# Last edited: 14 May 2020

# Last run: 14 May 2020

# Objective: convert Henry's algorithm from SAS to R

# run algorithms (neonate and child) on COMSA data

rm(list = ls())

# library(rJava)

# library(openVA)

library(readr)

library(dplyr)

# library(CrossVA)

library(readxl)

library(xlsx)

library(data.table)

library(tidyverse)

library(haven)

'%!in%' <- function(x,y)!('%in%'(x,y))

file <- getwd()

COMSAdata <- read.csv(file.path(file,"Data/20200128\_all\_WHO.csv"), stringsAsFactors = FALSE)

VASAvars <- read\_dta(file.path(file,"Data/20191028\_eng\_ext\_child\_VASA.dta"))

VASAvars <- as.data.frame(cbind(VASAvars$comsa\_id,VASAvars$c3101))

names(VASAvars)[names(VASAvars) == 'V1'] <- 'comsa\_id'

names(VASAvars)[names(VASAvars) == 'V2'] <- 'c3101'

VASAvars <- subset(VASAvars, comsa\_id != ".")

dim(COMSAdata)

dim(VASAvars)

COMSAdata <- merge(COMSAdata,VASAvars, by="comsa\_id", all.x=TRUE)

dim(COMSAdata)

table(COMSAdata$source)

names(COMSAdata) <- tolower(colnames(COMSAdata))

dim(COMSAdata) # 6177 cases

# Remove Stillbirths

table(COMSAdata$id10104, exclude = NULL)

table(COMSAdata$id10109, exclude = NULL)

table(COMSAdata$id10110, exclude = NULL)

COMSAdata$Stillbirth <- ifelse(COMSAdata$id10104 %in% c("no","dk") & COMSAdata$id10109 %in% c("no","dk") & COMSAdata$id10110 %in% c("no","dk"),1,0)

COMSAdata <- COMSAdata[COMSAdata$Stillbirth!=1,]

dim(COMSAdata)

# Make a matrix that will be filled with COMSA data

varlist <- c("malnutrition1", "malnutrition2", "AIDS1", "AIDS4", "diarrhea8", "dysentery8", "diardysn8", "possiblediar8\_4", "possibledysn8\_4", "possdiardysn8\_4","hemfever","malaria251",

"measles4","meningitis","pertussis","pneumoniafb2daysgr","possibleari3", "sepsis\_nomal251",

"residual\_infect\_slide15\_4", "malaria\_possible","injury","injury3\_slide15\_4","malaria251",

"congmalf2","bi5ba5","ba5","preterm\_all\_mo",

"allexpertdxs","allexpertdxs.alt")

# age at death in days

COMSAdata$ageatdeath <- NULL

COMSAdata$ageatdeath <- COMSAdata$ageindays

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$age\_neonate\_minutes) & !is.na(COMSAdata$isneonatal), COMSAdata$age\_neonate\_minutes/(60\*24), COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$age\_neonate\_hours) & !is.na(COMSAdata$isneonatal), COMSAdata$age\_neonate\_hours/(24), COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$ageindaysneonate) & !is.na(COMSAdata$isneonatal), COMSAdata$ageindaysneonate, COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$age\_child\_months) & !is.na(COMSAdata$ischild), COMSAdata$age\_child\_months\*30.4, COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$age\_child\_years) & !is.na(COMSAdata$ischild), COMSAdata$age\_child\_years\*365.25, COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$age\_adult) & !is.na(COMSAdata$isadult), COMSAdata$age\_adult\*365.25, COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$ageinyears) & !is.na(COMSAdata$isadult), COMSAdata$ageinyears\*365.25, COMSAdata$ageatdeath)

COMSAdata$ageatdeath <- ifelse(is.na(COMSAdata$ageatdeath) & !is.na(COMSAdata$ageinmonths) & !is.na(COMSAdata$isadult), COMSAdata$ageinmonths\*30.4, COMSAdata$ageatdeath)

table(is.na(COMSAdata$ageatdeath))

COMSAdata$age <- COMSAdata$ageatdeath

# COMSAdata.childu5 <- subset(COMSAdata, ischild==1 & COMSAdata$ageinyears <= 4 | COMSAdata$age\_child\_years <= 4 | COMSAdata$age\_child\_months <= 12)

COMSAdata$group[COMSAdata$ageatdeath >= 28 & COMSAdata$ageatdeath < 59\*30.4] <- "ch1\_59"

COMSAdata.childu5 <- subset(COMSAdata, group=="ch1\_59")

dim(COMSAdata.childu5)

COMSAdata <- COMSAdata.childu5

# COMSAdata$ID <- as.character(rownames(COMSAdata))

EAVA <- as.data.frame(matrix(data=NA,nrow=nrow(COMSAdata),ncol=length(varlist)))

names(EAVA) <- varlist

COMSAcols <- cbind(COMSAdata$comsa\_id,COMSAdata$source)

EAVA <- cbind(COMSAcols,EAVA,stringsAsFactors = FALSE)

head(EAVA)

names(EAVA)[names(EAVA) == '1'] <- 'ID'

names(EAVA)[names(EAVA) == '2'] <- 'source'

head(EAVA)

### Child - Algorithms

# \* Malnutrition1 (residual malnutrition â€“ placed at the bottom of the hierarchy);

# \* Expert VA and available in GC13 VA: Limbs became very thin during the fatal illness or had swollen legs or feet during the illness;

# if vacq4350=1 or vacq4360=1 then malnutrition1=1;

# else malnutrition1=2;

table(COMSAdata$id10244, exclude = NULL)

table(COMSAdata$id10249, exclude = NULL)

EAVA$malnutrition1 <- ifelse(COMSAdata$id10244 %in% "yes" | COMSAdata$id10249 %in% "yes",1,2)

table(EAVA$malnutrition1, exclude = NULL)

### check that EAVA is lining up with COMSA data correctly

test <- cbind(EAVA$ID, EAVA$malnutrition1, COMSAdata$comsa\_id, COMSAdata$id10244, COMSAdata$id10249)

head(test)

test <- as.data.frame(test)

test[,6] <- NA

head(test)

test[,6][test[,1]!=test[,3]]<-1

head(test)

###

# \* |++| Malnutrition2 |++|; (modified algorithm: when SA data on the order of symptoms appearance are not available);

# \* Expert VA and available in GC13 VA: Limbs became very thin during the fatal illness or had swollen legs or feet during the illness, and 1 of these was the first symptom of the illness (the last criterion requires using the SA data on the order of the symptoms);

# \* Expert VA and available in GC13 VA: Had swollen legs or feet during the illness AND the swelling duration was greater than or equal to the illness duration;

# if vacq4360=1 and vacq4370>=vacq1210 then malnutrition2=1;

# else malnutrition2=2;

# swelling duration

# COMSAdata$swelldur <- ifelse(COMSAdata$q5140a=="months",31\*COMSAdata$id10250,

# ifelse(COMSAdata$q5140a=="weeks",7\*COMSAdata$id10250,COMSAdata$id10250))

table(COMSAdata$id10250\_units,COMSAdata$id10250, exclude=NULL)

table(COMSAdata$id10250\_units,COMSAdata$id10250\_a, exclude=NULL)

table(COMSAdata$id10250\_units,COMSAdata$id10250\_b, exclude=NULL)

# illness duration

# COMSAdata$illdur <- NULL

# COMSAdata$illdur <- ifelse(COMSAdata$q2030a=="months",31\*COMSAdata$q2030,

# ifelse(COMSAdata$q2030a=="weeks",7\*COMSAdata$q2030,COMSAdata$q2030))

table(COMSAdata$id10120, exclude = NULL)

table(COMSAdata$illdur, exclude = NULL)

COMSAdata$illdur <- COMSAdata$illdur

COMSAdata$illdur[COMSAdata$illdur==99] <- NA

# EAVA$malnutrition2 <- ifelse(COMSAdata$id10249 %in% "yes" & (COMSAdata$swelldur >= COMSAdata$illdur),1,2)

EAVA$malnutrition2 <- ifelse(COMSAdata$id10249 %in% "yes" & (COMSAdata$id10250 >= COMSAdata$id10120) & !is.na(COMSAdata$id10250) & !is.na(COMSAdata$id10120),1,2)

table(EAVA$malnutrition2, exclude = NULL)

# \* AIDS1;

# \* Expert VA and available in GC13 VA: (Mother with positive HIV test or told by a health worker that she had AIDS) and

# >=2 of the following: (limbs became very thin, protruding belly, swelling in the armpits, a whitish rash inside

# the mouth/on the tongue, more frequent loose/liquid stools than usual >30 days, fever or a skin rash >30 days,

# or fast breathing or chest indrawing);

# if vacq4350=1 then a=1; else a=0; \* thin limbs;

# if vacq4400=1 then b=1; else b=0; \* protruding belly;

# if vacq4420=1 then c=1; else c=0; \* swelling in armpits;

# if vacq4430=1 then d=1; else d=0; \* whitish rash in mouth;

# if vacq4060=1 and ((vacq4080>30 and vacq4090=1) or (vacq4080-vacq4100>30)) and vacq4080^=99 then e1=1; else e1=0; \* liquid stools >30 days;

# if vacq4060=1 and ((vacq4080>14 and vacq4090=1) or (vacq4080-vacq4100>14)) and vacq4080^=99 then e2=1; else e2=0; \* liquid stools 15-30 days;

# if vacq4010=1 and vacq4020>30 and vacq4020^=99 then f1=1; else f1=0; \* fever >30 days;

# if vacq4010=1 and vacq4020>14 and vacq4020^=99 then f2=1; else f2=0; \* fever >14 days;

# if vacq4300=1 and vacq4330>30 and vacq4330^=99 then g1=1; else g1=0; \* skin rash >30 days;

# if vacq4300=1 and vacq4330>14 and vacq4330^=99 then g2=1; else g2=0; \* skin rash >14 days;

# if vacq4180=1 then h=1; else h=0; \* fast breathing;

# if vacq4200=1 then i=1; else i=0; \* chest indrawing;

# J=a+b+c+d+e1+f1+g1+h+i;

# K=a+b+e1+f1+g1+h+i;

# L=a+b+e2+f2+g2+h+i;

# if (vacq5180=1 or vacq5190=1) and J>1 then AIDS1=1; \* (Mother w/positive HIV test or told she had AIDS) + J>1;

# else AIDS1=2;

COMSAdata$a <- ifelse(COMSAdata$id10244 %in% "yes",1,0) # thin limbs

COMSAdata$b <- ifelse(COMSAdata$id10200 %in% "yes",1,0) # protruding belly

COMSAdata$c <- ifelse(COMSAdata$id10256 %in% "yes",1,0) # swelling in armpits

COMSAdata$d <- ifelse(COMSAdata$id10245 %in% "yes",1,0) # whitish rash in mouth

# COMSAdata$e1 <- ifelse(COMSAdata$id10181==1 & ((COMSAdata$id10182>30 & COMSAdata$id10185==1) | (COMSAdata$id10182-COMSAdata$q5052>30)) & COMSAdata$id10182!=99 ,1,0) # liquid stools >30 days

# COMSAdata$e2 <- ifelse(COMSAdata$id10181==1 & ((COMSAdata$id10182>14 & COMSAdata$id10185==1) | (COMSAdata$id10182-COMSAdata$q5052>14)) & COMSAdata$id10182!=99 ,1,0) # liquid stools 15-30 days

table(COMSAdata$id10185)

COMSAdata$e1 <- ifelse(COMSAdata$id10181 %in% "yes" & COMSAdata$id10182>30 & COMSAdata$id10182!=99 & !is.na(COMSAdata$id10182),1,0) # liquid stools >30 days

COMSAdata$e2 <- ifelse(COMSAdata$id10181 %in% "yes" & COMSAdata$id10182>14 & COMSAdata$id10182!=99 & !is.na(COMSAdata$id10182),1,0) # liquid stools 15-30 days

COMSAdata$f1 <- ifelse(COMSAdata$id10147 %in% "yes" & COMSAdata$id10148>30 & COMSAdata$id10148!=99 & !is.na(COMSAdata$id10148),1,0) # fever >30 days

COMSAdata$f2 <- ifelse(COMSAdata$id10147 %in% "yes" & COMSAdata$id10148>14 & COMSAdata$id10148!=99 & !is.na(COMSAdata$id10148),1,0) # fever >14 days

COMSAdata$g1 <- ifelse(COMSAdata$id10233 %in% "yes" & COMSAdata$id10234>30 & COMSAdata$id10234!=99 & !is.na(COMSAdata$id10234),1,0) # skin rash >30 days

COMSAdata$g2 <- ifelse(COMSAdata$id10233 %in% "yes" & COMSAdata$id10234>14 & COMSAdata$id10234!=99 & !is.na(COMSAdata$id10234),1,0) # skin rash >14 days

COMSAdata$h <- ifelse(COMSAdata$id10166 %in% "yes",1,0) # fast breathing

COMSAdata$i <- ifelse(COMSAdata$id10172 %in% "yes",1,0) # chest indrawing

COMSAdata$J=COMSAdata$a+COMSAdata$b+COMSAdata$c+COMSAdata$d+COMSAdata$e1+COMSAdata$f1+COMSAdata$g1+COMSAdata$h+COMSAdata$i

COMSAdata$K=COMSAdata$a+COMSAdata$b+COMSAdata$e1+COMSAdata$f1+COMSAdata$g1+COMSAdata$h+COMSAdata$i

COMSAdata$L=COMSAdata$a+COMSAdata$b+COMSAdata$e2+COMSAdata$f2+COMSAdata$g2+COMSAdata$h+COMSAdata$i

EAVA$AIDS1 <- ifelse(COMSAdata$id10127 %in% "yes" | COMSAdata$id10446 %in% "yes" & COMSAdata$J>1,1,2) # (Mother w/positive HIV test or told she had AIDS) + J>1

table(EAVA$AIDS1, exclude = NULL)

# \* Updated AIDS4 (different than Kalter, Perin, Black validation paper algorithm): includes vacq5165 as a sole required criterion);

# \* |++| AIDS4 |++| same as AIDS5 except, in addition to using both maternal HIV test or AIDS dx OR if both these are

# 'Refuse to answer' or 'Don't know', it also accepts a negative HIV test with a missing/DK/Refuse response for AIDS dx, or a negative AIDS dx with a missing/DK/Refuse response for HIV test (i.e., full info is not available for HIV test results and AIDS dx even though one of these variables is with a negative response), and then looks for swelling in the armpits or a whitish rash inside the mouth/on the tongue;

# \* Expert VA and available in GC13 VA: (Mother with positive HIV test or told by a health worker that she had AIDS)

# OR if both are 'refuse to answer' or 'DK' then (swelling in the armpits or a whitish rash inside the mouth/on

# the tongue) AND >=3 of the following: (limbs became very thin, protruding belly, more frequent loose/liquid stools than usual >30 days, fever or a skin rash >30 days, fast breathing, or chest indrawing);

# if (vacq5165=1 or ((vacq5190=1 or

# (vacq5190 in (.,8,9) and (vacq4420=1 or vacq4430=1))) and K>2))

# then AIDS4=1; \* child with HIV/AIDS before the fatal illness OR (+ maternal HIV test or Dx) or [if both Refuse to Answer or DK] then (swelling in armpits or whitish rash in mouth) + K>2;

# else AIDS4=2;

# EAVA$AIDS4 <- ifelse((COMSAdata$id10126=="yes" | COMSAdata$id10127=="yes") |

# ((COMSAdata$id10126 == c(".","dk","no","ref") & COMSAdata$id10127 == c(".","dk","no","ref")) |

# (COMSAdata$id10126 == c(".","dk","no","ref") & COMSAdata$id10127=="no") |

# (COMSAdata$id10126=="no") & COMSAdata$id10127 == c(".","dk","no","ref")) &

# (COMSAdata$id10256=="yes" | COMSAdata$id10245=="yes") & COMSAdata$K>2,1,2)

EAVA$AIDS4 <- ifelse(COMSAdata$id10127 %in% "yes" | # positive dx by HCS

(COMSAdata$id10446 %in% "yes" | (COMSAdata$id10446 %in% c("dk","ref") & #

(COMSAdata$id10256 %in% "yes" | COMSAdata$id10245 %in% "yes")) & COMSAdata$K)>2,1,2)

# child with HIV/AIDS before the fatal illness OR (+ maternal HIV test or Dx) or [if both Refuse to Answer or DK] then (swelling in armpits or whitish rash in mouth) + K>2;

table(EAVA$AIDS4,exclude = NULL)

# \* |++| Diarrhea8 |++| (acute diarrhea with >4 stools on worst day, or diarrhea >14 days [both without blood]);

# \* Expert VA and available in GC13 VA: more frequent loose or liquid stools than usual and >=5 stools on day with most

# stools and no blood in stools OR more frequent loose or liquid stools than usual >14 days and no blood in stools;

# if (vacq4060=1 and vacq4070>4 and vacq4070^=99 and vacq4110=2) or (vacq4060=1 and diardur>14 and vacq4110=2)

# then diarrhea8=1;

# else diarrhea8=2;

# table(COMSAdata$id10182)

# table(COMSAdata$id10185)

# table(COMSAdata$q5052)

# COMSAdata$diardur <- ifelse(COMSAdata$id10185=="yes",COMSAdata$id10182,COMSAdata$id10182-COMSAdata$q5052)

# COMSAdata$diardur <- COMSAdata$id10182

table(COMSAdata$id10182)

table(COMSAdata$id10184)

table(COMSAdata$id10184\_units, COMSAdata$id10184\_a)

table(COMSAdata$id10184\_units, COMSAdata$id10184\_b)

table(COMSAdata$id10184\_a, COMSAdata$id10184\_b, exclude = NULL)

table(COMSAdata$id10182, COMSAdata$id10184\_b, exclude=NULL)

COMSAdata$diardur <- COMSAdata$id10184\_b

COMSAdata$diardur[COMSAdata$diardur %in% c(99,9999)] <- NA

EAVA$diarrhea8 <- ifelse((COMSAdata$id10181 %in% "yes" & COMSAdata$id10183>4 & COMSAdata$id10183 %!in% 99 & !is.na(COMSAdata$id10183) & COMSAdata$id10186 %in% "no") |

(COMSAdata$id10181 %in% "yes" & COMSAdata$diardur>14 & COMSAdata$diardur %!in% 99 & !is.na(COMSAdata$diardur) & COMSAdata$id10186 %in% "no"),1,2)

table(EAVA$diarrhea8, exclude=NULL)

# \* |++| Dysentery8 |++| (bloody diarrhea with >4 stools on worst day, or bloody diarrhea >14 days);

# \* Expert VA and available in GC13 VA: more frequent loose or liquid stools than usual and >4 stools on day with most

# stools and blood in stools OR more frequent loose or liquid stools than usual >14 days and blood in stools;

# if (vacq4060=1 and vacq4070>4 and vacq4070^=99 and vacq4110=1) or (vacq4060=1 and diardur>14 and vacq4110=1)

# then dysentery8=1;

# else dysentery8=2;

EAVA$dysentery8 <- ifelse((COMSAdata$id10181 %in% "yes" & COMSAdata$id10183>4 & COMSAdata$id10183 %!in% 99 & !is.na(COMSAdata$id10183) & COMSAdata$id10186 %in% "yes") |

(COMSAdata$id10181 %in% "yes" & COMSAdata$diardur>14 & !is.na(COMSAdata$diardur) & COMSAdata$id10186 %in% "yes"), 1,2)

table(EAVA$dysentery8, exclude=NULL)

# if diarrhea8=1 or dysentery8=1 then diardysn8=1;

# else diardysn8=2;

EAVA$diardysn8 <- ifelse(EAVA$diarrhea8 %in% 1 | EAVA$dysentery8 %in% 1,1,2)

table(EAVA$diardysn8, exclude=NULL)

# \* |++| possible diarrhea |++|;

# \* Expert VA and available in GC13 VA: loose or liquid stools and (fever or convulsions or unconscious up till death) and no blood in the loose/liquid stools and no diarrhea8;

# if vacq4060=1 and (vacq4010=1 or vacq4250=1 or vacq4270 in (1, 2, 3)) and vacq4110=2 and diarrhea8=2 then possiblediar8\_4=1;

# else possiblediar8\_4=2;

EAVA$possiblediar8\_4 <- ifelse(COMSAdata$id10181 %in% "yes" & (COMSAdata$id10147 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes") & COMSAdata$id10186 %in% "no" & EAVA$diarrhea8 %in% 2, 1,2)

table(EAVA$possiblediar8\_4, exclude=NULL)

# \* |++| possible dysentery |++|;

# \* Expert VA and available in GC13 VA: loose or liquid stools and (fever or convulsions or unconscious up till death) and blood in the loose/liquid stools and no dysentery8;

# if vacq4060=1 and (vacq4010=1 or vacq4250=1 or vacq4270 in (1, 2, 3)) and vacq4110=1 and dysentery8=2 then possibledysn8\_4=1;

# else possibledysn8\_4=2;

EAVA$possibledysn8\_4 <- ifelse(COMSAdata$id10181 %in% "yes" & (COMSAdata$id10147 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes") & COMSAdata$id10186 %in% "yes" & EAVA$dysentery8 %in% 2, 1,2)

table(EAVA$possibledysn8\_4, exclude=NULL)

# \* |++| possible diarrhea/dysentery |++|;

# if possiblediar8\_4=1 or possibledysn8\_4=1 then possdiardysn8\_4=1;

# else possdiardysn8\_4=2;

EAVA$possdiardysn8\_4 <- ifelse(EAVA$possiblediar8\_4 %in% 1 | EAVA$possibledysn8\_4 %in% 1, 1,2)

table(EAVA$possdiardysn8\_4, exclude=NULL)

# \* Hemorrhagic fever;

# \* Expert VA and avaiable in GC13 VA: Fever and either (bled from anywhere, or had areas of the skin that turned black);

# if vacq4010=1 and (vacq4440=1 or vacq4460=1) then hemfever=1;

# \* if vacq4010=2 or (vacq4440=2 and vacq4460=2) then hemfever=2;

# else hemfever=2;

EAVA$hemfever <- ifelse(COMSAdata$id10147 %in% "yes" & (COMSAdata$id10241 %in% "yes" | COMSAdata$id10239 %in% "yes"), 1,2)

table(EAVA$hemfever, exclude=NULL)

# \* |++| Malaria251 |++|;

# \* Expert VA and available in GC13 VA: Fever that continued till death and was on and off and no stiff neck and no bulging fontanelle and (pallor, difficult breathing, convulsions, or unconscious till death) OR Fever that continued till death and was severe and no stiff neck and no bulging fontanelle and (pallor, convulsions, or unconscious till death);

# if (((vacq4010=1 and vacq4030=1 and vacq4050=2) and vacq4280=2 and vacq4290=2 and (vacq4410=1 or vacq4160=1 or vacq4250=1 or vacq4270 in (1, 2, 3))) or

# ((vacq4010=1 and vacq4030=1 and vacq4040=3) and vacq4280=2 and vacq4290=2 and (vacq4410=1 or vacq4250=1 or vacq4270 in (1, 2, 3)))) then malaria251=1;

# \* For malaria251 requiring bed net use to be included in the definition, un-edit out the following;

# \* if malaria251=1 and saq5b030 in (2, 3) then malaria251=1;

# else malaria251=2;

EAVA$malaria251 <- ifelse((COMSAdata$id10147 %in% "yes" & COMSAdata$id10149 %in% "yes" & COMSAdata$id10151 %in% "on\_and\_off" & COMSAdata$id10208 %in% "no" & COMSAdata$id10278 %in% "no" & (COMSAdata$id10268 %in% "yes" | COMSAdata$id10159 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes")) |

(COMSAdata$id10147 %in% "yes" & COMSAdata$id10149 %in% "yes" & COMSAdata$id10150 %in% "severe" & COMSAdata$id10208 %in% "no" & COMSAdata$id10278 %in% "no" & (COMSAdata$id10268 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes")), 1,2)

table(EAVA$malaria251)

# \* Measles4;

# \* Expert VA and available in GC13 VA: Child's age >=120 days and rash >=3 days and fever >=3 days and

# rash started on face);

# if ((vacq125n>3 and vacq125n^=99 and vacq125u=2) or vacq125u=3) and (vacq4010=1 and vacq4020>2 and vacq4020^=99) and

# (vacq4300=1 and vacq4330>2 and vacq4330^=99 and vacq4320=1) then measles4=1;

# else measles4=2;

EAVA$measles4 <- ifelse(COMSAdata$age>=120 & (COMSAdata$id10147 %in% "yes" & COMSAdata$id10148>2 & COMSAdata$id10148 %!in% 99 & !is.na(COMSAdata$id10148)) &

(COMSAdata$id10233 %in% "yes" & COMSAdata$id10234>2 & COMSAdata$id10234 != 99 & !is.na(COMSAdata$id10148) & COMSAdata$id10235 %in% c("face", "everywhere")), 1, 2)

table(EAVA$measles4, exclude=NULL)

# \* Meningitis/encephalitis;

# \* Expert VA and available in GC13 VA: Fever AND (stiff neck or bulging fontanelle);

# if vacq4010=1 and (vacq4280=1 or vacq4290=1) then meningitis=1;

# else meningitis=2;

EAVA$meningitis <- ifelse(COMSAdata$id10147 %in% "yes" & (COMSAdata$id10208 %in% "yes" | COMSAdata$id10278 %in% "yes"), 1,2)

table(EAVA$meningitis, exclude=NULL)

# \* |++| Pertussis |++|;

# \* Expert VA and available in GC13 VA: Cough>14 days and either (severe cough, vomited after coughing, or stridor);

# if (vacq4120=1 and vacq4130>14 and vacq4130^=99) and (vacq4140=1 or vacq4150=1 or vacq4220=1) then pertussis=1;

# else pertussis=2;

# EAVA$pertussis <- ifelse((COMSAdata$id10153=="yes" & COMSAdata$id10154>14 & COMSAdata$id10154!=99) & (COMSAdata$id10156=="yes" | COMSAdata$q5075=="yes" | COMSAdata$q5089=="yes"), 1,2)

EAVA$pertussis <- ifelse((COMSAdata$id10153 %in% "yes" & COMSAdata$id10154>14 & COMSAdata$id10154 %!in% 99 & !is.na(COMSAdata$id10154)) & (COMSAdata$id10156 %!in% "yes" | COMSAdata$id10158 %in% "yes" | COMSAdata$id10173 %in% "yes"), 1,2)

table(EAVA$pertussis, exclude=NULL)

# \* |++| VA\_pneumoniafb2daysgr |++|;

# \* Expert VA and available in GC13 VA: (cough>2 days or difficult breathing>2 days) and (fast breathing>2 days or chest indrawing or grunting);

# if ((vacq4120=1 and vacq4130>2) or (vacq4160=1 and vacq4170>2)) and ((vacq4180=1 and vacq4190>2) or vacq4200=1 or vacq4230=1) then pneumoniafb2daysgr=1;

# else pneumoniafb2daysgr=2;

EAVA$pneumoniafb2daysgr <- ifelse(((COMSAdata$id10153 %in% "yes" & COMSAdata$id10154>2 & !is.na(COMSAdata$id10154)) | (COMSAdata$id10159 %in% "yes" & COMSAdata$id10161>2 & !is.na(COMSAdata$id10161))) & ((COMSAdata$id10166 %in% "yes" & COMSAdata$id10167>2 & !is.na(COMSAdata$id10167)) | COMSAdata$id10172 %in% "yes" | COMSAdata$id10173 %in% "yes"), 1,2)

table(EAVA$pneumoniafb2daysgr, exclude = NULL)

# \* |++| possibleari3 |++|;

# \* Expert VA and available in GC13 VA: (Cough or difficult breathing or (fast breathing and (CI, stridor, grunting or wheezing))) and (severe cough or post-tussive vomiting or fast breathing or CI or grunting or stridor or wheezing or fever or convulsions or unconscious up till death) and no Pertussis and no pneumonia;

# if (vacq4120=1 or vacq4160=1 or (vacq4180=1 and (vacq4200=1 or vacq4220=1 or vacq4230=1 or vacq4240=1))) and

# (vacq4140=1 or vacq4150=1 or vacq4180=1 or vacq4200=1 or vacq4230=1 or vacq4220=1 or vacq4240=1 or

# vacq4010=1 or vacq4250=1 or vacq4270 in (1, 2, 3)) and pertussis=2 and pneumoniafb2daysgr=2 then possibleari3=1;

# else possibleari3=2;

# EAVA$possibleari3 <- ifelse(COMSAdata$id10153=="yes" | COMSAdata$id10159=="yes" | (COMSAdata$id10166=="yes" & (COMSAdata$id10172=="yes" | COMSAdata$q5089==1 | COMSAdata$q5090==1 | COMSAdata$q5091==1)) &

# (COMSAdata$id10156==1 | COMSAdata$q5075==1 | COMSAdata$id10166=="yes" | COMSAdata$id10172=="yes" | COMSAdata$q5090==1 | COMSAdata$q5089==1 | COMSAdata$q5091==1 |

# COMSAdata$id10147=="yes" | COMSAdata$id10219=="yes" | COMSAdata$id10216 == c(1, 2, 3)) & EAVA$pertussis==2 & EAVA$pneumoniafb2daysgr==2, 1,2)

EAVA$possibleari3 <- ifelse((COMSAdata$id10153 %in% "yes" | COMSAdata$id10159 %in% "yes" | (COMSAdata$id10166 %in% "yes" & (COMSAdata$id10172 %in% "yes" | COMSAdata$id10173 %in% "yes"))) &

(COMSAdata$id10156 %in% "yes" | COMSAdata$id10166 %in% "yes" | COMSAdata$id10172 %in% "yes" | COMSAdata$id10173 %in% "yes" |

COMSAdata$id10147 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes") & EAVA$pertussis %in% 2 & EAVA$pneumoniafb2daysgr %in% 2, 1,2)

table(EAVA$possibleari3, exclude=NULL)

# \* Sepsis wo/malaria251;

# \* This is sepsis without any localizing signs and without malaria251. The algorithm will be used in a hierarchy in which measles, pneumonia, meningitis and diarrhea will first be identified;

# \* This cause of death is no longer labeled â€˜sepsisâ€™. It is now â€˜other serious infectionâ€™;

# \* Expert VA and available in GC13 VA: Fever and (rash on trunk/abdomen/everywhere or convulsions or unconscious up till death) and no malaria251;

# if vacq4010=1 and (vacq4310 in (2, 4) or vacq4250=1 or vacq4270 in (1, 2, 3)) and malaria251=2 then sepsis\_nomal251=1;

# else sepsis\_nomal251=2;

EAVA$sepsis\_nomal251 <- ifelse(COMSAdata$id10147 %in% "yes" & (COMSAdata$id10235 %in% c("trunk\_or\_abdomen", "everywhere") | COMSAdata$id10219 %in% "yes" | COMSAdata$id10218 %in% "yes") & EAVA$malaria251 %in% 2, 1,2)

table(EAVA$sepsis\_nomal251, exclude=NULL)

# \* Residual infection;

# \* Expert VA and available in GC13 VA: fever and no other infectious diagnoses;

# if vacq4010=1 and (AIDS4=2 and measles4=2 and meningitis=2 and malaria251=2 and dysentery8=2 and diarrhea8=2 and

# pertussis=2 and pneumoniafb2daysgr=2 and sepsis\_nomal251=2 and possibledysn8\_4=2 and possiblediar8\_4=2 and

# possibleari3=2 and hemfever=2) then residual\_infect\_slide15\_4=1;

# else residual\_infect\_slide15\_4=2;

EAVA$residual\_infect\_slide15\_4 <- ifelse(COMSAdata$id10147 %in% "yes" & (EAVA$AIDS4 %in% 2 & EAVA$measles4 %in% 2 & EAVA$meningitis %in% 2 & EAVA$malaria251 %in% 2 & EAVA$dysentery8 %in% 2 & EAVA$diarrhea8 %in% 2 &

EAVA$pertussis %in% 2 & EAVA$pneumoniafb2daysgr %in% 2 & EAVA$sepsis\_nomal251 %in% 2 & EAVA$possibledysn8\_4 %in% 2 & EAVA$possiblediar8\_4 %in% 2 &

EAVA$possibleari3 %in% 2 & EAVA$hemfever %in% 2), 1,2)

table(EAVA$residual\_infect\_slide15\_4, exclude=NULL)

# \* Possible malaria;

# \* Use when including 'malaria\_possible' (includes SA bed net use criteria) instead of using 'residual\_infect\_slide15\_4' as possible malaria;

# \* Expert VA and available in GC13 VA with SA module: residual infection and slept under a bed net sometimes or never (vs. always slept under a bed net, which disqualifies residual infection as possible malaria);

# if residual\_infect\_slide15\_4=1 and saq5b030 in (2, 3) then malaria\_possible=1;

# else if ((residual\_infect\_slide15\_4=1 and saq5b030=1) or residual\_infect\_slide15\_4=2) then malaria\_possible=2;

# else malaria\_possible=.;

EAVA$malaria\_possible <- ifelse(EAVA$residual\_infect\_slide15\_4 %in% 1 & COMSAdata$c3101 %in% c(2,3), 1,2)

table(EAVA$source, EAVA$malaria\_possible, exclude=NULL)

# \* Injury;

# \* Expert VA and available in GC13 VA: Suffered from MVA, fall, drowning, poisoning, venomous bite/sting, burn,

# violence, other injury;

# if vacq4471=1 or vacq4472=1 or vacq4473=1 or vacq4474=1 or vacq4475=1 or vacq4476=1 or vacq4477=1 or vacq4478=1 then injury=1;

# else injury=2;

# EAVA$injury <- ifelse(COMSAdata$id10079=="yes" | COMSAdata$id10083=="yes" | COMSAdata$id10085=="yes" | COMSAdata$id10084=="yes" | COMSAdata$id10086=="yes" | COMSAdata$id10089=="yes" | COMSAdata$id10090=="yes" |

# COMSAdata$id10091=="yes" | COMSAdata$id10092=="yes" | COMSAdata$id10093=="yes" | COMSAdata$id10094=="yes" | COMSAdata$id10095=="yes" | COMSAdata$id10096=="yes" , 1,2)

EAVA$injury <- ifelse(COMSAdata$id10077 %in% "yes",1,2)

table(EAVA$injury, exclude=NULL)

# \* |++| Injury3\_slide15 |++| (modified algorithm: when SA data on the order of symptoms appearance are not available);

# \* Expert VA and available in GC13 VA: (Suffered from MVA, fall, drowning, poisoning, venomous bite/sting, burn,

# violence or other injury AND death <=1 day after the injury and the illness lasted <=1 day) OR (injury and

# no other VA diagnosis [except malnutrition allowed]) OR (injury and fever that lasted fewer days than the injury)

# (injury that was the first illness symptom and also had VA sepsis or fever) (the last criterion requires using the SA data on the order of the symptoms)

# (added 'illness lasted <=1 day' to 'survived <=1 day [and changed this from <24 hours]' because child could be injured well into the course of another illness from which s/he died and could die within 24 hours of the injury, but the death could be unrelated to the injury, while if the total illness duration was <=1 day then it's more likely that the injury caused the death);

# if ((injury=1 and vacq121u=1 and vacq121n<=1 and (vacq449u=1 or (vacq449u=2 and vacq449n<=1))) or (injury=1 and AIDS4^=1 and measles4^=1 and meningitis^=1 and dysentery8^=1 and diarrhea8^=1 and pneumoniafb2daysgr^=1 and malaria251^=1 and possibledysn8\_4^=1 and possiblediar8\_4^=1 and possibleari3^=1 and hemfever^=1 and sepsis\_nomal251^=1 and residual\_infect\_slide15\_4^=1) or (injury=1 and vacq4010=1 and vacq449u=2 and vacq4020 < vacq449n)) then injury3\_slide15\_4=1;

# else injury3\_slide15\_4=2;

# EAVA$injury3\_slide15\_4 <- ifelse((EAVA$injury==1 & COMSAdata$id10120=="days" & COMSAdata$illdur<=1 & (COMSAdata$q5403a=="hours" | (COMSAdata$q5403a=="days" & COMSAdata$q5403<=1))) |

# (EAVA$injury==1 & EAVA$AIDS4!=1 & EAVA$measles4!=1 & EAVA$meningitis!=1 & EAVA$dysentery8!=1 & EAVA$diarrhea8!=1 & EAVA$pneumoniafb2daysgr!=1 & EAVA$malaria251!=1 &

# EAVA$possibledysn8\_4!=1 & EAVA$possiblediar8\_4!=1 & EAVA$possibleari3!=1 & EAVA$hemfever!=1 & EAVA$sepsis\_nomal251!=1 & EAVA$residual\_infect\_slide15\_4!=1) |

# (EAVA$injury==1 & COMSAdata$id10147==1 & COMSAdata$q5403a=="days" & COMSAdata$id10148 < COMSAdata$q5403), 1,2)

EAVA$injury3\_slide15\_4 <- ifelse((EAVA$injury %in% 1 & COMSAdata$id10120<=1) |

(EAVA$injury %in% 1 & EAVA$AIDS4 %!in% 1 & EAVA$measles4 %!in% 1 & EAVA$meningitis %!in% 1 & EAVA$dysentery8 %!in% 1 & EAVA$diarrhea8 %!in% 1 & EAVA$pneumoniafb2daysgr %!in% 1 & EAVA$malaria251 %!in% 1 &

EAVA$possibledysn8\_4 %!in% 1 & EAVA$possiblediar8\_4 %!in% 1 & EAVA$possibleari3 %!in% 1 & EAVA$hemfever %!in% 1 & EAVA$sepsis\_nomal251 %!in% 1 & EAVA$residual\_infect\_slide15\_4 %!in% 1) |

(EAVA$injury %in% 1 & COMSAdata$id10147 %in% "yes"), 1,2)

table(EAVA$injury3\_slide15\_4, exclude=NULL)

table(EAVA$injury, EAVA$injury3\_slide15\_4, exclude=NULL)

#################################################### ADD NN CAUSES

# | + congenital malformation2 + | ;

# Expert VA: gross malformation present at birth;

# The â€œother abnormalitiesâ€ must be customized based on inspection of the dataset;

# Available in GC13 VA: head size very small or very large at birth or mass defect on back of head or spine or other abnormality;

# As an example, the following is from Niger;

# if (vacq3031=1 or vacq3032=1 or vacq3033=1 or vacq3034other="PIEDS ARQUES ET BOSSE AU LOMBA" or

# vacq3034other="FENTE SUR LA TETE" or vacq3034other="SPINA BIFIDA" or

# vacq3034other="L ESTOMAC ET LES INTESTINS") then congmalf2=1;else congmalf2=2;

# EAVA$congmalf2 <- ifelse(COMSAdata$id10373=="yes" | COMSAdata$id10372=="yes" | COMSAdata$id10371=="yes", 1,2) # vacq3034other="PIEDS ARQUES ET BOSSE AU LOMBA" or vacq3034other="FENTE SUR LA TETE" or vacq3034other="SPINA BIFIDA" or vacq3034other="L ESTOMAC ET LES INTESTINS") ,1,2)

EAVA$congmalf2 <- ifelse(COMSAdata$id10373 %in% "yes" | COMSAdata$id10372 %in% "yes" | COMSAdata$id10371 %in% "yes" | COMSAdata$id10370 %in% "yes", 1,2) # vacq3034other="PIEDS ARQUES ET BOSSE AU LOMBA" or vacq3034other="FENTE SUR LA TETE" or vacq3034other="SPINA BIFIDA" or vacq3034other="L ESTOMAC ET LES INTESTINS") ,1,2)

table(EAVA$congmalf2, exclude = NULL)

# |++| birth injury |++| (neonates alone - i.e., not with stillbirths) ;

# Expert VA and available in GC13 VA: Bruises or signs of injury on the baby's body at birth (vacq3010=live births);

# if vacq3010=1 then bi5=1;else bi5=2;

EAVA$bi5 <- ifelse(COMSAdata$id10115 %in% "yes", 1,2)

table(EAVA$bi5, exclude = NULL)

# |++| birth asphyxia5 |++|;

# Available in GC13 VA: (NRD [neonatal respiratory depression]: did not cry within 5 minutes after birth or did not

# breathe immediately after birth) + ((NE [neonatal encephalopathy]: not able to suckle normally in the first day

# of life or convulsions/spasms or lethargy) or 0 days old at death);

# Incorporates ba5 and possibleba5 (because 'possibleba5 actually describes BA with death in the first day of

# life, rather than possible BA);

# EAVA$ba5 <- ifelse(vacq3080 in (2, 3, 4) or vacq3040=2) and ((vacq3110=2 or vacq3250=1 or vacq3320=1) or vacq125n=0),1,2)

# EAVA$ba5 <- ifelse((COMSAdata$id10106 %!in% c("within\_5\_minutes") | COMSAdata$id10111 %in% "no") & ((COMSAdata$id10271 %in% "no" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10286 %in% "yes") | COMSAdata$age %in% 0) ,1,2)

# table(EAVA$ba5, exclude = NULL)

# COMSAdata$test1 <- ifelse(COMSAdata$age %in% 0, 1, 0)

# COMSAdata$test2 <- ifelse(COMSAdata$age == 0, 1, 0)

# COMSAdata$test1 <- ifelse(COMSAdata$id10106 %in% c(6:90), 1, 0)

# COMSAdata$test2 <- ifelse(COMSAdata$id10106 %!in% c(0:5), 1, 0)

table(COMSAdata$id10219, COMSAdata$id10275)

table(COMSAdata$id10219, COMSAdata$id10276)

EAVA$ba5 <- ifelse((COMSAdata$id10106 %in% c(6:90) | COMSAdata$id10111 %in% "no") & ((COMSAdata$id10271 %in% "no" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10275 %in% "yes" | COMSAdata$id10276 %in% "yes" | COMSAdata$id10286 %in% "yes") | COMSAdata$age %in% 0) ,1,2)

table(EAVA$ba5, exclude = NULL)

EAVA$bi5ba5 <- ifelse(EAVA$bi5 %in% 1 | EAVA$ba5 %in% 1, 1,2)

table(EAVA$bi5ba5, exclude = NULL)

# fast breathing

table(COMSAdata$age, COMSAdata$id10167)

COMSAdata$fb <- COMSAdata$age - COMSAdata$id10167

table(COMSAdata$fb, exclude = NULL)

# \* pregdur in months; \* months pregnancy is vac2020 in the variable translation

# if vacq2002>0 and vacq2002<11 then pregdur=vacq2002;

# if vacq2002>10 and vacq2002<301 then pregdur=vacq2002/30;

# if vacq2002=99 then pregdur=.;

# COMSAdata$pregdur <- COMSAdata$id10367

# \* |++| preterm\_all (in months) |++|;

# \* In Tanzania we also asked for weeks, but there were several with missing data;

# \* Does not use 'pregnancy ended early', i.e., uses just pregnancy duration in months;

# \* Expert VA and available in GC13 VA: (Pregnancy <9 months with RDS or pregnancy <8 months);

# \* RDS = (fast breathing starting on day 0 and no fever and no cold to touch);

# EAVA$preterm\_all\_mo <- ifelse((pregdur<9 and pregdur^=.) and (vacq3200=1 and vacq3210=0 and vacq3260=2 and vacq3290=2)) OR (pregdur<8 and pregdur^=.), 1, 2)

# EAVA$preterm\_all\_mo <- ifelse((EAVA$pregdur<9 & EAVA$pregdur %!in% "." & COMSAdata$id10166 %in% "yes" & COMSAdata$fb %in% 0 & COMSAdata$id10147 %in% "no" & COMSAdata$id10284 %in% "no") | (EAVA$pregdur<8 & EAVA$pregdur %!in% "."), 1, 2)

EAVA$preterm\_all\_mo <- ifelse((COMSAdata$id10367<9 & !is.na(COMSAdata$id10367) & COMSAdata$id10166 %in% "yes" & COMSAdata$fb == 0 & COMSAdata$id10147 %in% "no" & COMSAdata$id10284 %in% "no") | (COMSAdata$id10367<8 & !is.na(COMSAdata$id10367)), 1, 2)

table(EAVA$preterm\_all\_mo, exclude = NULL)

table(COMSAdata$id10367, exclude = NULL)

# Children (hierarchy)

# if injury3\_slide15\_4=1 then allexpertdxs="Injury";

# else if AIDS4=1 then allexpertdxs="AIDS";

# else if malnutrition2=1 then allexpertdxs="Malnutrition";

# else if measles4=1 then allexpertdxs="Measles";

# else if meningitis=1 then allexpertdxs="Meningitis/Encephalitis";

# else if diardysn8=1 then allexpertdxs="Diarrhea/Dysentery";

# else if pertussis=1 then allexpertdxs="Other infectious";

# else if pneumoniafb2daysgr=1 then allexpertdxs="Pneumonia";

# else if malaria251=1 then allexpertdxs="Malaria";

# else if possdiardysn8\_4=1 then allexpertdxs="Diarrhea/Dysentery";

# else if possibleari3=1 then allexpertdxs="Pneumonia";

# else if hemfever=1 then allexpertdxs="Other infections";

# else if sepsis\_nomal251=1 then allexpertdxs="Other infections";

# else if malaria\_possible=1 then allexpertdxs="Malaria";

# else if residual\_infect\_slide15\_4=1 then allexpertdxs="Other infections";

# else if malnutrition1=1 then allexpertdxs="Malnutrition";

# else allexpertdxs="Unspecified";

##############################################################################################################################

table(EAVA$source)

class(EAVA$source)

head(EAVA)

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & (EAVA$ba5==1 | EAVA$bi5==1)] <- "Intrapartum"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$congmalf2==1] <- "Malformation"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$injury3\_slide15\_4==1] <- "Injury"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$AIDS4==1] <- "AIDS"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$malnutrition2==1] <- "Malnutrition"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$measles4==1] <- "Measles"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$meningitis==1] <- "Meningitis/Encephalitis"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$diardysn8==1] <- "Diarrhea/Dysentery"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$pertussis==1] <- "Other infections"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$pneumoniafb2daysgr==1] <- "Pneumonia"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$malaria251==1] <- "Malaria"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$possdiardysn8\_4==1] <- "Diarrhea/Dysentery"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$possibleari3==1] <- "Pneumonia"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$hemfever==1] <- "Other infections"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$sepsis\_nomal251==1] <- "Other infections"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$malaria\_possible==1] <- "Malaria"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$residual\_infect\_slide15\_4==1] <- "Other infections"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$malnutrition1==1] <- "Malnutrition"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs) & EAVA$preterm\_all\_mo==1] <- "Preterm"

EAVA$allexpertdxs[is.na(EAVA$allexpertdxs)] <- "Unspecified"

table(EAVA$allexpertdxs, exclude = NULL)

EAVA$age <- COMSAdata$age

write.csv(EAVA,file.path(file,"/Data/eava\_child\_comsa.csv"), row.names = FALSE)