Questions on ZIKA dataset

* According to the metadata file, the Brazil\_RiodeJaneiro\_Cunha study includes women of reproductive age. This means it can also include nonpregnant women. Is this indeed the case? If yes, is there a variable indicating whether a woman was pregnant or not? All outcomes are only applicable to pregnant women so nonpregnant women should be excluded.
* Variable “loss\_etiology” – what do values of 4 mean?
* Variable “tobacco” – what do values of 3 mean?
* Check creation new outcome variables (see index.html)
* Variables “inf\_weight” (min 3.4 gram, max 23500 gram), “inf\_length” (values of 0, 4, 19 cm), inf\_head\_circ\_birth, inf\_head\_circ\_age\_fu1, age, all contain outliers -> can these be cleaned? zikv\_pcr\_ga\_1, zikv\_elisa\_ga\_1, zikv\_ga, symp\_ga, arb\_clindiag\_ga
* Outcome: ‘postnatal microcephaly’. We can use head-circumference variables to check for any changes. How? See if there is microcephaly at any time point and compare that to presence of microcephaly at birth?
* Help with passive imputation of variables.
* Effect modifier: “Genetic anomalies, metabolic disorders, perinatal brain injury” -> how can we construct this?
* Effect modifier: “Clinical/subclinical illness” -> how should we construct this?
* Effect modifier: “Presence and severity of maternal and infant clinical symptoms” -> how should we construct this?