

Anemia and Hb in Shuar

Alicia DeLouize

November 16, 2019

Revisions for first round of Revise and Resubmit to American Journal of Human Biology marked with (RandR2)

Data

```
set.seed(2323)

# Load packages
library(readxl)
library(psych)
library(lme4)

## Loading required package: Matrix

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.0.5      v dplyr   1.0.3
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x ggplot2::%+%( )    masks psych::%+%( )
## x ggplot2::alpha( )  masks psych::alpha( )
## x tidyr::expand( )   masks Matrix::expand( )
## x dplyr::filter( )   masks stats::filter( )
## x dplyr::lag( )      masks stats::lag( )
## x tidyr::pack( )     masks Matrix::pack( )
## x tidyr::unpack( )   masks Matrix::unpack( )

library(jtools)
library(scales)

##
## Attaching package: 'scales'
```

```

## The following object is masked from 'package:purrr':
##
##   discard

## The following object is masked from 'package:readr':
##
##   col_factor

## The following objects are masked from 'package:psych':
##
##   alpha, rescale

library(reshape2)

##
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':
##
##   smiths

library(ggpubr)

## Registered S3 methods overwritten by 'car':
##   method                      from
##   influence.merMod             lme4
##   cooks.distance.influence.merMod lme4
##   dfbeta.influence.merMod      lme4
##   dfbetas.influence.merMod     lme4

library(performance)
library(moments)
library(mgcv)

## Loading required package: nlme

##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
##   collapse

## The following object is masked from 'package:lme4':
##
##   lmList

## This is mgcv 1.8-33. For overview type 'help("mgcv-package")'.

```

```
library(voxel)
```

```
## Loading required package: lmerTest
```

```
##
```

```
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
```

```
##
```

```
##      lmer
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
##      step
```

```
library(lsr)
```

```
library(arsenal)
```

```
##
```

```
## Attaching package: 'arsenal'
```

```
## The following object is masked from 'package:mgcv':
```

```
##
```

```
##      negbin
```

```
## The following object is masked from 'package:scales':
```

```
##
```

```
##      ordinal
```

```
## The following object is masked from 'package:jtools':
```

```
##
```

```
##      %nin%
```

```
# attach data
```

```
load("Hb.RData")
```

Data Manipulation - Create Anemia Groups

```
Hb$AnemiaGroups <- Hb$AgeGroups
```

```
Hb$AnemiaGroups[which(Hb$AnemiaGroups == 3 & Hb$Sex == 0)] <- 5
```

```
Hb$AnemiaGroups[which(Hb$AnemiaGroups == 3 & Hb$Sex == 1 & Hb$Pregnant == 0)] <- 3
```

```
Hb$AnemiaGroups[which(Hb$AnemiaGroups == 3 & Hb$Sex == 1 & Hb$Pregnant == 1)] <- 4
```

```
describe(Hb$AnemiaGroups)
```

```
##      vars      n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1      1 1637  2.2 1.57      2      2.1 1.48   0  5     5 0.59    -0.84 0.04
```

```
describe(Hb$AgeGroups)
```

```
##      vars      n mean  sd median trimmed  mad min max range  skew kurtosis   se
## X1      1 1637 1.86 1.05      2    1.92 1.48   0   3    3 -0.11    -1.5 0.03
```

Data Manipulation - Create Anemia

```
Hb$Anemia <- 0
Hb$Anemia <- ifelse(Hb$AnemiaGroups == 0 & Hb$Hb_AltAdj < 11.0, 1,
ifelse(Hb$AnemiaGroups == 1 & Hb$Hb_AltAdj < 11.5, 1,
ifelse(Hb$AnemiaGroups == 2 & Hb$Hb_AltAdj < 12.0, 1,
ifelse(Hb$AnemiaGroups == 3 & Hb$Hb_AltAdj < 12.0, 1,
ifelse(Hb$AnemiaGroups == 4 & Hb$Hb_AltAdj < 11.0, 1,
ifelse(Hb$AnemiaGroups == 5 & Hb$Hb_AltAdj < 13.0, 1, 0))))))
```

Data Management - Anemia Mild Moderate Severe

```
# Creating Mild Category
Hb$AnemiaMild <- 0
Hb$AnemiaMild <- ifelse(Hb$AnemiaGroups == 0 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 10.00, 1,
ifelse(Hb$AnemiaGroups == 1 & Hb$Hb_AltAdj < 11.5 & Hb$Hb_AltAdj >= 11.00, 1,
ifelse(Hb$AnemiaGroups == 2 & Hb$Hb_AltAdj < 12.0 & Hb$Hb_AltAdj >= 11.00, 1,
ifelse(Hb$AnemiaGroups == 3 & Hb$Hb_AltAdj < 12.0 & Hb$Hb_AltAdj >= 11.00, 1,
ifelse(Hb$AnemiaGroups == 4 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 10.00, 1,
ifelse(Hb$AnemiaGroups == 5 & Hb$Hb_AltAdj < 13.0 & Hb$Hb_AltAdj >= 11.00, 1, 0))))))
```

```
# Creating Moderate Category
Hb$AnemiaModerate <- 0
Hb$AnemiaModerate <- ifelse(Hb$AnemiaGroups == 0 & Hb$Hb_AltAdj < 10.0 & Hb$Hb_AltAdj >= 7.00, 1,
ifelse(Hb$AnemiaGroups == 1 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 8.00, 1,
ifelse(Hb$AnemiaGroups == 2 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 8.00, 1,
ifelse(Hb$AnemiaGroups == 3 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 8.00, 1,
ifelse(Hb$AnemiaGroups == 4 & Hb$Hb_AltAdj < 10.0 & Hb$Hb_AltAdj >= 7.00, 1,
ifelse(Hb$AnemiaGroups == 5 & Hb$Hb_AltAdj < 11.0 & Hb$Hb_AltAdj >= 8.00, 1, 0))))))
```

```
# Creating Severe Category
Hb$AnemiaSevere <- 0
Hb$AnemiaSevere <- ifelse(Hb$AnemiaGroups == 0 & Hb$Hb_AltAdj < 7.0, 1,
ifelse(Hb$AnemiaGroups == 1 & Hb$Hb_AltAdj < 8.0, 1,
ifelse(Hb$AnemiaGroups == 2 & Hb$Hb_AltAdj < 8.0, 1,
ifelse(Hb$AnemiaGroups == 3 & Hb$Hb_AltAdj < 8.0, 1,
ifelse(Hb$AnemiaGroups == 4 & Hb$Hb_AltAdj < 7.0, 1,
ifelse(Hb$AnemiaGroups == 5 & Hb$Hb_AltAdj < 8.0, 1, 0))))))
```

Distributions

```
table(Hb$Anemia)
```

```
##  
##      0      1  
## 1451   172
```

```
skewness(Hb$Anemia, na.rm = T)
```

```
## [1] 2.560192
```

```
kurtosis(Hb$Anemia, na.rm = T)
```

```
## [1] 7.554585
```

```
table(Hb$Sex)
```

```
##  
##      0      1  
## 746   904
```

```
skewness(Hb$Sex, na.rm = T)
```

```
## [1] -0.1923993
```

```
kurtosis(Hb$Sex, na.rm = T)
```

```
## [1] 1.037017
```

```
describe(Hb$Age)
```

```
##   vars    n mean    sd median trimmed mad  min max range skew kurtosis  se  
## X1     1 1637 19.2 16.73     12   16.42 8.9 0.67  86 85.33 1.42     1.37 0.41
```

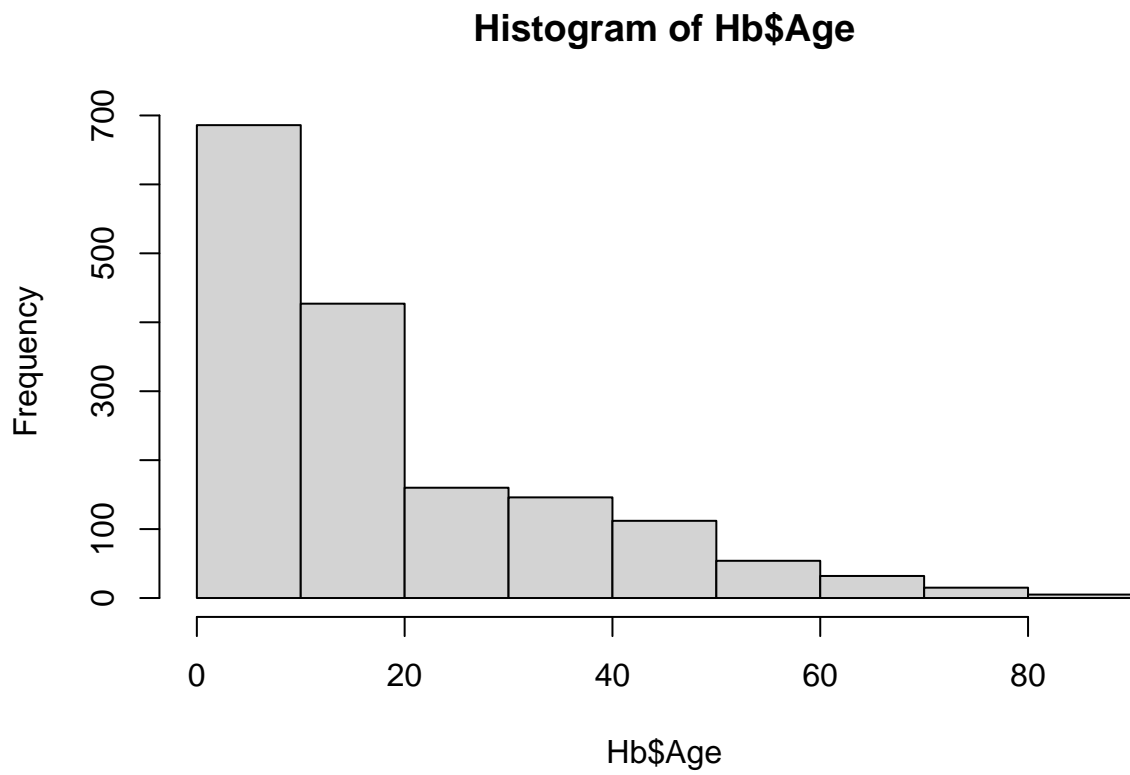
```
skewness(Hb$Age, na.rm = T)
```

```
## [1] 1.417802
```

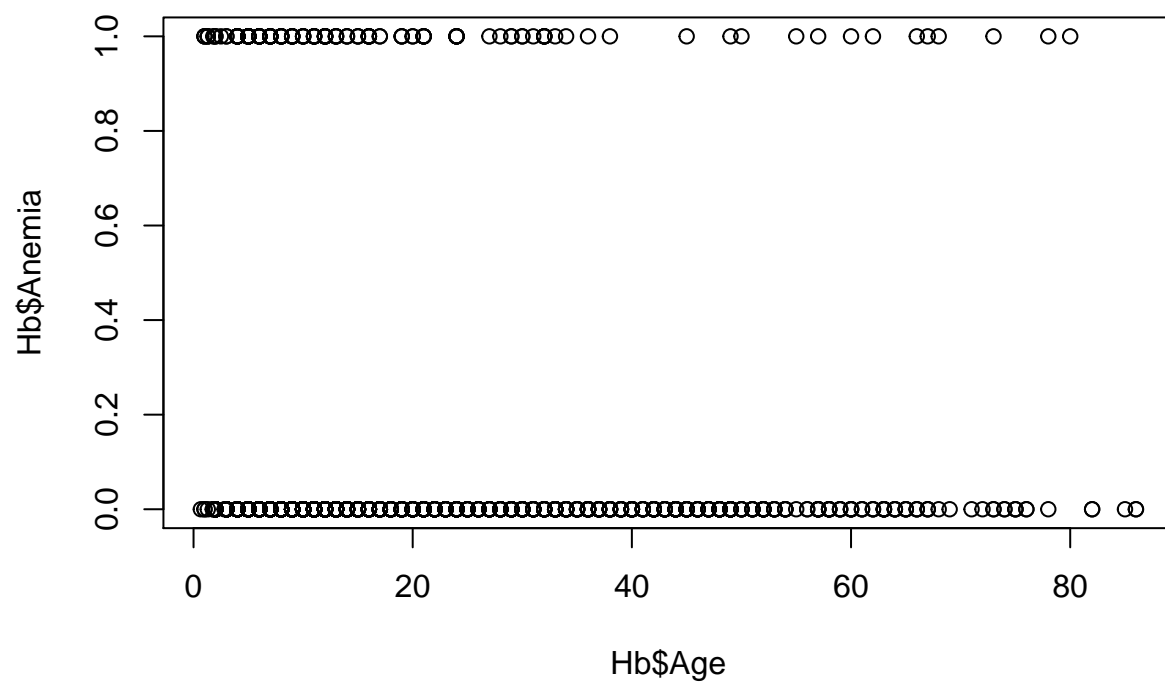
```
kurtosis(Hb$Age, na.rm = T)
```

```
## [1] 4.378749
```

```
hist(Hb$Age)
```



```
plot(Hb$Age, Hb$Anemia)
```



Participants Section

```
#table(Hb$Region, Hb$Community)
min(Hb$Date)
```

```
## [1] "2008-08-26 UTC"
```

```
max(Hb$Date)
```

```
## [1] "2017-11-01 UTC"
```

```
describe(Hb$Age)
```

```
##      vars      n mean    sd median trimmed mad  min max range skew kurtosis   se
## X1      1 1637 19.2 16.73     12   16.42 8.9 0.67  86 85.33 1.42      1.37 0.41
```

```
table(Hb$Sex)
```

```
##
##      0      1
## 746 904
```

```
table(Hb$Shuar, exclude = NULL)
```

```
##
##      2
## 1650
```

```
table(Hb$Region, exclude = NULL)
```

```
##
##      0      1      2
## 1061  588      1
```

```
table(Hb$Sex, Hb$Region)
```

```
##
##           0      1      2
##      0 464 282      0
##      1 597 306      1
```

```
table(Hb$Pregnant)
```

```
##
##      0      1
## 1624    26
```

Descriptives Table 1

```
# Overall Anemia rates
mytable <- table(Hb$Anemia, Hb$AnemiaGroups)
mytable
```

```
##
##           0      1      2      3      4      5
##      0 101 582 146 342  23 251
##      1  26  71  18  40   3  14
```

```
margin.table(mytable, 1)
```

```
##
##      0      1
## 1445  172
```

```
prop.table(mytable, 2)
```

```
##
##           0           1           2           3           4           5
##      0 0.79527559 0.89127106 0.89024390 0.89528796 0.88461538 0.94716981
##      1 0.20472441 0.10872894 0.10975610 0.10471204 0.11538462 0.05283019
```



```
# Mild Anemia rates
```

```
table1 <- table(Hb$AnemiaMild, Hb$AnemiaGroups)
table1
```

```
##
##      0    1    2    3    4    5
##  0 107 613 153 358  24 256
##  1   20  40  11  24   2   9
```

```
margin.table(table1, 1)
```

```
##
##      0    1
## 1511  106
```

```
prop.table(table1, 2)
```

```
##
##      0          1          2          3          4          5
##  0 0.84251969 0.93874426 0.93292683 0.93717277 0.92307692 0.96603774
##  1 0.15748031 0.06125574 0.06707317 0.06282723 0.07692308 0.03396226
```

```
# Moderate Anemia Rates
```

```
table2 <- table(Hb$AnemiaModerate, Hb$AnemiaGroups)
table2
```

```
##
##      0    1    2    3    4    5
##  0 121 622 158 368  25 261
##  1   6  31   6  14   1   4
```

```
margin.table(table2, 1)
```

```
##
##      0    1
## 1555  62
```

```
prop.table(table2, 2)
```

```
##
##      0          1          2          3          4          5
##  0 0.95275591 0.95252680 0.96341463 0.96335079 0.96153846 0.98490566
##  1 0.04724409 0.04747320 0.03658537 0.03664921 0.03846154 0.01509434
```

```
# Severe Anemias Rates
```

```
table3 <- table(Hb$AnemiaSevere, Hb$AnemiaGroups)
table3
```

```
##
##      0   1   2   3   4   5
##  0 127 653 163 380 26 264
##  1   0   0   1   2   0   1
```

```
margin.table(table3, 1)
```

```
##
##      0   1
## 1613   4
```

```
prop.table(table3, 2)
```

```
##
##              0              1              2              3              4              5
##  0 1.000000000 1.000000000 0.993902439 0.994764398 1.000000000 0.996226415
##  1 0.000000000 0.000000000 0.006097561 0.005235602 0.000000000 0.003773585
```

```
# Mean Hb by group
describe.by(Hb$Hb_AltAdj, group = Hb$AnemiaGroups)
```

```
## Warning: describe.by is deprecated. Please use the describeBy function
```

```
##
## Descriptive statistics by group
## group: 0
##   vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1      1 127 11.75 1.03   11.8   11.77 0.89   8 14    6 -0.43    0.65 0.09
## -----
## group: 1
##   vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1      1 653 12.61  1   12.6   12.6 1.04 8.8 18   9.2 0.16    1.24 0.04
## -----
## group: 2
##   vars   n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1      1 164 13.3 1.44  13.45   13.4 1.11  5 17.7 12.7 -1.52    6.82 0.11
## -----
## group: 3
##   vars   n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1      1 382 13.31 1.33   13.4   13.39 1.19  5 16.5 11.5 -1.24    4.79 0.07
## -----
## group: 4
##   vars   n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1      1  26 12.1 1.06   12.25   12.15 0.82 9.5 14.4  4.9 -0.46    0.21 0.21
## -----
## group: 5
##   vars   n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1      1 265 15.08 1.36   15.2   15.16 1.19 7.9 18.9  11 -1.1    3.72 0.08
```

Descriptives Table 2

```
mean(Hb$Hb_AltAdj[which(Hb$AnemiaMild == 1)])
```

```
## [1] 11.27547
```

```
mean(Hb$Hb_AltAdj[which(Hb$AnemiaModerate == 1)])
```

```
## [1] 10.19677
```

```
mean(Hb$Hb_AltAdj[which(Hb$AnemiaSevere == 1)])
```

```
## [1] 6.3
```

```
Hb$Children <- ifelse(Hb$AnemiaGroups == 0 | Hb$AnemiaGroups == 1 | Hb$AnemiaGroups == 2, 1, 0)
Hb$Women <- ifelse(Hb$AnemiaGroups == 3 | Hb$AnemiaGroups == 4, 1, 0)
Hb$Men <- ifelse(Hb$AnemiaGroups == 5, 1, 0)
```

```
mean(Hb$Hb_AltAdj[which(Hb$Children == 1 & Hb$Region == 1)], na.rm=TRUE)
```

```
## [1] 12.50253
```

```
mean(Hb$Hb_AltAdj[which(Hb$Children == 1 & Hb$Region == 0)], na.rm=TRUE)
```

```
## [1] 12.66858
```

```
mean(Hb$Hb_AltAdj[which(Hb$Women == 1 & Hb$Region == 1)], na.rm=TRUE)
```

```
## [1] 13.16144
```

```
mean(Hb$Hb_AltAdj[which(Hb$Women == 1 & Hb$Region == 0)], na.rm=TRUE)
```

```
## [1] 13.27529
```

```
mean(Hb$Hb_AltAdj[which(Hb$AnemiaGroups == 5 & Hb$Region == 1)], na.rm=TRUE)
```

```
## [1] 15.21058
```

```
mean(Hb$Hb_AltAdj[which(Hb$AnemiaGroups == 5 & Hb$Region == 0)], na.rm=TRUE)
```

```
## [1] 14.99814
```

```
table(Hb$Anemia, Hb$Children, Hb$Region )
```

```
## , , = 0
##
##
##      0  1
##  0 381 552
##  1  35  76
##
## , , = 1
##
##
##      0  1
##  0 235 277
##  1  22  39
##
## , , = 2
##
##
##      0  1
##  0  0  0
##  1  0  0
```

```
table(Hb$Anemia, Hb$Women, Hb$Region)
```

```
## , , = 0
##
##
##      0  1
##  0 705 228
##  1  84  27
##
## , , = 1
##
##
##      0  1
##  0 375 137
##  1  45  16
##
## , , = 2
##
##
##      0  1
##  0  0  0
##  1  0  0
```

```
table(Hb$Anemia, Hb$Men, Hb$Region)
```

```
## , , = 0
##
##
##      0  1
##  0 780 153
##  1 103   8
##
```

```
## , , = 1
##
##
##      0  1
##  0 414 98
##  1  55  6
##
## , , = 2
##
##
##      0  1
##  0  0  0
##  1  0  0
```

```
table(Hb$Anemia, Hb$AnemiaGroups, Hb$Region)
```

```
## , , = 0
##
##
##      0  1  2  3  4  5
##  0 64 392 96 216 12 153
##  1 20 44 12 25  2  8
##
## , , = 1
##
##
##      0  1  2  3  4  5
##  0 37 190 50 126 11 98
##  1  6 27  6 15  1  6
##
## , , = 2
##
##
##      0  1  2  3  4  5
##  0  0  0  0  0  0  0
##  1  0  0  0  0  0  0
```

```
table(Hb$Anemia, Hb$Children)
```

```
##
##      0  1
##  0 616 829
##  1  57 115
```

```
table(Hb$Anemia, Hb$Children, Hb$Sex)
```

```
## , , = 0
##
##
##      0  1
##  0 251 402
##  1  14  65
```

```
##
## , , = 1
##
##
##      0    1
##  0 365 427
##  1  43  50
```

```
table(Hb$Anemia, Hb$Women)
```

```
##
##      0    1
##  0 1080 365
##  1  129  43
```

```
table(Hb$Anemia, Hb$Men)
```

```
##
##      0    1
##  0 1194 251
##  1  158  14
```

```
table(Hb$Anemia, Hb$AnemiaGroups)
```

```
##
##      0    1    2    3    4    5
##  0 101 582 146 342  23 251
##  1  26  71  18  40   3  14
```

Inferentials: Linear

```
# Preparing data
Hb$Anemia <- factor(Hb$Anemia)
Hb_glmer <- filter(Hb, !is.na(Age), !is.na(Sex), !is.na(Community), !is.na(Anemia), !is.na(Year))
Hb_glmer$Community <- factor(Hb_glmer$Community)

# Controlling for community and year
## Check to make sure there are enough groups
unique(Hb$Year)
```

```
## [1] 2008 2013 2009 2010 2103 2016 2012 2011 2014 2017
```

```
## Anemia and sex
```

```
## Adults 15+
```

```
logit_sex_1 <- glmer(formula = Anemia ~ Sex + Pregnant + (1|Community) + (1|Year), data = Hb_glmer, fa
summary(logit_sex_1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Anemia ~ Sex + Pregnant + (1 | Community) + (1 | Year)
## Data: Hb_glmer
## Subset: AgeGroups == 3
##
##      AIC      BIC    logLik deviance df.resid
##    390.5    413.0   -190.2    380.5      668
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.4885 -0.3265 -0.2878 -0.2134  5.0379
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Community (Intercept) 0.23995  0.4898
## Year      (Intercept) 0.04734  0.2176
## Number of obs: 673, groups: Community, 34; Year, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.0567     0.3346  -9.134  <2e-16 ***
## Sex           0.8059     0.3243   2.485   0.013 *
## Pregnant      0.1819     0.6422   0.283   0.777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Sex
## Sex          -0.742
## Pregnant     -0.012 -0.125
```

```
icc(logit_sex_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.080
##      Conditional ICC: 0.077
```

```
plogis(0.8059)
```

```
## [1] 0.6912351
```

```
## Children Combined
```

```
logit_sex_2 <- glmer(formula = Anemia ~ Sex + (1|Community) + (1|Year), data = Hb_glmer, family = "binomial")
summary(logit_sex_2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Anemia ~ Sex + (1 | Community) + (1 | Year)
```

```
## Data: Hb_glmmer
## Subset: AgeGroups != 3
##
##      AIC      BIC    logLik deviance df.resid
##    681.3    700.7   -336.6    673.3     940
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.9761 -0.3828 -0.3018 -0.2469  4.8143
##
## Random effects:
##  Groups      Name      Variance Std.Dev.
##  Community (Intercept) 0.66173  0.8135
##  Year      (Intercept) 0.09174  0.3029
## Number of obs: 944, groups:  Community, 25; Year, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.7748     0.2700  -6.573 4.94e-11 ***
## Sex          -0.3458     0.2069  -1.672  0.0946 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## Sex -0.343
```

```
icc(logit_sex_2)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.186
##      Conditional ICC: 0.185
```

```
plogis(-0.3458)
```

```
## [1] 0.4144013
```

Inferentials: GAM

```
# Data prep for GAM
Hb_gam <- as.data.frame(Hb_glmmer)
Hb_gam$Community <- as.numeric(factor(Hb_gam$Community))
Hb_gam <- filter(Hb_gam, !is.na(Age), !is.na(Sex), !is.na(Community), !is.na(Anemia), !is.na(Year))

# GAM Model
gam_age <- gamm(Anemia ~ s(Age) + Sex + Pregnant, data = Hb_gam, family = binomial, method = "REML", ran

##
## Maximum number of PQL iterations: 20
```



```
## iteration 1
```

```
## iteration 2
```

```
## iteration 3
```

```
## iteration 4
```

```
## iteration 5
```

```
## iteration 6
```

```
## Model Results
```

```
summary(gam_age)
```

```
##      Length Class Mode
```

```
## lme 20      lme  list
```

```
## gam 31      gam  list
```

```
summary(gam_age$gam)
```

```
##
```

```
## Family: binomial
```

```
## Link function: logit
```

```
##
```

```
## Formula:
```

```
## Anemia ~ s(Age) + Sex + Pregnant
```

```
##
```

```
## Parametric coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.30157    0.19963  -11.529  <2e-16 ***
## Sex          0.06575    0.17029   0.386    0.699
## Pregnant     0.53960    0.65320   0.826    0.409
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Approximate significance of smooth terms:
```

```
##              edf Ref.df      F  p-value
## s(Age) 4.737    4.737 7.807 8.56e-06 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## R-sq.(adj) =  0.0184
```

```
##   Scale est. = 1          n = 1617
```

```
summary(gam_age$lme)
```

```
## Linear mixed-effects model fit by maximum likelihood
```

```
## Data: data
```

```
##      AIC      BIC    logLik
```

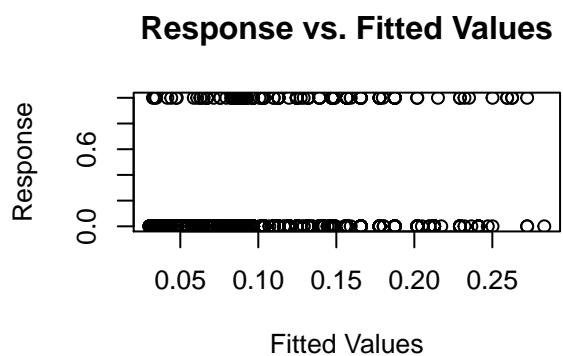
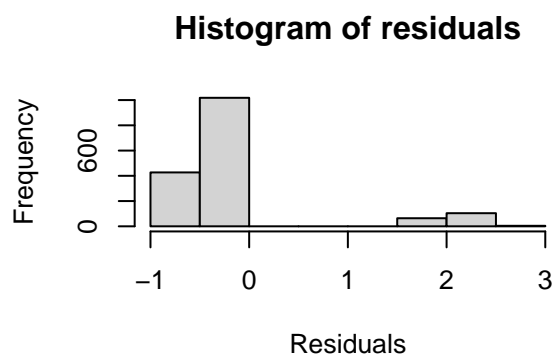
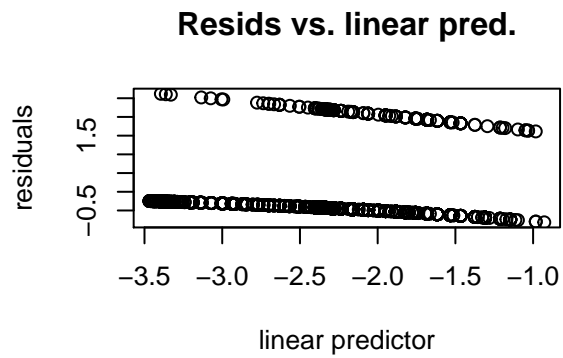
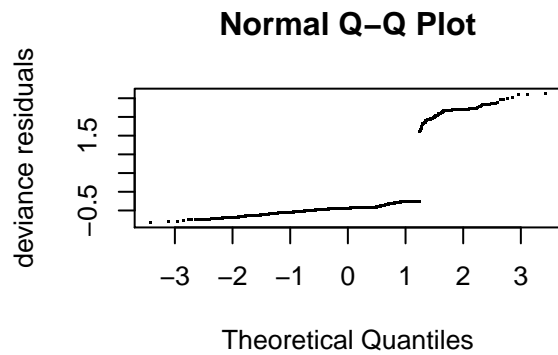
```

##      8627.507 8665.226 -4306.754
##
## Random effects:
## Formula: ~Xr - 1 | g
## Structure: pdIdnot
##           Xr1      Xr2      Xr3      Xr4      Xr5      Xr6      Xr7      Xr8
## StdDev: 5.316716 5.316716 5.316716 5.316716 5.316716 5.316716 5.316716 5.316716
##
## Formula: ~1 | Community %in% g
##           (Intercept)
## StdDev: 0.4177743
##
## Formula: ~1 | Year %in% Community %in% g
##           (Intercept) Residual
## StdDev: 0.5905442      1
##
## Variance function:
## Structure: fixed weights
## Formula: ~invwt
## Fixed effects: list(fixed)
##           Value Std.Error  DF    t-value p-value
## X(Intercept) -2.3015732 0.1996939 1561 -11.525506 0.0000
## XSex          0.0657524 0.1703461 1561 0.385993 0.6996
## XPregnant     0.5395988 0.6534021 1561 0.825830 0.4090
## Xs(Age)Fx1    -1.1297546 0.8826776 1561 -1.279918 0.2008
## Correlation:
##           X(Int) XSex  XPrgnn
## XSex        -0.479
## XPregnant   -0.022 -0.110
## Xs(Age)Fx1  0.026 -0.034 0.005
##
## Standardized Within-Group Residuals:
##           Min      Q1      Med      Q3      Max
## -0.8576200 -0.3608735 -0.2803738 -0.2298034 6.8939795
##
## Number of Observations: 1617
## Number of Groups:
##           g      Community %in% g
##           1      34
## Year %in% Community %in% g
##           53

```

Assumptions checks

```
gam.check(gam_age$gam)
```



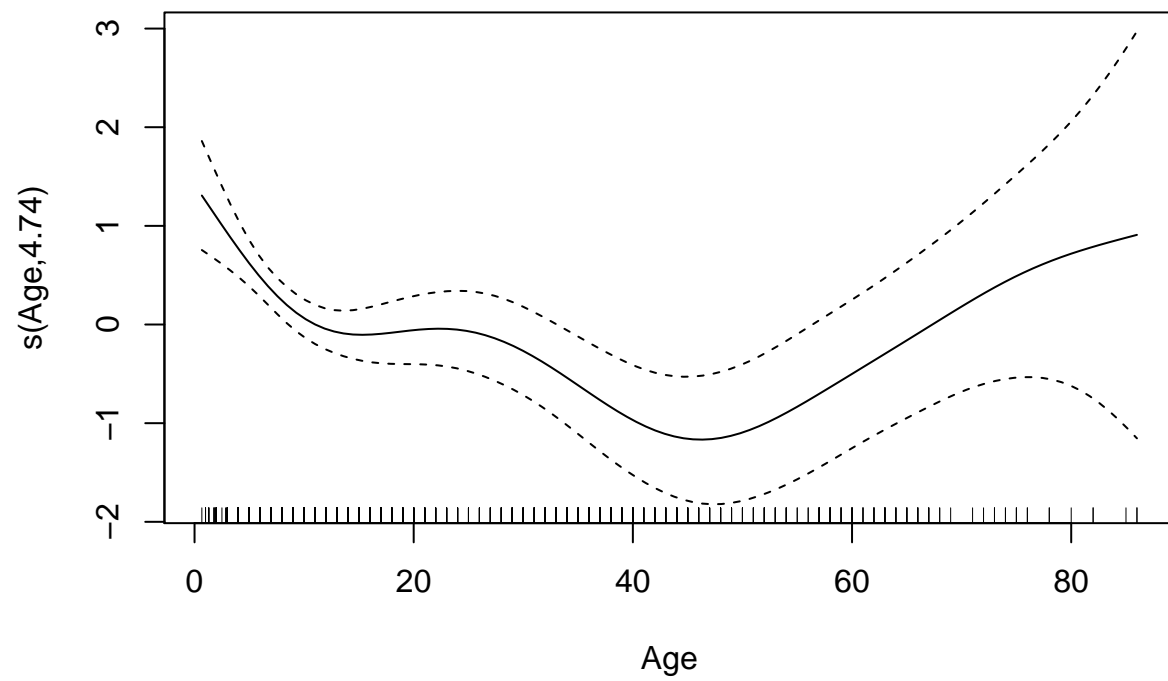
```
##
## 'gamm' based fit - care required with interpretation.
## Checks based on working residuals may be misleading.
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'   edf k-index p-value
## s(Age) 9.00 4.74   0.91  0.045 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
concurvity(gam_age$gam)
```

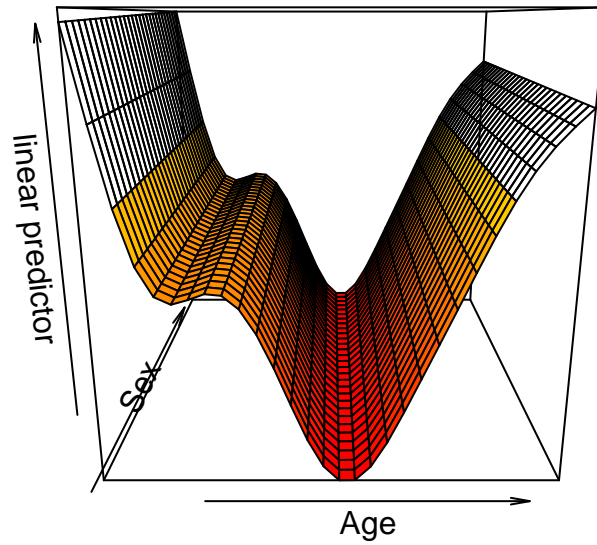
```
##          para      s(Age)
## worst    0.55005 0.048885183
## observed 0.55005 0.006589522
## estimate 0.55005 0.005837093
```

```
gam <- gam_age$gam
lme <- gam_age$lme
# Plot log odds

plot(gam, all.terms=F)
```



```
vis.gam(gam, view = c("Age", "Sex"))
```



```
mean(gam$model$Year)
```

```
## [1] 2012.563
```

```
# PLOT
```

```
# Create new datatable from model
```

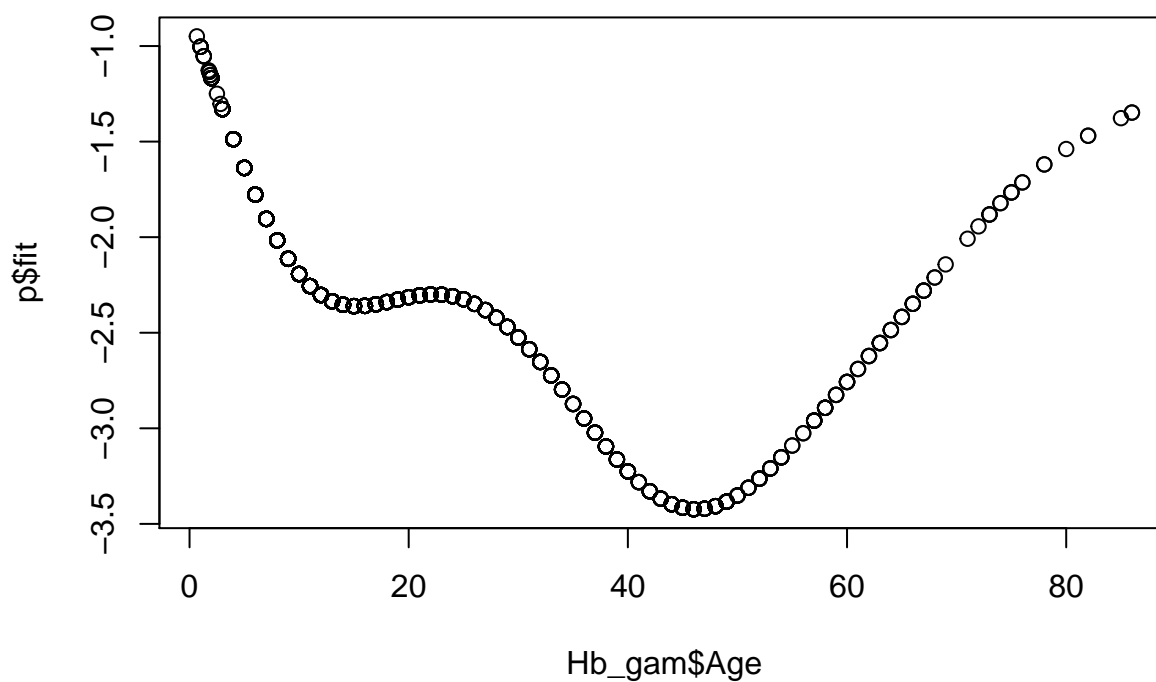
```
testdata = data.frame(Age=Hb_gam$Age, Year=mean(gam$model$Year), Community=mean(gam$model$Community), S
```

```
# Predict to get fit values and se
```

```
p<-predict(gam_age$gam,se.fit=TRUE,newdata=testdata)
```

```
# Basic plot of fit
```

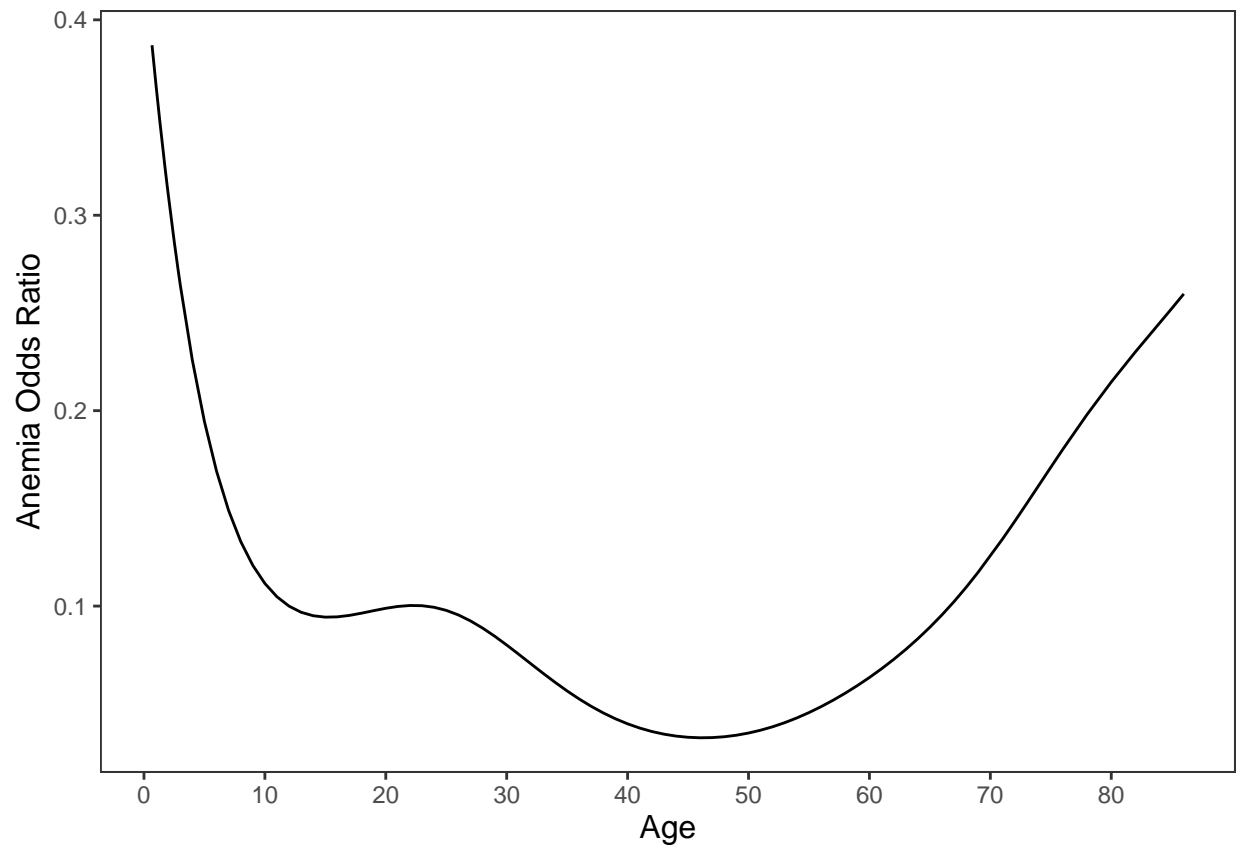
```
plot(Hb_gam$Age, p$fit)
```



```
# Add original data to fit values
p2 <- do.call("cbind", c(testdata, p))
p2 <- data.frame(p2)

# Transform from log using exp
p3 <- mutate(p2, exp_fit = exp(fit), exp_se = exp(se.fit), lower = exp_fit-exp_se, upper = exp_fit+exp_se)

# Plot
ggplot(p3, aes(Age, exp_fit)) +
  geom_line() +
  scale_x_continuous(breaks = c(0, 10, 20, 30, 40, 50, 60, 70, 80)) +
  labs(y = 'Anemia Odds Ratio', x = "Age") +
  theme_apapa()
```



```
# Finding Inflection points
p4 <- p3[order(p3$Age), ]
infl <- c(FALSE, diff(diff(p4$fit)>0)!=0)
infl2 <- c(FALSE, infl)
infl3 <- cbind(p4, infl2)
infl4 <- filter(infl3, infl2 == TRUE)

infl2 <- c(FALSE, diff(diff(p4$fit)<0)!=0)
infl2_2 <- c(FALSE, infl2)
infl2_3 <- cbind(p4, infl2_2)
infl2_4 <- filter(infl2_3, infl2_2 == TRUE)
```

Data Manipulation - Create BMI (RandR2)

```
#Convert height to m from cm
Hb$Height <- Hb$Height/100

# Calculate BMI
Hb$BMI <- Hb$Weight/(Hb$Height^2)
Hb <- filter(Hb, BMI < 185 | is.na(BMI))
```

```
# (RandR1)
describe(Hb$BMI)
```

```
##      vars      n mean   sd median trimmed  mad   min    max range skew kurtosis   se
## X1      1 1566 20.31 4.53  18.62   19.86 4.13 9.89 43.41 33.51 0.86    0.35 0.11
```

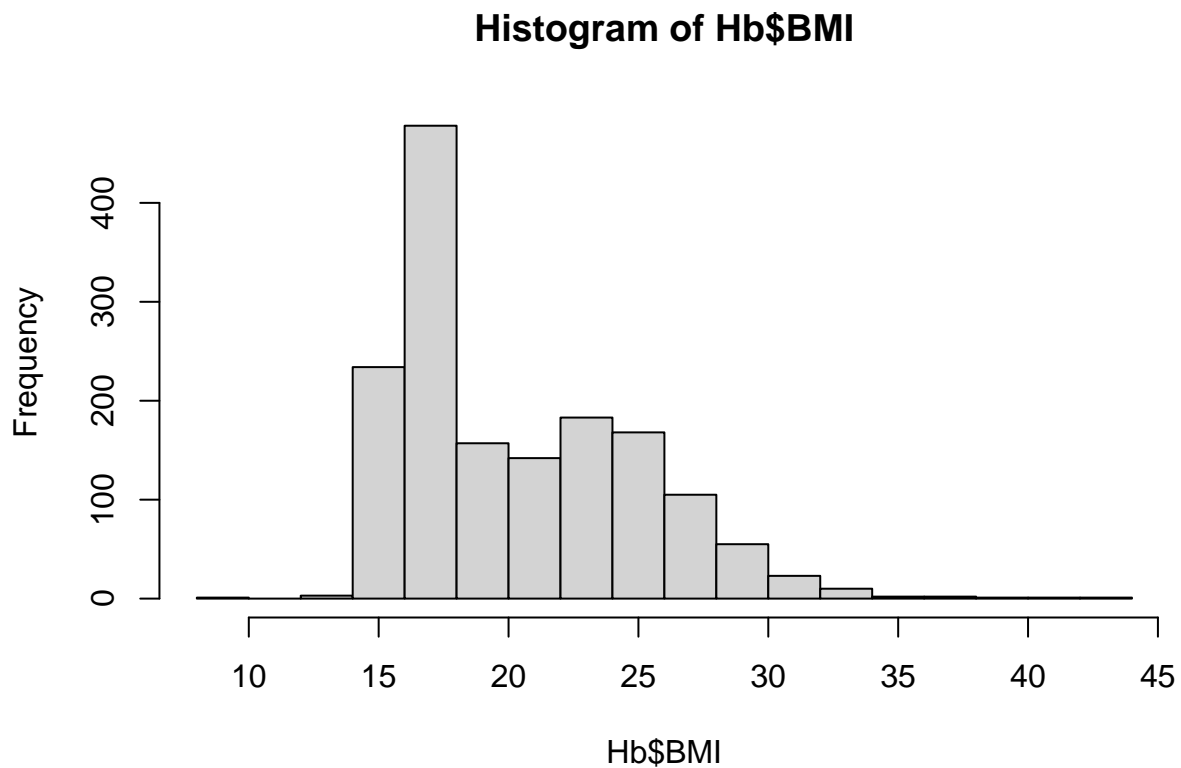
```
skewness(Hb$BMI, na.rm = T)
```

```
## [1] 0.8606539
```

```
kurtosis(Hb$BMI, na.rm = T)
```

```
## [1] 3.350275
```

```
hist(Hb$BMI)
```



Inferentials regional comparison for revision I (RandR1)

```
Hb_comp <- Hb %>%
  filter(!is.na(Hb_AltAdj), !is.na(Region)) %>%
```



```

group_by(Region) %>%
  summarize(AvgHb = mean(Hb_AltAdj), N = n()) %>%
  print

## # A tibble: 2 x 3
##   Region AvgHb      N
## *   <dbl> <dbl> <int>
## 1     0  13.2  1046
## 2     1  13.2   580

print(Hb_comp)

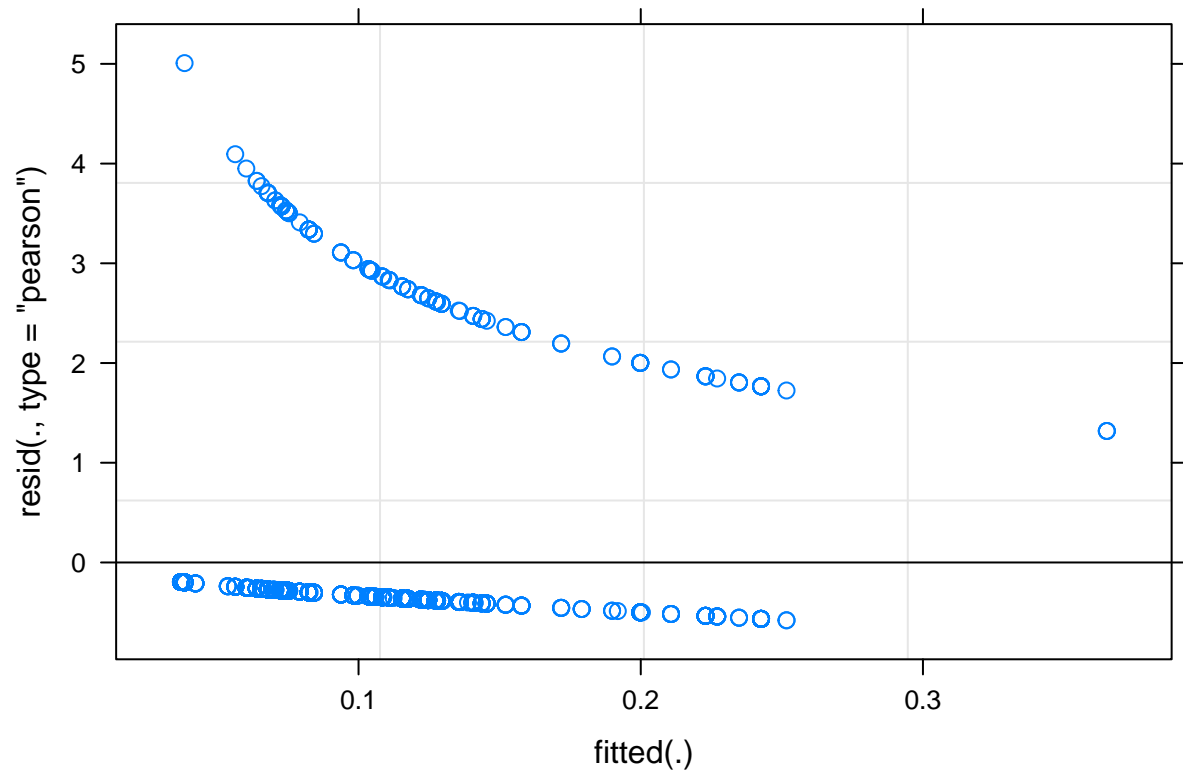
## # A tibble: 2 x 3
##   Region AvgHb      N
## *   <dbl> <dbl> <int>
## 1     0  13.2  1046
## 2     1  13.2   580

glmerRegion <- glmer(Anemia ~ Region + (1|AnemiaGroups) + (1|Year), data = Hb, family = "binomial")
summary(glmerRegion)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Anemia ~ Region + (1 | AnemiaGroups) + (1 | Year)
## Data: Hb
##
##      AIC      BIC   logLik deviance df.resid
##  1077.0   1098.6   -534.5   1069.0     1610
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.5799 -0.3729 -0.2996 -0.2697  5.0075
##
## Random effects:
##   Groups      Name      Variance Std.Dev.
##   Year      (Intercept)  0.2536   0.5036
##   AnemiaGroups (Intercept) 0.2077   0.4557
## Number of obs: 1614, groups: Year, 9; AnemiaGroups, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.0828     0.2869  -7.261 3.84e-13 ***
## Region       -0.1351     0.2108  -0.641  0.522
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## Region -0.252

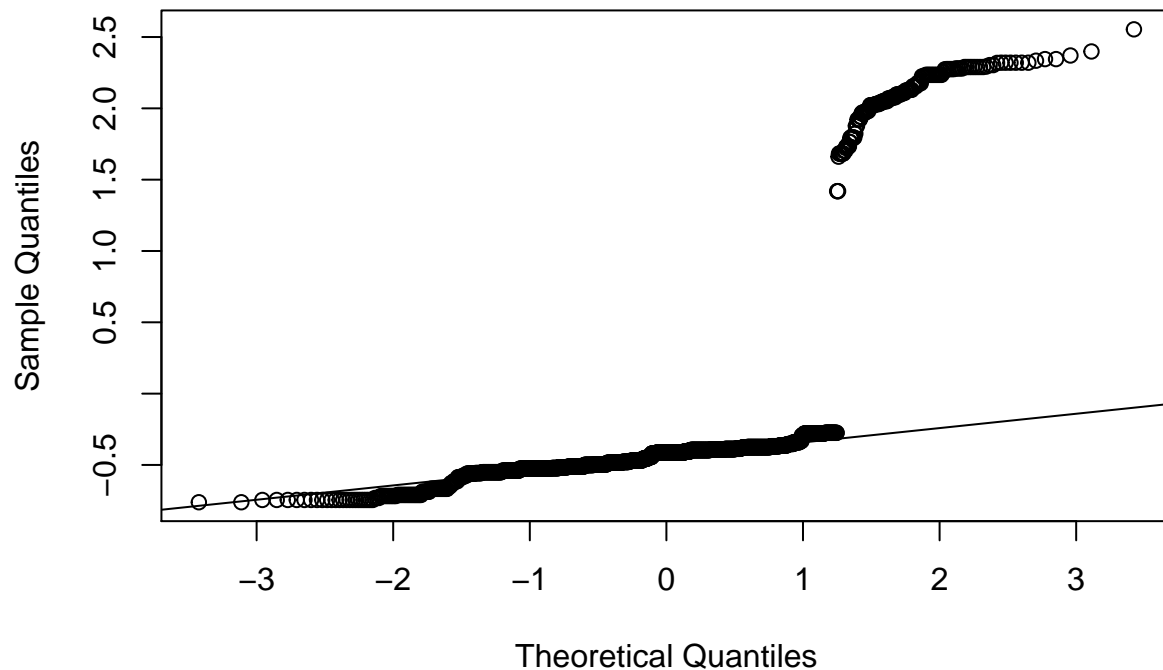
```

```
plot(glmerRegion)
```



```
qqnorm(resid(glmerRegion))  
qqline(resid(glmerRegion))
```

Normal Q-Q Plot



```
icc(glmerRegion)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.123
##      Conditional ICC: 0.123
```

```
plogis(-0.1351)
```

```
## [1] 0.4662763
```

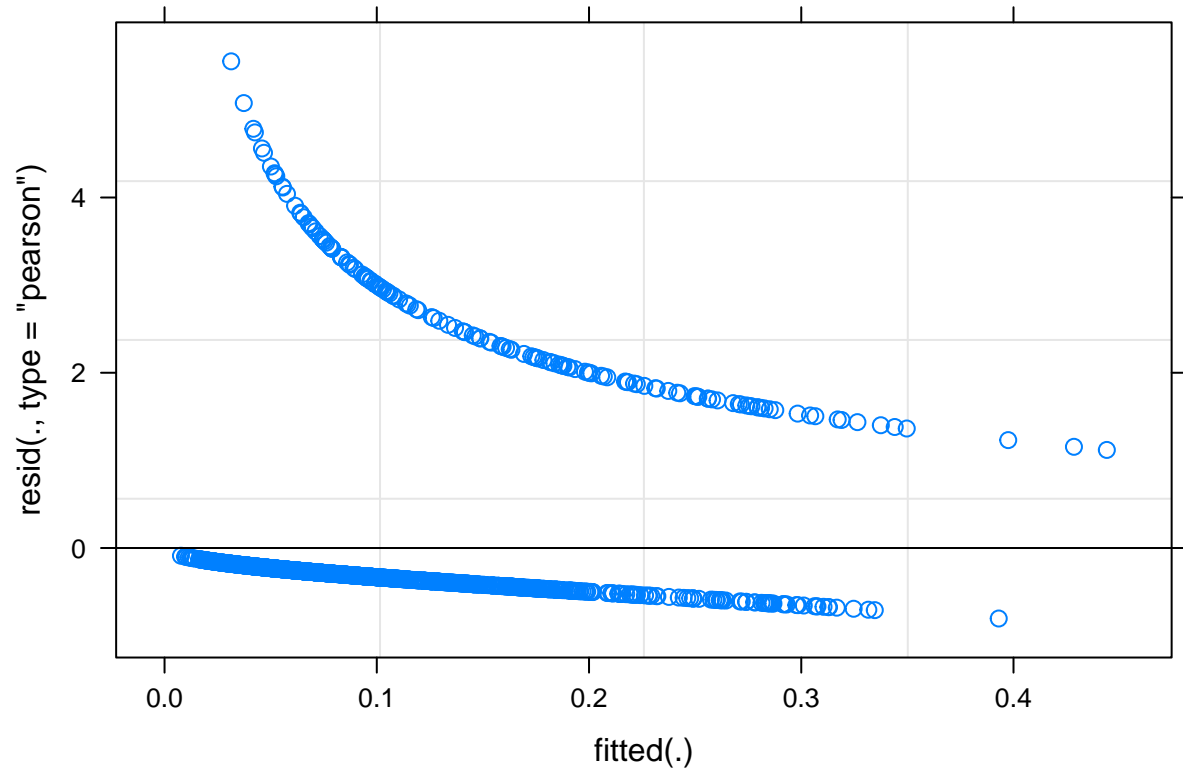
Inferentials BMI association for revision I (RandR1)

```
glmerBMI <- glmer(Anemia ~ BMI + (1|AnemiaGroups) + (1|Community) + (1|Year), data = Hb, family = "binomial")
summary(glmerBMI)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial ( logit )
##   Formula: Anemia ~ BMI + (1 | AnemiaGroups) + (1 | Community) + (1 | Year)
##   Data: Hb
```

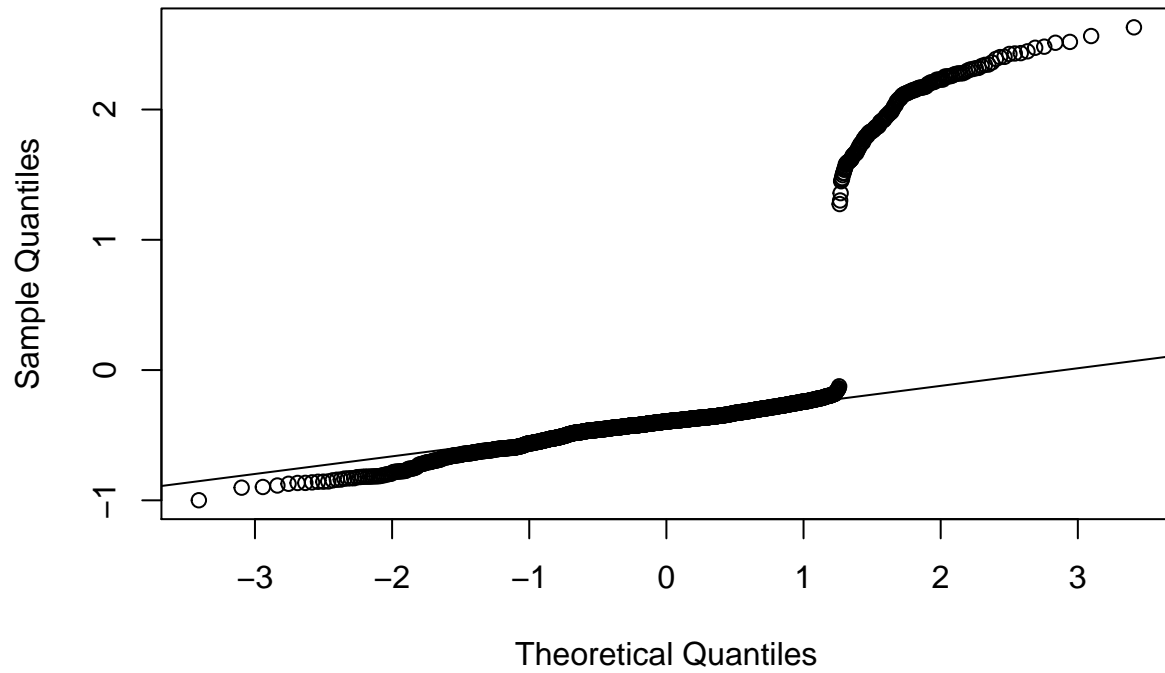
```
##
##      AIC      BIC   logLik deviance df.resid
##    982.5    1009.2   -486.3    972.5     1530
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.8047 -0.3518 -0.2853 -0.2150  5.5537
##
## Random effects:
##   Groups      Name      Variance Std.Dev.
## Community (Intercept) 0.39455  0.6281
## Year      (Intercept) 0.15432  0.3928
## AnemiaGroups (Intercept) 0.09169  0.3028
## Number of obs: 1535, groups: Community, 34; Year, 9; AnemiaGroups, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.17413    0.70631  -0.247  0.80526
## BMI          -0.10213    0.03256  -3.137  0.00171 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## BMI -0.926
```

```
plot(glmerBMI)
```



```
qqnorm(resid(glmerBMI))  
qqline(resid(glmerBMI))
```

Normal Q-Q Plot



```
icc(glmerBMI)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.163
##      Conditional ICC: 0.155
```

```
plogis(-0.10213)
```

```
## [1] 0.4744897
```