Analytical plan for *In vivo* module

## Objectives

* 1. To determine the efficacy of IPTp-SP in clearing peripheral parasiatemias and prevent new infections in asymptomatic parasitaemic pregnant women
	2. To determine the average duration of post-treatment prophylaxis
	3. To determine which parasite genotypes recrudesce, or cause new infections during the post-treatment prophylactic period
	4. To assess the main determinants of the treatment response
	5. To determine whether sparse sampling can replace the standard weekly sampling scheme

## Preparation of the analytical data set

### Protocol violations

Inclusion criteria consist of parasitaemia as defined by the study sites: e.g. by RDT-pLDH positive, smear confirmed, or by blood smear positive in the absence of pLDH-based RDTs.

Exclusion/Enrolment deviations consist in excluding screening failures:

* + - 1. HIV positive women
			2. Hb <5g/dl or Ht <15% or severe anaemia on day 0
			3. No confirmed parasitaemia on day 0 by microscopy

### Definition of loss-to follow up

If more than 3 weeks without blood smear or RDT results, patient is considered lost to follow up on day of last visit (censored data).

For the purpose of reporting efficacy outcomes, a visit occurring 3 days before or after the day of scheduled visit is classified as having occurred on the day of the scheduled visit:

* 1. 7 day blocks thereafter: day 4-10 = day 7, day 11-17 = 14, 18-24=21, etc…
	2. where day 0 = day of treatment, day 1=1st day after SP, day2 and day 3 are the 2nd and 3rd day after treatment
	3. if two visits occur per time block, and both are unscheduled visits, the latest visit is used.

If no blood smear or RDT results are available within the day 39 and day 45 period in a day 42 follow-up study, or 32-38 day period in a 35 day follow-up study, the patient is considered lost to follow up on day of last visit. The patient is considered lost to follow up and censored on the day of last valid visit (censored data). In other words, data will be excluded if it was collected beyond these periods

* 1. Example 1: 42-Day study: a patient comes on day 35, and then comes next on day 48. This patient will be censored on day 35, and the day 48 information will be ignored
	2. Example 2: 42-day study: a patient comes on day 28, 34, and 38. All are ‘scheduled visits’ (not unscheduled visits due to acute illness). The day 38 was supposed to be the day-42 visit, but the patient came a few days early because she had to travel. The values on days 28, 38 will be used, but the value on day 34 will be ignored as day 38 falls outside the 39-45 day time window, but within the day-35 week time window. The patient is censored on day 38.
	3. Example 3: 42-day study: a patient comes on day 28, 34, and 38. The first 2 are ‘scheduled visits’, and day 38 is an unscheduled visits due to acute illness. The patient was negative on day 38, but does not return again. The values on days 28, 34 will be used as scheduled visits, and the value on day 38 as unscheduled. Day 38 does not carry over to day 42 window. The patient is censored on day 38.
	4. Example 3: 42-day study: a patient comes on day 28, 34, and 38. The first 2 are ‘scheduled visits’, and day 38 is an unscheduled visits due to acute illness. The patient was positive on day 38 and is treated again. The values on days 28, 34 will be used as scheduled visits, and the value on day 38 as failure day. Day 38 does not carry over to day 42 window. The patient is treated as an event (failure) that occurred on day 38.

## Definition of efficacy outcomes

Clinical failure: this will not be used as symptomatic women were excluded from the study

Parasitological failure (WHO definitions)

1. Before day 7 (i.e. day 4), treatment failure is not defined unless the patient required rescue therapy because of clinical illness or as determined by study team;
2. Late parasitological failure (LPF) is defined as the presence of parasitaemia on any day between day 4 and last day of follow-up (e.g. day 38 for 5 week follow-up studies (day 35 + 3 days), and day 45 (day 42 + 3 days) for 6 week follow-up studies;
3. Adequate parasitological response (ACPR) is the absence of parasitaemia on the last day of follow up

Treatment failure (day 4 to 45) is measures according to the diagnostic tool used

* 1. Microscopy: use microscopy only
	2. RDT: use RDT only
	3. RDT-pLDH results. If RDT-pLDH result is missing, use blood smear result.
	4. Blood smear results between day 7 and 10 and by RDT-pLDH results from day 11.

## Definition of resistance

Molecular markers to consider

1. Double mutant *dhps*  540, 437
2. Triple mutant *dhfr* 51, 59, 140
3. Quintuple mutant *dhps/dhfr*
4. *dhps* 540 (proxy for quintuple mutation)
5. *dhps* 437
6. *dhps* 581
7. Quintuple mutations plus *dhps* 581

## Statistical analysis

### Descriptive analysis

Baseline data overall and by gravidity includes

1. women age
2. gestational age
3. number of ANC visits
4. employment (employed/unemployed)
5. schooling level (categorized: 0-4, 5-8, 9+, and missing category)
6. wealth status/ SES (derived by principle component analysis (PCA) at site level)
7. use of mosquito net

Proportion of women with treatment failures by each week of follow-up using week blocks

### Survival analysis

1. Uses the actual day of follow-up (not the week blocks)
2. Status is a dichotomous variable (0/1)
3. Survival time is the time between the day of inclusion (day 0) and the day of failure or censoring or end of follow-up. Survival analysis will be reported for days 14, 28, 35, and 42 to allow multiple measures of comparison across sites and to historical data.
4. If genotyping tests are done to distinguish between reinfection and recrudescence, 2 analyses are performed
	1. PCR-Unadjusted analysis: recurrent parasitaemias after day 4 are considered as treatment failures
	2. PCR-Adjusted: recurrent parasitaemias are considered failures, and new infections are censored. If PCR result is unavailable (undetermined result, PCR not done, missing sample), two strategies are explored: (1) recurrent parasitaemias are censored on the day of recurrent parasitaemia, (2) missing values are imputed
5. Failure rates
6. Are calculated using Kaplan-Meier product limit estimate for day 7, 14, 21, 28 and every 7 days until the end of follow-up period (i.e. day 42 for 42-day study)
7. Confidence intervals are based on the asymptomatic variance of the log-log transformed survival function (see Kalbfleish and Prentice, 2002)
8. Failure rates are presented graphically on a Kaplan-Meier survival curve and in a table using the actual days of visit for the step function
9. Failure rates are shown at day 28, day 35 and day 42 to take into account the different lengths of the studies included
10. Cox regression
	1. Uses the actual days of visit rather than the week aggregates
	2. Include a random effect term (frailty) to account for study site heterogeneity
	3. Use an exact method for ties: Efron preferred over Breslow?
	4. Examination of possible confounders
		1. Country/site
		2. Net use/ownership
		3. Gravidity
		4. Prevalence of resistance: quintuple mutation *dhfr*/*dhps* (from *In vivo* at day 0or OPD modules)
	5. Adjusted analysis
		1. Covariates
			1. Initial model include all relevant variables such as net, gravidity, gestational age, net use/ownership, socio-economic status, education level, rural/urban, country/site, prevalence of resistance (quintuple *dhfr*/*dhps*, *dhps* 581, 540)
			2. Full model always includes gravidity because some studies stratified recruitment by gravidity
	6. Subgroup analysis
		1. By gravidity as primigravidae/secundigravidae (G1-2) versus multigravidae (G3+)
11. Compare rich data versus sparse sampling analyses to address objective 1e
	1. Rich data consists of pooling all *in vivo* studies
	2. Sparse sampling consists of taking a random sample from the complete database

### Molecular markers of antimalarial resistance

Level of antimalarial resistance is assessed by determining the prevalence of each molecular marker or combination of markers using genotyped samples at day 0.

## Tables

### Table 1: Distribution of women characteristics



### Table 2: Number of LPF and ACPR and failure rates by week of follow-up



### Table 3: Frequency of molecular markers of resistance

|  |  |
| --- | --- |
|  | ***In vivo* module (n=34)\*\*\*** |
| **Alleles, % mixed/mutant (no.)** |
| ***dhfr,*** *no. loci* |  |
| 51 |  |
| 59 |  |
| 108 |  |
| 164 |  |
| ***dhps,*** *no. loci* |  |
| 437 |  |
| 540 |  |
| 581 |  |
| **Haplotypes, % mixed/mutant (no.)** |
| ***dhfr,*** *no. loci* |  |
| Wildtype |  |
| Single |  |
| Double |  |
| Triple |  |
| ***dhps,*** *no. loci* |  |
| Wildtype |  |
| Single |  |
| Double |  |
| Triple |  |
| **Combined *dhfr-dhps*,** *no. loci* |  |
| Wildtype |  |
| Single |  |
| Double |  |
| Triple |  |
| Quadruple |  |
| Quintuple |  |
| Sextuple |  |

## Figures

### Figure 1: Flowchart



### Figure 2: Distribution of failures by weeks and gravidity- example based on hypothetical data



### Figure 3: Kaplan-Meier cumulative parasite failure curve by gravidity - example based on hypothetical data



## References

* Clinical Module, Worldwide Antimalarial Resistance Network (WWARN), 2011. Data Management and Statistical Analysis Plan.
* World Health Organization 2009. Methods for surveillance of antimalarial drug efficacy. ISBN 978 92 4 159753 1
* Kalbfleish, Prentice. The statistical analysis of failure time datat. New-York: John Wiley & Sons, 2002.