greater than 450 is considered to be severe disease (Best et al., 1976).

The patient will be asked to complete a symptom diary for a week prior to assessment of the CDAI; this cannot be taken in the week preceding a colonoscopy. Patients should finish the diary prior to starting bowel prep for colonoscopy.

The CDAI score is calculated by summing the below components after adjustment with a weighting factor:

Clinical or laboratory variable	Weighting factor
Number of liquid or very soft stools each day for seven days	x 2
Abdominal pain (graded from 0-3 on severity) each day for seven days	x 5
General well being, subjectively assessed from 0 (well) to 4 (terrible) each day for seven days	x 7
Presence of complications (one point for each of a – arthritis/arthralgia, b - Iritis/uveitis, erythema nodosum or pyoderma gangrenosum, c - Aphthous ulcers, d - Anal fissure, fistula, abscess, e - Other fistula, f - Fever over 100 F (37.8°C))	x 20
Taking Lomotil or opiates for diarrhoea (1=yes)	x 30
Presence of an abdominal mass (0 as none, 2 as questionable, 5 as definite)	x 10

Hematocrit of 47-crit in men and 42-crit in women	x 6
(Standard body weight – actual weight) / standard weight * 100 (kg)	x 1

10.3 PRO2

The patient reported outcome (PRO2) measure is based on two patient reported components (stool frequency and abdominal pain) of the CDAI (Khanna et al., 2015). It is calculated as:

$$PRO2 = 2 \times (Number\ of\ liquid\ or\ very\ soft\ stools\ averaged\ over\ 7\ days) + 5 \times (Abdominal\ pain\ (graded\ 0-3) averaged\ over\ 7\ days).$$

The PRO2 score will not be calculated if less than 5 days of data are available for each of the components. If 5 (or 6) days of data are available for a component the average will be taken across the 5 (or 6) days of available data.

10.4 Harvey Bradshaw Index

The Harvey Bradshaw Index is calculated by summing scores for general wellbeing (0 very well to 4 terrible), abdominal pain (0 non to 3 severe), number of liquid stools per day, abdominal mass (0 none to 3 definite and tender) and presence of 8 pre-specified complications (1 point for each) (Harvey & Bradshaw, 1980). The Harvey Bradshaw Index ranges from 0 to more than 20, a score of 4 or less can be considered as clinical remission. Data for all items must be available for the index to be valid, and hence no imputation of missing items will be conducted.

10.5 SES CD

The Simple Endoscopic Score for Crohn's Disease is calculated based on findings from the colonoscopy. Findings from five segments (Ileum, right colon, transverse colon, left colon and rectum) on four variables are recorded and given a 0-3 score as follows:

Variable	0	1	2	3
Size of ulcers (cm)	None	Aphthous ulcers (diameter 0.1-0.5)	Large ulcers (diameter 0.5-2)	Very large ulcers (diameter >2)
Ulcerated surface	None	<10%	10-30%	>30%
Affected surface	Unaffected segment	<50%	50-75%	> 75%
Presence of narrowings	None	Single, can be passed	Multiple, can be passed	Cannot be passed

The SES-CD score is calculated as a sum of the four variables over the five explored segments (Daperno et al., 2004); it ranges from 0 to 60 with higher scores representing more disease activity. If

one or more segments are resected, a sum of the scores for the remaining segments is calculated. If a segment is not explored but it has NOT been resected, data from the MRI reviewed by a blinded radiologist can be used to impute scores for the missing segment. If the SES-CD score is available for a segment then it will not be superseded by information from the MRI. Note that for the primary outcome, the use of MRI data to impute SES-CD score will be used for sensitivity analysis only (see section 9.1).

10.6 IBDQ

The Inflammatory Bowel Disease Questionnaire (IBDQ) measures health related quality of life for patients with inflammatory bowel disease; it consists of 32 questions each scored on a 7-point scale. It consists of a total score and four dimension scores;

- Bowel systems (mean score of 10 bowel questions)
- Emotion health (mean score of 12 emotion questions)
- Systemic systems (mean score of 5 systemic questions)
- Social function (mean score of 5 social questions)

Dimension scores range from 1-7, a higher score represents better function in that area.

The total score ranges from 32 to 224 and is calculated by summing all 32 questions, higher scores indicate better quality of life. If one item is missing from a dimension, we will impute the mean score for the other items of the dimension, if more than one item is missing the dimensional score cannot be calculated. When calculating the total score, missing items will be imputed as the mean score from other items in the dimension. The total score can be calculated if:

- no more than 4 items are missing, and
- no dimension contains more than 2 missing items, and
- at least 3 of the four dimensions can be calculated.

The scoring rules have been taken from the IBDQ developer package (version 7). Participants with a stoma will complete the IBDQ-Stoma, the scoring is identical. All dimension scores and total score will be presented.

10.7 IBD-Control

The IBD-Control is a measure of disease control from the patients perspective (Bodger, Ormerod, Shackcloth, & Harrison, 2014). The IBD-Control-8 score is based on a sum of responses to eight items (questions 1a, 1b, 3a to 3f), each allocated a score of 0-2. If any item is missing the total score will not be calculated. The IBD-Control-8 ranges from 0 (worst control) to 16 (best control). The IBD-Control VAS ranges from 0 (worst possible control) to 100 (best possible control).

10.8 EQ-5D-5L

The EQ-5D-5L consists of five questions (measuring mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) each with five possible answers. The scores are assigned utility values and combined to give a value index score. The score ranges from –0.22 to 1.00 (a score of zero means death, 1 is full health and a negative score is a state worse than death). The score will not be calculated if any items are missing. The algorithm for scoring the EQ-5D-5L can be found in the Database Specification. The EQ-5D VAS your health state today is measured on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

10.9 PGIC

The Patients' Global Impression of Change (PGIC) scale consists of two scores;

- one score 1 (no change) – 7 (a great deal better) and
- one score 0 (much better) to 10 (much worse) (Hurst & Bolton, 2004).

10.10 Scoring summary table

Name	No. of items	Score range	Description	Interpretation of score
CDAI	8	0 – 600+	Measures degree of illness in Crohn's disease	Below 150 is considered remission, greater than 450 severe disease
PRO2	2	0-15+	Weighted combination of abdominal pain and stool frequency averaged over 7 days	Higher score represents more severe disease
Harvey Bradshaw Index	5	0-20+	Measures Crohn's disease activity	Higher scores represent more disease activity. Scores of 4 or less represents clinical remission
SES-CD	20	0-60	Endoscopic score of disease activity	Higher scores represent more disease activity
IBDQ	32	32-224	Measures health related quality of life for patients with inflammatory bowel disease	Higher scores indicate better quality of life
IBD-Control	8	0-16	Measures disease control from patient perspective	0 represents worst control, 16 best control
IBD-Control VAS	1	0-100		0 represents worst possible control, 100 best possible control
EQ-5D-Y value index	5	-0.594-1	Measure of health status. 5 domains include mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.	A score of zero means death, 1 is full health, negative score is a state worse than death
EQ-5D-Y VAS	1	0- 100	Measure of health status.	A score of zero means worst health and 100 means best health.

PGIC	2	1-7, and 0-10	Measures patients impression of change	1 (no change) to 7 (a great deal better) 0 (much better) to 10 (much worse)

11 Implementation of the original analysis plan

This SAP will be used as a work description for the statistician involved in the trial. All analyses should ideally be performed by the same statistician (under the supervision of senior trial statisticians Mike Bradburn and Prof. Richard Emsley) and consequently none of the investigators involved in the trial will perform any of the statistical analyses.

Initially, the data manager will provide blinded data for preliminary checks by the statistician. Following database freeze, unblinded data will be delivered to the statistician to define analysis sets and test statistical programs. Any queries will be communicated to the data manager prior to database lock, and any changes to the database during this time will be documented. The database will be locked after agreement between the statistician, data manager and study manager. It is expected that no data amendments should be required following database lock. However if an amendment is required, the process is documented in CTRU SOP DM012.

12 Modifications to the original protocol analysis statement

Not applicable

1 References

Best, W. R., Becktel, J. M., Singleton, J. W., & Kern, F. (1976). Development of a Crohn's Disease Activity Index: National Cooperative Crohn's Disease Study. *Gastroenterology*, *70*(3), 439–444. https://doi.org/10.1016/S0016-5085(76)80163-1

Bodger, K., Ormerod, C., Shackcloth, D., & Harrison, M. (2014). Development and validation of a rapid, generic measure of disease control from the patient's perspective: The IBD-control questionnaire. *Gut*, *63*(7), 1092–1102. https://doi.org/10.1136/gutjnl-2013-305600

Carpenter, J. R., Kenward, M. G., & White, I. R. (2007). Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Statistical Methods in Medical Research*, *16*(3), 259–275. https://doi.org/10.1177/0962280206075303

Cattaneo, M. D., & Jansson, M. (2018). Kernel-Based Semiparametric Estimators: Small Bandwidth Asymptotics and Bootstrap Consistency. *Econometrica*, *86*(3), 955–995. https://doi.org/10.3982/ECTA12701

Conference, I., Harmonisation, O. N., Technical, O. F., For, R., Of, R., For, P., & For, R. (1999). ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Statistics in Medicine*, *18*(15), 1905–1942. https://doi.org/10.1002/(SICI)1097-0258(19990815)18:15<1903::AID-SIM188>3.0.CO;2-F

Daperno, M., D'Haens, G., Van Assche, G., Baert, F., Bulois, P., Maunoury, V., … Rutgeerts, P. (2004). Development and validation of a new, simplified endoscopic activity score for Crohn's disease: The SES-CD. *Gastrointestinal Endoscopy*, *60*(4), 505–512. https://doi.org/10.1016/S0016-5107(04)01878-4

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot and Feasibility Studies*, *2*(1), 64. https://doi.org/10.1186/s40814-016-0105-8

Evans, S. R., & Hosmer, D. W. (2004). Goodness of fit tests in mixed effects logistic models characterized by clustering. *Communications in Statistics - Theory and Methods*, *33*(5), 1139–1155. https://doi.org/10.1081/STA-120029829

Gamble, C., Krishan, A., Stocken, D., Lewis, S., Juszczak, E., Doré, C., … Loder, E. (2017). Guidelines for the content of statistical analysis plans in clinical trials. *JAMA - Journal of the American Medical Association*, *318*(23), 2337–2343. https://doi.org/10.1001/jama.2017.18556

Harvey, R. F., & Bradshaw, J. M. (1980). A SIMPLE INDEX OF CROHN'S-DISEASE ACTIVITY. *The Lancet*, *315*(8167), 514. https://doi.org/10.1016/S0140-6736(80)92767-1

Hippel, P. T. Von. (2015). Regression with Missing Ys: an Improved Strategy for Analyzing Multiply Imputed Data. *Direct*.

Hurst, H., & Bolton, J. (2004). Assessing the clinical significance of change scores recorded on subjective outcome measures. *Journal of Manipulative and Physiological Therapeutics*, *27*(1), 26–35. https://doi.org/10.1016/j.jmpt.2003.11.003

ICH Harmonised Tripartite Guideline. (1996). Guideline for good clinical practice E6(R1). *ICH Harmonised Tripartite Guideline*, *1996*(4), i-53. https://doi.org/10.1056/NEJMp1012246

Khanna, R., Zou, G., D'Haens, G., Feagan, B. G., Sandborn, W. J., Vandervoort, M. K., … Levesque, B. G. (2015). A retrospective analysis: The development of patient reported outcome measures for the assessment of Crohn's disease activity. *Alimentary Pharmacology and Therapeutics*, *41*(1), 77–86. https://doi.org/10.1111/apt.13001

Labopin, B. M., & Iacobelli, S. (2008). Statistical Guidelines for Ebmt. *Statistical Guidelines for Ebmt*, *48*(5), 462–463. https://doi.org/10.1093/rheumatology/kep025

Lindsay JO, Allez M, Clark M, Labopin M, Ricart E, Satsangi J, Rogler G, Rovira M, Farge D, Hawkey C, on behalf of A. T. G. (2017). Autologous stem - cell transplantation in treatment - refractory Crohn's disease - an analysis of pooled data from the ASTIC trial. *Lancet Gastroenterology & Hepatology*, 1–5.

Little, R. J., Long, Q., & Lin, X. (2009). A Comparison of Methods for Estimating the Causal Effect of a Treatment in Randomized Clinical Trials Subject to Noncompliance. *Biometrics*, *65*(2), 640–649. https://doi.org/10.1111/j.1541-0420.2008.01066.x

NHS. (n.d.). Smoking.

Peng, Y., Little, R. J. A., & Raghunathan, T. E. (2004). An extended general location model for causal inferences from data subject to noncompliance and missing values. *Biometrics*, *60*(3), 598–607. https://doi.org/10.1111/j.0006-341X.2004.00208.x

Royston, P., & Altman, D. G. (1994). Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Applied Statistics*. https://doi.org/10.2307/2986270

Satsangi, J., Silverberg, M. S., Vermeire, S., & Colombel, J. F. (2006). The Montreal classification of inflammatory bowel disease: Controversies, consensus, and implications. *Gut*, *55*(6), 749–753. https://doi.org/10.1136/gut.2005.082909

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, *63*(8), 834–840. https://doi.org/10.1016/j.jclinepi.2010.02.005

StataCorp. (2017). Stata Statistical Software: Release 15. *2017*. https://doi.org/10.2307/2234838

UK Statutory Instruments. (2004). The Medicines for Human Use (Clinical Trials) Regulations 2004. *Statutory Instruments*, (1031), 1–86.

Trial Documents

Title	Version	Date	Location
Study Protocol	6	20th September 2018	X:\ScHARR\PR_ASTIC-LITE\General\1 Study Documents\1.1 Approved protocol and amendments\Protocol
Data Validation Specification	1	23rd October 2018	X:\ScHARR\PR_ASTIC-LITE\Data_Management\Data validation\DVS

CTRU Standard Operating Procedures

Title	Version	Date	Location
ST001 The Statistical Analysis Plan	5	9 Jan 2018	M:\Junctions\DFS_ScHARR\projects\CTRU\Quality Assurance\SOPs\Current SOPs
ST003 Data Evaluation	5		
ST006 Undertaking a Statistical Analysis	2	11 May 2017	
ST007 Randomisation	2	14 March 2018	
DM005 Central Data Validation	5	21 Jun 2018	
DM012 Study database lock and retention	4	28 Mar 2017	

2 Appendix

2.1 Late effects

Table 2*: Long term screening for late effects of HSCT, assessed as part of standard care following a stem cell transplant*

Recommended timing of assessment	**3 months**	**6 months**	**1 year**
Corresponding study visit	**Week 14**	**Week 24**	**Week 48**
General			
Weight	1	1	1
Blood pressure	1	1	1
Performance status (Karnofsky/Lansky)	1	1	1
Haematology			
FBC	1	1	1
Renal			
Renal function	1	1	1
Urine protein (dipstick)	1	1	1
Liver			
Liver function	1	1	1
Iron studies		1	1
Endocrine			
Thyroid function TSH, Free T4	1	1	1
Gonadal function FSH, LH, oestradiol, Progesterone (women <=50 years) FSH, LH, Testosterone (men)	1	1	1
Sexual function assessment (as per patient report)		1	1
Bone			
Bone profile	1	1	1
Bone density scan Women and men with evidence of hypogonadism Patients on prolonged corticosteroids or calcineurin inhibitors			1
Respiratory			
Clinical assessment	1	1	1
Pulmonary function test			1
Chest radiograph		*	*
Counselling re: smoking cessation	1	1	1
Nervous system			
Neurological assessment			1
Vascular			
Cardiovascular risk factors			1
Echocardiogram			1
HbA1c		1	1
Lipid profile and abdominal girth		1	1
Immune System			
CD4 subsets	1	1	1
Immunoglobulin levels	1	1	1

Antimicrobial prophylaxis as per local protocol	1	1	1
Immunisation and antibody levels as per local protocol			1
Oral complications			
Dental assessment		1	1
Ocular			
Cataracts assessment	1	1	1
Second cancers			
Mammograms (Women >40 years)			1
Vigilance and self-examination		1	1
Second autoimmune diseases			
Second autoimmune diseases		1	1
Psychosocial			
Psychosocial/psychosexual issues, by standard holistic needs assessment	1	1	1

1 = recommended for all transplant patients

* = reassessment recommended if previously abnormal

2.2 Example Tables and Figures

This section includes example tables and figures. The lists of data displayed in the tables are not comprehensive and are included only as an example.

Figure 1: CONSORT diagram: participant flow through the study

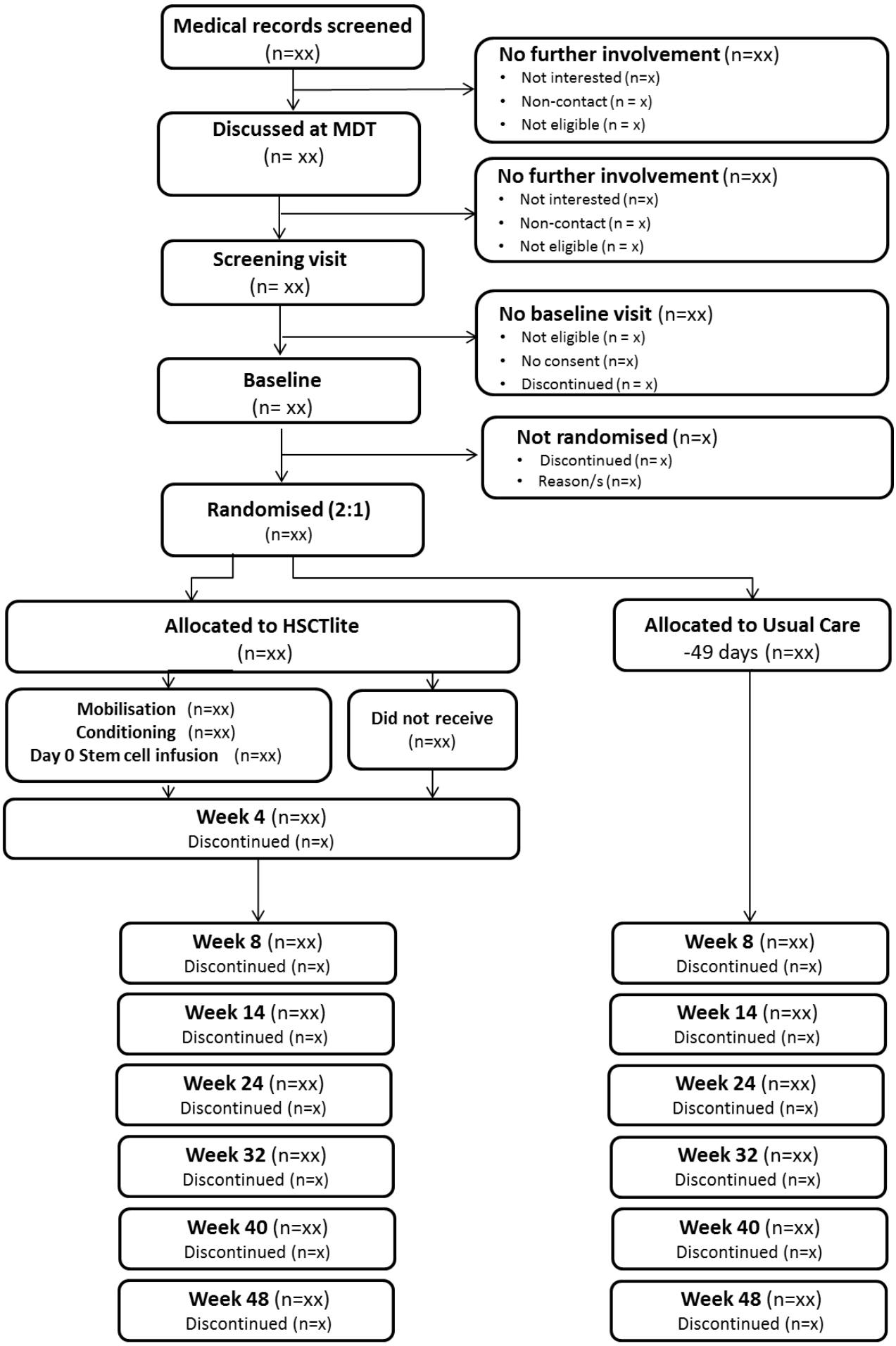

Medical records screened
(n=xx)
No further involvement (n=xx)
Not interested (n=x)
Non-contact (n = x)
Not eligible (n = x)
Discussed at MDT
(n= xx)
No further involvement (n=xx)
Not interested (n=x)
Non-contact (n = x)
Not eligible (n = x)
Screening visit
(n= xx)
No baseline visit (n=xx)
Not eligible (n = x)
No consent (n=x)
Discontinued (n = x)
Baseline
(n= xx)
Not randomised (n=x)
Discontinued (n= x)
Reason/s (n=x)
Randomised (2:1)
(n=xx)
Allocated to HSCTlite
(n=xx)
Allocated to Usual Care
-49 days (n=xx)
Mobilisation (n=xx)
Conditioning (n=xx)
Day 0 Stem cell infusion (n=xx)
Did not receive
(n=xx)
Week 4 (n=xx)
Discontinued (n=x)
Week 8 (n=xx)
Discontinued (n=x)
Week 14 (n=xx)
Discontinued (n=x)
Week 24 (n=xx)
Discontinued (n=x)
Week 32 (n=xx)
Discontinued (n=x)
Week 40 (n=xx)
Discontinued (n=x)
Week 48 (n=xx)
Discontinued (n=x)
Week 8 (n=xx)
Discontinued (n=x)
Week 14 (n=xx)
Discontinued (n=x)
Week 24 (n=xx)
Discontinued (n=x)
Week 32 (n=xx)
Discontinued (n=x)
Week 40 (n=xx)
Discontinued (n=x)
Week 48 (n=xx)
Discontinued (n=x)

Table 3: Demographics of participants

NB: Tables of other screening/baseline variables will be presented in a similar manner to demographics below.

Variable	Scoring	HSCTlite (n=xx)	Usual Care (n=xx)	All (N=xx)
Centre	Barts Health	xx(xx%)	xx(xx%)	xx(xx%)
	Sheffield	xx(xx%)	xx(xx%)	xx(xx%)
	Nottingham	xx(xx%)	xx(xx%)	xx(xx%)
	Cambridge	xx(xx%)	xx(xx%)	xx(xx%)
	…	xx(xx%)	xx(xx%)	xx(xx%)
Sex	Male	xx(xx)	xx(xx)	xx(xx)
	Female	xx(xx)	xx(xx)	xx(xx)
Age (years)	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
Ethnicity [a]	White [b]	xx(xx%)	xx(xx%)	xx(xx%)
	Mixed/multiple ethnic groups [c]	xx(xx%)	xx(xx%)	xx(xx%)
	Asian/Asian British [d]	xx(xx%)	xx(xx%)	xx(xx%)
	Black/African/Caribbean/Black British [e]	xx(xx%)	xx(xx%)	xx(xx%)
	Other ethnic group [f]	xx(xx%)	xx(xx%)	xx(xx%)
	Prefer not to say	xx(xx%)	xx(xx%)	xx(xx%)
BMI	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx

[a] Main ethnic groups could be collapsed depending on the observed distribution. [b] White: English/Welsh/Scottish/Northern Irish/British, Irish, Gypsy or Irish Traveller, and Any other White background; [c] Mixed/multiple ethnic groups: White and Black Caribbean, White and Black African, White and Asian, and Any other mixed/multiple ethnic groups background; [d] Asian/Asian British: Indian, Pakistani, Bangladeshi, Chinese, and Any other Asian background; [e] Black/African/Caribbean/Black British: African, Caribbean, and Any other Black/African/Caribbean/Black British background; [f] Other ethnic group: Arab, and Any other ethnic group.

Table 4: Continuous baseline characteristics by treatment group and missing data status

NB: variables included in the table are given as an example and are not a comprehensive list of the variables that will be included in the analysis

Variable	Summary Statistic	Completers			Non-completers		
		HSCTlite (n=XX)	Usual Care (n=XX)	All (n=XX)	HSCTlite (n=xx)	Usual Care (n=xx)	All (n=xx)
Age (yrs)	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
BMI	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
Duration of CD	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
Age at CD onset	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
CDAI	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
IBDQ	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)

Table 5: Categorical baseline characteristics by treatment group and missing primary outcome status

Variable	Scoring	Completers			Non-completers		
		HSCTlite (n=xx)	Usual Care (n=xx)	All (n=xx)	HSCTlite (n=xx)	Usual Care (n=xx)	All (n=xx)
Sex	Male	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Female	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Ethnicity	White	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Mixed/multiple ethnic groups	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Asian/Asian British	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Black/African/Caribbean/Black British [e]	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Other ethnic group	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Prefer not to say	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Disease location	L1 Ileal	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	L2 Colonic	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	L3 Ileocolonic	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	L3 Isolated upper disease	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Perianal CD	Yes	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	No	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)

NB: variables included in the table are given as an example and are not a comprehensive list of the variables that will be included in the analysis

Table 6: Treatment summaries and compliance

Variable	Scoring	HSCTlite (n=xx)
Mobilisation		
Cyclophosphamide dose (g)	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Number of days GCF	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Days between cyclophosphamide and harvest of stem cells	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Number of days GCF	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Number of cells harvested $CD34^{+}$	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Disease activity after mobilisation		
Karnofsky performance status	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Harvey Bradshaw Index	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
CRP	Mean(SD) Median(IQR) Min to max	xx(xx) xx(xx to xx) xx to xx
Conditioning and transplantation		
	…	

Table 7: Primary and sensitivity effectiveness analysis: Regression of ileo-colonic ulceration at 48 weeks

Primary outcome Regression of ileo-colonic ulceration	**HSCTlite**	**Usual care**	**Unadjusted Odds ratio (95% CI)**	**Adjusted Odds ratio (95% CI)** [a]	
	N (%)	**N (%)**			p-value*
Intention to treat	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Sensitivity analysis on primary outcome					
CACE	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Removing participants that have died	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Removing participants that had surgery	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Adjusting for baseline predictors of missingness [b]	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Imputation of "worst case" treatment failure	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx
Imputation using 24 week colonscopy data	xx(xx)	xx(xx)	xx (xx to xx)	xx (xx to xx)	x.xx

Usual care is the reference group; [a] adjusted for baseline SES-CD ulcer subscale score (fixed effect) and site (random effect)

Figure 2: Forest plot of primary and sensitivity effectiveness analysis: OR for primary outcome between HSCTlite and Usual Care

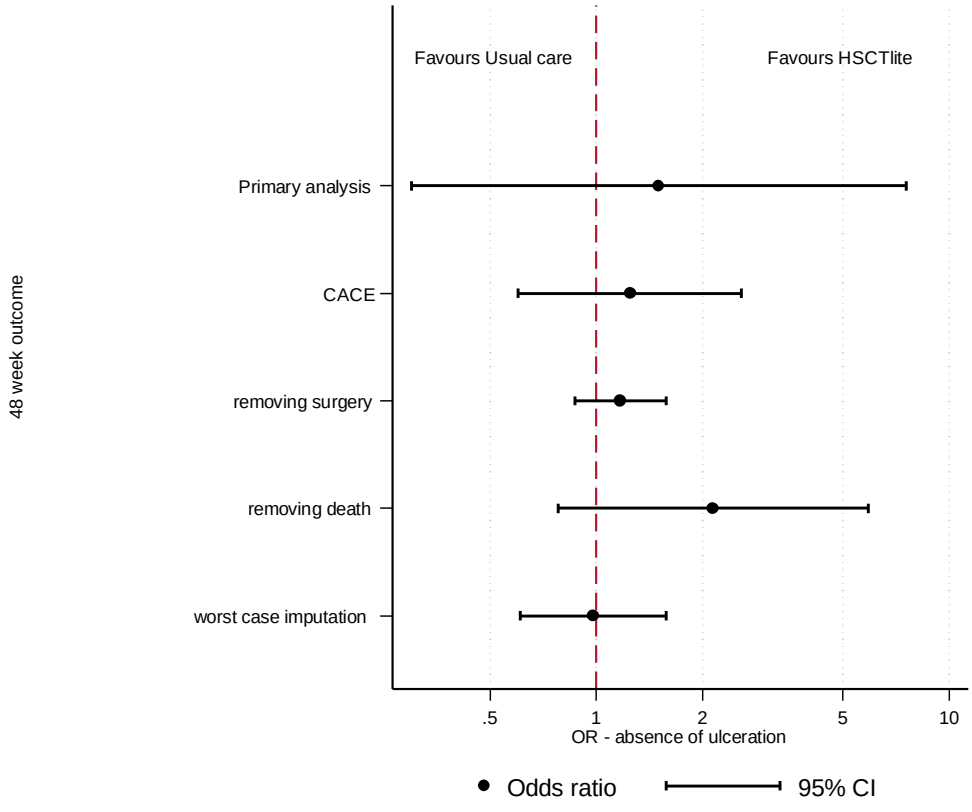


Table 8: Effectiveness analysis: secondary continuous outcomes at 48 weeks

Secondary continuous outcomes	**HSCTlite**		**Usual Care**		**Adjusted mean difference (95% CI)** [a]	**p-value**
	n	**Mean(SD)**	**n**	**Mean(SD)**		

CDAI		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
CDAI change from baseline		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
SES-CD		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
SES-CD change from baseline		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
MaRIA score		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
IBDQ	Bowel systems	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
	Emotion health	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
	Systemic systems	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
	Social function	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
	Total score	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
IBD-control		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D-5L		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D VAS		Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx

Usual care is the reference group; [a] adjusted for baseline assessment of outcome (fixed effect) and site (random effect)

Table 9: Effectiveness analysis: categorical outcomes at 48 weeks

Secondary categorical outcomes	**HSCTlite**	**Usual care**	**Adjusted Odds ratio (95% CI) [a]**	**p-value**
	N (%)	**N (%)**		
Clinical remission (CDAI<150)	xx(xx)	xx(xx)	xx (xx to xx)	x.xx
Steroid free clinical remission (CDAI <150)	xx(xx)	xx(xx)	xx (xx to xx)	x.xx
Clinical remission (Harvey Bradshaw Index ≤4)	xx(xx)	xx(xx)	xx (xx to xx)	x.xx
Clinical remission (PRO2– abdominal pain ≤1, stool frequency ≤1.5)	xx(xx)	xx(xx)	xx (xx to xx)	x.xx
Complete endoscopic remission (SES CD score=0)	xx(xx)	xx(xx)	xx (xx to xx)	x.xx

; [a] adjusted for baseline assessment of outcome (fixed effect) and site (random effect)

Figure 3: Spaghetti plot of CDAI over time for HSCTlite patients

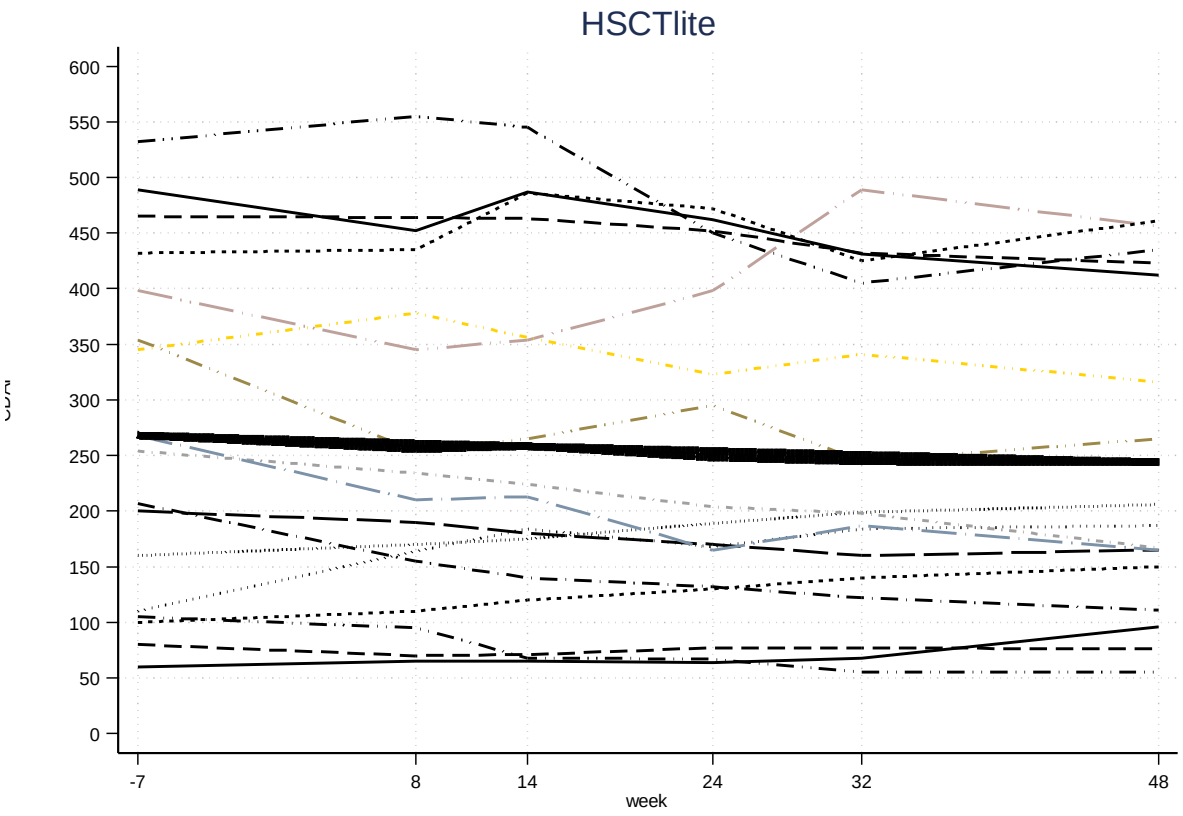


Table 10: Reintroduction of anti TNF therapy efficacy and safety (presented for those receiving maintenance anti-TNF therapy).

Variable	Scoring	Maintenance anti-TNF therapy		No maintenance anti-TNF therapy
		(n=xx)	(95% CI)	(n=xx)
Efficacy				
CDAI (change at 6 weeks after initiation)	Mean(SD)	xx(xx)	(xx to xx)	
CDAI (change at 22 weeks after initiation)	Mean(SD)	xx(xx)	(xx to xx)	
CDAI<150	6 weeks after initiation	xx(xx%)	(xx to xx)	
	22 weeks after initiation	xx(xx%)	(xx to xx)	
Safety				
Number (%) participants	experiencing ≥1 AE	xx(xx)		xx(xx)

Number of all AEs including repeat events		xx(xx)	xx(xx)
AE category	Infectious	xx(xx)	xx(xx)
	GI	xx(xx)	xx(xx)
	Haematological	xx(xx)	xx(xx)
AE related to anti-TNF		xx(xx)	xx(xx)
Number (%) participants experiencing ≥1 SAE		xx(xx)	xx(xx)
Number of all SAEs including repeat events		xx(xx)	xx(xx)
SAE category	Infectious	xx(xx)	xx(xx)
	GI	xx(xx)	xx(xx)
	Haematological	xx(xx)	xx(xx)

Table 11: Safety outcomes: Adverse events by treatment group and timing of event

NB: a similar table to the below will be presented for Serious Adverse Events

Safety outcomes – Adverse Events	Mobilisation phase[a]		Transplant phase[b]		Follow up phase [c]		Total		
	HSCTlite (n=XXX)	Usual Care (n=XXX)	HSCTlite (n=XXX)	Usual Care (n=XXX)	HSCTlite (n= XXX)	Usual Care (n=XXX)	HSCTlite (n= XXX)	Usual Care (n=XXX)	All (n=XXX)
Number (%) of participants who experienced ≥1 AE	XX (xx%)	XX (xx%)	XX (xx%)	XX (xx%)	XXX (xx %)	XXX (xx%)	XXX (xx %)	XXX (xx%)	XXX (xx%)
Number (%) participants experiencing ≥1 AE by relationship to intervention									
Cyclophosphamide	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Filgrastim	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Fludarabine	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Rabbit ATG	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
…	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Number of all AEs (including repeated events)	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
AE Details									
Relationship to intervention									
Cyclophosphamide	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Filgrastim	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Fludarabine	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Rabbit ATG	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
…	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)

Total		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Category										
Infectious		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Viral	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Sepsis	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Localised	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
GI		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Disease flare	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Non-flare symptoms	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
…		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Total		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)

a - period from start of mobilisation to start of conditioning (Randomisation to day 0 for usual care participants)
b - period from start of conditioning up to 100 days from day 0 i.e. day of autologous transplant/infusion (Day 0 to day 100 for usual care participants)
c - from +100 days post- transplant phase to one year assessment (Day 100 to one year assessment in usual care participants.)

Table 12 Safety Outcomes: AEs and SAEs by CTCAE grade

CTCAE Grade:	Grade 1-2		Grade 3		Grade 4	
Safety outcomes – Adverse Events	HSCTlite (n=XXX)	Usual Care (n=XXX)	HSCTlite (n=XXX)	Usual Care (n=XXX)	HSCTlite (n= XXX)	Usual Care (n=XXX)
Number (%) of participants who experienced ≥1 AE	XX (xx%)	XX (xx%)	XX (xx%)	XX (xx%)	XXX (xx %)	XXX (xx%)
Number (%) of participants who experienced ≥1 SAE	XX (xx%)	XX (xx%)	XX (xx%)	XX (xx%)	XXX (xx %)	XXX (xx%)

Number of all AEs (including repeated events)		XXX	XXX	XXX	XXX	XXX	XXX
Number of all SAEs (including repeated events)		XXX	XXX	XXX	XXX	XXX	XXX
AE by Category							
Infectious		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Viral	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Sepsis	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Localised	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
…		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Total		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
SAE by Category							
Infectious		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Viral	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Sepsis	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Localised	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Gi		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Disease flare	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
	Non-flare symptoms	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
…		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Total		xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)

Table 13: Late effects of HSCTlite

Late effects		HSCTlite (n=xx)	Usual Care (n=xx)
General	Weight	xx(xx.x%)	xx(xx.x%)
	Blood pressure	xx(xx.x%)	xx(xx.x%)
	Performance status (Karnofsky/Lansky)	xx(xx.x%)	xx(xx.x%)
Haematology	FVC	xx(xx.x%)	xx(xx.x%)
Renal	Renal function	xx(xx.x%)	xx(xx.x%)
	Urine protein (dipstick)	xx(xx.x%)	xx(xx.x%)
Liver	Liver function	xx(xx.x%)	xx(xx.x%)
	Iron studies (after week 24)	xx(xx.x%)	xx(xx.x%)
…			

Table 14: Exploratory effect of HSCTlite intervention by pre-specified subgroup: primary outcome at 48 weeks

Subgroup	Classification	HSCTlite		Usual Care		Adjusted Odds ratio (95% CI)	P-value
		N	Absence of ulceration n (%)	N	Absence of ulceration n (%)		
Disease Location	Ileal	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Colonic	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Ileocolinic	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
Perianal Crohn's Disease	Yes	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	No	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
Current smoker	Yes	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	No	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx

Figure 4: Example lowess smoother of probability of response in relation to covariate

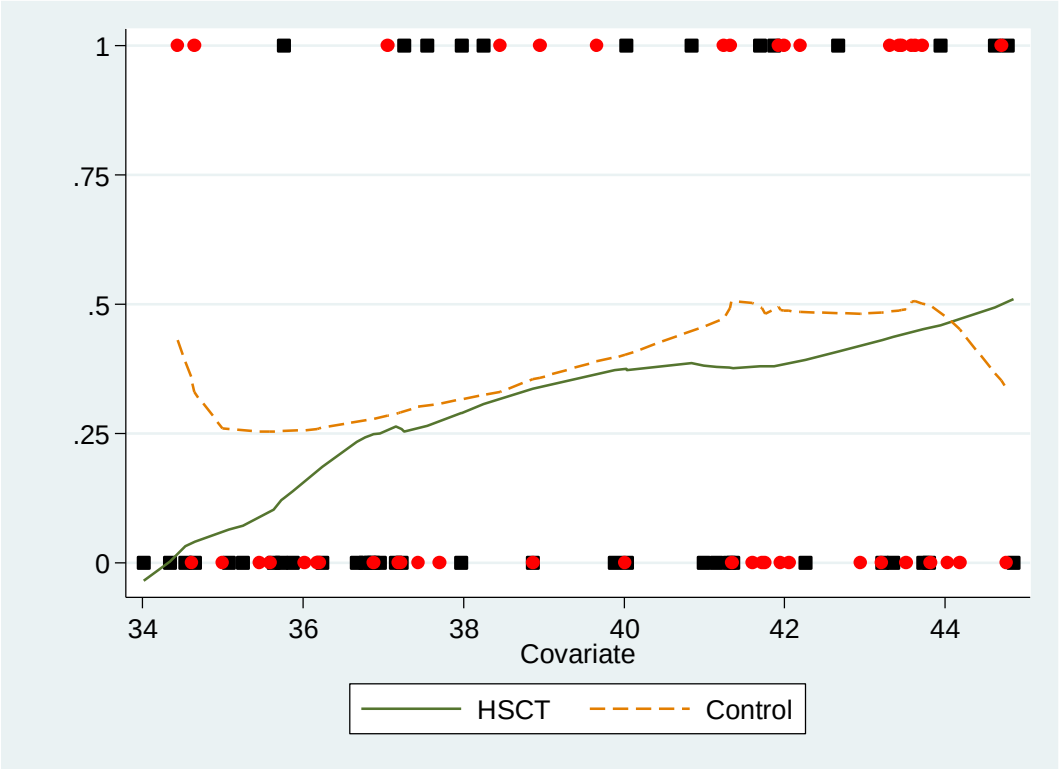


Figure 5: Example kernel estimator of treatment effect in relation to covariate