Data dictionary

**combined\_data.csv:**

|  |  |
| --- | --- |
| T | Time since treatment |
| N\_I\_tripl\_new | New infections with *dhps* AKA |
| N\_I\_quadr\_new | New infections with *dhps* **G**KA |
| N\_I\_quint\_new | New infections with *dhps* **GE**A |
| N\_I\_sext\_new | New infections with *dhps* **GEG** |
| N\_I\_undeterm\_new | New infections with undetermined or multiple mixed *dhps* genotype |
| N\_I\_tripl | Cumulative new infections with *dhps* AKA |
| N\_I\_quadr | Cumulative new infections with *dhps* **G**KA |
| N\_I\_quint | Cumulative new infections with *dhps* **GE**A |
| N\_I\_sext | Cumulative new infections with *dhps* **GEG** |
| N\_I\_undeterm | Cumulative new infections with undetermined or multiple mixed *dhps* genotype |
| N\_uninf | Number of participants that remain at risk of new infection (uninfected) |

Notes:

* dhps codons: A437**G**/K540**E**/A581**G** (mutant forms shown in **bold**)
* New infections with each genotype refer to new infections occurring between T-1 and T.
* Cumulative number of new infections with each genotype refer to all infections occurring between T=0 and T.
* Some of the studies in the code and data are labelled with a study code allocated by WWARN, e.g., RZENT, NMMSB etc. (for more information see first column of Supplementary Table 1)

**day0\_frequencies**

|  |  |
| --- | --- |
| triple\_day0 | Day 0 frequency of *dhps* AKA |
| quadruple\_day0 | Day 0 frequency of *dhps* **G**KA |
| quintuple\_day0 | Day 0 frequency of *dhps* **GE**A |
| sextuple\_day0 | Day 0 frequency of *dhps* **GEG** |
| other\_day0 | Day 0 frequency of any other *dhps* genotypes |

Note:

We define the frequency of a genotype as the proportion of samples with that genotype among samples with unmixed infections in the 3 codons of interest (437,540 and 581)