Parasite Clearance Estimator Report

Automated analysis and report of dataset "ExampleData"

Generated 29-Apr-2013

Contents

1 Background

The emergence and spread of resistance to antimalarial drugs threatens the efficacy of existing drug treatments. Parasite clearance rate (the rate at which parasitaemia declines) is an important measure of drug efficacy. It can be used to assess the in vivo responses to treatment with artemisinin derivatives, evaluate new antimalarial drugs and assess therapeutic response in severe malaria and hyperparasitaemia. A lag phase, which often precedes a fall in the parasite count, complicates the estimation of parasite clearance. This period must be identified by the observer, introducing subjectivity that may influence both the accuracy and consistency of results.

Clear guidelines do not currently exist that define a consistent method of identifying resistance. This increases the need for an accurate and consistent method of clearance estimation, to allow data comparison between populations (see Figure 1). Only by comparison with other populations can prolonged clearance, an important early warning sign of resistance to artemisinin derivatives, be identified. To address this need, WWARN has developed the Parasite Clearance Estimator (PCE). Our aim is to provide the community with an accurate and consistent method of clearance estimation. PCE uses a new approach to identify the lag phase, clean the data and fit the best model available to estimate parasite clearance. Several measures of parasite clearance are calculated: clearance rate constant, slope half-life, PC50, PC90, PC95 and PC99 (see Glossary).

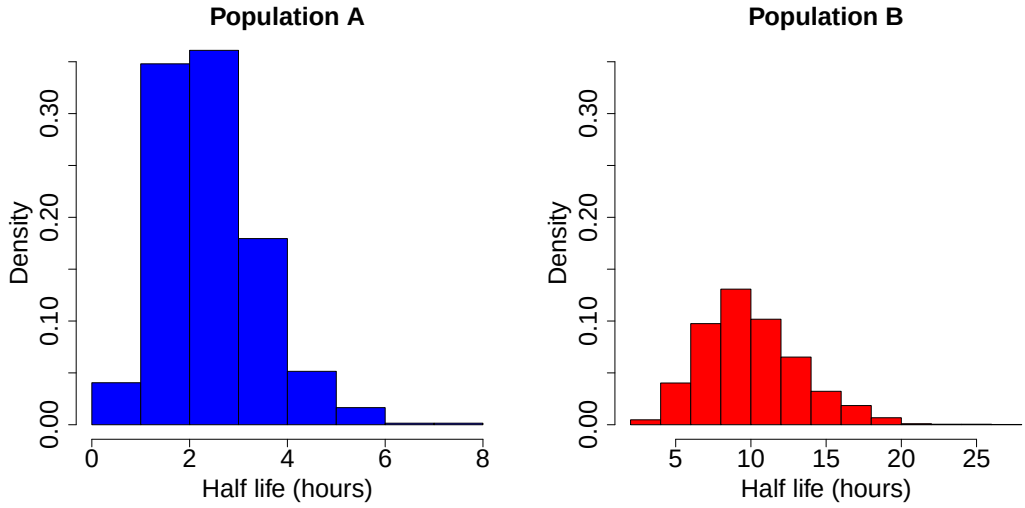


Figure 1: Stylised graphs showing the distribution of parasite clearance half lives for two populations: population B shows evidence of prolonged clearance, when compared to population A.

2 Methods

WWARN's PCE estimates several measures of parasite clearance for each individual patient, based on the linear part of the log_e parasitaemia-time profile. The tool also identifies lag phase, tail and outlier observations. See Glossary for explanation of terms.

Figure 2 illustrates how parasite clearance estimates may be affected by these factors, underlining the importance of a consistent, reliable method to identify and accommodate them.

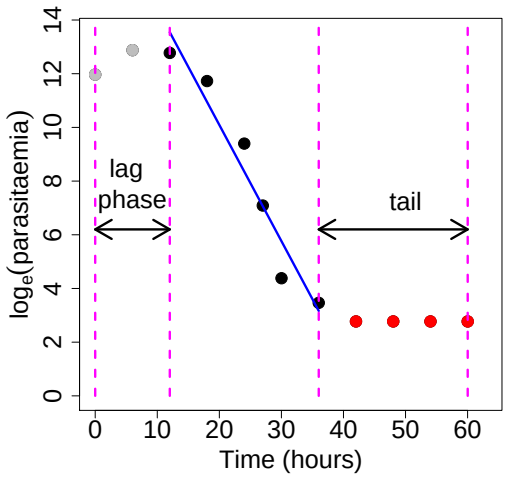


Figure 2: Stylised graph showing the effect of lag phase and tail exclusion on the calculation of the clearance rate constant.

The parasite clearance measures are estimated by a four-stage process:

1. Data cleaning: data relating to the outliers and the tail are removed.
2. Model fitting: the curve of log_e parasite count versus time is described mathematically by a polynomial model. Tobit regression models are used to account for parasitaemias below the level of detection. For details of tobit regression see Tobin, J. (1958). "Estimation of relationships of limited dependent variables." Econometrica 26: 24-36.
3. Lag phase identification from the mathematical model.
4. Clearance rate estimation: if a lag phase is not identified, the clearance rate constant is estimated as the absolute value of the slope of the linear regression model fitted to all data. If a lag phase is identified, the clearance rate constant is estimated as the absolute value of the slope from the linear part of the predicted profile.

3 Summary of results

The limit of detection for this report was 40 parasites/microlitre.

3.1 Individual data profiles

Measure	Number of profiles
In dataset	20
Analysed	20
With lag phase detected	5
With tail detected	0
Profiles with outliers removed	0

3.2 Data profiles not analysed

Reason for exclusion	Number of profiles	Profile ID
Too few data points	0	
Parasitaemia too low	0	
Zero replaced and last positive parasitaemia exceeds 1000 and the zero is not informative	0	
No zero replaced and last positive parasitaemia > 1000	0	

3.3 Estimated clearance rate constant

Summary of clearance rate constant (/hour)

Statistic	value
Median	0.2102
Range*	(0.106,0.3157)
IQR**	(0.1763,0.2436)

*Range = (minimum, maximum)

**IQR = Inter Quartile Range; (25th centile, 75th centile)

Clearance rate constant distribution

Clearance rate (/hour)*	N	%	Cumulative %
0.00 to 0.05	2	10.00	10.00
0.05 to 0.10	7	35.00	45.00
0.10 to 0.15	7	35.00	80.00
0.15 to 0.20	3	15.00	95.00
0.20 to 0.25	1	5.00	100.00

*each interval includes the left-hand side value and does not include the right-hand side value

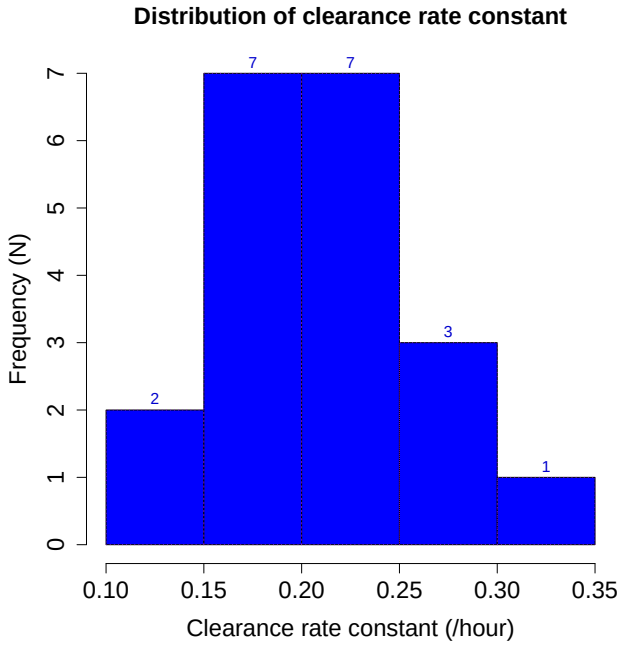


Figure 3: Distribution of clearance rate constant

3.4 Slope half life (hours)

Summary of slope half life

Statistic	value
Median	3.299
Range*	(2.195,6.537)
IQR**	(2.846,3.945)

*Range = (minimum, maximum)
**IQR = Inter Quartile Range; (25th centile, 75th centile)

Slope half life distribution

Slope half life (hours)*	N	%	Cumulative %
0 to 3	7	35	35
3 to 4	8	40	75
4 to 5	4	20	95
5 to 6	0	0	95
6 to 7	1	5	100

*each interval includes the left-hand value and excludes the right-hand value

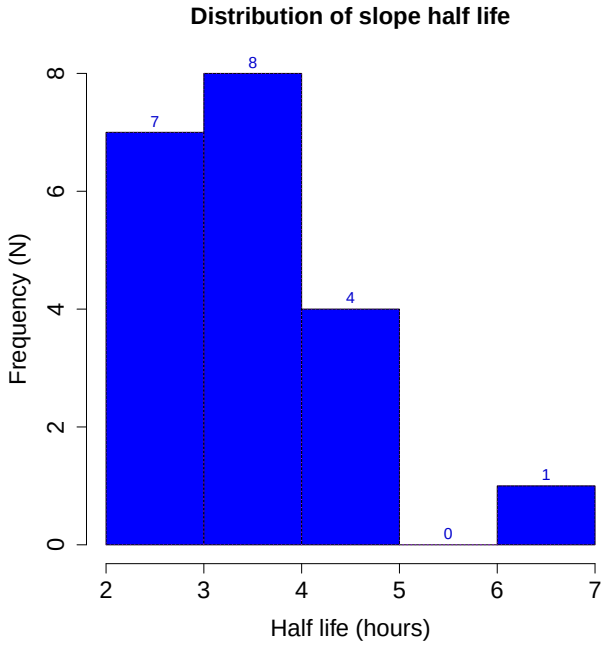


Figure 4: Distribution of slope half life

3.5 Duration of lag phase

Summary of individual patient lag phase (hours)

Statistic	
Number of profiles with lag phase	5
Duration of lag phase (hours)	
Median	19.58
Range*	(4.17,20.17)
IQR**	(7.92,20)

*Range = (minimum, maximum)
**IQR = Inter Quartile Range; (25th centile, 75th centile)

Summary of time to clear 50%, 90%, 95% and 99% of parasitaemia (hours)

	Statistic		
	Median	Range*	IQR**
PC50	6.396	(0.03271,33.38)	(4.774,9.877)
PC90	13.06	(7.184,41.71)	(11.13,20.69)
PC95	16.47	(10.04,45.3)	(13.8,25.58)
PC99	24.31	(15.16,53.63)	(21.14,36.34)

*Range = (minimum, maximum)
**IQR = Inter Quartile Range; (25th centile, 75th centile)

4 Explanation of output

Two .CSV files with estimation results are provided. Tables 1 and 2 provide file column definitions. Figures are provided as .pdf files.

Table 1: Results file Estimates_short.csv column definitions

Name	**Description**	**Coding/remarks**
Id	Patient id	As given by investigator
Estimation summary	Indication of whether clearance estimation was successful	0 = lag estimation attempted 1 = no lag estimation attempted 2 = no estimation
Excluded observations	Indication if there were any data-points excluded from estimation	0 = all points included 1 = outlier detected 2 = tail detected 3 = both tail and outlier detected
No Fit	Reason for not fitting a model	1 = Not enough data points 2 = First parasitaemia < 1000 3 = Last positive parasitaemia > 1000
Tlag	Duration of lag phase in hours	
Clearance rate constant	Estimated clearance rate constant (K) (1/hours)	Clearance rate constant = minus slope of the final model after exclusion of outliers, lag phase and tail.
SE_clearance	Standard error of clearance rate constant	
Intercept_tlag	Intercept at time = Tlag	
Slope_half_life	Estimated time in hours for parasitaemia to decrease by half	
R2	R^2 statistic from the linear (or tobit) regression on datapoints used to estimate clearance (i.e. lag phase and tail are excluded) and excluding measurements below the level of detection.	
PC50	Estimated time in hours for parasitaemia to reduce by 50% of initial value	
PC90	Estimated time in hours for parasitaemia to reduce by 90% of initial value	
PC95	Estimated time in hours for parasitaemia to reduce by 95% of initial value	
PC99	Estimated time in hours for parasitaemia to reduce by 99% of initial value	

Table 2: Definition of columns in results file Estimates.csv

Name	Description	Coding/remarks
Id	Patient id	As given by investigator
Time	Time of measurement in hours	As given by investigator
Para	Measured parasitaemia (per microliter	As given by investigator
Lpara	log_e measured parasitaemia	Calculated from the investigator values
Detect	Detection limit used	As given by investigator
Outlier	Outlier/tail detection	0 = data is included (no outlier detected) 1 = extreme value 2 = outlier 3 = tail 4 = recurrence episode 5 = final sequence of zeros
Estimation summary	Indication of whether clearance estimation was successful	0 = lag estimation attempted 1 = no lag estimation attempted 2 = no estimation
No Fit	Reason for not fitting a model	1 = Not enough data points 2 = First parasitaemia < 1000 3 = Last positive parasitaemia > 1000
Tlag	Duration of lag phase in hours	
Clearance rate constant	Estimated clearance rate constant (K) (1/hours)	Clearance rate constant = - slope of the final model after exclusion of outliers, lag phase and tail.
SE_clearance	Standard error of clearance rate constant	
Intercept_tlag	Intercept at time = Tlag	
Slope_half_life	Estimated time in hours it takes for the parasitaemia to decrease by half (50%)	
R2	R^2 statistic from the linear (or tobit) regression on datapoints used to estimate clearance (i.e. lag phase and tail are excluded) and excluding measurements below the level of detection.	
PC50	Estimated time in hours for parasitaemia to be reduced by 50% of its initial value	
PC90	Estimated time in hours for parasitaemia to be reduced by 90% of its initial value	
PC95	Estimated time in hours for parasitaemia to be reduced by 95% of its initial value	
PC99	Estimated time in hours for parasitaemia to be reduced by 99% of its initial value	
Predicted	Predicted log parasitaemia from the final model (excluding tail and lag-phase if identified)	
Linear_slope	Slope of the linear regression or tobit linear regression model after exclusion of tails and outliers but not measurements in the lag phase.	
SE_linear_slope	Standard error of Linear slope	
Intercept_linear	Intercept of the linear regression or tobit linear regression model after exclusion of tails and outliers but not measurements in the lag phase.	
R2_linear	R^2 statistic from linear regression; measurements below detection limit are excluded	
Predicted_linear	Predicted log parasitaemia from the linear regression model or tobit linear regression model	

5 Acknowledgements

The statistical models used to estimate the parasite clearance measures and lag phase duration were fitted using the Parasite Clearance Estimator developed by the WorldWide Antimalarial Resistance Network (WWARN).

If clearance estimates generated using WWARN PCE are included in publications or presentations, the following acknowledgement should appear:

> "The statistical models used to estimate the parasite clearance measures and lag phase duration were fitted using the Parasite Clearance Estimator developed by the WorldWide Antimalarial Resistance Network (WWARN)."

The following citation should appear in publications that use the Harvard system of referencing:

> WorldWide Antimalarial Resistance Network (WWARN), 2012. Parasite Clearance Estimator. [online] Available at: <http://www.wwarn.org/research/parasite-clearance-estimator> [Accessed 30 May 2012]

Suggested citation: Clinical Module, WWARN. 2011. Parasite Clearance Report.

For more information, contact:

pce@wwarn.org

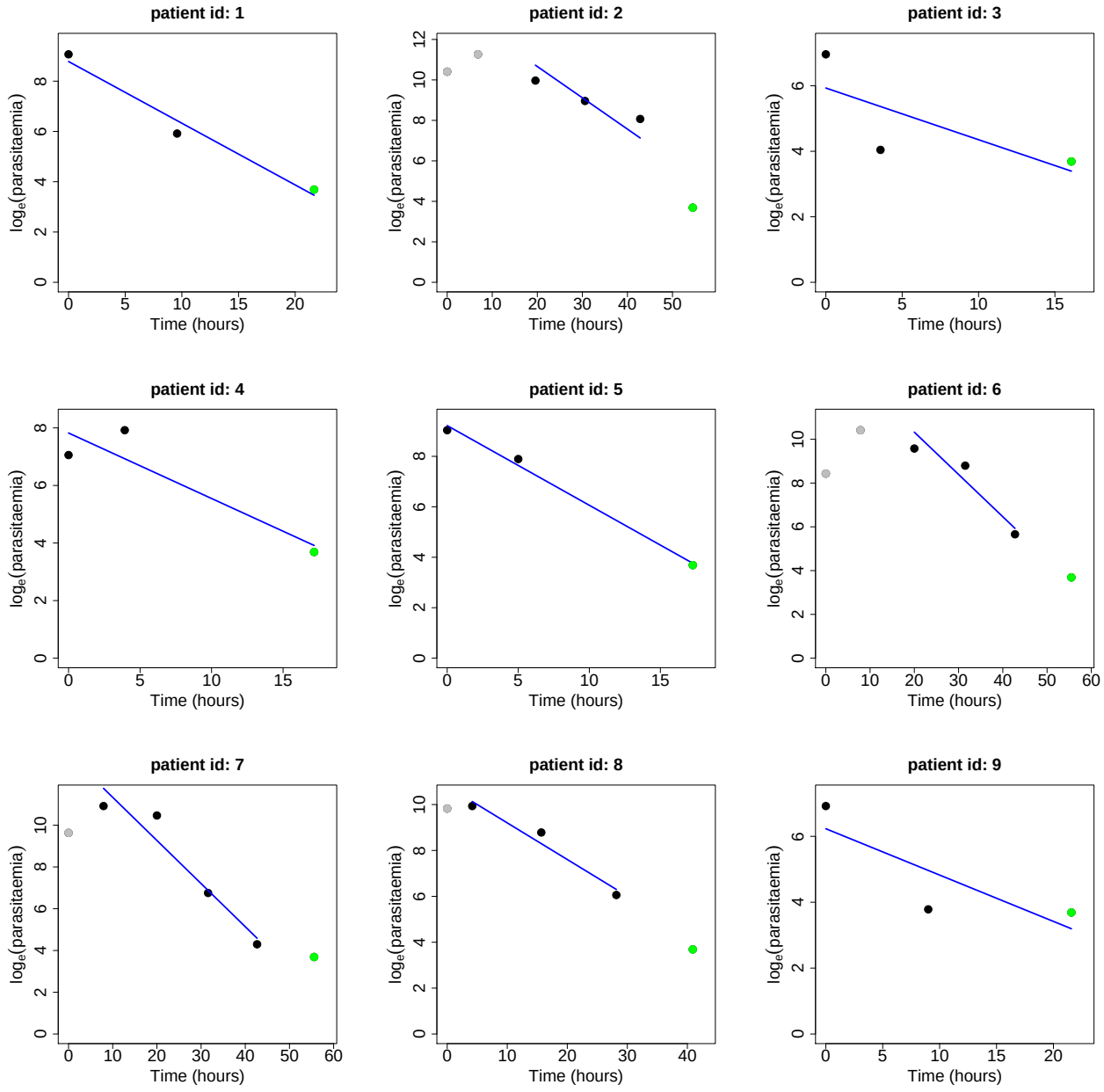
WorldWide Antimalarial Resistance Network (WWARN)

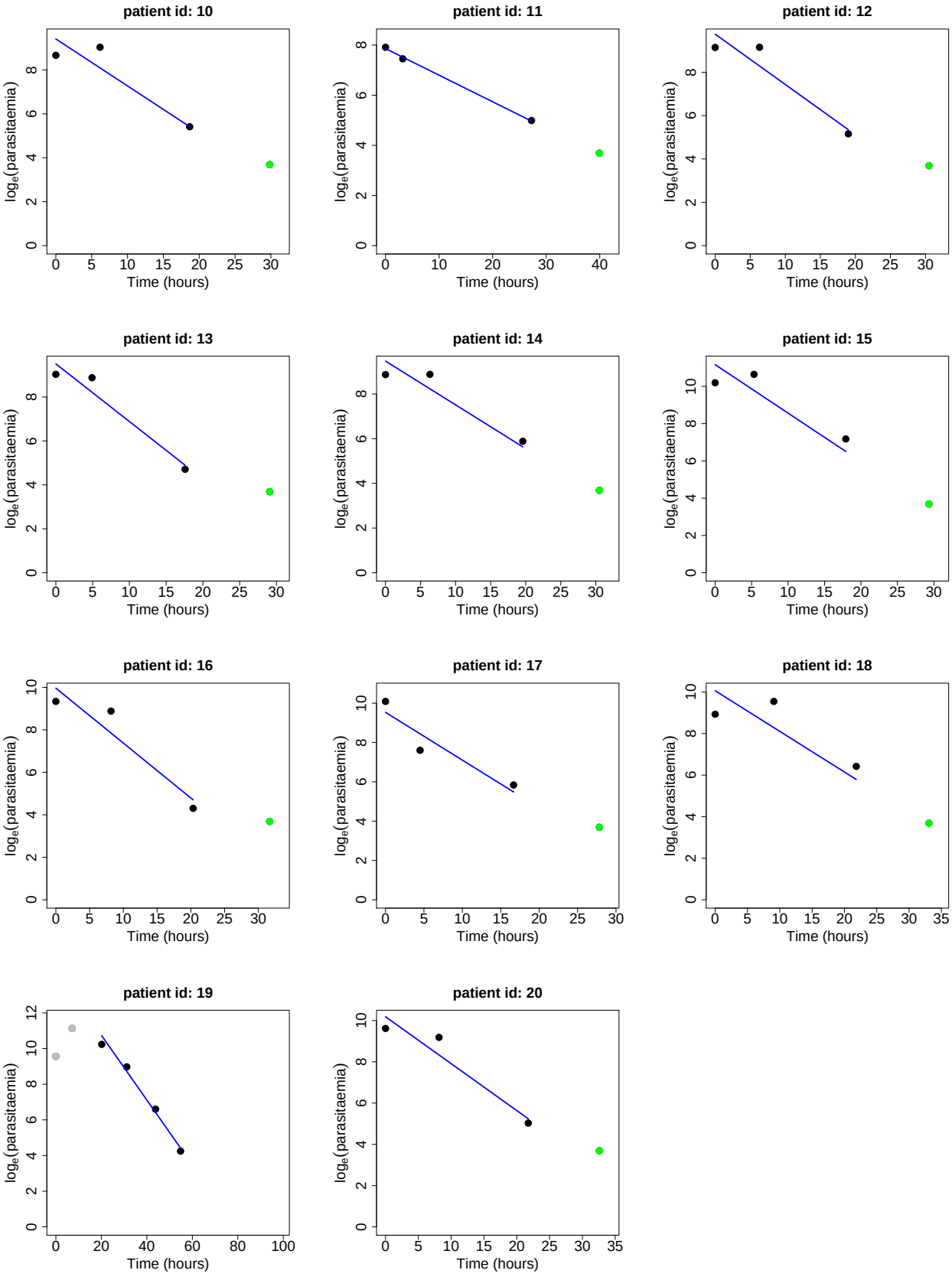
www.wwarn.org

A Parasite clearance estimation for individual patient

The limit of detection for this report was 40 parasites/microlitre. The individual patient parasite clearance estimation charts use these colours:

- measurements during the lag phase
- measurements used in estimation of clearance rate
- outliers excluded from calculations
- measurements below the level of detection
- linear regression model (i.e. predicted log parasitaemia) used to derive clearance rate




patient id: 10
patient id: 11
patient id: 12
patient id: 13
patient id: 14
patient id: 15
patient id: 16
patient id: 17
patient id: 18
patient id: 19
patient id: 20
log$_e$(parasitaemia)
Time (hours)

B Glossary

Definitions of terms used in relation to the Parasite Clearance Estimator.

Detection limit the thick blood smear is used for low parasite counts. The number of parasites (asexual forms only i.e. trophozoites) are counted against the number of white blood cells (usually 200 or 500). The detection limit will depend on the number of white blood cells (wbc) counted. To estimate parasitaemia per microlitre the following formula is used:

Parasitaemia per microlitre = number of parasites per slide × white blood cell count / x

where x is the number of white cells counted

Ideally the wbc is measured by an automated cell counter or manually. If this is unavailable, the counts are assumed to be 8,000/uL. For an assumed wbc of 8,000 for all patients, the detection limit is 40/uL (counting per 200 wbc), 16/uL (counting per 500 wbc).

Negative parasite slide if no parasites are seen while the full number of white cells have been counted, the parasite count is recorded as zero. This means that the count is below the limit of detection.

Outliers parasite counts which are not biologically possible or are highly unlikely based on other parasite measurements in the same individual. These often result from transcription errors.

Lag phase the initial flatter part of the parasite clearance profile. Note: a lag phase is not observed in all profiles. For details of the process by which the lag phase is estimated, see the documentation:

http://www.wwarn.org/research/parasite-clearance-estimator

Tail terminal part of the parasite clearance profile where parasitaemia remains close to the detection limit and does not decrease over a number of measurements' time-points. Tails are not observed in all profiles.

Clearance rate constant the predominant portion of the parasitaemia clearance slope follows a first order process. Therefore the fraction by which parasite count falls per unit time is constant. If parasitaemia at time t is given by $P_t = P_0 e^{-Kt}$, where P_0 is the initial parasitaemia, then the parasite fractional reduction is equal to

$$\frac{P_{t=1} - P_0}{P_0} = e^{-K} - 1.$$

The parameter K is the clearance rate constant and can be shown to be equal to the minus slope of the log_e parasitaemia–time linear relationship (that is, $K > 0$). Parasitaemia declines according to the first order process:

$$P_t = P_0 e^{-Kt}$$

Taking log_e of each side of the equation:

$$log_e(P_t) = log_e(P_0) - Kt$$

Note: the clearance rate constant is calculated from the subset of data after the tail and lag phase are removed.

Slope half life the time needed for parasitaemia to reduce by half. This is a constant independent of the starting value of parasitaemia since reduction in parasitaemia follows the first order process (after excluding tail and lag phase). Half life is calculated with the formula

$$T_{1/2} = log_e 2/K = 0.692/K$$

where K is the clearance rate constant.

PCx time taken for parasitaemia to reduce by x% of the admission parasitaemia, based on the linear model fitted to the part of the profile identified as having linear clearance. We provide values for PC50, PC90, PC95 and PC99. Since PCx is estimated from the fitted model, in rare cases, when the intercept for the linear model is much lower then the measured initial parasitaemia, PC50 may not be estimable. Note: PC50 is equal to the slope half life only in the absence of a lag phase and when the fitted value at the initial time point is equal to the initial parasitaemia.