SMAART analysis plan

Date: 20th October 2024

Enrolment

- Number enrolled by site and study arm (case, control and background)
- Line graph showing cumulative enrolment by site over time separately for cases, controls and background.

Admission characteristics

The following will be tabulated by cases and controls, and by site for cases only. Continuous variables will be compared using means and standard deviations (or median and IQR where non normal). Categorical variables will be tested for a difference using chi-squared tests, or Fisher's exact test if any category has a frequency less than 5.

- Age at admission (months): mean (SD) or median (IQR)
- Sex: n (%) male, female
- Axillary temperature (°C): mean (SD) or median (IQR)
- Haemoglobin (g/dL): mean (SD) or median (IQR)
- Assessment of severity:
    - impaired consciousness, two or more convulsions, respiratory distress, compensated shock, decompensated shock, clinical evidence of jaundice, dark or coca coloured urine, severe anaemia, known HIV: n (%) yes, no
- Number of severity features [cases only, but check other arms have 0]: n (%) 1-9
- How sick child is clinically (1-10): mean (SD) or median (IQR)
- Malaria details:
    - pfHRP2 test: n (%) positive, negative, invalid, not done
    - RDT: n (%) positive, negative
    - Blood film: n (%) positive, negative, invalid, not done
- Sickle cell: n (%) previously positive, previously negative, not previously tested

Clinical history of this illness

- History of cough, difficulty breathing, vomiting, diarrhoea (>3 loose motions in last 24 hours), fits in this illness, hand/foot pain (dactylitis), bone/joint swelling (bone infection): n (%) yes, no, don't know
- Bloody diarrhoea: n (% of those with diarrhoea) yes, no, don't know
- Fits lasting more than 30 mins: n (% of those with fits in this illness) yes, no, don't know

Past history – before this illness

- Two or more hospital admissions in the last year, previously received a blood transfusion ever, epilepsy, able to sit without support before this illness, able to walk without help before this illness: n (%) yes, no, don't know

- Number of occasions received a blood transfusion: n (% of those with previous blood transfusion ever) 1, 2, 3-4, 5+
- On ARVs: n (% of those previously diagnosed with HIV) yes, no

About the child's family

- Is the mother the main caregiver: n (%) yes, no, don't know
- Parents alive: n (%) both alive, one alive, both died
- Age of child's mother if alive: n (%) < 18, 18-25, 26-35, >35 years
- Number of other children mother looks after: n (%) 0, 1-2, 3-5, 6+
- Any siblings with sickle cell anaemia: n (%) yes, no, don't know
- Homestead: n (%) urban, semi-urban, rural
- Child sleep under bed net/mosquito net: n (%) yes, no, don't know

Clinical examination

- Indrawing, deep breathing, sunken eyes, decreased skin turgor, cold hands or feet only, liver size >2cm below costal margin, very severe wasting: n (%) yes, no, not assessed
- Crackles: n (%) unilateral, bilateral, none
- Splenomegaly: n (%) not palpable, enlarged, gross
- Signs of kwashiorkor: n (%) none, pretibial, hands/legs, generalised
- Severe pallor: n (%) none, palmar, tongue or conjunctive, all

Neurological

- Inability to sit up right unsupported (prostrate), fitting currently, neck stiffness, bulging fontenelle (infants only), divergent gaze, posturing, bruxism, developmental delay/cerebral palsey: n (%) yes, no, not assessed
- Pupil symmetry: n (%) equal, unequal

Details of admission

The following will be summarised by study arm (cases and controls, and background where specified) and site.

- Duration of admission (including backgrounds): median (IQR)
- Interventions during admission:
    - Received blood transfusion (including backgrounds): n (%)
        - If yes, number of transfusions: n (%) 1, 2, 3+
    - Received fluids: n (%)
        - IV bolus: n (%)
        - Maintenance fluids: n (%)
        - Type of maintenance fluid: n (%) 0.18% saline, dextrose saline, normal saline, Ringer's, Darrow's
    - Received oxygen: n (%)
    - Received glucose: n (%)
        - If yes: n (%) 5%, 10%, other
- Medications prescribed during admission:
    - Number prescribed any antibiotic: n (%)
    - Number prescribed each antibiotic: n (%) for the most frequently prescribed
    - Timing of prescription: n (%) on admission, after admission, at discharge

    - Duration of prescription: (1 day, 5 days, 7 days etc)
    - Total duration by time prescribed: median (IQR) for prescribed at admission, after admission, at discharge
- Number that developed severe malaria during admissions (controls and backgrounds only)
- Outcome of admission: n (%) discharged, absconded, died in hospital (including backgrounds)

iStat and other lab data

The following baseline characteristics will be summarised by study arm (cases or control) and site (for cases). A sensitivity analysis will remove values where there were concerns about quality of the samples, before re-training.

Missing data will be imputed using random forest imputation or multivariate imputation by chained equations (MICE) with predictive mean matching. Both the results based on observed data and after imputation will be presented.

- iStat completeness
- Number of participants with iStat done: n (%)
- Sodium, potassium, chloride, calcium, tco2, BUN, creatinine, HCT, anion gap, base excess, PH, PCO2, HCO3: number with data (% where iStat was done), mean (SD) or median (IQR)
- Lactate, glucose, CRP, PCT: number with data (%), mean (SD) or median (IQR)
- WBC, RBC, Hb, HCT, platelets, lymphocytes, neutrophils, monocytes: number with data (%), mean (SD) or median (IQR)

Describing phenotype of malaria at each site.

Pre-defined phenotypes

The number of cases meeting each of these criteria will be tabulated by site.

- Lactate > 2mmol/L
- BUN > 20mmol/L
- Blantyre Coma Score <=2
- Positive RDT and Hammersmith colour scale >=6
- Teule criteria (temperature >38°C or <36°C, plus at least one of prostration, respiratory distress, Hb<5g/dL or HIV)

Hierarchical clustering

We will use hierarchical clustering to investigate whether there are other phenotypes present in the data. The clustering model will include clinical factors only initially. Depending on the findings, we will consider whether there is added value in including laboratory values. Where there are a high proportion of missing lab values, we will explore imputing the missing values using random forest imputation or predictive mean matching. The number of clusters will be determined using the Nb_clust package in R, and bootstrapping methods to investigate cluster stability.

Finite mixture models

We will use finite mixture models to explore whether there is evidence of subpopulations with different distributions of key lab values such as haemoglobin, lactate and BUN. We will compare models with 1-6 subgroups and compare these using AIC to select the best model.

Phenotypes as a percentage of malaria admissions

Once the set of phenotypes identified through these methods has been defined, the proportion of severe malaria admissions of each phenotype will be summarised by site. It will be explored whether probability weighting can be used to estimate the burden of each phenotype as a percentage of all malaria admissions.

Lactate/parasite clearance

The following will be summarised for Soroti and Kalongo.

- Time from enrolment to lactate < 2mmol/L (of those with >2 at enrolment): median (IQR) hours
- Time to negative malaria slide: median (IQR) hours

Follow up

The following will be tabulated by study arm (cases and controls) and site.

- Visits considered complete, defined as attended or died before the visit took place, at day 28 and day 180: n (%)
- Child status at day 28 and day 180: n (%) visit done, died, withdrawn, lost to follow up, missed visit

Outcomes

The following will be analysed by study arm for cases and controls.

- 6 month mortality
    - n (%) died by 6 months
    - Kaplan Meier
    - Cox proportional hazards model
- Readmission to hospital:
    - n (%) readmitted
        - Poisson regression or negative binomial if there is evidence of over-dispersion
    - n (%) readmitted with positive RDT
        - Poisson regression or negative binomial if there is evidence of over-dispersion
- New episodes of malaria defined by any one of: antimalarial use, positive RDT, febrile illnesses:
    - n (%)
    - Poisson regression or negative binomial if there is evidence of over-dispersion

Predictive factors for outcomes

A model will be built to find factors associated with each outcome: death, readmission, readmission with positive RDT. Candidate predictors will include data from physical examination and assessment

of severity at screening, clinical history of the illness and past history before this illness. Backwards elimination with exit p=0.1 used for variable selection. Interpretation of the effects will focus on those with p<0.05.

Time matching

Enrolment rates between cases and controls over time will be compared and where appropriate any case control comparison may need to use logistic regression including calendar date as covariate (continuous), but still have problem in period when cases only as no temporal controls.

This calendar date adjustment needs work and thought, given lack of temporal overlap between cases and controls – need to explore interactions between case/control and calendar date, overall and separately by site, and gain thorough understanding of plausible variation using splines with lots of knots before adding any other covariates.

Evaluation of the new RDT pHRP2 test

Results from the RDT will be compared to HRP2 results from stored plasma, and the sensitivity and specificity of the test will be calculated.