University of Cape Town

Longitudinal Data Analysis

Comparing the effect of two different Malaria treatments on Haemoglobin

Author:
Andomei Smit

Student Number:
SMTAND051

August 31, 2025

Contents

1 Introduction

Haemoglobin (Hb) is a key clinical indicator in malaria studies, as malaria infection leads to the destruction of red blood cells and consequently lowers Hb levels, contributing to anemia. Severe anemia is a major cause of morbidity and mortality, particularly in children. Effective treatment not only clears parasitemia but also facilitates the recovery of Hb levels, thereby reducing the risk of complications.

In this study, Hb is used as a secondary outcome measure of treatment efficacy. The ideal therapeutic response is characterized by a rapid increase in Hb following treatment and the maintenance of higher Hb levels throughout the follow-up period. Comparisons between treatment arms therefore focus on whether patients in one arm recover Hb more quickly, achieve higher levels overall, or sustain those levels for longer. The treatment that produces faster and greater Hb recovery is considered superior with respect to this endpoint.

Note, however, in malaria infections, Hb levels typically decline shortly after the initiation of treatment before rebounding during recovery. Several studies have documented this characteristic Hb trajectory. For example, [1] Zwang et al. (2017) found that in children under five, Hb levels dropped by approximately 12.8% around 1.5 days following treatment initiation—falling from 9.8 to 8.7 g/dL—before rising linearly at a rate of ≈0.6% per day, ultimately exceeding baseline by ≈13.8% by day 28. For the purposes of this study, baseline Hb at day 0 was included in exploratory plots to show the full course of Hb, but excluded from the modelling since we were interested in recovery trajectories post-treatment rather than pre-treatment baseline differences.

2 Data and Methods

2.1 Data Description

The given malaria dataset consists of the following variables:

- Hb: Haemoglobin reading of a subject.
- arm: the treatment arm that a subject was subscribed to, namely SP (Sulfadoxine and Pyrimethamine) or SP/ACT which adds Artemisinin to SP.
- site: four sites where subjects were treated, namely Boane, Catuane, Magude, and Namaacha.
- pid: the unique patient ID used to identify different subjects.

- pday: the day on which the measurement of Hb was taken, made up of days 0, 1, 2, 3, 7, 14, 21, 28 and 42. Day 0 was kept to serve as a baseline measurement from where the Hb trajectories could be compared. Note that there were no observations on day 1 or 2, thus it was treated as if the only measurements were taken on day 0, 3 and onwards.
- gender: classification of each subject as Male or Female.
- weight: the weight of each subject at the start of the treatment.
- age: the age of each subject at the start of the treatment.

One important aspect that was not stated explicitly at the outset of the study was the clinical protocol across the different sites and if that was homogenous. This project will assume that the protocols were homogenous and that differences in patient weights and ages would trigger the same dose adjustments of the drug concentrations across different sites. This assumption will be formally tested by investigating if there is an interaction between treatment arm and site (or region as will be explained later on).

2.2 Data Cleaning and Preparation

It was noted early on that many of the variables had missing values. For the purpose of this paper, all subjects that had fewer than two readings of Hb were removed from the data altogether. Furthermore, one subject was removed that had a Hb reading of 0 for one of the measurement days. This assumes that the reading was faulty and thus that the subject's measurements should be removed, which seemed like a reasonable assumption as a Hb reading of zero is physiologically impossible and would mean with certainty that the person would be dead. The remaining data consisted of 286 unique subjects.

The original dataset had a Year and Country variable that was the same for all subjects, thus was removed. Other variables available in the original dataset such as the concentration of different drugs or parasite densities were ignored as the focus of this project is to only take into account demographic variables.

2.3 Exploratory Data Analysis

Various checks were done to ensure that the data could be used to make inference on the differences between the two different treatment arms. Some initial checks were done at a high-level to get an understanding of the distribution of treatment arm, subjects and gender between the different sites and to spot any anomalies or issues.

2.3.1 Initial Investigations

Preliminary exploratory checks were performed to evaluate whether the data exhibited substantial skewness or structural complexities. The findings were as follows:

- Age and weight were found to be constant within subject for the entire treatment duration, thus these were assumed to be the age and weight at day 0.
- Roughly 60% of all subjects were female, meaning that overall the study was not dominated by one gender.
- The treatment arms were equally split among subjects (49% received SP and 51% received SP/AP).
- An imbalance was noticed between the number of subjects within each site. The smallest site was Catuane with only 37 subjects, compared to the largest site, Magude with 124 subjects. Namaacha and Boane had 48 and 77 subjects, respectively. Figure 1 shows that Magude is to the North of Maputo, whereas Naamacha, Boane and Magude are to the South. Due to the imbalance of the number of subjects within the sites, it was decided to combine the three sites to the south and compare the overall difference in treatment between the Northern and Southern sites. In total, the Northern sites had 124 subjects and the Southern sites 162.
- The treatment arms were roughly equally split among subjects within each region, meaning that there wasn't a region that predominantly administered one treatment arm.
- Within each region, both treatment arms included adequate numbers of male and female subjects. This indicates that there was no systematic bias toward assigning a particular treatment to one gender. Although the proportion of female participants varied across regions from 50% to 65%, these figures are broadly consistent with the overall sample, in which females represented 60% of participants.
- pday (measurement day) 1 and 2 had no measurements taken on those days. These two days were ignored in the remainder of the project.

2.3.2 Correlation between age and weight

It was noted that age and weight had a correlation of 0.85. This does make sense biologically. In order to avoid multicollinearity, only one of these will be used in the model. The correlation between age and weight with Hb was 0.28 and 0.37, respectively. In practice, the weight of a subject is what usually determines the dose of the treatment, thus given this clinical factor and the higher correlation of

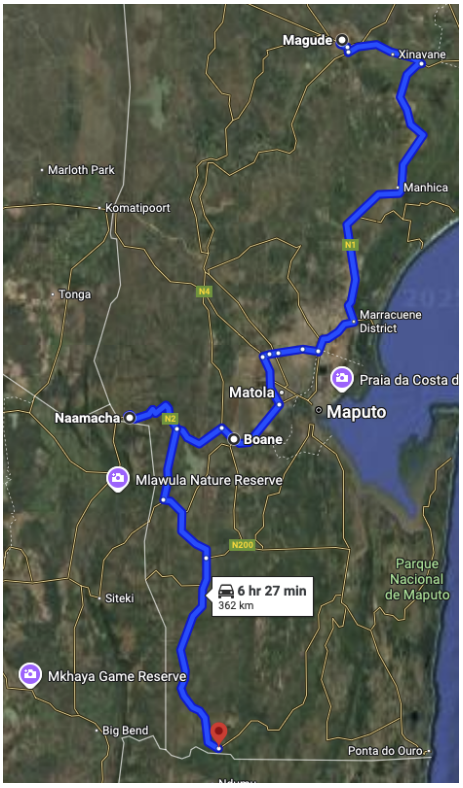


Figure 1: Locations of Boane, Catuane, Magude, and Namaacha in Maputo Province, Mozambique. Catuane can be see as the location at the bottom of the map with the red location pointer. The regions to the South will be grouped together, whereas Magude in the North is the only site making up the Northern region.

weight with Hb, weight will be used for the remainder of the analysis and age will be removed.

Furthermore, this simplifying assumption does seem reasonable as weight could be used as a proxy for age for other physiological differences between adults and children. It is intuitive that there could be changes in treatment effect in children versus adults, but since these effects would not change discretely, but rather continuously, the use of weight to account for these differences seems fair. For example, one would not expect there to be much different between a subject aged 17 years and 11 months versus a subject aged 18 simply because the latter is technically classified as an adult. In the same breath, one would assume their weights would be similar on average, holding all other things constant.

2.3.3 Distribution of age, weight and Hb

The distribution of age and weight across treatment arm and region is plotted in Figure 2. This plot shows that the age and weight distributions between the two regions are very similar, with the South having slightly older and heavier subjects. The subjects in the north are mostly younger than 20, whereas the south sees a higher number of older subjects. Overall, these distributions between regions does not identify clear differences in treatment protocol triggered by different subject ages or weights. As noted above, weight will be used in the remainder of the analysis due to the high correlation between age and weight and weight having a stronger correlation with Hb.

The distribution of weight can further be seen in the two histograms in Figure 3. Figure 3b shows that the log of weight improves the very skewed weight distribution. Thus, the log of weight will be used for the analysis.

Finally, the distribution of Hb can be seen in Figure 4. As expected, this distribution follows that of a normal distribution, which is the underlying assumption of the Mixed-Effects model that was built in this project.

2.3.4 Interactions within the data

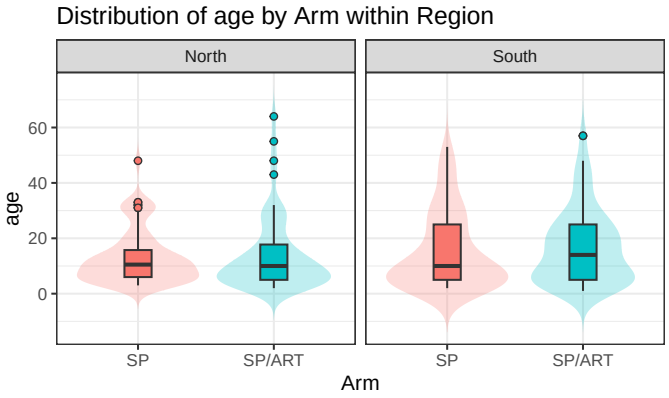
Various interactions were investigated to be considered in the model later on. The aim of this section is to do a visual inspection for obvious results that should or should not be tested formally in the modelling section later on. The interactions can broadly be divided into two groups, namely

1. Treatment arm related interactions and
2. Time related interactions

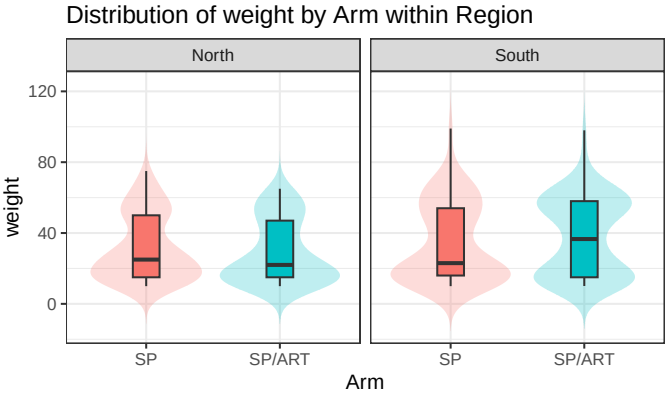
Treatment arm related interactions considered the interactions between treatment arm (arm) and the variables listed below, each with the relevant question that the interaction is considering.

- measurement day (pday): Does Hb trajectory differ between treatment arms? This is, of course, the main analysis question in this project.
- Gender: Does the treatment effect differ between males and females, on average?
- Weight: Does the treatment effect vary by patient weight?

Because weight is continuous, interactions involving weight are less easily visualized in a simple plot. These will therefore be formally tested in the modelling stage.



(a) The distribution of age within region for the two different treatment arms.

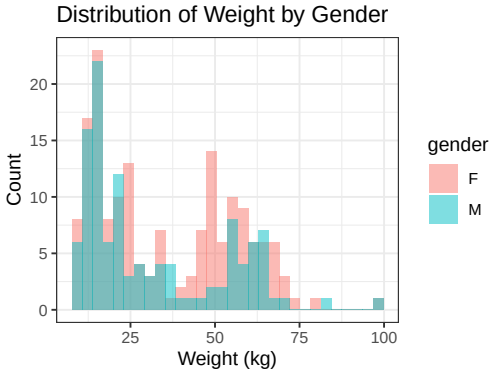


(b) The distribution of weight within region for the two different treatment arms.

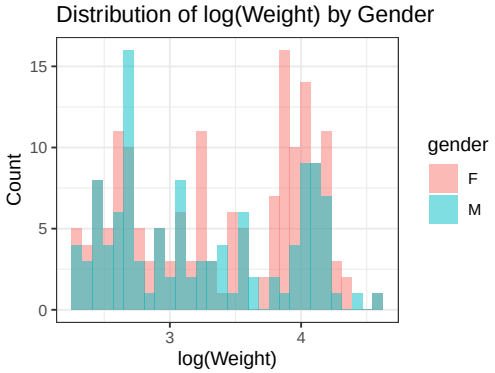
Figure 2: The distribution of age and weight by region and treatment arm. It can be seen that the ages and weights do not differ significantly between the regions. It can also be noted that the distributions of age look similar to those of weight.

Figure 5a determines if there is an interaction between treatment arm and measurement day (pday) and if the trajectory of a the change in Hb is different for the different treatment arms. From this figure we can see that the rate of change (or slope of the graph) differs for some of the days, for example from pday 3 to 7. Overall the rate of change seems very similar between treatment arms. However, since this is so closely related to the main research question of this project, the interaction will be formally tested later on.

Figure 5b investigates the interaction between treatment arm and gender and tries to find obvious signs that the effect that the treatment has on a subject is different



(a) A histogram of the distribution of weight for males and females.



(b) A histogram of the distribution of the log of weight for males and females.

Figure 3: The distribution of weight and the log of weight for males and females. It is clear that the distribution of weight is highly skewed. It seems that the log of weight is an improvement.

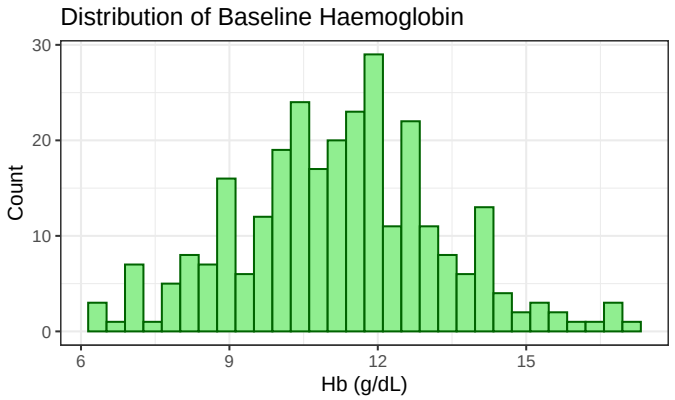
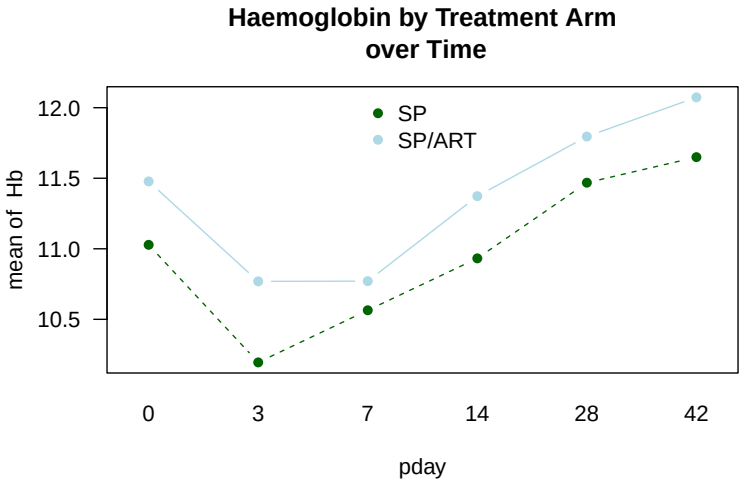


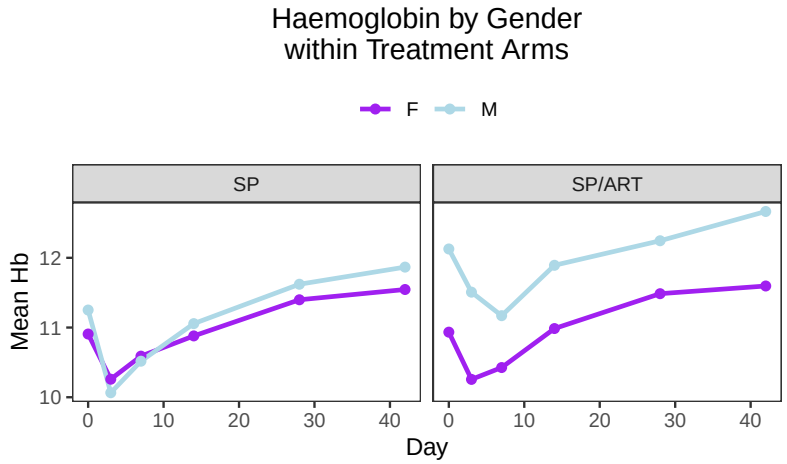
Figure 4: A histogram of the values of Haemoglobin (Hb). The shape follows that from a normal distribution.

for the two genders. From this graph it seems that the treatment effect is the same between males and females for SP, whereas there seems to be a difference for SP/ART for some of the changes in Hb, although these differences are still minor. See for example the change from day 3 to 7 again, where for males Hb decreased, but for females Hb increased. Since the visual inspection is alluding to some differences, this interaction term will formally be tested in the modelling section.

Next, the **Time related interactions** are interactions with pday and the variables listed below, each with the relevant question that the interaction is considering.



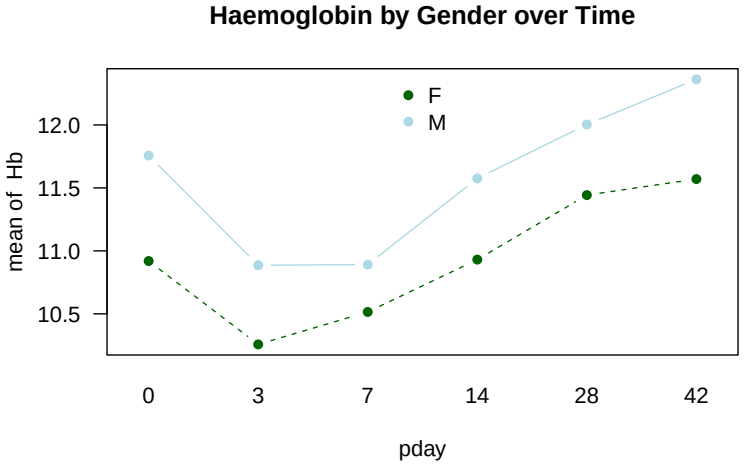
(a) The change in mean Hb for the two different treatment arms. This plot investigates if there is an interaction between treatment arm and measurement day (pday).



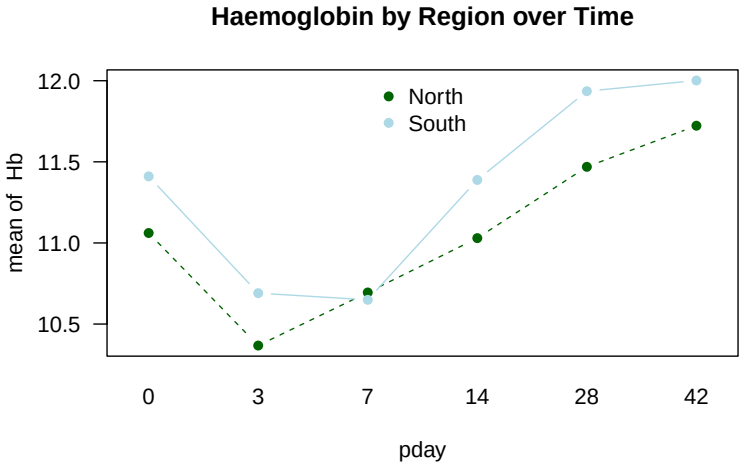
(b) The change in mean Hb for males and females for the two different treatment arms. This plot investigates if there is an interaction between gender and treatment arm.

Figure 5: A visual inspection of two different treatment interactions, an interaction with pday at the top and an interaction with gender at the bottom.

- Gender: Do males and females have different Hb trajectories over time, regardless of arm?
- Region: Do Hb trajectories differ across regions?



(a) The change in mean Hb for the different genders. This plot investigates if there is an interaction between measurement day (pday) and gender.



(b) The change in mean Hb for the northern and southern regions for the two different treatment arms. This plot investigates if there is an interaction between region and measurement day (pday).

Figure 6: A visual inspection of two different time interactions, an interaction with gender at the top and an interaction with region at the bottom.

- Weight: Do heavier/lighter patients recover Hb at different rates?

Figure 6 investigates these interactions visually. From Figure 6a there only seems to be very minor differences between the rate of change of mean Hb between males and females, thus the interaction between gender and pday will not be tested further. In contrast, Figure 6b shows more drastic differences between the rate of change of mean Hb between the Northern and Southern Regions. Specifically, when looking at the rate of change (slope) from days 3 to 7 where mean Hb declines in the Souther Region and increases in the Northern Region and the much larger rate of increase in the Southern Region from day 7 to 14. Thus, this interaction will be investigated further in the modelling section.

In summary, **the interactions that will be formally tested are**:

- treatment arm and measurement day
- treatment arm and gender
- treatment arm and weight
- measurement day and region
- measurement day and weight

2.3.5 Subject Profiles

In order to inform the decision of including random intercepts or slopes, the subject profiles (repeated measures of Hb) were plotted for each treatment arm in Figure 7. From this plot it is clear that there is a difference in intercept and slope between subjects, thus a random slope and intercept will be considered in the modelling section.

2.3.6 Dropout Rates of Subjects

Finally, the dropout rates of subjects was compared between Regions (Northern and Southern) and between treatment arms to see if there are any clear issues within a region or treatment arm. Figure 8 shows the dropout rate between the two treatments arms in Figure 8a and between the two regions in Figure 8b.

There is not a large difference in the dropout rate between the Northern and Southern region with both regions losing subjects at relatively the same rate and ending the study window with the same difference in the number of subjects within the regions as at day 0. Whereas the rate of dropout for subjects receiving the SP treatment seems much higher at various stages throughout the treatment schedule. However, even with the larger dropout rate, the SP treatment arm still is left with about 110 subjects at the end of the study period, meaning that there wasn't an alarmingly high dropout rate, just a dropout rate faster than that of the SP/ART

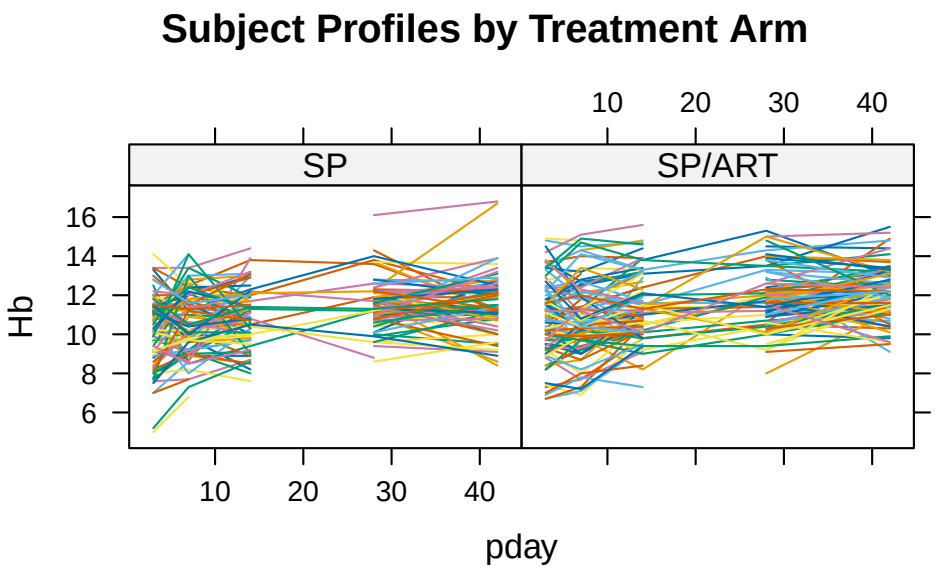
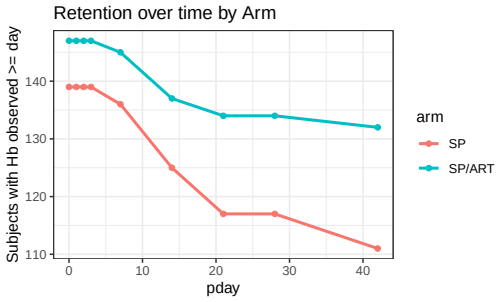
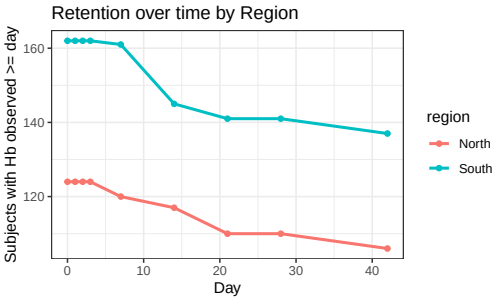


Figure 7: Subject profiles (repeated measures of Hb) for each treatment arm. It is clear that there is a difference in intercept and slope between subject, but this will still formally be tested in the modelling section.



(a) Rate of dropout of subjects by treatment arm.



(b) Rate of dropout of subjects by region.

Figure 8: A plot to show the difference in dropout rates for different treatment arms (left) and regions (right).

stream. The magnitude of the dropout rates do not seem to clearly point to a systemic issue as the number of dropouts still fall within reasonable rates.

2.4 Model Specification

Since this project used repeated measures of Hb on the same subjects, a Linear Mixed Effects model was used to model the difference in treatment efficacy between the two treatment arms. This modelling framework allows for correlation within subject from the repeated measures.

One of the first considerations taken into account was whether region should be treated as a fixed or random effect. The type of question the model would answer differs for the two different model specifications. If we model region as a fixed effect, we would be asking if the treatment effect differs between the different regions, which is very closely related to the more broad question this project is trying to answer (if the treatment effect on Hb differs between the treatment arms).

Further, if we assume that the regions are actually a random sample from a whole population of regions and thus model region as a random effect, we would no longer be estimating the difference between regions but rather estimating the overall variance that there is between regions. This would only work if we have a sufficient amount of regions so that the variance can be estimated. Since there are only two regions, North and South, any variance estimate between the regions would be extremely unstable. It is for these reasons that it was decided to model region as a fixed effect.

As mentioned in the introduction, at this point in the project all observations for day 0 (and technically day's 1 and 2 as well, although both these days only had missing values) were removed and the Hb trajectory from day 3 onwards were modelled since the aim is to model the recovery of Hb levels after the initial drop that is expected from malaria treatment. Figure 9 shows exactly what was modelled, namely the recovery of Hb for different treatment arms.

The approach was to first fit a basic model that had one fixed effect, namely measurement day (pday), and only a random intercept. Thereafter the other fixed effects, namely region, arm and the demographic covariates gender, weight and their various interactions were added one at a time and tested for significance. As noted earlier, the log of weight was used in all the models. Since these elements changes the fixed effects, the Maximum Likelihood method was used as estimation method. Once the fixed effect structure was finalized, the same final model was refit using Restricted Maximum Likelihood (REML) and the remainder of the model fitting was done with REML.

Thereafter, a random slope was added and tested for significance. This was to test if the Hb trajectories for each subject differed. Using this final model, the variance structure was investigated to see if it needed further modelling. Finally, the

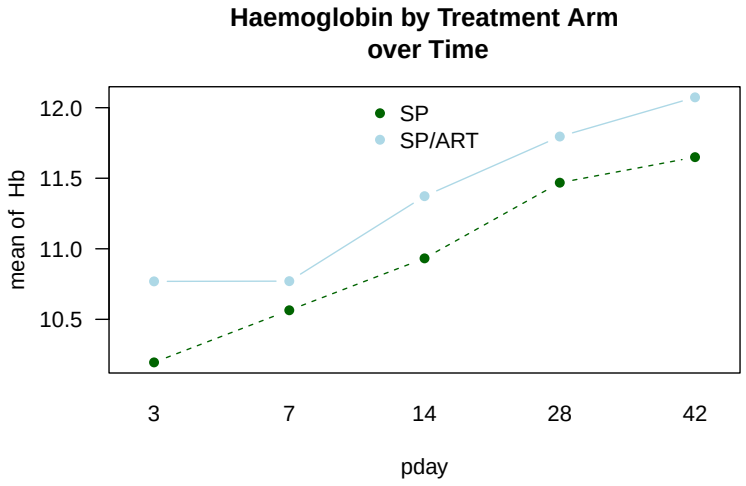


Figure 9: The recovery of Hb after day 0 was removed for the two treatment arms.

correlation structure was also investigated to see if the assumption of uncorrelated error terms was valid.

3 Model Fitting Results

3.1 Building the initial model

As mentioned above, the model fitting began by fitting the Base model as defined in Equation 1. Here i represents the subject and j the measurement day, pday.

$$Hb_{ij} = \beta_0 + \beta_1 \cdot \text{pday}_{ij} + b_{0i} + \epsilon_{ij} \tag{1}$$

with the following distributional assumptions

$$b_{0i} \sim N(0, \sigma_b^2), \quad \epsilon_{ij} \sim N(0, \sigma^2 I), \quad b_{0i} \perp \epsilon_{ij}$$

Here, the fixed effects part is captured by $\beta_0 + \beta_1 \cdot \text{pday}_{ij}$ and the random effects by b_{0i}, representing the random intercept fitted for each subject. ϵ_{ij} is simply the residuals.

From this model, various fixed effects and covariates were added one at a time and tested using a Likelihood ratio test to determine if it significantly improves the model fit. Thereafter the interactions mentioned in the section above were added and tested for significance in the same way. This was done in a sequential manner and the procedure and outcome of each step is described in Table 1. Note that any term that separates two variables with a ':' refers to an interaction between those two terms.

Thus, after the modelling procedure as outlined in Table 1, the best model was model 9 (m9), which is formulated in Equation 2. As noted above, these models were fit using ML since they change the fixed effects in the model. At this point the model 9 was refit using REML so that Likelihood ratio tests could be performed for the remaining of the model fitting procedure. It is worth noting that the estimated of the fixed effects had almost no change in value when moving from the ML to REML procedure.

$$Hb_{ij} = \beta_0 + \beta_1 \cdot \text{arm}_i + \beta_2 \cdot \text{pday}_{ij} + \beta_3 \cdot \text{weight}_{ij} + \beta_4 \cdot (\text{pday}_{ij} \cdot \text{weight}_{ij}) + \beta_5 \cdot \text{gender}_i + b_{0i} + \epsilon_{ij} \tag{2}$$

$$b_{0i} \sim N(0, \sigma_b^2), \quad \epsilon_{ij} \sim N(0, \sigma^2 I), \quad b_{0i} \perp \epsilon_{ij}$$

From the modelling procedure in Table 1 and the resulting model in Equation 2 we can conclude the following:

Table 1: Summary of Fixed Effect Model Building Procedure

Model	**ANOVA p-value**	**Superior Model**
m0: Base model (Equation 1)	NA	NA
m1: m0 + arm	m0 vs m1: p-value of 0.01	m1
m2: m1 + weight	m1 vs m2: p-value of < 0.0001	m2
m3: m2 + gender	m2 vs m3: p-value of < 0.0001	m3
m4: m3 + region	m3 vs m4: p-value of 0.59	m3
m5: m3 + arm:day	m3 vs m5: p-value of 0.53	m3
m6: m3 + arm:gender	m3 vs m6: p-value of 0.11	m3
m7: m3 + arm:weight	m3 vs m7: p-value of 0.32	m3
m8: m3 + pday:region	m3 vs m8: p-value of 0.78	m3
m9: m3 + pday:weight	m3 vs m9: p-value of < 0.0001	m9

- Most noticeably, treatment arm was not significant in the model. This will be discussed in more detail later on.
- weight and gender were significant in modelling the Hb trajectory, whereas region was not.
- The significant interaction between weight and measurement day (pday) indicates that the trajectory of haemoglobin over time varies depending on a subject's weight.

Finally, a random slope was added to the model as described in Equation 2 and tested for significance again using a Likelihood ratio test. The result was a p-value of < 0.0001, meaning that the random slope significantly improved the model fit. Thus, the resulting model was as described in Equation xx.

$$\begin{aligned} Hb_{ij} = \beta_0 + \beta_1 \,\text{arm}_i + \beta_2 \,\text{pday}_{ij} + \beta_3 \,\text{weight}_{ij} \\ + \beta_4 \,(\text{pday}_{ij} \cdot \text{weight}_{ij}) + \beta_5 \,\text{gender}_i + b_{0i} + b_{1i} \,\text{pday}_{ij} + \epsilon_{ij} \end{aligned} \tag{3}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{pmatrix} \right), \qquad \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2 I), \qquad \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \perp \epsilon_{ij}.$$

where b_{0i} and b_{1i} are the random intercept and slopes respectively, τ_{00} is the variance of the random intercepts, τ_{11} is the variance of the random slopes and τ_{01} is the covariance between the random slope and intercept.

The model described in Equation 3 does make the following assumptions about the underlying variance and correlation structures:

- Variance structure assumption: this model assumes that the residual variance is homoscedastic across all observations, i.e. $\text{Var}(\epsilon_{ij}) = \sigma^2 \quad \forall i, j$. This means that every measurement of Hb is assumed to have the same residual variance, regardless of treatment arm, time or covariates.
- Correlation structure assumption: this model assumes that the correlation within the data is fully accounted for by the random effects, i.e. the residuals are assumed to be independent within subject once the random effects are accounted for.

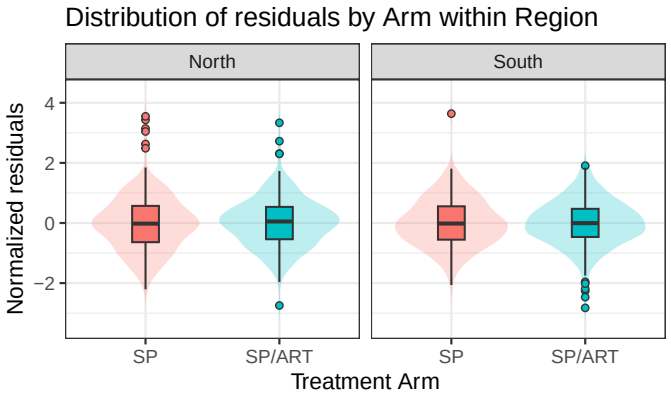
Both of these assumptions will be tested in the two sections that follow.

3.2 Investigation the variance structure

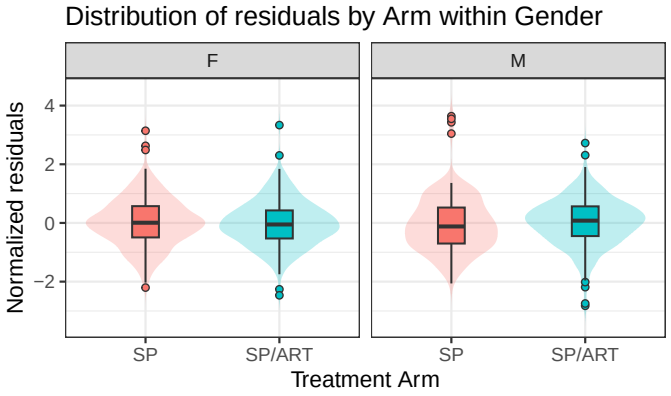
To determine if the variance is homoscedastic, the residuals of the model as in Equation 3 were plotted by treatment arm and region (Figure 10a) and gender (Figure 10b) as well as the residuals for each measurement day (Figure 10c). From all of these plots it is clear that the variance is homoscedastic across different regions, genders, treatment arms and time (i.e. the residuals do not increase at later measurement days).

Furthermore, Figure 11 shows that the standardised residuals plotted against the fitted values form a random scatter with no fanning effect seen as the fitted values increase. For these reasons, the assumption of the current variance structure holds by visual inspection of these plots.

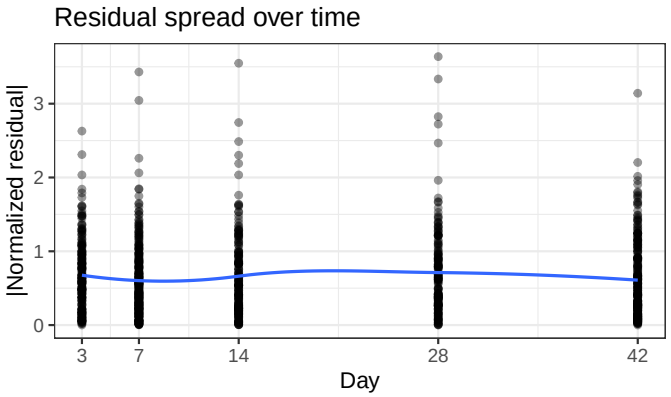
However, for the sake of completeness, a model that has variance that increases or decreases as a power of the mean was also fit and compared with the model in Equation 3. This variance structure is commonly used if the residual spread grows with fitted values of Hb (which is common in biology). The resulting Likelihood ratio test has a p-value of 0.18, which means that we could confidently conclude that the existing variance structure is sufficient to accurately model the data.



(a) The distribution of residuals for the two regions within each treatment arm.



(b) The distribution of residuals for males and females within each treatment arm.



(c) The distribution of residuals within each measurement day.

Figure 10: The distribution of residuals of the model as in Equation 3 by various facets.

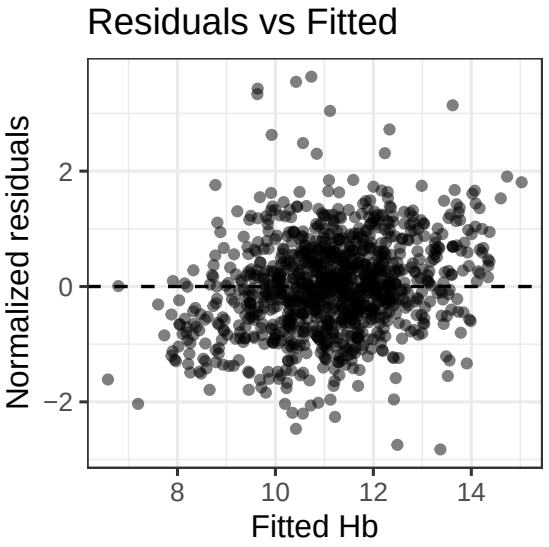


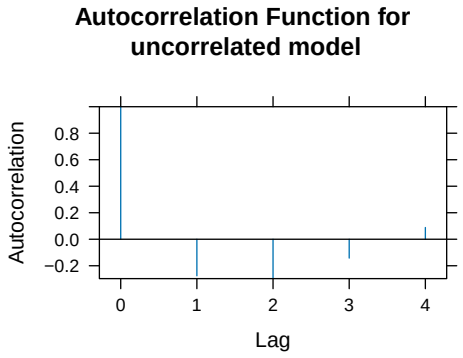
Figure 11: A plot of the standardised results against the fitted values from the model in Equation 3.

3.3 Investigating the correlation structure

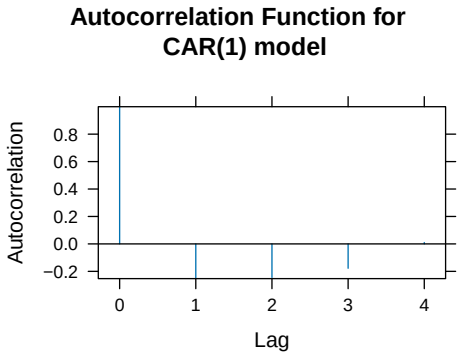
The Autocorrelation Functions for both this model and the model in Equation 3 were plotted in Figure 12. From this figure it is clear that the correlation among residuals was weak, with no autocorrelations exceeding an absolute value of 0.2. This suggests that while some temporal dependence may be present, it is minimal in practice.

Given this result, the gain from introducing a continuous AR(1) structure is limited. Model comparison confirmed this: the CAR(1) model did not provide a statistically significant improvement in fit relative to the simpler random intercept and slope model. Therefore, the additional complexity of modelling residual autocorrelation was not warranted, and the final specification retained the assumption of independent residuals after accounting for random effects.

In summary, although repeated haemoglobin measurements could in principle exhibit autocorrelation, the empirical diagnostics indicated that the magnitude was too small to materially affect inference. Thus, the model in Equation 3 was deemed adequate.



(a) The autocorrelation function for model as in Equation 3.



(b) The autocorrelation function for the CAR(1) model.

Figure 12: Autocorrelation functions to investigate if the underlying correlation structures of the models capture the correlations in the data.

3.4 Model Checking

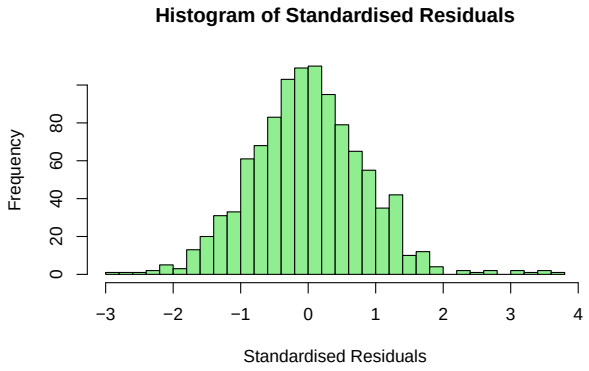
1. **Normality of residuals**
   The within-subject residuals ϵ_{ij} are assumed to be normally distributed with mean zero. This assumption was confirmed by a residual histograms (Figure 13a) and a Q-Q plot of residuals (Figure 13b). The histogram shows that the residuals follow the shape of a normal distribution with some potential outlying observations. Similarly, the Q-Q plot shows that there is a slight deviation from normality at the extreme ends.

2. **Homoscedasticity (constant variance)**
   The residual variance is assumed to be constant across levels of the predictors and over time. This assumption was already checked in a previous section.

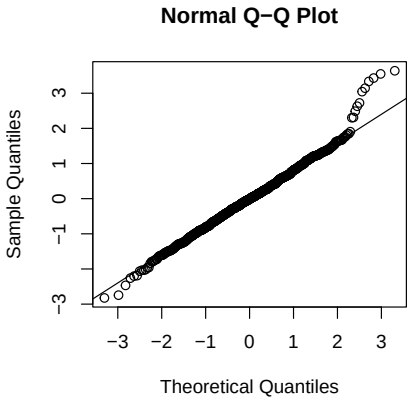
3. **Independence of residuals**
   After accounting for random effects, residuals are assumed independent within and across subjects. This assumption was already checked with autocorrelation function (ACF) plots in Figure 12.

4. **Random effects distribution**
   The random effects (b_{0i}, b_{1i}) are assumed to follow a multivariate normal distribution with mean zero and variance-covariance matrix τ. This assumption was checked with Q-Q plots of estimated random effects in Figure 14a and 14b for the intercept and slope, respectively. From these plots it is clear that the distributions of both follow that of a normal distribution.

Overall, the diagnostic checks indicated that the assumptions of the final model were

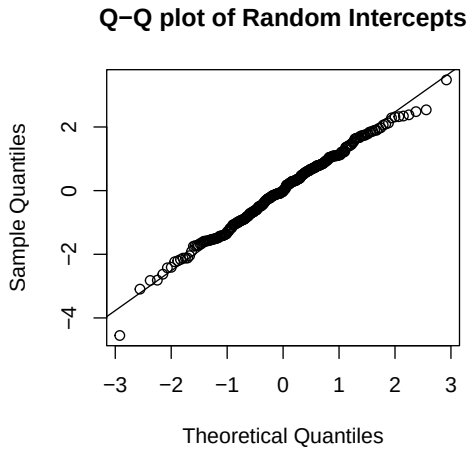


(a) The histogram of standardised residuals for the model as in Equation 3.

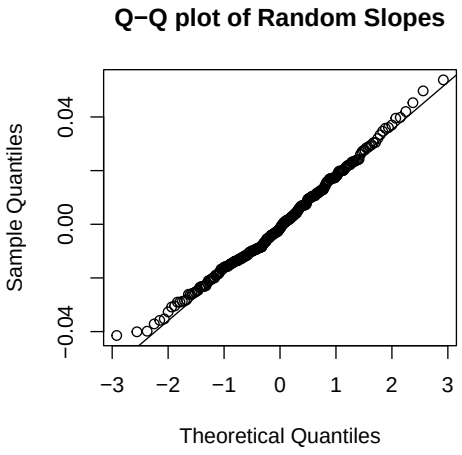


(b) A Q-Q plot of the standardised residuals for the model as in Equation 3.

Figure 13: Figures that help to check the assumption of normality of residuals.



(a) A Q-Q plot of the estimated random intercepts for the model as in Equation 3.



(b) A Q-Q plot of the estimated random slopes for the model as in Equation 3.

Figure 14: Q-Q plots that investigate the distributional assumptions of the random effects in the final model.

reasonably satisfied, with no evidence of major violations or concerning deviations. The next section therefore turns to the substantive question of whether haemoglobin trajectories differed across treatment arms.

4 Conclusion and Discussion

We fitted a linear mixed-effects model with random intercepts and slopes by patient, where the final model equation is as in Equation 4 below. After adjusting for weight and gender, there was a significant difference in mean haemoglobin between treatment arms ($\beta_1 = 0.32$ g/dL, 95% CI = [0.04, 0.60], $p < 0.0001$). This means that Treatment arm SP/ART subjects maintained higher Hb across the 42 days than arm SP, by about 0.32 g/dL on average.

$$\begin{aligned} Hb_{ij} = \beta_0 + \beta_1 \,\mathrm{arm}_i + \beta_2 \,\mathrm{pday}_{ij} + \beta_3 \,\mathrm{weight}_{ij} \\ + \beta_4 \,(\mathrm{pday}_{ij} \cdot \mathrm{weight}_{ij}) + \beta_5 \,\mathrm{gender}_i + b_{0i} + b_{1i} \,\mathrm{pday}_{ij} + \epsilon_{ij} \end{aligned} \tag{4}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{pmatrix} \right), \qquad \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2 I), \qquad \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \perp \epsilon_{ij}.$$

The arm:pday interaction was not significant, indicating that the change in haemoglobin over time was similar between treatment arms. Thus, the main effect of treatment was a shift in overall Hb level, rather than a differential recovery slope.

Although initial Hb levels differed between arms, after day 3 the model showed no evidence of different recovery slopes over the 42-day follow-up. In other words, once baseline differences were accounted for, both regimens showed similar rates of Hb recovery and maintained comparable Hb levels over time. This suggests that the apparent differences were driven by starting values rather than treatment efficacy with respect to Hb.

Appendix: R Code Repository

The full R code used in this analysis is available at:

`github.com/Annie0619/LDA-Assignment-1-2025`

References

[1] Julien Zwang, Umberto D'Alessandro, Jean-Louis Ndiaye, Abdoulaye A Djimé, Grant Dorsey, Andreas A Mårtensson, Corine Karema, Piero L Olliaro, et al. Haemoglobin changes and risk of anaemia following treatment for uncomplicated falciparum malaria in sub-saharan africa. *BMC Infectious Diseases*, 17(1):443, 2017.