Uncomplicated malaria is defined as symptomatic malaria without evidence of complicated malaria.

Complicated malaria is defined as symptomatic malaria with the presence of any “danger signs” or meeting criteria for severe malaria.

Danger signs are as follows:

1. Less than 3 convulsions over 24 hour period
2. Inability to sit up or stand
3. Vomiting everything
4. Unable to breastfeed or drink
5. Lethargy

Criteria for severe malaria include any of the following:

1. Impaired consciousness or unarousable coma not attributable to any other cause (cerebral malaria)
2. Generalized convulsions (> 3 convulsions over 24 hours period)
3. Severe normocytic anemia (Hb < 5 gm/dL)
4. Hypoglycemia (blood glucose < 2.2 mmol/L or < 40 mg/dL)
5. Metabolic acidosis (plasma bicarbonate < 15 mmol/L)
6. Hemoglobinuria
7. Hyperlactatemia (lactate > 5 mmol/L)
8. Acute renal failure (serum creatinine > 265 µmol/L)
9. Acute pulmonary edema and adult respiratory distress syndrome (ARDS)
10. Circulatory collapse or shock (systolic blood pressure < 70 mm Hg in adults and < 50 mm in children)
11. Abnormal spontaneous bleeding
12. Jaundice plus evidence of other vital organ dysfunction

Other comments in reference to your variable names and definitions for your consideration:

1) "ave\_anopheles", "num collections", "totalanopheles" - The entomology databases is structured such that we have one observation per house per night of collection (using CDC light traps). The main variable from this database is "totalanopheles" which is simply the total number of female anopheles mosquitoes collected in that house, that night. This data will be most useful as a "incident density measure", that is the number of something per unit time (analogous to the incidence of malaria when is the number of episodes of malaria per calendar time observed). For the entomology data the most useful density measure would be the number of female anopheles collected per number of nights of collection (this could be covered to an annual human biting rate by multiplying this by 365 days per year). The density measure could be over any period of time (we collect mosquitoes once a month) and for an individual house or a group of houses (i.e. all houses in one of the 3 study sites). Thus what I think you are looking for is one derived variable which would be the " the number of female anopheles collected per number of nights of collection" (the person doing the search can specify over what time frame and what house(s)).

2) "ave anopheles positive", "ave anopheles tested", "ave percent anopheles positive" - analogous to the above, the entomology database is structured such that for each house and each night of collection we report the number of female anopheles mosquitoes that were tested for sporozoites (this may be all mosquitoes collected or subset collected) and the number that tested positive for sporozoites. I think we are just interested in one derived variable: the "sporozoite rate", which is the proportion of mosquitoes tested for sporozoites which test positive. Again the person doing the search can specify over what time frame and what house(s) they are interested in.

3) asymptomatic parasitemia should be define as the presence of asexual parasites by microscopy in the absence of subjective fever or documented elevated temperature

4) BSdich - is whether or not a blood smear was positive for asexual parasites by microscopy

5) eaves - for our purposes eaves are define as a gap between the roof and the walls (important for mosquito entry into a home)

6) fever - defined as a subjective fever as reported by the study participant or their parent/guardian

7) gender - your definition seemed to be unnecessarily complex

8) housetype - this is a composite variable we derived with two categories: modern house (cement/wood/metal wall; tiled/metal roof; and closed eaves) and traditional house (all other houses)

9) parasite density should be defined as the number of asexual parasites per uL