Analysis Plan for the updated Best Available Treatment Study (BATS)

Background

THE BATS study was established in 2020, during the early months of the SARS-COV-2 pandemic, after first recognition of the new Multisystem Inflammatory syndrome in Children (MIS-C), also known as Paediatric Inflammatory multisystem syndrome temporally associated with SARS-CoV-2 infection (PIMS-TS). The study aimed to collect data on children with MIS-C treated in countries world-wide. As there was no proven treatment for this new disorder, paediatricians caring for children with this new syndrome were adopting treatments based on similarities between MIS-C and other inflammatory diseases, often influenced by availability of specific agents, or national guidance in each country. BATS aimed to compare the outcome associated with the different treatment used, using propensity score weighting to correct for differences in severity and other factors between each treatment group.

The first analysis of >500 children enrolled in BATS by 24th February 2021 was reported in July 2021[1] (McArdle et al NEJM). Recruitment has continued with the aim of reporting the treatment outcomes of a larger number of patients than included in the preliminary report. The initial report was underpowered to detect differences between the treatment groups, and it was hoped by continuing enrolment, a second analysis would have adequate power to detect differences between the most commonly used treatments.

Full details of the BATS study, including criteria for inclusion in the study, data collection methodology, and statistical approach were reported in the previous publication, available here: https://www.nejm.org/doi/full/10.1056/NEJMoa2102968. No changes have been made in the overall study methodology, inclusion criteria or aims of the study. However, minor modifications to the exclusion criteria and the statistical analysis plan have been made, based on the larger numbers of patients now enrolled, and on new data on the disease and other published information. This document aims to describe all changes made in the previously published analysis plan, and to explain the reasons for these. This updated analysis plan should be seen as an addendum to the previously published study description and analysis plan.

The BATS cohort now contains over 2000 patients, and recruitment numbers will soon be finalised. Before drafting this updated analysis plan, we performed a scoping analysis of the primary treatments received by this larger cohort, to determine which treatment groups we would have reasonable power to compare. This was undertaken with no examination of outcomes by treatment group, and hence did not compromise comparison of different treatments in any way.

Study Team:

Imperial College London (Study Management group):

Study Coordinator: Professor Michael Levin

Co-investigators: Dr Claire Broderick, Dr Aubrey Cunnington, Dr Jethro Herberg, Dr Myrsini Kaforou, Dr Andrew McArdle, Dr Ruud Nijman, Dr Harsita Patel, Dr Eleanor Seaby, Dr Priyen Shah, Ms Ortensia Vito, Dr Elizabeth Whittaker, Dr Clare Wilson, Dr Michael Carter, Dr Samuel Channon-Wells

Statisticians: Dr Clive Hoggart, Dr Myrsini Kaforou, Dr Andrew McArdle, Dr Oleg Blyuss, Dr Samuel Channon-Wells

Data management: Dr Tisham De, Ms Ortensia Vito

International Advisory Board:

Dr Daniel Munblit[1], Dr Adriana Tremoulet[2], Dr Rolando Ulloa-Gutierrez[3]

[1]Department of Pediatrics and Pediatric Infectious Diseases, Institute of Child's Health, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia; [2]Department of Pediatrics, Rady Children's Hospital - San Diego, 3020 Children's Way, San Diego, CA 92123; [3]Servicio de Infectología Pediátrica, Hospital Nacional de Niños "Dr. Carlos Sáenz Herrera", C.C.S.S., San José, Costa Rica

Data preparation and definitions

Data preparation

Data are entered in RedCap version 6.14.2. We will include all patients admitted to hospital before 1st March 2022. The final list of included patients will be finalised on 25th April 2022, with subsequent data changes restricted to correction of obvious errors and missing data.

Subsequent processing and analysis will be undertaken in R version 4.1.2, using the packages WeightIt, Cobalt and Survey. Data will be processed such that repeated clinical, laboratory and treatment variables are represented in a table with one row per patient-day.

Inclusions

Clinicians were asked to include patients on their judgement of the patient meeting one or more of the international definitions for MIS-C as defined in the initial protocol. These are:

- Royal College of Paediatrics and Child Health definition of PIMS-TS[2]
- WHO case definition of MIS-C[3]
- CDC case definition of MIS-C[4]

Exclusions

Patients will be excluded based on the following criteria:

- Missing key data fields required for determining treatment arms or outcomes:
  - No admission date
  - No data on treatment form
  - No discharge date and absence of any daily data (level of care, KD features, blood results, cardiac investigations)
  - Unclear date of first immunomodulatory therapy
- Age <1-month. This is updated from the initial protocol as despite the now well-known phenomenon of neonatal PIMS-TS/MIS-C we believe this small group of patients is not reflective of the general child with MIS-C and will require more individualized treatment.
- Absence of outcome data – if patients have no relevant outcome available for a specific outcome analysis, they will be excluded from that specific analysis.
- Patients who only receive immunomodulator therapy with low dose oral steroids will be excluded from the steroid treatment arm and reclassified as receiving no treatment for the purposes of establishing treatment effects. Low-dose steroids are defined as an equivalent dose (as per BNFc[5]) of less than 1mg/kg or 40mg-total of prednisolone, whichever is lower.

Inclusion for weighted analysis

Only patients treated from the day of admission or transfer will contribute outcomes for weighted analyses, as determination of matching covariates and outcomes is not possible for patients treated before transfer. Where recruiting sites enter complete data (enabling both covariate balancing and outcome assessments) from the referring centre we will treat this as one single admission for weighted analysis.

Missing data, interpolation, and imputation

To reduce the volume of missing data we will communicate with the recruiting sites to address missing/inconsistent data for key fields required for the analysis (Appendix A). Level-of-care and supportive treatment variables, including respiratory support and inotropes, and the clinical variable fever will be interpolated for missing daily data where preceding and following values are identical. Where missing data for respiratory support and inotropes follow a final value, if the final value indicates no support was needed, subsequent daily values will be interpolated as no support needed.

Further, where total number of days of invasive ventilation, non-invasive ventilation, oxygen, and inotropic support are available, missing data will be entered assuming no discontinuous periods of treatment (supported by a low frequency of multiple episodes of inotropes, ventilation, or oxygen usage in complete data in the preliminary analysis). In addition, for a particular level of care, if the final outcome data reports a patient had no support (e.g., no days of respiratory support) and the daily data for each day is either missing or entered as no support needed, then we will interpolate the patient to have no support needed on all days of daily data whenever missing. Finally, where missing values of respiratory support follow a period of ventilatory support, with the reported individual days of ventilatory support identical to the reported total days of ventilatory support, we will impute the missing respiratory support values to be no support if the total days of oxygen therapy are reported as zero, otherwise we will impute to a value of “Oxygen or air”.

The imputation of covariates for analysis is reported below in the corresponding analysis section.

Merging consecutive admissions

Where multiple hospitals within one location (city/town/country) report patients, we will inspect plots of admissions and ages to identify possible consecutive admissions of the same patient. More detailed comparison of age, sex, weight, reported ethnicity, admission periods and laboratory and clinical variables will be used to confirm, and recruiting sites contacted if uncertainty remains. Consecutive admissions will be merged into a single record by splicing daily data and taking preliminary admission baseline data and final admission outcomes, with original records excluded.

Assessing for misalignment

During our scoping analysis there has been a suspicion of data misalignment. This can occur because dates of treatments are recorded by date, but timings of other data are recorded by days relative to admission, and the dates are calculated for comparison with treatments. In our definition “day-0” corresponds to the day of admission, and “day-1” to the first full day in hospital. If sites enter admission data as day 1 then misalignment can occur. Excess daily data for some patients alerted us to this issue. We will inspect each patient for discrepancies between the entered daily data and admission length, and overly similar daily

data for distinct days, to assess for potential misalignment. We will then contact recruiting sites for clarification of data entry processes, to ensure we correct for any misalignment.

Laboratory values

Each site was asked to report laboratory variables in units prespecified in the data collection tool, or with alternative units. Conversion to the same units will be undertaken. Manual inspection of result distributions from individual sites will be used to identify and correct incorrect or discrepant units. Extreme outliers will be inspected on a per individual basis and corrected manually when the value is clearly discrepant with the rest of the biomarker time course, and it is clear how to resolve the discrepancy (e.g., a single haemoglobin value out by a factor of 10, indicating an error in recording units). Extreme outliers are those visibly far outside the range of most results. Where it is not possible to correct these outliers, we will contact recruiting sites for confirmation, and exclude results which cannot be resolved.

Definition of clinical severity scale

For each day of admission, clinical severity will be calculated on an ordinal scale:

1. Ventilated (invasive or non-invasive) and on inotropic support
2. Ventilated (invasive or non-invasive)
3. Inotropic support
4. Receiving oxygen
5. No supportive therapy, last CRP ≥ 50
6. No supportive therapy, last CRP < 50

} 5.5 No supportive therapy, CRP unknown

7. Discharged

Additional levels will be added for graphical presentation: death, ECMO and transferred. This ordinal scale was used in the preliminary analysis and was developed by clinical consensus because there were no existing clinical severity scales for this condition when the BATS clinical database form was developed. As previously described, it would be inappropriate to use scales intended for acute COVID-19, which is initially a respiratory illness progressing to systemic disease, whereas MIS-C is a systemic illness with cardiovascular compromise predominating, and secondary respiratory compromise in the majority of patients. Our scale considers escalating levels of clinical support and, in those not on support, differentiates by level of CRP and admission status. This accords with clinical priorities when caring for patients: for those receiving organ support, coming off support is a key sign of improvement. For those not receiving organ support, improvement in inflammation is particularly important, and following that being fit for discharge.

Demographics & baseline clinical data

Age is collected in years and additional months. Where additional months are missing, they will be assumed to be zero. If the child is under 2-years old, we will contact recruiting sites to clarify age in months. If age in years is missing and the data cannot be obtained, the child's age will be replaced with the predicted age based on sex and weight of other children in the cohort using multiple linear regression for imputation.

Patients' weight-for-age Z scores will be calculated from the WHO reference data using the RCPCH Growth API[6]. The World Bank lending group classification will be used for country economic status.

Significant past medical history will be defined identically to the preliminary analysis. This will be regarded as primary or secondary immunodeficiency, HIV, autoimmune disease, chronic lung disease, congenital heart disease, chronic neurological disorder, or malignancy.

Treatment definitions

Primary treatment will be defined as the immunomodulatory agent(s) initiated on the same calendar day before any other treatments. Where two agents are commenced on the same calendar day this will be considered a combination treatment and the effects of the combined agents will be evaluated together. Thus, primary treatment can be either single agent or multi-agent therapy. Immunomodulators administered on subsequent days will be considered secondary treatments.

Low dose IV hydrocortisone is commonly used as an adjunct to inotropic therapy in sick children with PIMS-TS. We will therefore classify courses of IV hydrocortisone as non-steroid therapy where the administered dose is low, as defined above.

Statistical power

Our scoping analysis determined that Intravenous Immunoglobulin (IVIG) alone (over 650 patients), Steroids alone (over 650), or both therapies (over 450) contribute the majority of primary treatments, before considering exclusions. Other therapies, or combinations, account for very small numbers of primary treatments. Our updated power calculations (Appendix B) indicate that these numbers are large enough to detect a 30% reduction in our 1st primary outcome, and an 80% hazard ratio in our 2nd primary outcome, using a Bonferroni-Holm correction with a family-wise error rate of 0.05, adjusting for two primary outcomes and two comparisons of both IVIG with steroids and IVIG with combination therapy (Steroids & IVIG).

Therefore, those receiving IVIG alone, steroids alone or IVIG and steroids in combination as primary treatment, will remain the primary treatment comparison groups in this analysis, with IVIG alone as the reference category. Secondary exploratory analyses will be undertaken to determine the subsequent effect of adding additional therapies, including biologic agents. We will also perform a descriptive analysis of the number of each different immunomodulator used as primary, secondary, and additional therapies. Exploratory analyses will also be undertaken comparing outcomes between treated and untreated patients (those not receiving any immunomodulator therapy.

Primary Outcome Definitions

1. **Inotropic support or ventilation or death (dichotomous)**
   Inotropic support or ventilation (invasive or non-invasive) at any time from the second day post-treatment, or death at any time. Inotropic support and ventilation will be regarded as not available if the patient was transferred or died on day one or two, without report of support being received on day 2. If the patient was discharged on day 1 or 2, the outcome will be regarded as negative. Death will be regarded as missing for all transferred patients, and as negative for all patients whose destination was not recorded.

2. **Time to improvement in clinical severity (continuous)**
   In the previous report reduction in disease severity was used as a second primary outcome, based on a seven-point ordinal scale between day 0 and day 2 described in detail above. We have moved this from primary to secondary outcomes, as it captures a similar outcome to that of time to clinical improvement (new primary outcome). The use of the continuous rather than dichotomous outcome was felt to be more clinically relevant and provides a more comprehensive picture of the effect of treatment on speed of recovery. In addition, this approach may provide additional power compared with the other dichotomous primary outcomes, as demonstrated in our preliminary analysis.
   For each patient, the time to improvement in clinical severity is calculated in days as the time to improve by one or more points in the above ordinal clinical severity scale. This equates to improvement being defined by:
   - Time to come off ventilator or inotropes for patients receiving both therapies
   - Time to come off ventilator for patients ventilated
   - Time to come off inotropes for patients receiving inotropes
   - Time to come off oxygen for patients receiving oxygen
   - Time for CRP to fall below 50 mg/l for patients with final CRP on day of treatment or earlier of greater than or equal to 50 mg/l
   - Time until discharge for all patients, where preceding other event

<u>Secondary Outcome Definitions</u>

1. **No improvement at day 2 (dichotomous)**
   To unify the direction of treatment effect across primary outcomes we have inverted the previously defined outcome from “Improvement” to “No improvement” at day 2. This will aid both reporting and visualization of treatment effect.
   Improvement at day 2 will be defined relative to day 0 as:
   - Any patient who was discharged on or before day 2
   - Patients stepped down from ventilation or inotropic support
   - Patients not ventilated or on inotropes who stepped down from oxygen
   - Patients not receiving organ support whose CRP fell from above 50 mg/l on or before the day of treatment to below 50 mg/l.

   Improvement will be regarded as unknown if a patient was transferred on or before day 2, and negative for a patient who died on or before day 2.

2. **Failure/escalation of primary treatment**
   Defined as the addition of any new immunomodulator from the first day after primary treatment, <u>or</u> an additional dose of IVIG after primary treatment which includes IVIG, <u>or</u> an escalation in steroid therapy after primary treatment which includes steroids. For patients receiving corticosteroids within primary treatment, an escalation of more than 5 mg/kg prednisolone equivalent in total daily dose will be required for further steroid usage to class as failure. If transferred to another hospital before the fifth day following primary treatment, failure will be regarded as not available.

3. **Death or Inotropic support or Ventilation**

We will consider the individual components of the composite primary outcome as individual secondary outcomes, as defined above.

4. **Fever**
   Presence of fever at any point from day 2. If no fever reported, but missing data, the outcome will be regarded as not available.

5. **Increase in level of support**
   This is based on any commencement of:
   - ECMO for patients not on ECMO on day 0
   - Any ventilation for patients not ventilated on day 0
   - Invasive ventilation for patients receiving non-invasive ventilation on day 0
   - Inotropic support for patients not on inotropes on day 0
   - Oxygen for patients not on oxygen on day 0

   Where none of the above led to classification of deterioration, death is regarded as deterioration and transfer is regarded as the outcome being unavailable. Patients discharged home or with unreported discharge destination were regarded as not having increased support.

6. **Persisting/new coronary artery dilation**
   The presence of a coronary artery with Lopez z-score ≥ 2.5 or a report of aneurysm without z-score on the final in-hospital echocardiogram, undertaken on the second or subsequent days following treatment. Will be regarded as not available if no echocardiogram reported, and negative if echocardiogram reported with no aneurysm or z-score < 2.5. Presence of pre-treatment coronary artery dilatation will be added as a balancing covariate.

7. **Persisting Coronary artery dilatation > 6-weeks after discharge**
   This was not included in the original report. In view of the importance in establishing whether treatment regimens, not including Immunoglobulin (the proven therapy for Kawasaki disease), is associated with persistent coronary artery dilation or aneurysms, patients with coronary artery aneurysms detected during admission will have follow up information requested from their treating Paediatric team, and the presence of persisting aneurysms 6-weeks or more after discharge will be reported by treatment group.

8. **Inflammatory and other biomarkers**
   Time courses for inflammatory markers will be plotted as percentages of the peak value, per patient, throughout their admission. Line plots will be weighted by covariate-balancing propensity scores as described below. Smoothed curves with confidence intervals will be plotted using a generalized additive model (geom_smooth from the ggplot2 package in R). Depending on degree of missingness, we will assess changes in markers of inflammation (e.g., CRP, Ferritin, LDH, Neutrophil-Lymphocyte-ratio) and markers of organ damage/dysfunction (e.g., troponin creatinine, liver enzymes). Comparisons will also be made within each treatment group for age and for patients who fulfill the

2017 AHA criteria for Kawasaki Disease. Patients whose treatment commenced on day 7 of admission or beyond will be excluded, as time courses principally represent the natural course.

9. **Complications of drug therapy**
   Complications deemed by the treating clinician to be the result of immunomodulatory treatment, including but not limited to: allergy/anaphylaxis, cataracts, gastric perforation, gastric ulceration, hip necrosis, hyperglycaemia, hyperlactataemia, opportunistic infection, profound bradycardia, psychosis, and steroid-induced hypertension. These will be reported descriptively.

10. **Left ventricular dysfunction**
    The presence of left ventricular dysfunction on any echocardiogram 24-hours after commencement of primary immunomodulatory treatment. For this analysis, the presence of left ventricular dysfunction, prior to starting immunomodulatory treatment, will be added as an additional covariate for calculation of propensity scores to control for confounding due to potential differences in pre-treatment prevalence in each of the treatment arms.

Analyses

Study enrolment data

A descriptive analysis will be performed to summarize the number of countries and sites within each country that have enrolled onto the study, along with the number of patients from each country and site.

Demographic data

Demographic data will be summarized in a table, including: sex (proportion female), age (mean and standard deviation), ethnicity, weight for age (z-scores), proportion with significant comorbidities, and country classification by income level (as defined by the World Bank classification). In addition to presenting the data for the population, further stratification will be done by diagnostic groups. This will include patients who do not have all the WHO criteria for MIS-C, and patients excluded from WHO criteria due to bacteraemia or reported toxic shock syndrome, and patients meeting complete Kawasaki disease criteria who are aged <6-years.

Clinical features, laboratory markers & echocardiogram findings:

Salient clinical features during the patient's admission will be summarized and tabulated, including: fever, sore throat, cough, respiratory distress, abdominal pain, diarrhoea, vomiting, headache, encephalopathy, irritability, lethargy, and SARS-CoV-2 PCR status. We further present mucocutaneous/dermatologic features during admission including the presence of: rash, red lips, mucosal membrane changes, conjunctival injection, oedema, skin peeling, lymphadenopathy, and BCG scar reactivity.

Laboratory markers will be summarized and tabulated showing mean and standard deviation and include: CRP, troponin, lactate dehydrogenase, D-Dimer, ferritin, white cell count, lymphocyte count, neutrophil count, platelets, creatinine, BNP, haemoglobin, prothrombin

time, activated partial thromboplastin time, and fibrinogen. These clinical and laboratory marker data will be presented for the entire cohort and further stratified by diagnostic groups.

We will perform a descriptive analysis of the number of patients with coronary artery aneurysms (CAA) at any time and persisting to discharge. Depending on data availability after discharge, we will perform secondary analysis of patients with persistent CAA at follow-up after discharge, stratified by primary treatment group.

Confounding

All primary outcomes, sensitivity analyses, subgroup analyses, and secondary outcomes (excluding death, secondary diagnoses, and complications) will undergo analysis following weighting by multinomial covariate-balanced propensity scores[7] to control for baseline confounding factors, as implemented by WeightIt version 0.12.0, using the "just-identified" approach. The Average Treatment Effect (ATE) will be estimated, except when comparing inflammatory markers between treated and untreated patients, when the Average Treatment Effect in the Treated (ATT) will be calculated with the untreated group as the reference due to the likely dissimilarity of a smaller untreated group and the need to preserve the full sample.

The following variables will be considered for balancing for analysis of the primary outcomes:

1. Age, continuous
2. Sex, binary
3. Weight-for-age z-score greater than 2, binary (imputation will be undertaken for missing values: patients with "severe obesity" checked in the list of comorbidities will be assigned to weight-for-age z-score >2, whereas patients without this comorbidity will be assigned to the <2 group)
4. Significant comorbidity, binary
5. Resource group, three categories will be considered: High income, Upper-middle income, Low and Lower-middle income (low and lower-middle are grouped as very few sites from low-income countries recruited patients to BATS). High income as the largest category, will be baseline, with two binary covariates coding the other two categories
6. KD features, binary (meet criteria for complete KD at any time up to treatment day
7. Requiring inotropes up to treatment day, binary
8. Requiring mechanical ventilation or ECMO up to treatment day, binary
9. Maximum CRP up to treatment day, continuous (we plan to use single imputation where possible, using baseline clinical/severity data and inflammatory markers)

This will be reduced based on data availability and area of common support across treatment groups. Important covariates will be added for certain secondary analyses as described above. Balancing will be repeated for every analysis on the population providing the outcome. No imputation for missing outcome data will be undertaken except for that already described above for daily level-of-care and support variables. When comparing patients receiving and not receiving immunomodulator therapy, variables reporting features up to the day of treatment will be replaced with corresponding variables on admission due to a lack of corresponding first treatment day for those not receiving any immunomodulator.

We will aim for absolute standardised mean differences of 0.1 and below in continuous variables, and Kolmogorov-Smirnov distances of 0.1 and below. Love plots will be used to examine the extent of imbalance and consider the potential impact. We will tolerate some deviation since covariates are also included in outcome models.

Models

As described in the preliminary analysis, we will use weighted logistic-regression methods to analyse dichotomous outcomes. Time-to-event analyses will be performed using weighted Cox regression methods[8], with weights determined by inverse probability according to covariate-balancing propensity scores to account for baseline differences among the three major treatment groups. Robust sandwich standard errors will be used, with dichotomous outcomes to be analysed using the survey package, adding all covariates used in covariate balancing, to produce doubly-robust estimates. To account for overdispersion quasibinomial regression with a logit link function (a generalised linear model) will be used to estimate odds ratios and 95% confidence intervals. Time to event analyses will be undertaken using weighted Cox proportional hazards model[8] estimated average hazard ratios. This allows for violation of the proportional hazards assumption.

E-values will be presented for primary outcomes as per the method of VanderWeele and Ding[9].

Correction for multiple testing

This will be undertaken using the Bonferroni-Holm method for the two primary outcomes. All other outcomes will be presented with 95% confidence intervals alone.

Clinical severity over time

Clinical severity over time will be presented as proportional column charts from two days before treatment to 10 days after treatment. Only patients treated after day 1 will contribute severity data for preceding days, since patients treated on day 1 of admission will provide no severity data for pre-treatment days. The charts will be presented both unadjusted and weighted by the covariate-balanced propensity score.

Subgroup analyses

The following planned sensitivity analyses will be undertaken for all primary outcomes:

- Patients fully meeting the WHO criteria for MIS-C (as per preliminary analysis)
- Meeting WHO criteria except for presence of bacteremia (as per initial protocol)
- Missing WHO classification by one criterion (as per initial protocol)
- Missing WHO classification by >1 criterion (as per initial protocol)
- Patients from High- and Upper-middle-income countries (New addition)
- Patients from Low- and Lower-middle-income countries (New addition)
- Stratified by age-group as follows: Under 6-years; 6-11-years; Over 11-years (New addition)
- Restricting to patients without significant comorbidities (New addition)
- Stratifying patients based on degree of inflammatory response, separating by peak CRP before treatment into tertiles.

For the 2nd primary outcome, time to improvement in clinical severity, we will perform a subgroup analysis for each category on the ordinal scale. For example, we will analyse the time to improvement for patients who are ventilated at the time of starting primary treatment as one subgroup analysis.

A secondary analysis will also be performed comparing primary treatments of steroids alone with IVIG & steroids, using steroids as the reference group the same methodology described above.

Sensitivity analyses

The following planned sensitivity analyses will be undertaken for all primary outcomes:

- Defining primary treatment as all immunomodulatory treatments administered over two consecutive days (day 0-1) (as per preliminary analysis)
- Repeating covariate balancing without imputed CRP values
- Using propensity matching model rather than covariate-balancing propensity score weighted analysis. Details to be determined

For the 2nd primary outcome, time to improvement in clinical severity, we will undertake an additional sensitivity analysis requiring a 2-point improvement in clinical severity on the ordinal scale. We will also undertake a sensitivity analysis for the secondary outcome 9 (Left ventricular dysfunction) including maximal troponin to treatment day as an additional covariate, excluding samples without a troponin measurement before treatment.

Summary of changes from preliminary analysis

For reference, changes from preliminary analysis to second analysis are defined as:

- Study team updated
- Date of data finalization updated
- Exclusion criteria broadened to include neonates and patients only receiving low-dose oral steroids
- Description of process for contacting sites to reduce missingness of key fields, defined in Appendix A
- Additional description of handling missing level-of-care data based on total days of supportive therapy
- Description of process for detecting and correcting cases where incorrect alignment of daily-data and treatment dates may exist
- Power calculations updated
- Recategorization of primary and secondary outcomes
- Clarification of time to improvement outcome based on ordinal severity scale
- Clarification of the failure/escalation of primary treatment secondary outcome
- Additional secondary outcome described
- Additional blood tests added to those considered for time-course plotting
- Alterations to variates considered for covariate-balanced propensity score estimation
- Description of imputation of weight z-score for covariate balancing purposes
- Intention to impute CRP values used for covariate balancing, with additional sensitivity analyses without CRP imputation
- Expansion of subgroup analyses to those described in the original protocol with new subgroups described
- Addition of sensitivity analyses, including using propensity matching methodology; without CRP imputation; time-to-improvement calculated as 2-point improvement on ordinal scale

References

1. McArdle, A., Vito, O., Patel, H. et al. Treatment of Multisystem Inflammatory Syndrome in Children. *N Engl J Med. 385, 11-22 (2021)*
2. RCPCH. Paediatric multisystem inflammatory syndrome temporally associated with COVID-19 (PIMS). Available from: *https://www.rcpch.ac.uk/resources/paediatricmultisystem-inflammatory-syndrome-temporally-associated-covid-19-pims-guidance*
3. WHO. Multisystem inflammatory syndrome in children and adolescents temporally associated with COVID-19. Available from: *https://www.who.int/publications/i/item/multisystem-inflammatory-syndrome-inchildren-and-adolescents-with-covid-19*
4. CDC C for DC and P. Information for Healthcare Providers about Multisystem Inflammatory Syndrome in Children (MIS-C). Available from: *https://www.cdc.gov/misc/hcp/*
5. NICE Equivalent anti-inflammatory doses of oral corticosteroids. Available from: *https://cks.nice.org.uk/topics/corticosteroids-oral/background-information/equivalent-anti-inflammatory-doses/*
6. RCPCH growth app. Available from*: https://www.rcpch.ac.uk/resources/growthcharts*
7. Imai, K. & Ratkovic, M. Covariate balancing propensity score. *J. R. Stat. Soc. Ser. B Stat. Methodol. 76, 243–263 (2014).*
8. Schemper, M., Wakounig, S. & Heinze, G. The estimation of average hazard ratios by weighted Cox regression. *Stat. Med. 28, 2473–2489 (2009).*
9. VanderWeele, T. J. & Ding, P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med. 167, 268–274 (2017).*

Appendix A – Key fields required for full analysis

Field category	**Fields**
Admission	Admission date; treatments received prior to admission at recruiting hospital
Discharge	Date of discharge or death; place of discharge
Demographics	Age (in years & months); weight; sex-at-birth
Clinical history	Presenting symptoms; presence of comorbidities
Therapy	Start dates of immunomodulator therapies; steroid name, dose, and route
Daily level-of-care & support	Level-of-care; supportive therapies, including oxygen, ventilation, inotropes, renal replacement therapy and Extracorporeal membrane oxygenation
Investigations	CRP & other blood results by day; echocardiogram results, including presence of coronary artery aneurysms

Appendix B – Power Calculation - BATS Analysis

Primary outcome 1 - Inotropic support or ventilation or death (dichotomous)

Assumptions:

1. 30% of patients in the control arm will have an outcome (estimated from the weighted numbers seen in preliminary analysis)
2. Aim to detect a 30% relative decrease in outcomes in the experimental arm (clinically appreciable improvement)
3. Aim for power of 0.8
4. Aim for type-1 error rate of 0.013 – this equates to global type-1 error rate of 0.05 across four comparisons (2 treatment comparisons, 2 primary outcomes)
5. Using a two-sided test
6. Assume equal proportion of patients in each arm

Under these assumptions, we would require **516 patients** in each group.

If we assume instead that there are 600 patients in our largest group, we will require 452 **patients** in the smaller group to have the above power to detect the assumed effect size above.

Varying effect size

Below is a table demonstrating how changing effect size (relative decrease) alters the number of patients needed in each arm, when assuming equal proportions.

Relative Decrease	N
20%	1208
25%	758
30%	516
35%	370

Primary outcome 2 - Time to improvement (continuous)

Assumptions:

1. 90% of patients in both arms will have an outcome (estimated from the numbers seen in preliminary analysis)
2. Aim to detect a postulated hazard ratio of 0.8 between treatment arms (clinically appreciable improvement)
3. Aim for power of 0.8
4. Aim for type-1 error rate of 0.013 – this equates to global type-1 error rate of 0.05 across four comparisons
5. Assume equal proportion of patients in each arm

Under these assumptions, using the Cox Proportional-Hazards Model we would require **498 patients** in each group.

If we assume instead that there are 50% more patients in our largest group, we will require **595 patients** in the larger group and **397 patients** in the smaller group to have the above power to detect the assumed effect size above.

Notes:

These calculations have been performed using the packages pwr[1] and powerSurvEpi[2].

References:

1. Stephane Champely (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. https://CRAN.R-project.org/package=pwr
2. Weiliang Qiu, Jorge Chavarro, Ross Lazarus, Bernard Rosner and Jing Ma. (2021). powerSurvEpi: Power and Sample Size Calculation for Survival Analysis of Epidemiological Studies. R package version 0.1.3. https://CRAN.R-project.org/package=powerSurvEpi