# Last edited: 20 Apr 2020

# Last run: 26 Apr 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)

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

file <- getwd()

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

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

dim(COMSAdata)

# 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)

varlist <- c("nnt1", "congmalf2", "bi5", "ba5", "preterm\_rds\_mo", "pregdur", "preterm\_all\_mo", "meningitis451", "meningitis451\_nonnt1", "diarrhea8","pneumo157sign1","pneumo157sign2",

"pneumo157sign3","pneumo157sign4","pneumo157sign5","pneumo157signs","pneumonia157", "sepsisfvr","sepsisfvr2\_2", "sepsisfvr2\_2\_nonnt1","sepsisfvr2sign1","sepsisfvr2sign3",

"sepsisfvr2sign5","sepsisfvr2sign7","sepsisfvr2sign9","sepsisfvr2sign10","sepsisfvr2sign12","possiblediar8\_8","possiblepneumonia9",

"jaundice2","hemorrhageNN","suid","allexpertdxs1","allexpertdxs2")

# age at death

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$age\_neonate\_days) & !is.na(COMSAdata$isneonatal), COMSAdata$age\_neonate\_days, COMSAdata$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)

COMSAdata$age <- COMSAdata$ageatdeath

COMSAdata.neonatal <- subset(COMSAdata, age<28)

COMSAdata <- COMSAdata.neonatal

dim(COMSAdata)

head(COMSAdata)

COMSAdata$ID <- COMSAdata$comsa\_id

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

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

names(EAVA) <- varlist

EAVA <- cbind(COMSAdata$comsa\_id,EAVA,stringsAsFactors = FALSE)

head(EAVA)

names(EAVA)[names(EAVA) == 'COMSAdata$comsa\_id'] <- 'ID'

### Neonates - Algorithms

# |++| neonatal tetanus1 |++|;

# Expert VA and available in GC13 VA: Age 3-27 days at death and convulsions or spasms and either ((able to suckle normally during the first day of life and stopped being able to suckle normally) or (cried within 5 minutes after birth and stopped being able to cry));

# if vacq125n>2 and vacq125n^=99 and vacq3250=1 and ((vacq3110=1 and vacq3130=1) or ((vacq3070=1 or vacq3080=1) and vacq3090=1)) then nnt1=1;

# else nnt1=2;

# EAVA$nnt1 <- ifelse(COMSAdata$age > 2 & COMSAdata$age %!in% 99 & COMSAdata$id10219 %in% "yes" & ((COMSAdata$id10271 %in% "yes" & COMSAdata$id10273 %in% "yes") | (COMSAdata$id10106 %in% c(0:5) & COMSAdata$id10107 %in% "yes")) ,1,2)

# table(EAVA$nnt1, exclude = NULL)

EAVA$nnt1 <- ifelse(COMSAdata$age > 2 & COMSAdata$age %!in% 99 & COMSAdata$id10219 %in% "yes" & ((COMSAdata$id10271 %in% "yes" & COMSAdata$id10273 %in% "yes") | (COMSAdata$id10106 %in% c(0:5) & COMSAdata$id10107 %in% "yes")),1,2)

table(EAVA$nnt1, exclude = NULL)

# | + congenital malformation2 + | ;

# Expert VA: gross malformation present at birth;

# The “other abnormalities” 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)

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)

# \* |++| preterm with rds (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 and (fast breathing starting

# on day 0 and no fever and no cold to touch);

# EAVA$preterm\_rds\_mo <- ifelse(vacq2020<9 and vacq2020^=.) and (vacq3200=1 and vacq3210=0 and vacq3260=2 and vacq3290=2),1,2)

# fast breathing

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

table(COMSAdata$fb, exclude = NULL)

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

table(EAVA$preterm\_rds\_mo, 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)

# \* |++| meningitis451 |++| ;

# \* Evidence (Kenya study) shows that fever but not hypothermia is strongly associated with meningitis;

# \* Expert VA and available in GC13 VA: Fever and (bulging fontanelle or convulsions) and

# (lethargic or unresponsive/unconscious);

# if vacq3260=1 and (vacq3340=1 or vacq3250=1) and (vacq3320=1 or vacq3330=1) then meningitis451=1;

# else meningitis451=2;

# EAVA$meningitis451 <- ifelse(COMSAdata$id10147 %in% "yes" & (COMSAdata$id10278 %in% "yes" | COMSAdata$id10219 %in% "yes") & (COMSAdata$id10286 %in% "yes" | COMSAdata$id10282 %in% "yes" | COMSAdata$id10283 %in% "yes"), 1,2)

EAVA$meningitis451 <- ifelse(COMSAdata$id10147 %in% "yes" & (COMSAdata$id10278 %in% "yes" | COMSAdata$id10219 %in% "yes" | COMSAdata$id10275 %in% "yes" | COMSAdata$id10276 %in% "yes") & (COMSAdata$id10286 %in% "yes" | COMSAdata$id10281 %in% "yes"), 1,2)

table(EAVA$meningitis451, exclude = NULL)

# \* |++| meniningitis451 without nnt1 |++| (used to avoid unrealistic comorbidity with neonatal tetanus);

# if meningitis451=1 and nnt1=2 then meningitis451\_nonnt1=1;

# else meningitis451\_nonnt1=2;

EAVA$meningitis451\_nonnt1 <- ifelse(EAVA$meningitis451 %in% 1 & EAVA$nnt1 %in% 2, 1, 2)

table(EAVA$meningitis451\_nonnt1, exclude = NULL)

# \* |++| diarrhea8 |++| (for NNs: acute diarrhea with >4 stools on worst day);

# \* Expert VA and available in GC13 VA: More frequent loose or liquid stools than usual and >4 stools

# on the day diarrhea was most frequent;

# if vacq3440=1 and vacq3450>4 and vacq3450^=99 then diarrhea8=1;

EAVA$diarrhea8 <- ifelse(COMSAdata$id10181 %in% "yes" & COMSAdata$id10183>4 & COMSAdata$id10183 %!in% 99 & !is.na(COMSAdata$id10183), 1,2)

table(EAVA$diarrhea8, exclude = NULL)

# \* |++| pneumonia157 |++|;

# \* Expert VA and available in GC13 VA: (fast breathing lasting 1 day or more or difficult breathing lasting 1 day or

# more and lasting until death) and 2 or more of the following signs: (chest indrawing or grunting or no/poor cry);

# \* 'no/poor cry' is also general to sepsis, but the structure of the pneumonia157 algorithm requires that the child had

# the pneumonia-specific signs of either fast breathing or difficult breathing and either chest indrawing or grunting;

# \* It can also be seen that 'pneumonia157' is more specific than 'pneumonia156' and the only difference between the two

# is that 'pneumonia157' uses 'no/poor cry' while 'pneumonia156' uses 'poor suckle';

# if vacq3200=1 then pneumo157sign1=1; else pneumo157sign1=0; # fast breathing;

# if vacq3230=1 then pneumo157sign2=1; else pneumo157sign2=0; # CI;

# if vacq3240=1 then pneumo157sign3=1; else pneumo157sign3=0; # grunting;

# if vacq3110=2 or vacq3130=1 then pneumo157sign4=1; else pneumo157sign4=0; # not able to suckle on day 1 or stopped suckling;

# if vacq3080=4 or vacq3090=1 then pneumo157sign5=1; else pneumo157sign5=0; # never cried or stopped crying;

# pneumo157signs=pneumo157sign2+pneumo157sign3+pneumo157sign5;

# if vacq3180+vacq3190=vacq125n then dblasttilldeath=1;

# else dblasttilldeath=2;

#

# EAVA$pneumonia157 <- ifelse((vacq3200=1 and vacq3220>0) or (vacq3170=1 and vacq3190>0 and dblasttilldeath=1)) and pneumo157signs>1, 1, 2)

EAVA$pneumo157sign1 <- ifelse(COMSAdata$id10166 %in% "yes", 1,0) # fast breathing;

EAVA$pneumo157sign2 <- ifelse(COMSAdata$id10172 %in% "yes", 1,0) # CI;

EAVA$pneumo157sign3 <- ifelse(COMSAdata$id10173 %in% "yes", 1,0) # grunting;

EAVA$pneumo157sign4 <- ifelse(COMSAdata$id10271 %in% "no" | COMSAdata$id10273 %in% "yes", 1,0) # not able to suckle on day 1 or stopped suckling;

EAVA$pneumo157sign5 <- ifelse(COMSAdata$id10104 %in% "no" | COMSAdata$id10107 %in% "yes", 1,0) # never cried or stopped crying;

EAVA$pneumo157signs <- EAVA$pneumo157sign2 + EAVA$pneumo157sign3 + EAVA$pneumo157sign5

# COMSAdata$dblasttilldeath <- ifelse(COMSAdata$id10161==COMSAdata$age, 1, 2)

EAVA$pneumonia157 <- ifelse((COMSAdata$id10166 %in% "yes" & COMSAdata$id10167>0 & !is.na(COMSAdata$id10167) | COMSAdata$id10159 %in% "yes" & COMSAdata$id10161>0 & !is.na(COMSAdata$id10161)) & EAVA$pneumo157signs>1, 1, 2)

table(EAVA$pneumonia157, exclude = NULL)

# \* |++| sepsisfvr |++|;

# \* Expert VA and available in GC13 VA: Fever or cold to touch (used as part of sepsisfvr2\_2);

# EAVA$sepsisfvr <- ifelse(vacq3260=1 or vacq3290=1, 1,2)

EAVA$sepsisfvr <- ifelse(COMSAdata$id10147 %in% "yes" | COMSAdata$id10284 %in% "yes", 1,2)

table(EAVA$sepsisfvr, exclude = NULL)

# \* |++| sepsisfvr2\_2 |++| ;

# \* Sepsis 2 (2 or more of the following signs, i.e., does not require fever or cold to touch, instead they become 2 of the possible signs included to accept sepsis as a diagnosis);

# \* The addition of sepsisfvr (as below, i.e., if sepsisfvr=1 or sepsisfvr2signs>1) enables sepsis to be diagnosed in the

# presence of fever alone (as well as in the presence of any 2 of the following 8 signs);

# \* sepsisfvr2\_2: Expert VA and available in GC13 VA (took out "never cried," based on low RR in "Clinical signs that predict severe illness in children under age 2 months: a multicentre study. Lancet 2008,371:135-42". This change should eliminate much of the overlap of sepsis with birth asphyxia since ba5 requires "not able to breathe normally at birth" or "not able to cry at birth": 2 or more of the following 8 signs: (fever-3260 or cold to touch-3290), stopped being able to cry-3090, (not able to suckle normally on day1-3110=2 or stopped being able to suckle-3130), spasms/convulsions-3250, vomited everything-3460, (yellow skin-3470 or yellow eyes-3480), (lethargic-3320 or unconscious-3330), (chest indrawing-3230 or grunting-3240);

# \* Also could check on LBI as source of the sepsis: (umbilical pus-3350, umb redness-3360, umb red extend to skin-3370,

# pustules-3380, 3390-ulcers/pits, red/swollen skin-3400==>combine into one variable if any of the 6);

# if vacq3260=1 or vacq3290=1 then sepsisfvr2sign1=1; else sepsisfvr2sign1=0; \* (fever/cold-to-touch);

# if vacq3110=2 or vacq3130=1 then sepsisfvr2sign3=1; else sepsisfvr2sign3=0; \* (no normal day1 suckle/stopped suckle);

# if vacq3250=1 then sepsisfvr2sign5=1; else sepsisfvr2sign5=0; \* (convulsion);

# if vacq3460=1 then sepsisfvr2sign6=1; else sepsisfvr2sign6=0; \* (vomited everything);

# if vacq3470=1 or vacq3480=1 then sepsisfvr2sign7=1; else sepsisfvr2sign7=0; \* (jaundice);

# if vacq3090=1 then sepsisfvr2sign9=1; else sepsisfvr2sign9=0; \* (stopped cry);

# if vacq3320=1 or vacq3330=1 then sepsisfvr2sign10=1; else sepsisfvr2sign10=0; \* (lethargic/unconscious);

# if vacq3230=1 or vacq3240=1 then sepsisfvr2sign12=1; else sepsisfvr2sign12=0; \* (CI/grunt);

# if vacq3350=1 or vacq3360=1 or vacq3370=1 or vacq3380=1 or vacq3390=1 or vacq3400=1 then LBI=1; else LBI=2;

# if LBI=1 then sepsisfvr2sign14=1; else sepsisfvr2sign14=0; \* with jaundice;

# # \* without jaundice (took out of sepsis definition to allow more cases with jaundice to be assigned to NN jaundice

# # without sepsis, which is lower down in the hierarchy);

# # \* with LBI;

# # \* without LBI;

# # o Fever OR cold to touch OR 2 or more of the following 7 signs: (fever OR cold to touch, did not suckle normally

# on the first day of life OR stopped suckling, convulsions, vomited everything, stopped crying, lethargic OR unconscious, chest indrawing OR grunting)

# sepsisfvr2signs=sepsisfvr2sign1+sepsisfvr2sign3+sepsisfvr2sign5+sepsisfvr2sign6+sepsisfvr2sign9+

# sepsisfvr2sign10+sepsisfvr2sign12;

# if sepsisfvr=1 or sepsisfvr2signs>1 then sepsisfvr2\_2=1;

# else sepsisfvr2\_2=2;

EAVA$sepsisfvr2sign1 <- ifelse(COMSAdata$id10147 %in% "yes" | COMSAdata$id10284 %in% "yes", 1,0) # (fever/cold-to-touch);

EAVA$sepsisfvr2sign3 <- ifelse(COMSAdata$id10271 %in% "no" | COMSAdata$id10273 %in% "yes", 1,0) # (no normal day1 suckle/stopped suckle);

EAVA$sepsisfvr2sign5 <- ifelse(COMSAdata$id10219 %in% "yes", 1,0) # (convulsion);

# COMSAdata$sepsisfvr2sign6 <- ifelse(COMSAdata$id10188=="yes", 1,0) # (vomited everything);

EAVA$sepsisfvr2sign7 <- ifelse(COMSAdata$id10289 %in% "yes" | COMSAdata$id10265 %in% "yes", 1, 0) # (jaundice);

EAVA$sepsisfvr2sign9 <- ifelse(COMSAdata$id10107 %in% "yes", 1,0) # (stopped cry);

EAVA$sepsisfvr2sign10 <- ifelse(COMSAdata$id10286 %in% "yes" | COMSAdata$id10215 %in% "yes", 1,0) # (lethargic/unconscious);

EAVA$sepsisfvr2sign12 <- ifelse(COMSAdata$id10172 %in% "yes" | COMSAdata$id10173 %in% "yes", 1,0) # (CI/grunt);

# COMSAdata$LBI <- ifelse(COMSAdata$q5030==1 | COMSAdata$id10287=="yes" | COMSAdata$q5033==1 | COMSAdata$id10227=="yes" | COMSAdata$id10240=="yes", 1,2) # with jaundice;

# COMSAdata$sepsisfvr2sign14 <- ifelse(COMSAdata$LBI==1,1,0)

EAVA$sepsisfvr2signs = EAVA$sepsisfvr2sign1 + EAVA$sepsisfvr2sign3 + EAVA$sepsisfvr2sign5 + EAVA$sepsisfvr2sign7 + EAVA$sepsisfvr2sign9 + EAVA$sepsisfvr2sign10 + EAVA$sepsisfvr2sign12

EAVA$sepsisfvr2\_2 <- ifelse(EAVA$sepsisfvr %in% 1 | EAVA$sepsisfvr2signs>1, 1,2)

table(EAVA$sepsisfvr2\_2, exclude = NULL)

# \* |++| sepsisfvr2\_2 without neonatal tetanus |++|;

# \* Used to avoid unrealistic comorbidity with neonatal tetanus;

# if sepsisfvr2\_2=1 and nnt1=2 then sepsisfvr2\_2\_nonnt1=1;

# else sepsisfvr2\_2\_nonnt1=2;

EAVA$sepsisfvr2\_2\_nonnt1 <- ifelse(EAVA$sepsisfvr2\_2 %in% 1 & EAVA$nnt1 %in% 2, 1, 2)

table(EAVA$sepsisfvr2\_2\_nonnt1, exclude = NULL)

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

# \* more frequent loose or liquid stools than usual AND VA sepsis AND no VA diarrhea;

# if vacq3440=1 and sepsisfvr2\_2=1 and diarrhea8=2 then possiblediar8\_8=1;

# else possiblediar8\_8=2;

EAVA$possiblediar8\_8 <- ifelse(COMSAdata$id10181 %in% "yes" & EAVA$sepsisfvr2\_2 %in% 1 & EAVA$diarrhea8 %in% 2, 1,2)

table(EAVA$possiblediar8\_8, exclude = NULL)

# \* |++| possible pneumonia |++|;

# \* Difficult breathing AND VA sepsis AND No VA pneumonia;

# if vacq3170=1 and sepsisfvr2\_2=1 and pneumonia157=2 then possiblepneumonia9=1;

# else possiblepneumonia9=2;

EAVA$possiblepneumonia9 <- ifelse(COMSAdata$id10159 %in% "yes" & EAVA$sepsisfvr2\_2 %in% 1 & EAVA$pneumonia157 %in% 2, 1, 2)

table(EAVA$possiblepneumonia9, exclude = NULL)

# \* |++| jaundice |++|;

# \* jaundice (yellow skin or eyes) without septicemia, but with signs of severe illness, i.e., = neonatal jaundice;

# \* 'without septicemia' determined by placing this below septicemia in the hierarchy and by requiring no fever or hypothermia;

# \* (yellow skin or eyes) plus (stopped being able to suckle normally or lethargic or unresponsive/unconscious) plus (no fever or hypothermia);

# if (vacq3470=1 or vacq3480=1) and (vacq3130=1 or vacq3320=1 or vacq3330=1) and (vacq3260^=1 and vacq3290^=1) then jaundice2=1;

# else jaundice2=2;

EAVA$jaundice2 <- ifelse((COMSAdata$id10289 %in% "yes" | COMSAdata$id10265 %in% "yes") & (COMSAdata$id10273 %in% "yes" | COMSAdata$id10286 %in% "yes" | COMSAdata$id10215 %in% "yes" | COMSAdata$id10282 %in% "yes" | COMSAdata$id10283=="yes") & (COMSAdata$id10147 %in% "no" & COMSAdata$id10284 %in% "no"), 1, 2)

table(EAVA$jaundice2, exclude = NULL)

# \* |++| hemorrhagic disease of the newborn |++|;

# \* bleeding from anywhere AND no fever or hypothermia;

# if vacq3420=1 and (vacq3260=2 and vacq3290=2) then hemorrhageNN=1;

# else hemorrhageNN=2;

EAVA$hemorrhageNN <- ifelse(COMSAdata$id10241 %in% "yes" & COMSAdata$id10147 %in% "no" & COMSAdata$id10284 %in% "no", 1, 2)

table(EAVA$hemorrhageNN, exclude = NULL)

# \* |++| SUID (sudden unexplained infant death) |++|;

# \* appeared healthy and then died suddenly;

# if vacq3490=1 and vacq3010^=1 and vacq3020^=1 and vacq3040^=2 and vacq3050^=1 and vacq3060^=1 and

# (vacq3070^=2 or vacq3080=1) and vacq3090^=1 and vacq3110^=2 and vacq3130^=1 and vacq3170^=1 and

# vacq3200^=1 and vacq3230^=1 and vacq3240^=1 and vacq3250^=1 and vacq3260^=1 and vacq3290^=1 and

# vacq3320^=1 and vacq3330^=1 and vacq3340^=1 and vacq3350^=1 and vacq3360^=1 and vacq3380^=1 and

# vacq3390^=1 and vacq3400^=1 and vacq3410^=1 and vacq3420^=1 and vacq3440^=1 and vacq3460^=1 and

# vacq3470^=1 and vacq3480^=1 then suid=1;

# else suid=2;

EAVA$suid <- ifelse(COMSAdata$id10290 %in% "yes" & COMSAdata$id10115 %in% "no" & COMSAdata$id10370 %in% "no" & COMSAdata$id10111 %in% "yes" & COMSAdata$id10112 %in% "no" & COMSAdata$id10113 %in% "no" &

(COMSAdata$id10105 %in% "yes" | COMSAdata$id10106 %in% c(0:5)) & COMSAdata$id10107 %in% "no" & COMSAdata$id10271 %in% "yes" & COMSAdata$id10273 %in% "no" & COMSAdata$id10159 %in% "no" &

COMSAdata$id10166 %in% "no" & COMSAdata$id10172 %in% "no" & COMSAdata$id10173 %in% "no" & COMSAdata$id10219 %in% "no" & COMSAdata$id10147 %in% "no" & COMSAdata$id10284 %in% "no" &

COMSAdata$id10286 %in% "no" & COMSAdata$id10215 %in% "no" & COMSAdata$id10282 %in% "no" & COMSAdata$id10283 %in% "no" & COMSAdata$id10278 %in% "no" & COMSAdata$id10287 %in% "no" &

COMSAdata$id10288 (ulcers/pits) %in% "no" & COMSAdata$id10233 %in% "no" & COMSAdata$id10240 %in% "no" & COMSAdata$id10239 %in% "no" & COMSAdata$id10241 %in% "no" & COMSAdata$id10181 %in% "no" & COMSAdata$id10189 %in% "no" &

COMSAdata$id10289 %in% "no" & COMSAdata$id10265 %in% "no", 1, 2)

table(EAVA$suid, exclude = NULL)

# Neonates (hierarchies)

#

# RDS near top of the hierarchy and all other preterm at the bottom (“Kalter” hierarchy in the KPB validation paper: according to ICD-10 rules, and yields a much lower estimate of preterm delivery than the WHO estimates)

# if nnt1=1 then allexperdxs1="NNT";

# else if congmalf2=1 then allexpertdxs="Malformation";

# else if (ba5=1 or bi5=1) then allexpertdxs="Intrapartum";

# else if preterm\_rds\_mo=1 then allexpertdxs="Preterm";

# else if meningitis451\_nonnt1=1 then allexpertdxs="Meningitis";

# else if diarrhea8=1 then allexpertdxs="Diarrhea";

# else if pneumonia157=1 then allexpertdxs="Pneumonia";

# else if possiblediar8\_8=1 then allexpertdxs="Diarrhea";

# else if possiblepneumonia9=1 then allexpertdxs="Pneumonia";

# else if sepsisfvr2\_2\_nonnt1=1 then allexpertdxs="Sepsis”;

# else if jaundice2=1 then allexpertdxs="Jaundice";

# else if hemorrhageNN=1 then allexpertdxs="Hemorrhage";

# else if suid=1 then allexpertdxs="SUID";

# else if preterm\_all\_mo=1 then allexpertdxs="Preterm";

# else allexpertdxs="Unspecified";

# Option 1 - Compromise Hierarchy

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$nnt1==1] <- "NNT"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$congmalf2==1] <- "Malformation"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & (EAVA$ba5==1 | EAVA$bi5==1)] <- "Intrapartum"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$preterm\_all\_mo==1] <- "Preterm"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$meningitis451\_nonnt1==1] <- "Meningitis"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$diarrhea8==1] <- "Diarrhea"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$pneumonia157==1] <- "Pneumonia"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$possiblediar8\_8==1] <- "Diarrhea"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$possiblepneumonia9==1] <- "Pneumonia"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$sepsisfvr2\_2\_nonnt1==1] <- "Sepsis"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$jaundice2==1] <- "Other"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$hemorrhageNN==1] <- "Other"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1) & EAVA$suid==1] <- "Other"

EAVA$allexpertdxs1[is.na(EAVA$allexpertdxs1)] <- "Unspecified"

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

# Option 2 - All preterm at the bottom and preterm with rds at top

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$nnt1==1] <- "NNT"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$congmalf2==1] <- "Malformation"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & (EAVA$ba5==1 | EAVA$bi5==1)] <- "Intrapartum"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$preterm\_rds\_mo==1] <- "Preterm"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$meningitis451\_nonnt1==1] <- "Meningitis"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$diarrhea8==1] <- "Diarrhea"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$pneumonia157==1] <- "Pneumonia"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$possiblediar8\_8==1] <- "Diarrhea"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$possiblepneumonia9==1] <- "Pneumonia"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$sepsisfvr2\_2\_nonnt1==1] <- "Sepsis"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$jaundice2==1] <- "Other"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$hemorrhageNN==1] <- "Other"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$suid==1] <- "Other"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2) & EAVA$preterm\_all\_mo==1] <- "Preterm"

EAVA$allexpertdxs2[is.na(EAVA$allexpertdxs2)] <- "Unspecified"

EAVA$age <- COMSAdata$age

table(EAVA$allexpertdxs1, EAVA$allexpertdxs2)

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