The University Of Sheffield.

Clinical Trials Research Unit

MAGIC

The MAGIC trial (Melatonin for Anxiety prior to General anaesthesia In Children): A Multicentre, Parallel Randomised Controlled Trial of Melatonin Versus Midazolam in the Premedication of Anxious Children Attending for Elective Dental, Ophthalmologic or ENT Surgery Under General Anaesthesia.

Statistical Analysis Plan

Version 3
23rd January 2023

Protocol Version: 4.1

Authored by:

Esther Herbert
Statistician, Sheffield CTRU
23 / 01 / 2023
Date

Approved by:

Mike Bradburn
Statistician, Sheffield CTRU
23 / 01 / 2023
Date

Professor Anne Schilder
TSC Chair. Professor of Paediatric Otolaryngology,
23 / 01 / 2023
Date

Professor Chris Deery
Chief Investigator, School of Clinical Dentistry, Sheffield
23 / 01 / 2023
Date

Table of Contents

List of Tables

List of Figures

Abbreviations

ASA	American Society of Anaesthesiologists
CHU9D	Child Health Utility 9D
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
CTRU	Clinical Trials Research Unit
DMEC	Data Monitoring and Ethics Committee
FPS-R	Faces Pain Scale – Revised
HRA	Health Research Authority
IQR	Interquartile range
ITT	Intention-to-treat
MAGIC	Melatonin for Anxiety prior to General anaesthesia In Children
MHRA	Medicines and Healthcare products Regulatory Agency
mYPAS-SF	Modified Yale Preoperative Anxiety Scale – Short Form
PAED	Paediatric Anaesthesia Emergence Delirium
PHBQ-AS	Post Hospitalization Behaviour Questionnaire for Ambulatory Surgery
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
SAP	Statistical Analysis Plan
SD	Standard deviation
SOP	Standard Operating Procedure
STAI	State-Trait Anxiety Inventory
TAU	Theatre Admission Unit
TSC	Trial Steering Committee
VSRS	Vancouver Sedation Recovery Scale

1 Introduction

This statistical analysis plan (SAP) is written in conjunction with the International Conference on Harmonisation topic E9 [1], applicable standard operating procedures (SOPs) from the University of Sheffield Clinical Trials Research Unit (CTRU) and trial documents (Protocol, case report form (CRF) and Data Validation Specifications).

The 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 economics evaluation (which will be described elsewhere).

1.1 Background and Rationale

The hospital anaesthetic room is a worrying place for a child, and reducing their distress leads to a better overall experience, and also improves recovery from the anaesthetic, reduces pain after surgery and avoids unnecessary reappointments and delays to operations. Currently, those children with high levels of distress are recommended a "premedication"; that is, a medicine to reduce anxiety ahead of surgery. Midazolam – the current premedication for an anxious child needing an anaesthetic – is effective, although it has many side-effects including loss of coordination and risks to breathing. Midazolam can also have unpredictable effects on anxiety, with some children becoming overexcited rather than being calmed. Melatonin, which also has anxiolytic properties, offers an alternative calming medicine, has shown promise as it avoids midazolam's side-effects.

1.2 Objectives

1.2.1 Feasibility objectives

To undertake an internal pilot trial to determine the feasibility of a full-scale trial, in terms of:

- Recruitment,
- Retention,
- Allocation concealment and blinding.

1.2.2 Clinical Objectives

Efficacy

To evaluate if melatonin, in relation to midazolam is:

- Non-inferior in dealing with pre-operative anxiety evaluated by the modified Yale Preoperative Anxiety Scale – Short Form (mYPAS-SF) over the following three standard preoperative time points recommended for the scale:
    - Start of transfer to theatre
    - Entry to anaesthetic room
    - Administration of anaesthesia
- Superior in dealing with secondary efficacy outcomes (anaesthetic turnaround time, recovery time)
- Non-inferior in dealing with secondary efficacy outcomes (anaesthetic failure rate)

Harms and Safety

- To evaluate if melatonin, in relation to midazolam is superior in dealing with secondary safety outcomes (Paediatric Anaesthesia Emergence Delirium [PAED] score, Vancouver Sedation Recovery Scale [VSRS], Faces Pain Scale – Revised [FPS-R], analgesia requirements, Post Hospitalization Behaviour Questionnaire for Ambulatory Surgery [PHBQ-AS], adverse events [AEs], orientation and cognitive/psychomotor function)
- To describe Serious Adverse Events data (summarised both at participant level and event level) and report listings between the different arms.

2 Study Methods

2.1 Trial Design

MAGIC is a parallel group, double blind, randomised controlled trial (RCT) to evaluate the non-inferiority of melatonin against midazolam in dealing with pre-operative anxiety in children undergoing surgery. The study will be conducted in twenty large NHS trusts and participants will be randomised to receive either midazolam or melatonin in the ratio 1:1. The trial has been designed with an internal pilot phase during the first 6 months of active recruitment which will assess the feasibility aspects of conducting the main trial as guided by pre-planned STOP/GO criteria (see Section 2.5).

2.2 Randomisation

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, on a 1:1 basis, to receive either midazolam or melatonin. 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 delegated person performing the randomisation will access the web-based system, participant details (ID, date of birth and stratification variables) will be entered, and the treatment allocation and randomisation number will be returned. An email confirming randomisation will be sent to the PI, research nurses and trial pharmacist. Randomisation will be stratified by centre, surgical speciality (dental/Ear, Nose and Throat [ENT]/ophthalmology) and gender (male/female) using permuted blocks of random size. Participants will be allocated a treatment pack by pharmacy that relates to the relevant investigational medicinal product (IMP). Pharmacy will blind-label the pack before dispensing.

2.2.1 Changes to the Randomisation System During the Course of the Trial

As a result of the internal pilot phase changes were made to the inclusion criteria to boost recruitment. One of these changes opened up participation in the trial to children undergoing gastroenterology, radiology, plastic, orthopaedic, urology or other general surgeries under general anaesthesia. Under the existing randomisation system this would increase the number of lists from 6 per site to 18 per site. As a result of this and a need to potentially increase the number of sites it was decided that a move to a minimisation system was necessary to avoid imbalance between the randomisation groups. This

decision was approved by the Trial Steering Committee (TSC) and the Data Monitoring and Ethics Committee (DMEC) and was included in a major amendment (SA5) to the protocol which was approved by the Research Ethics Committee (REC), Health Research Authority (HRA) and Medicines and Healthcare products Regulatory Agency (MHRA) in June 2020.

Allocation data from participants randomised under the first randomisation system will be uploaded into the minimisation system, this will be considered as a burn-in period for the minimisation. Participant allocation will be based on probabilistic minimisation using the following factors:

- Site
- Gender (male/female)
- Surgical specialty (head and neck [dental, ophthalmological, ENT], gastroenterology and radiology, other [plastic, orthopaedic, urology, other general surgery])

2.3 Sample Size

A sample size calculation has been based on the primary outcome of mYPAS-SF scores at three time points (start of transfer to Theatre Admission Unit [TAU], entry to TAU and administration of anaesthesia), whilst adjusting for baseline scores, on a non-inferiority basis. Following a review of current literature and expert opinion, it was suggested that a minimum clinically important difference would be 12.96 on the mYPAS-SF scale. A non-inferiority margin of one-third of this will be used, giving a value of 4.3 (0.172 standardised effect size). This represents less than one-point change on any one of the domains within the mYPAS-SF outcome. A standard deviation of 25.03 [2], a correlation of 0.54 [3] between baseline and follow-up measures, a one-sided test with an alpha of 2.5% and power of 90% assuming no difference between the drugs, results in a sample size requirement of 592 participants (296 per arm). Assuming 5% attrition, 624 children will need to be recruited to the trial. One to two parents/carers per child will also be recruited into the trial (therefore 624-1248 parents/carers).

2.4 Framework

There are different frameworks used within this trial. The analysis of the primary outcome is on a non-inferiority basis. The analysis of the secondary outcomes are on either superiority or non-inferiority, see Section 1.2 for more details.

2.5 Statistical Interim Analyses and Stopping Guidance

The following committees will be established:

- Data Management and Ethics Committee (DMEC) – established with an independent chair that will adhere to the Standard Operating Procedure of the CTRU.
- Trial Steering Committee (TSC) - consist of an independent chair, dental professionals, ENT clinician and anaesthetists and one patient representative. The committee will meet every 6 months from the start of the trial.
- Trial Management Group (TMG) – oversee the day-to-day management of the trial and will comprise the core members of the team.

2.5.1 Internal Pilot

There will be an internal pilot which will occur after 6 months of recruitment where it is expected that 195 children will have been recruited and 160 of these will have completed their 2-week follow-up.

Site Set-up

Red:	Less than 5 sites open
Amber:	5-10 sites open
Green:	More than 10 sites open

Recruitment

Red:	Recruitment of less than 78 participants (40% target for pilot; 12.5% of target for full trial)
Amber:	Recruitment of between 79 and 155 participants (40% < pilot target < 80%; 12.5% < full trial target < 25%)
Green:	Recruitment of 156 or more participants (80% of the pilot target; 25% of the full trial target)

Retention (adverse events reporting and PHBQ follow-up)

Red:	Retention of less than 64 participants (approx. 40% of those completed 2-week follow-up)
Amber:	Retention of between 65 and 127 participants
Green:	Retention of 128 or more participants (approx. 80% of those completed 2-week follow-up)

Allocation concealment and blinding

Data on confirmed and suspected unblinding will be presented to the TSC and DMEC at the end of the internal pilot phase.

2.6 Timing of Final Analysis

Prior to database freeze the statistician will have access to blinded data in which any links to treatment group (including the suspected unblinding) are anonymised. At that point the statistician will check the blind data and issues will be resolved where possible. At database freeze, data will then be unblinded in order to perform final data review prior to lock, after which the statistician will conduct the analyses. Further details on this process are given in the MAGIC Data Management Plan.

2.7 Timing of Outcome Assessments

Assessments	Screening	Baseline	30 minutes pre-transfer to theatre	Start of transfer to theatre	Entry to anaesthetic room	Induction of anaesthesia	During surgery	Arrival at PACU	Post-surgery	Discharge	14 days post-surgery
Eligibility assessment	X										
Consent		X									
Current medication		X									
Demographics		X									
ASA score		X									
STAI [a]		X									
mYPAS-SF		X		X	X	X					
QoL – CHU9D		X									X
Vital signs		X							X [b]	X	
Cooperation score		X							X [b]		
IMP administration			X								
Adverse events			X	X	X	X	X	X	X	X	X
Clinical outcome data [c]					X	X	X	X	X	X	
Concomitant medications					X	X	X		X		X
FPS-R									X [b]		
PAED index									X [b]		
VSR									X [b]		
PHBQ-AS											X
Resource Use		X									X

[a] To be completed by one or both parents/legal guardians.

[b] To be repeated every 15 minutes for 2 hours unless the child is ready for discharge earlier than this.

[c] Failure to progress with anaesthesia, time of entry to anaesthetic room, time of anaesthetic induction, point of unconsciousness, use of sevoflurane, local anaesthetic type, amount and concentration, time of surgery completion, time of entry in PACU, time to discharge readiness, and time to actual discharge.

3 Statistical Principles

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.

3.1 Confidence Intervals and P-values

The 95% confidence intervals for the difference between treatment groups will be reported as well as the associated p-value. Non-inferiority will be declared for the primary endpoint if the upper limit of 95% confidence interval on the difference (melatonin vs midazolam) does not exceed 4.3. Superiority will be declared for secondary endpoints (anaesthetic turnaround time, recovery time, PAED, VSRS, FPS-R, analgesia requirements, PHBQ-AS, orientation and cognitive/psychomotor function) if the associated p-value is less than 0.05. Non-inferiority for the secondary endpoint anaesthetic failure rate will be declared if the upper limit of the 95% confidence interval on the difference does not exceed 1.7%.

3.2 Adherence and Protocol Deviations

As this is a single-dose study, dose modifications will not be permitted. If a child spits out their drug, then they will be observed to see if the drug has taken effect. If so, they shall continue with their assessments as per the protocol. If the child is not fit for surgery following spitting out the drug, then the child will be discontinued from trial treatment and treated as a withdrawn participant (see Section 4.4).

All incidents of spitting out IMP will be recorded on the unblinding CRF. Proportions of children spitting out the drug and other protocol deviations (recorded on the non-compliance CRF) will be summarised by treatment group and overall.

There is a potential for discrepancies in IMP dose calculations due to rounding errors. Doses within ±1ml of the correctly calculated dose will be considered as per-protocol (PP). However, all doses outside of this range will constitute as a major protocol deviation and those participants will be excluded from the PP populations (see Section 3.3). Instances of incorrect doses (> ±1ml of the correct dose) and any associated (AEs) will be summarised.

Due to the half-life of both treatments the IMP should ideally be administered between 20-75 minutes prior to the start of transfer to theatre. However, this is often impractical, and delays may occur which prolong the time to surgery. Where the start of transfer to surgery occurs outside of this window (i.e., < 20 minutes or > 75 minutes after IMP administration) the participant will be excluded from the PP populations (see Section 3.3).

Time from IMP administration to start of transfer to theatre will be summarised by treatment group and overall. The impact of transfer window on the treatment effect will be assessed as described in Section 5.2.1.

All major protocol deviations will be assessed by the study team to determine if they have the potential to impact treatment (e.g. IMP dose >±1ml from correct dose). Where this is the case, these participants will be excluded from both PP populations (see Section 3.3). Major protocol deviations which do not impact on the

administration of treatment (e.g. failure to report an SAE within 24 hours) will be summarised but the participants will be included in the PP populations. These decisions will be recorded on the log of protocol deviations maintained by the study manager.

3.3 Analysis Populations

The following analysis populations will be created:

Intention-to-treat Population (ITT)	The ITT population will be composed of all participants that are randomised regardless of drug intake, non-compliance, protocol deviations or withdrawals that occur post-randomisation. Participants will be analysed based on the treatment they were randomisation to. Patients will be included within the analysis as long as they have completed the baseline measure and at least one of the three follow-up measures.
Per-protocol Population 1 (PP1)	The first per-protocol population will be composed of all participants that are randomised, took the whole dose of the study drug, within the allowed time window (20-75 minutes), and had no major protocol deviations which altered adherence to trial treatment as defined above. The participants will be analysed based on the treatment they received as defined by the protocol, as there cannot be any cross-over within the study between treatments, this will always be the treatment group that the participant was randomised to.
Per-protocol Population 2 (PP2)	As with per-protocol population 1 but did not need to take the whole dose of the drug but received at least some of the drug.
Safety Population	The safety population will be composed of all participants who received at least some of the study drug. The participants will be analysed based on treatment they were receiving.

4 Trial Population

4.1 Screening Data

Wherever possible, potential participants will be identified at the time of pre-operative assessment before the day of surgery and will be approached with an information sheet. Where this is not possible it will be done on the day of surgery. Eligibility cannot be confirmed until the day of surgery, but some participants may be identified as ineligible at this stage. On the day of surgery potential participants will be approached to enrol in the study.

A screening log will be maintained at each site to collect:

- Surgical speciality;
- Gender and age;
- Whether the child and parent/carer were interested in the study;
- Ineligibility with reason linked to inclusion and exclusion criteria (see Section 4.2).

Screening logs will be summarised as in Table 8.1, surgical speciality and demographic data will be used to describe the representativeness of the trial sample.

4.2 Eligibility

Patients will be eligible if the meet the following inclusion and exclusion criteria.

4.2.1 Inclusion Criteria

1. Children aged 3-14 years
2. Children undergoing elective dental, ophthalmological, ENT, gastroenterology, radiology, plastic, orthopaedic, urology or other general surgery under general anaesthesia.
3. Pragmatically assessed by healthcare professionals as requiring premedication as per local standard care for high/expected high levels of preoperative distress prior to surgery.
4. American Society of Anaesthesiologists (ASA) grades I & II.
5. Parent or person with parental responsibility able to give written, informed consent.

4.2.2 Exclusion Criteria

1. Not undergoing elective, day-case surgery under general anaesthesia.
2. Not displaying level of anxiety that would usually warrant premedication under the standard NHS care pathway.
3. Reason for premedication other than anxiety.
4. Current prescription of melatonin, midazolam or other non-permitted drug (see Section 7.11.2 of the protocol).
5. Obstructive sleep apnoea.
6. ASA grades III, IV & V.
7. Severe learning disability rendering child unable to communicate even with specialised support.
8. Child verbally declines to participate in the trial.

Reasons for ineligibility will be summarised for participants listed as not eligible on the screening log and for those who are listed as interested but are deemed ineligible later.

4.3 Recruitment

A CONSORT style flow diagram will be presented to summarise the recruitment and 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 8.1 Figure 8.1. For the purpose of the CONSORT diagram, data completeness will be based on the primary outcome as defined in Section 1.2.2.

Monthly recruitment by site and overall will be summarised graphically in cumulative recruitment graphs and tables.

4.4 Withdrawal/Follow-up

A table giving detailed reasons for withdrawal by treatment group and stage will be presented alongside the CONSORT flow diagram. This table will also summarise who withdrew consent/assent (the child, the primary caregiver, the secondary caregiver).

Participants may withdraw from surgery during the trial. Participants for whom the decision to withdraw from surgery is made before receiving the blinded IMP will not be followed up but will be included in the ITT analyses of available outcomes. Participants withdrawing from surgery after receiving the blinded IMP will be followed-up to 14 days. Surgery abandonment will be summarised as in Table 8.8.

4.5 Baseline Participant Characteristics

Summaries of baseline variables will be presented by treatment group and overall for the three analysis populations described in Section 3.3 (see Section 8.2: Table 8.2, Table 8.3, Table 8.4, Table 8.5, Table 8.6 and Table 8.7). The baseline data will be assessed for comparability between groups, any noted differences will be described and considered for adjustment in sensitivity analyses to assess the robustness of the primary analysis [4]. No statistical testing will be undertaken on baseline data. The following summaries will be presented:

Child demographics	Categorical variables • Centre • Gender • Ethnicity • ASA Physical Status (ASA I or II [a]) Continuous variables • Age • Height (cm) • Weight (kg)
Parent/carer demographics (primary and secondary where applicable)	Categorical variables • Gender • Relationship to child Continuous variables • Age
Procedure details	Categorical variables • Speciality • Specific surgery type (see Table 8.4)
Child anxiety/health at baseline	Continuous variables • mYPAS-SF • Cooperation score • CHU9D
Parent/carer anxiety at baseline (primary and secondary where applicable)	Continuous variables • STAI
Vital Signs	Continuous variables • Heart rate (bpm) • Respiratory rate (Breaths/minute) • Blood pressure (mmHg) • SPO_2 (%)

[a] ASA grades III to V are an exclusion criterion (see Section 4.2.2)

5 Analysis

5.1 Outcome Definitions

5.1.1 Primary Outcome

Modified Yale Preoperative Anxiety Scale – Short Form (mYPAS-SF)

The mYPAS-SF [5] measures a child's preoperative distress and is completed by a trained member of the research team. The mYPAS-SF will be measured at three time points: prior to theatre transfer, on entry into anaesthetic room and on induction of anaesthesia. The questionnaire consists of four questions, three of which are scored 1:4, the fourth is scored 1:6. The total mYPAS-SF score is calculated by dividing each item rating by

the highest possible rating (i.e., 6 for the "vocalizations" item and 4 for all other items), adding all the produced values, dividing by 4, and multiplying by 100. The range of the score is 22.92-100 with higher scores indicating more anxiety.

5.1.2 Secondary Outcomes

Anaesthetic Failure Rate

The anaesthetic failure rate is the proportion of participants for whom surgery is abandoned before the point of unconsciousness, reasons for abandonment will also be collected as stated in Section 4.3.

Anaesthetic Turnaround Time

Time of entry to the anaesthetic room and the time at which the anaesthetist leaves the participant in recovery will be recorded, anaesthetic turnaround time will be calculated as the duration in minutes between these two times.

Recovery Time

Time of surgery completion and discharge readiness will be recorded in the participants' CRFs. Recovery time will be calculated as the duration in minutes between these two times. If the surgery is abandoned after the point of unconsciousness, then the abandonment time will be recorded and recovery time will be calculated as the duration in minutes between abandonment and discharge readiness. If the surgery is abandoned before the point of unconsciousness, then the participant will be removed from summaries and analysis of recovery time. Recovery time for completers and abandoners will be analysed separately and overall. In the unlikely event that discharge readiness time is missing then discharge time will be used instead and it will be assumed that there was no delay after discharge readiness. There should be no censoring as all participants should have been discharged by the end of the study.

Paediatric Anaesthesia Emergence Delirium Scale (PAED)

The PAED [6] scale measures emergence delirium[1] in children, it will be recorded by a research nurse every 15 minutes after surgery for two hours unless the child is ready for discharge earlier than this. The scale is made up of 5 items each of which is rated 4 = not at all, 3 = just a little, 2 = quite a bit, 1 = very much, or 0 = extremely. Scoring for items 4 and 5 are reversed. The scores of each item are summed to obtain the PAED scale score which ranges from 0-20. Higher scores indicate higher emergence delirium.

If three or more items are missing the score will not be calculated, if one or two items are missing then a weighted sum will be computed [7] [8].

Vancouver Sedation Recovery Scale (VSRS)

The VSRS [9] measures recovery from sedation and will be recorded by a research nurse every 15 minutes post-surgery for two hours unless the child is ready for discharge earlier than this. The scale is made up of 12 items each with between 2-5 options rated from 0 upwards. e.g. item F has three options 2 = Accommodates, 1 = No

[1] Mental disturbance when recovering from general anaesthesia

attempt to accommodate and 0 = unable to accommodate. The total score is the sum of the items' scores and ranges from 0 (unable to arouse) to 22 (fully awake).

Faces Pain Scale – Revised (FPS-R)

The FPS-R [10] is a self-reported measure to assess children's pain levels, it consists of six faces scored 0-2-4-6-8-10 where 0 is no pain and 10 is a lot of pain. A child is asked to point to the face that shows how much they hurt. The FPS-R will be recorded every 15 minutes post-surgery for two hours unless the child is ready for discharge earlier than this.

Analgesia Requirements

Concomitant medications received in, and post hospital (up to 14 day follow-up) will be reviewed to determine whether participants received analgesics to relieve pain after surgery up to when they are ready for discharge from hospital.

Post Hospitalization Behaviour Questionnaire for Ambulatory Surgery (PHBQ-AS)

The PHBQ-AS [11] asks parents to compare their child's behaviour before surgery to their current behaviour, 14 days after surgery. The questionnaire has 11 behaviours each scored from 1 (much less than before) to 5 (much more than before). An option of "not-applicable" is also available for each item and these incidences will be scored as 3 (no change) and the proportions of individuals using this for this item will be reported. The total score is calculated as an average of the scores for the 11 questions, this produces a continuous scale between 0 (improvement in behaviour) and 5 (maladaptive behavioural changes) with a score of 3 indicating no change.

Scores will be imputed for missing items where 10% or fewer of the items are missing for an individual. This will be done by calculating the mean value for the missing item across all participants with complete data.

Co-operation Score

The co-operation score [12] will be measured at baseline and every 15 minutes post-surgery for two hours unless the child is ready for discharged earlier than this. The score has 5 different elements, two are scored between 0-2 and the other three are scored between 0-1. The total score is the sum of the elements and rages from 0 (no cooperation) to 7 (all tasks performed correctly).

State-Trait Anxiety Inventory (STAI)

The STAI [13], recorded at baseline, is a self-reported measure of a parent's current symptoms of anxiety and their general propensity to be anxious. There are two subscales: the state anxiety scale (S-Anxiety, questions 1-20) and the trait anxiety scale (T-Anxiety, questions 21-40). Item scores are added to obtain subtest total scores. Scoring should be reversed for anxiety-absent items (19 items of the total 40). Both subsets result in scores between 20-80 and higher scores indicate more anxiety in both domains.

Child Health Utility 9D (CHU9D)

The CHU9D [14] is a paediatric, generic, preference based measure of health related quality of life. For children aged 5-6 the questionnaire is completed by the parent/carer, children aged 7-14 complete the form themselves. The questionnaire consists of 9 dimensions each with 5 possible responses. A utility score is calculated by

assigning utility values to each response and then summing these values. A CHU-9D score can only be calculated if all questions have been answered. The CHU-9D utility score ranges from 0.33 to 1 where higher scores represent greater health related quality of life.

5.2 Analysis Methods

5.2.1 Primary Outcome

The primary outcome (mYPAS-SF score) will be analysed using a linear mixed model:

Fixed effects:
- Treatment
- Time point (start of transfer to theatre, entry to anaesthetic room, induction of anaesthesia)
- Baseline mYPAS-SF score
- Gender
- Speciality group (head and neck, gastro and MRI, other)
- Site

Random effects:
- Participant (random intercept)

The model will be fitted using restricted maximum likelihood (REML) estimation.

The point estimate and its related 95% confidence intervals for the difference between treatment groups will be reported as well as the associated p-value.

This model will be fitted for the ITT, PP1 and PP2 analysis populations. The ITT and PP1 populations will be co-primary, we will require both the PP1 and ITT analyses to demonstrate statistically significant evidence of non-inferiority to declare that the treatment is non-inferior. If the results of the analyses are discrepant (e.g. the ITT rejects the null of inferiority but the PP1 analysis does not, or vice versa) then we will report the conflicting results from both analyses highlighting the inconclusive nature of the results. Since both analyses are required to be significant to declare non-inferiority there is no need for multiplicity adjustment [15].

Model Checks

Model goodness of fit will be investigated via graphical methods (e.g. histograms of residuals and scatterplots of residuals vs. covariates). If residuals do not look to be normally distributed then a Poisson/negative-binomial model or appropriate transformations (e.g., a log transform) may be used. Influential observations and outliers will also be investigated and sensitivity analyses at the discretion of the trial statistician will be undertaken and reported. The model will also be fitted with a random slope which will be compared to the model with a random intercept only using the generalised likelihood ratio test. This model will have an unstructured covariance matrix.

Sensitivity Analyses

As this analysis assumes that there is going to be a consistent treatment effect over time, a sensitivity analysis will be completed using the same model as outlined above but including the interaction term of time and treatment to evaluate if this assumption is valid.

The effect of the transfer window

A further sensitivity analysis will assess the impact of the transfer window (the time from IMP administration to the start of transfer to theatre) on the primary outcome. An interaction term between treatment and transfer

window (as a continuous variable) will be added to the model and a graph of adjusted, average, mYPAS-SF scores vs time window, by treatment group will be produced.

5.2.2 Secondary Outcomes

Continuous, longitudinal secondary outcomes will be analysed in the same way as the primary outcome. Binary, longitudinal outcomes will be analysed using a generalised linear mixed model with logit link function, similar to that for the continuous outcomes.

Post-surgery outcomes collected repeatedly until discharge readiness will be analysed by modelling time to discharge and the longitudinal outcome jointly [16] using the JM [17] package in R [18].

Outcomes measured at a single post-baseline time point will be analysed with a linear model or a generalised linear model as appropriate. The models will include treatment, baseline measurement (where applicable), gender, speciality group.

In some cases, the post-surgery measurements taken every 15 minutes until 2 hours or discharge readiness may be recorded for longer than is necessary. Extra measurements (i.e., after 2 hours or after the discharge ready time) will be excluded from the analyses of these outcomes.

Differences between treatment groups for binary outcomes will be reported as odds ratios with associated 95% confidence intervals and p-values.

Secondary outcomes will be analysed for the ITT and PP1 populations.

Model Checks

Model goodness of fit will be investigated via graphical methods (e.g., histograms of residuals and scatterplots of residuals vs. covariates). If residuals do not look to be normally distributed then appropriate transformations (e.g., a log transform) may be used. Influential observations and outliers will also be investigated and sensitivity analyses at the discretion of the trial statistician will be undertaken and reported.

Sensitivity Analyses

FPS-R has been shown to be appropriate for use with children from age 4 or 5 onward. Therefore, a sensitivity analysis will be undertaken, analysing the FPS-R excluding children aged 3.

5.3 Missing Data

Missing items on a score will be handled as described in the Section in 5.1 specific to the score's rules.

In the case of missing data on the primary endpoint, the missing data mechanism will be explored and multiple imputation may be applied as a sensitivity analysis as appropriate. Other sensitivity analyses will be performed in order to evaluate the robustness of the primary analysis as detailed in Section 5.2.1).

5.4 COVID-19

Due to the global COVID-19 pandemic recruitment to the trial paused after the internal pilot phase. As the trial has a short follow-up period all participants recruited had been followed-up prior to the pause. However, patient experience upon commencement of recruitment will be different and this is likely to have an impact on

children's anxiety. Guidance on adjustment to analyses due to the impact of COVID-19 on trials has been developed and we will follow the recommendations outlined by the European Medicines Agency (EMA)[19].

Sub-group analyses of the primary outcome for those recruited before and after the pause in recruitment are expected to determine whether treatment effects differed between the two groups.

5.5 Subgroup Analyses

Subgroup analyses will be performed for the primary outcome. The following subgroups of interests for exploratory analyses have been pre-specified:

1) Gender (male/female)
2) Surgical speciality (head and neck, gastro and MRI, other)

Heterogeneity will be explored through an overall interaction test by fitting an interaction term between the intervention arm and subgroup indicator into the primary model (see Section 5.2.1). The intervention effects (point estimates and associated 95% CIs) will be obtained in each category of the subgroup and visually displayed using a forest plot [20]. The overall interaction test (intervention arm × subgroup) rather than calculating separate p-values within each category of the subgroup will be used to examine the strength of evidence for treatment heterogeneity across subgroups [21]–[23]. This analysis will be undertaken for the ITT and PP1 analysis populations.

5.6 Harms and Safety

Safety information will be summarised for the safety population (see Section 3.3). The following figures will be presented overall and by treatment received:

- The number of AEs
- The number and percentage of participants reporting an AE
- The number and percentage of participants reporting a SAE
- The number and percentage of participants reporting a treatment related AE
- A list of all AEs

For all AEs the intensity, relation to study drug, actions taken, and outcomes will be summarised by overall and by treatment received (see Table 8.13). For SAEs the seriousness, frequency and expectedness will also by summarised (see Table 8.14).

5.7 Analysis following early stopping

In the event that the trial is stopped early for feasibility, all analyses will be performed with the following exceptions:

- P-values will not be calculated but CIs will still be included (Section 3.1)
- Outcomes will not be assessed for superiority or non-inferiority (Section 2.4)
- Subgroup analyses will not be performed (Section 5.5)

- Sensitivity analyses will not be performed including the analysis looking at the effect of the transfer window on the primary outcome (Sections 5.2.1 and 5.2.2)
- Outcomes will be analysed for the ITT and PP1 populations only (Section 3.3)
- Multiple imputation will not be used to assess the impact of missing data (Section 5.3)

5.8 Statistical Software

All statistical analyses will be performed in a validated statistical software package such as R [18] or STATA [24].

6 Document History

Version	Date Approved	Modifications (with section)	Prior to/After blind review Prior to/After database lock
1.0	18-10-2019	-	Prior to blind review and database lock.
2.0	14-09-2020	**Section 2.2** - Details of move to a minimisation system for randomising participants. **Section 3.2** - Details of the required timing of IMP administration. **Section 4.2** - Updated inclusion criteria to include 3-4 year olds and extra surgical specialities. - Updated exclusion criteria to exclude children who verbally decline to participate in the trial. **Section 5.2** - Updated model covariates to take into account the new specialities - Added a sensitivity analysis for the FPS-R. **Section 5.4** - New section detailing plans for sensitivity analyses due to COVID-19. **Section 7** - Updated protocol version	Prior to blind review and database lock.
3.0	As signed	**General changes** - Changed discharge to discharge readiness throughout **Section 3.2** - Clarification on dose errors - Details about transfer windows - Details about major protocol deviations **Section 3.3** - Clarify the PP1 analysis set **Section 4.5** - State which analysis populations will be presented **Section 5.1.2** - Give details on how to handle missing discharge readiness time **Section 5.2.1** - More model detail given - Analysis populations specified - Further analysis to investigate the impact of transfer windows **Section 5.2.2** - More model detail given	After blind review and prior to database lock.

		- Handling of frequent scores/measurements **Section 5.5** - New section on subgroup analyses **Section 8.1** - Clearer CONSORT diagram **Section 8.2** - Clearer and additional tables	

7 References

Trial Documents

Title	Version	Date	Location
Study Protocol	4.1	28th September 2020	X:\ScHARR\PR_MAGIC\General\01 Study Documents\1.1 Approved Protocol
Data Management Plan	3	16th January 2023	X:\ScHARR\PR_MAGIC\General\09 Study Management\9.5 Data Management Arrangements\DMP

CTRU Standard Operating Procedures

Title	Version	Date	Location
ST001 The Statistical Analysis Plan	6	10th December 2021	X:\ScHARR\CTRU_SOPs_Current
ST007 Randomisation using the CTRU randomisation system	3	9th December 2021	
DM012 Study database lock and retention	6	22nd March 2022	

[1] International Conference on Harmonisation, "International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9," 2010.

[2] Z. N. Kain *et al.*, "Family-centered Preparation for Surgery Improves Perioperative Outcomes in ChildrenA Randomized Controlled Trial," *Anesthesiology*, vol. 106, no. 1, pp. 65–74, Jan. 2007.

[3] L. Frison and S. J. Pocock, "Repeated measures in clinical trials: Analysis using mean summary statistics and its implications for design," *Stat. Med.*, vol. 11, no. 13, pp. 1685–1704, Jan. 1992.

[4] European Medicines Agency (EMA), "Guideline on adjustment for baseline covariates in clinical trials." 2015.

[5] B. N. Jenkins, M. A. Fortier, S. H. Kaplan, L. C. Mayes, and Z. N. Kain, "Development of a short version of the modified Yale Preoperative Anxiety Scale.," *Anesth. Analg.*, vol. 119, no. 3, pp. 643–50, Sep. 2014.

[6] N. Sikich and J. Lerman, "Development and psychometric evaluation of the pediatric anesthesia emergence delirium scale.," *Anesthesiology*, vol. 100, no. 5, pp. 1138–45, May 2004.

[7] N. J. J. F. Janssen *et al.*, "On the utility of diagnostic instruments for pediatric delirium in critical illness: An evaluation of the Pediatric Anesthesia Emergence Delirium Scale, the Delirium Rating Scale 88, and

the Delirium Rating Scale-Revised R-98," *Intensive Care Med.*, vol. 37, no. 8, pp. 1331–1337, Aug. 2011.

[8] R. J. Blankespoor, N. J. J. F. J. F. Janssen, A. M. H. H. Wolters, J. Van Os, and J. N. M. M. Schieveld, "Post-hoc revision of the pediatric anesthesia emergence delirium rating scale: clinical improvement of a bedside-tool?," *Minerva Anestesiol.*, vol. 78, no. 8, pp. 896–900, Aug. 2012.

[9] A. J. Macnab, M. Levine, N. Glick, L. Susak, and G. Baker-Brown, "A research tool for measurement of recovery from sedation: the Vancouver Sedative Recovery Scale.," *J. Pediatr. Surg.*, vol. 26, no. 11, pp. 1263–7, Nov. 1991.

[10] C. L. Hicks, C. L. Von Baeyer, P. A. Spafford, I. Van Korlaar, and B. Goodenough, "The Faces Pain Scale - Revised: Toward a common metric in pediatric pain measurement," *Pain*, vol. 93, no. 2. No longer published by Elsevier, pp. 173–183, 01-Aug-2001.

[11] B. N. Jenkins *et al.*, "Revisiting a measure of child postoperative recovery: Development of the Post Hospitalization Behavior Questionnaire for Ambulatory Surgery," *Paediatr. Anaesth.*, vol. 25, no. 7, pp. 738–745, Jul. 2015.

[12] J. C. RAEder and G. MisvAEr, "Comparison of propofol induction with thiopentone or methohexitone in short outpatient general anaesthesia," *Acta Anaesthesiol. Scand.*, vol. 32, no. 8, pp. 607–613, Nov. 1988.

[13] L. J. Julian, "Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A).," *Arthritis Care Res.*, vol. 63 Suppl 1, no. SUPPL. 11, pp. S467-72, Nov. 2011.

[14] K. Stevens, "Valuation of the Child Health Utility 9D Index," *Pharmacoeconomics*, vol. 30, no. 8, pp. 729–747, Aug. 2012.

[15] European Medicines Agency (EMA), "Guideline on multiplicity issues in clinical trials (draft)," 2016.

[16] R. Henderson, P. Diggle, and A. Dobson, "Joint modelling of longitudinal measurements and event time data," *Biostatistics*, vol. 1, no. 4, pp. 465–480, 2000.

[17] D. Rizopoulos, "JM: An R package for the joint modelling of longitudinal and time-to-event data," *J. Stat. Softw.*, vol. 35, no. 9, pp. 1–33, Jul. 2010.

[18] R Core Team, "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, 2018.

[19] European Medicines Agency (EMA), "Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials," 2020.

[20] J. Cuzick, "Forest plots and the interpretation of subgroups," *Lancet*, vol. 365, no. 9467. Elsevier, p. 1308, 09-Apr-2005.

[21] R. Wang, S. W. Lagakos, J. H. Ware, D. J. Hunter, and J. M. Drazen, "Statistics in Medicine — Reporting of Subgroup Analyses in Clinical Trials," *N. Engl. J. Med.*, vol. 357, no. 21, pp. 2189–2194, Nov. 2007.

[22] S. J. Pocock, S. E. Assmann, L. E. Enos, and L. E. Kasten, "Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems," *Stat. Med.*, vol. 21, no. 19, pp. 2917–2930, Oct. 2002.

[23] S. F. Assmann, S. J. Pocock, L. E. Enos, and L. E. Kasten, "Subgroup analysis and other (mis)uses of baseline data in clinical trials," *Lancet*, vol. 355, no. 9209, pp. 1064–1069, Mar. 2000.

[24] StataCorp, "Stata Statistical Software." StataCorp LP, College Station, TX, USA, 2017.

8 Appendix

8.1 CONSORT Flow Diagram

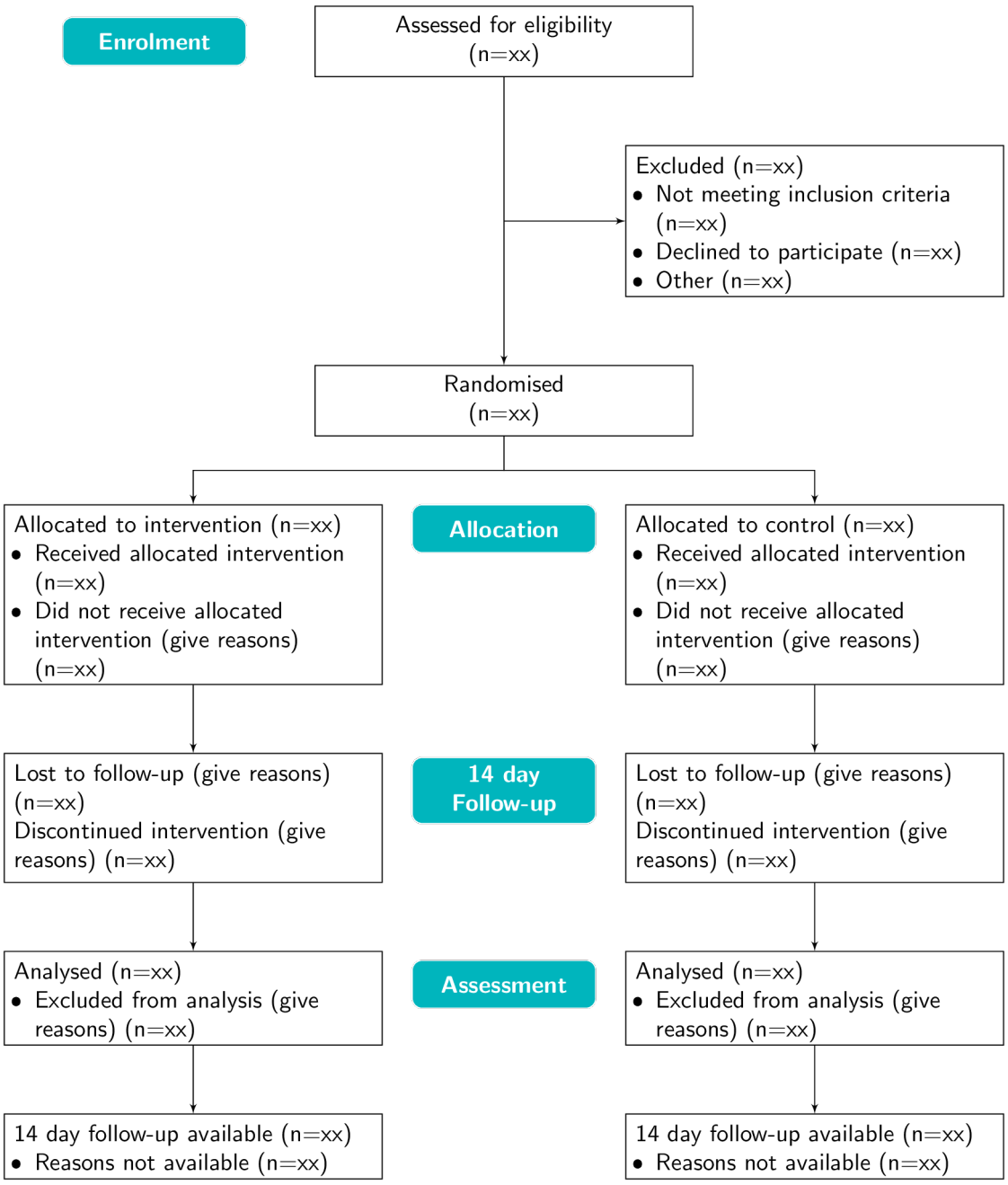

Figure 8.1 Example of the CONSORT flow diagram

8.2 Example Tables

Table 8.1 Summary of screening logs.

		N = xx
Speciality	n	n
	Dental	x (x.x%)
	ENT	x (x.x%)
	Ophthalmology	x (x.x%)
	Gastroenterology	x (x.x%)

		N = xx
	Radiology	x (x.x%)
	Plastic	x (x.x%)
	Orthopaedic	x (x.x%)
	Urology	x (x.x%)
	Other general surgery	x (x.x%)
Gender	n	n
	Male	x (x.x%)
	Female	x (x.x%)
Age (years)	n	n
	Mean (SD)	x (xx)
Outcome	n	n
	Interested	x (x.x%)
	Parent not interested/unable	x (x.x%)
	Child not interested/unable	x (x.x%)
	Not approached	x (x.x%)
	Not Eligible	x (x.x%)
	Other	x (x.x%)

Table 8.2 Demographics and characteristics of child participants

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Age	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Gender	n	n	n	n
	Male	x (x.x%)	x (x.x%)	x (x.x%)
	Female	x (x.x%)	x (x.x%)	x (x.x%)
Height (cm)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Weight (kg)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Ethnicity [a]	n	n	n	n
	White [b]	x (x.x%)	x (x.x%)	x (x.x%)
	Mixed/multiple ethnic groups [c]	x (x.x%)	x (x.x%)	x (x.x%)
	Asian/Asian British [d]	x (x.x%)	x (x.x%)	x (x.x%)
	Black/African/Caribbean/Black British [e]	x (x.x%)	x (x.x%)	x (x.x%)
	Other ethnic group [f]	x (x.x%)	x (x.x%)	x (x.x%)
	Prefer not to say	x (x.x%)	x (x.x%)	x (x.x%)
ASA Physical Status	n	n	n	n
	ASA I	x (x.x%)	x (x.x%)	x (x.x%)
	ASA II	x (x.x%)	x (x.x%)	x (x.x%)

[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 8.3 Demographics and characteristics of the primary and secondary caregivers

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Age	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Gender	n	n	n	n
	Male	x (x.x%)	x (x.x%)	x (x.x%)
	Female	x (x.x%)	x (x.x%)	x (x.x%)
Relationship to child	n	n	n	n
	Biological mother	x (x.x%)	x (x.x%)	x (x.x%)
	Biological father	x (x.x%)	x (x.x%)	x (x.x%)
	Biological parent's partner (living together)	x (x.x%)	x (x.x%)	x (x.x%)
	Foster Carer	x (x.x%)	x (x.x%)	x (x.x%)
	Step-parent	x (x.x%)	x (x.x%)	x (x.x%)
	Grandparent	x (x.x%)	x (x.x%)	x (x.x%)
	Adoptive parent	x (x.x%)	x (x.x%)	x (x.x%)
	Other	x (x.x%)	x (x.x%)	x (x.x%)

Table 8.4 Procedure details

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Dental	Total	n (x.x%)	n (x.x%)	n (x.x%)
	No extraction	n (x.x%)	n (x.x%)	n (x.x%)
	Soft tissue procedure (no extraction)	n (x.x%)	n (x.x%)	n (x.x%)
	Primary tooth extraction	n (x.x%)	n (x.x%)	n (x.x%)
	Permanent tooth extraction	n (x.x%)	n (x.x%)	n (x.x%)
	Surgical extraction	n (x.x%)	n (x.x%)	n (x.x%)
	Management of intrabony lesion	n (x.x%)	n (x.x%)	n (x.x%)
	Other	n (x.x%)	n (x.x%)	n (x.x%)
ENT	Total	n (x.x%)	n (x.x%)	n (x.x%)
	Tonsillectomy	n (x.x%)	n (x.x%)	n (x.x%)
	Gromits	n (x.x%)	n (x.x%)	n (x.x%)
	Other	n (x.x%)	n (x.x%)	n (x.x%)
Ophthalmology	Total	n (x.x%)	n (x.x%)	n (x.x%)
	Lactrymal probing	n (x.x%)	n (x.x%)	n (x.x%)
	Excision & drainage	n (x.x%)	n (x.x%)	n (x.x%)
	Examination	n (x.x%)	n (x.x%)	n (x.x%)
	Squint surgery	n (x.x%)	n (x.x%)	n (x.x%)
	Other	n (x.x%)	n (x.x%)	n (x.x%)
Gastroenterology	*	n (x.x%)	n (x.x%)	n (x.x%)
Radiology	*	n (x.x%)	n (x.x%)	n (x.x%)
Plastic	*	n (x.x%)	n (x.x%)	n (x.x%)
Orthopaedic	*	n (x.x%)	n (x.x%)	n (x.x%)
Urology	*	n (x.x%)	n (x.x%)	n (x.x%)
Other general surgery	*	n (x.x%)	n (x.x%)	n (x.x%)

* Surgery details for these specialities will be collected in free text and categorised retrospectively.

Table 8.5 Baseline measurements of anxiety and health and child participants

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
mYPAS-SF	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Cooperation Score	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
CHU9D (7-14)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
CHU9D (3-6)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx

Table 8.6 Baseline vital signs

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Heart rate (bpm)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Systolic blood pressure (mmHg)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Diastolic blood pressure (mmHg)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
SPO_2 (%)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx

Table 8.7 Parent/carer anxiety at baseline

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
STAI	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx

Table 8.8 Abandonment of surgery

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Was surgery abandoned?	Yes	x (x.x%)	x (x.x%)	x (x.x%)
	No	x (x.x%)	x (x.x%)	x (x.x%)
At what point	Before IMP administration	x (x.x%)	x (x.x%)	x (x.x%)

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
	Between IMP administration and start of transfer to theatre	x (x.x%)	x (x.x%)	x (x.x%)
	Between start of transfer to theatre and entry to anaesthetic room	x (x.x%)	x (x.x%)	x (x.x%)
	Between entry to anaesthetic room and start of cannulation/application of mask	x (x.x%)	x (x.x%)	x (x.x%)
	Between start of cannulation/application of mask and point of unconsciousness	x (x.x%)	x (x.x%)	x (x.x%)
	Between point of unconsciousness and surgery start	x (x.x%)	x (x.x%)	x (x.x%)
	Between surgery start and surgery completion	x (x.x%)	x (x.x%)	x (x.x%)
Time from IMP administration and abandonment (mins)	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx

Table 8.9 Summary of mYPAS-SF scores at each time point by treatment and overall.

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Baseline	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Start of transfer to theatre	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Entry to anaesthetic room	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx
Induction of anaesthesia	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx

Table 8.10 Primary analysis of the mYPAS-SF scores for the ITT and PP analysis populations.

	Midazolam Mean (SD)	**Melatonin** Mean (SD)	**Adjusted Treatment Effect (95% CI)**	**p-value**
ITT	n	n		
Start of transfer to theatre	x (xx)	x (xx)	x (xx, xx)	p=xxx
Entry to anaesthetic room	x (xx)	x (xx)		
Induction of anaesthesia	x (xx)	x (xx)		
PP	n	n		
Start of transfer to theatre	x (xx)	x (xx)	x (xx, xx)	p=xxx
Entry to anaesthetic room	x (xx)	x (xx)		
Induction of anaesthesia	x (xx)	x (xx)		

Table 8.11 Post-surgery measurements (first hour)

		15mins			**30mins**			**45mins**			**1hr**		
		Midazolam	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**
		N = xx	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**
Face pain scale score	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
Cooperation score	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
PAED	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
VSRS	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
Heart rate (bpm)	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	n	n	n	n	n	n	n	n	n	n	n	n	n

		15mins			30mins			45mins			1hr		
		Midazolam	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**	**Midazolam**	**Melatonin**	**All**
		N = xx	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**	**N = xx**
Respiratory rate (breaths/minute)	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
Systolic blood pressure (mmHg)	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
Diastolic blood pressure (mmHg)	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
SPO_2 (%)	n	n	n	n	n	n	n	n	n	n	n	n	n
	Mean (SD)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)	x (xx)
	Median (IQR)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)	x (xx, xx)
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx

Table 8.12 Summary of suspected and formal unblinding

		Midazolam	Melatonin	All
		N = xx	N = xx	N = xx
Unblindings	Total	x (x.x%)	x (x.x%)	x (x.x%)
	Formal	x (x.x%)	x (x.x%)	x (x.x%)
	Suspected	x (x.x%)	x (x.x%)	x (x.x%)
Who was unblinded?	First research nurse [a]	x (x.x%)	x (x.x%)	x (x.x%)
	Second research nurse	x (x.x%)	x (x.x%)	x (x.x%)
	Health economist	x (x.x%)	x (x.x%)	x (x.x%)
	TSC	x (x.x%)	x (x.x%)	x (x.x%)
	CI	x (x.x%)	x (x.x%)	x (x.x%)
	Statistician	x (x.x%)	x (x.x%)	x (x.x%)
	Study manager	x (x.x%)	x (x.x%)	x (x.x%)
	Data manager	x (x.x%)	x (x.x%)	x (x.x%)
	Participant	x (x.x%)	x (x.x%)	x (x.x%)
	PI	x (x.x%)	x (x.x%)	x (x.x%)
	Other	x (x.x%)	x (x.x%)	x (x.x%)
Source of unblinding	Participant spat out the IMP	x (x.x%)	x (x.x%)	x (x.x%)
	Participant reacted in some way to the IMP	x (x.x%)	x (x.x%)	x (x.x%)
	SCRAM	x (x.x%)	x (x.x%)	x (x.x%)
	Other	x (x.x%)	x (x.x%)	x (x.x%)
Reason for unblinding	Accidental	x (x.x%)	x (x.x%)	x (x.x%)
	Emergency	x (x.x%)	x (x.x%)	x (x.x%)
	SUSAR reporting	x (x.x%)	x (x.x%)	x (x.x%)
	Other	x (x.x%)	x (x.x%)	x (x.x%)
Was assessment potentially affected?	Yes	x (x.x%)	x (x.x%)	x (x.x%)
	No	x (x.x%)	x (x.x%)	x (x.x%)

[a] IMP administrator

Table 8.13 Adverse events

		Midazolam		Melatonin		All	
		Events	Individuals	Events	Individuals	Events	Individuals
All AEs		n	n (%)	n	n (%)	n	n (%)
Intensity	Mild	n	n (%)	n	n (%)	n	n (%)
	Moderate	n	n (%)	n	n (%)	n	n (%)
	Severe	n	n (%)	n	n (%)	n	n (%)
Relation to study drug	Reasonable possibility	n	n (%)	n	n (%)	n	n (%)
	No reasonable possibility	n	n (%)	n	n (%)	n	n (%)
	Not Assessable	n	n (%)	n	n (%)	n	n (%)
Action taken	None	n	n (%)	n	n (%)	n	n (%)
	Reduce dose	n	n (%)	n	n (%)	n	n (%)
	Drug withdrawn	n	n (%)	n	n (%)	n	n (%)
	Specific treatment	n	n (%)	n	n (%)	n	n (%)
	Other	n	n (%)	n	n (%)	n	n (%)
Outcome	Recovered	n	n (%)	n	n (%)	n	n (%)
	Recovering	n	n (%)	n	n (%)	n	n (%)
	Not Recovered	n	n (%)	n	n (%)	n	n (%)
	Recovered with sequelae	n	n (%)	n	n (%)	n	n (%)

		Midazolam		Melatonin		All	
		Events	Individuals	Events	Individuals	Events	Individuals
	Fatal	n	n (%)	n	n (%)	n	n (%)
	Unknown	n	n (%)	n	n (%)	n	n (%)

Table 8.14 Serious adverse events

		Midazolam		Melatonin		All	
		Events	Individuals	Events	Individuals	Events	Individuals
All SAEs		n	n (%)	n	n (%)	n	n (%)
Intensity	Mild	n	n (%)	n	n (%)	n	n (%)
	Moderate	n	n (%)	n	n (%)	n	n (%)
	Severe	n	n (%)	n	n (%)	n	n (%)
Relation to study drug	Reasonable possibility	n	n (%)	n	n (%)	n	n (%)
	No reasonable possibility	n	n (%)	n	n (%)	n	n (%)
	Not Assessable	n	n (%)	n	n (%)	n	n (%)
Action taken	None	n	n (%)	n	n (%)	n	n (%)
	Reduce dose	n	n (%)	n	n (%)	n	n (%)
	Drug withdrawn	n	n (%)	n	n (%)	n	n (%)
	Specific treatment	n	n (%)	n	n (%)	n	n (%)
	Other	n	n (%)	n	n (%)	n	n (%)
Outcome	Recovered	n	n (%)	n	n (%)	n	n (%)
	Recovering	n	n (%)	n	n (%)	n	n (%)
	Not Recovered	n	n (%)	n	n (%)	n	n (%)
	Recovered with sequelae	n	n (%)	n	n (%)	n	n (%)
	Fatal	n	n (%)	n	n (%)	n	n (%)
	Unknown	n	n (%)	n	n (%)	n	n (%)
Seriousness	Death	n	n (%)	n	n (%)	n	n (%)
	Life threatening	n	n (%)	n	n (%)	n	n (%)
	Inpatient hospitalisation	n	n (%)	n	n (%)	n	n (%)
	Prolongs hospitalisation	n	n (%)	n	n (%)	n	n (%)
	Persistent of significant disability or incapacity	n	n (%)	n	n (%)	n	n (%)
	Congenital abnormality or birth defect	n	n (%)	n	n (%)	n	n (%)
	Considered medically significant by the investigator	n	n (%)	n	n (%)	n	n (%)
Frequency	Isolated	n	n (%)	n	n (%)	n	n (%)
	Intermittent	n	n (%)	n	n (%)	n	n (%)
	Continuous	n	n (%)	n	n (%)	n	n (%)
	Unknown	n	n (%)	n	n (%)	n	n (%)
Was the SAE expected?	Yes	n	n (%)	n	n (%)	n	n (%)
	No	n	n (%)	n	n (%)	n	n (%)