```
###########################################################
#####     Probabilistic estimates of Daily Intake    #######
#####       Using LitStream data                  #######
###########################################################

###########################################################
##### IMPORTANT!!!!!!!!!!!!!!!!!!!                 ######
##### Please first run the R code of               ######
#####  "4.Probablistic results_litstream_1.BCEP.R" ######
###########################################################
#####                                         #######
#####     This code is only for BCIPP + BCIPHIPP     #######
#####                                         #######
###########################################################

###################################################
### Author: Dr. Wei-Wen Hsu           #############
### University of Cincinnati          #############
### Division of Biostatistics and Bioinformatics ###
### hsuwe@uc.edu                      #############
### Latest modification: 3/5/2024     #####
#########################################
#### Updated the FUE values     ##########
#########################################

rm(list=ls(all=TRUE))

library(dplyr)
library(tidyr)
```

```
library(openxlsx)

library("writexl")

library("readxl")

library(basictabler)

library(psych)

library(goft)

library(ggplot2)

#install.packages("goft")

#install.packages("EnvStats")

library(EnvStats);


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


## Location for the results

path_results ="D:\\UC-OneDrive\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Results";

#path_results ="C:\\Users\\hsuwe\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Results";


## As the results in the file named:

exceldata_name = "\\4. LitStream_Distributional_data_results.xlsx";




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

####  FUE parameters  ####

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


# Old values

# c(best estimate, high, low)
```

```
#FUE_BCEP = c(0.2, 0.41, 0.04)
#FUE_BCIPP = c(0.39, 0.78, 0.08)
#FUE_BDCIPP = c(0.42, 0.5, 0.08)
#FUE_BCIPP_BCIPHIPP = c(0.78, 0.94, 0.16)

## New values as of 2/24/2024
FUE_BCEP   = c(0.13, 0.27, 0.03)
FUE_BCIPP  = c(0.23, 0.47, 0.05)
FUE_BDCIPP = c(0.23, 0.28, 0.05)
FUE_BCIPP_BCIPHIPP = c(0.58, 0.90, 0.12)
FUE_BCIPHIPP = c(0.35, 0.69, 0.07)



############################
#### UFRBW parameters   ####
############################
# From NHANES 2017-2018 data
# NOTE: the position of the value is important

groups    = c("3-5 yrs","6-11 yrs","12-17 yrs","18+ yrs")
UFRBW_GM  = c(0.063917438,0.024195805,0.016092702,0.01293934)
GSD = c(2.639638606,2.241928384,2.129355164,2.072937724)

parameters_UFRBW = as.data.frame(cbind(groups, UFRBW_GM,GSD) )
parameters_UFRBW$UFRBW_GM = as.numeric(parameters_UFRBW$UFRBW_GM)
parameters_UFRBW$GSD = as.numeric(parameters_UFRBW$GSD)
```

```
########################################

## Estimated Chemicals fron LitStream ##

## By age groups                       ##

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

# NOTE: the first two age groups have BCIPP + BCIPHIPP; the third one has BCIPHIPP only

# study = c("29-Phillips-2018","144-Hoffman-2017","39-Hoffman-2021")


age_groups =c("BCIPP + BCIPHIPP: 3-6 years", "BCIPP + BCIPHIPP: 20-39 years", "BCIPHIPP: 20-60+ years");


# sqrt(180/(181*2-2)*(2.4254267^2+2.73362244^2))=2.584123 ##GSD for 29-Phillips-2018

# sqrt(349/(349*2-2)*(2.3688981^2+2.02455534^2))=2.206627 ##GSD for 144-Hoffman-2017


GM_conc_BCIPP  = c(1.72    , 1.2     , 0.16)

GSD_conc_BCIPP = c(2.584123, 2.206627, 1.93760132 )



#install.packages("EnvStats")

library(EnvStats); # a package for generating random values from a triangular distribution


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

### By Chemicals         ###

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


# a random sample of 1,000

n = 1000



## BCIPP + BCIPHIPP
```

```
BCIPP_matrix=matrix(,1000,3); ## for 3 age-groups

yr=0

for (agrp in c(1,4,4)){   ## select the first (3-5), 4th (18+) age groups from NHANES 2017-18 cycle

yr= yr+1

if (yr<=2){  ## the first two studies are BCIPP + BCIPHIPP
FUE_BCIPP_rand = rtri(n, min=FUE_BCIPP_BCIPHIPP[[3]], max=FUE_BCIPP_BCIPHIPP[[2]], mode=FUE_BCIPP_BCIPHIPP[[1]])
} else{ ## for the last study, we use BCIPHIPP only
FUE_BCIPP_rand = rtri(n, min=FUE_BCIPHIPP[[3]], max=FUE_BCIPHIPP[[2]], mode=FUE_BCIPHIPP[[1]])
}

UFRBW_BCIPP_rand = rlnorm(n, meanlog = log(parameters_UFRBW$UFRBW_GM[[agrp]]), sdlog = log(parameters_UFRBW$GSD[[agrp]]))

GM_conc_BCIPP_rand = rlnorm(n, meanlog = log(GM_conc_BCIPP[[yr]]), sdlog = log(GSD_conc_BCIPP[[yr]]))

DI_BCIPP = (GM_conc_BCIPP_rand * UFRBW_BCIPP_rand) / FUE_BCIPP_rand;  ## pay attention to the unit

BCIPP_matrix[,yr]=DI_BCIPP
}

BCIPP_matrix = as.data.frame(BCIPP_matrix)
names(BCIPP_matrix)= age_groups
```

```
#hist(BCIPP_matrix[[5]])

BCIPP_summary = BCIPP_matrix %>%   summarise(

    mean_grp1 = (mean(BCIPP_matrix[[1]], na.rm = TRUE)),
    mean_grp2 = (mean(BCIPP_matrix[[2]], na.rm = TRUE)),
    mean_grp3 = (mean(BCIPP_matrix[[3]], na.rm = TRUE)),
    SD_grp1 = (sd(BCIPP_matrix[[1]], na.rm = TRUE)),
    SD_grp2 = (sd(BCIPP_matrix[[2]], na.rm = TRUE)),
    SD_grp3 = (sd(BCIPP_matrix[[3]], na.rm = TRUE)),
    Gmean_grp1 = exp(mean(log(BCIPP_matrix[[1]]), na.rm = TRUE)),
    Gmean_grp2 = exp(mean(log(BCIPP_matrix[[2]]), na.rm = TRUE)),
    Gmean_grp3 = exp(mean(log(BCIPP_matrix[[3]]), na.rm = TRUE)),
    GSD_grp1 = exp(sd(log(BCIPP_matrix[[1]]), na.rm = TRUE)),
    GSD_grp2 = exp(sd(log(BCIPP_matrix[[2]]), na.rm = TRUE)),
    GSD_grp3 = exp(sd(log(BCIPP_matrix[[3]]), na.rm = TRUE)),
    p25_grp1 = quantile(BCIPP_matrix[[1]],0.25),
    p25_grp2 = quantile(BCIPP_matrix[[2]],0.25),
    p25_grp3 = quantile(BCIPP_matrix[[3]],0.25),
    p50_grp1 = quantile(BCIPP_matrix[[1]],0.5),
    p50_grp2 = quantile(BCIPP_matrix[[2]],0.5),
    p50_grp3 = quantile(BCIPP_matrix[[3]],0.5),
    p75_grp1 = quantile(BCIPP_matrix[[1]],0.75),
    p75_grp2 = quantile(BCIPP_matrix[[2]],0.75),
    p75_grp3 = quantile(BCIPP_matrix[[3]],0.75),
    p90_grp1 = quantile(BCIPP_matrix[[1]],0.9),
    p90_grp2 = quantile(BCIPP_matrix[[2]],0.9),
    p90_grp3 = quantile(BCIPP_matrix[[3]],0.9),
```

```
        p95_grp1 = quantile(BCIPP_matrix[[1]],0.95),
        p95_grp2 = quantile(BCIPP_matrix[[2]],0.95),
        p95_grp3 = quantile(BCIPP_matrix[[3]],0.95),
)

names(BCIPP_summary) = c(paste0("Mean_",age_groups),paste0("SD_",age_groups),paste0("GM_",age_groups),paste0("GSD_",age_groups),
                  paste0("p25_",age_groups), paste0("p50_",age_groups),paste0("p75_",age_groups),paste0("p90_",age_groups),paste0("p95_",age_groups))




##########################################
#### Save the summary results     ########
##########################################

study = c("29-Phillips-2018","144-Hoffman-2017","39-Hoffman-2021")
biomarker = c("BCIPP and BCIPHIPP","BCIPP and BCIPHIPP","BCIPHIPP")
Age_grp = c("3-6 years", "20-39 years", "20-60+ years");
Mean = c(BCIPP_summary[1:3])
SD  = c(BCIPP_summary[4:6])
GM  = c(BCIPP_summary[7:9])
GSD = c(BCIPP_summary[10:12])
p25 = c(BCIPP_summary[13:15])
p50 = c(BCIPP_summary[16:18])
p75 = c(BCIPP_summary[19:21])
p90 = c(BCIPP_summary[22:24])
p95 = c(BCIPP_summary[25:27])
```

```
NH_result_matrix = as.data.frame(cbind(study, biomarker, Age_grp, Mean, SD, GM, GSD, p25, p50, p75,
p90, p95))

names(NH_result_matrix) = c("Study","Biomarker", "Age Group", "Arithmetic mean", "Arithmetic SD",
"Geometric Mean", "Geometric SD", "25th percentile", "50th percentile", "75th percentile", "90th
percentile", "95th percentile");

NH_result_matrix

#######################################################
### Save results as a worksheet in the Excel file  ###
#######################################################

wb <- loadWorkbook(paste0(path_results,exceldata_name));
addWorksheet(wb,"Sim_BCIPP+BCIPHIPP")
writeData(wb, sheet = "Sim_BCIPP+BCIPHIPP", x =BCIPP_matrix, borders = "columns", startCol = 1,
startRow = 1)

### graphs
BCIPP_data_long <- gather(BCIPP_matrix, age_groups, BCIPP, age_groups, factor_key=TRUE)
g2.s = ggplot(BCIPP_data_long, aes(x=BCIPP, color=age_groups, fill=age_groups)) +
geom_histogram(alpha=0.6, show.legend = FALSE) +
  scale_x_continuous(name = "BCIPP+BCIPHIPP or BCIPHIPP level") +
  scale_y_continuous(name = "Frequency")+
facet_wrap(~ age_groups, scale="free", ncol=2) +
  theme(
      axis.text.y = element_text(size=10),
```

```
        axis.title.y = element_text(size=10),
        axis.text.x = element_text(size=10),
        axis.title = element_text( size = 10, face = "bold" ),
        strip.text = element_text(size = 10),
        axis.title.x = element_text(size=10)
        )
#g2.s

plot(g2.s)
insertPlot(wb, "Sim_BCIPP+BCIPHIPP", startCol = 6, startRow = 2)

### log-normal distribution test
BCIPP_lnorm_test_results1 = lnorm_test(BCIPP_matrix[["BCIPP + BCIPHIPP: 3-6 years"]])
BCIPP_lnorm_test_results2 = lnorm_test(BCIPP_matrix[["BCIPP + BCIPHIPP: 20-39 years"]])
BCIPP_lnorm_test_results3 = lnorm_test(BCIPP_matrix[["BCIPHIPP: 20-60+ years"]])

BCIPP_lnorm = rbind(

cbind(BCIPP_lnorm_test_results1$method,BCIPP_lnorm_test_results1$data.name,BCIPP_lnorm_test_results1$p.value),

cbind(BCIPP_lnorm_test_results2$method,BCIPP_lnorm_test_results2$data.name,BCIPP_lnorm_test_results2$p.value),

cbind(BCIPP_lnorm_test_results3$method,BCIPP_lnorm_test_results3$data.name,BCIPP_lnorm_test_results3$p.value)
            )

colnames(BCIPP_lnorm) = c("Method:", "Data:","p-value=")
BCIPP_lnorm=transform(BCIPP_lnorm)
BCIPP_lnorm[[3]]=as.numeric(BCIPP_lnorm[[3]])
```

```
summary_stat_BCIPP= rbind(describe(BCIPP_matrix[["BCIPP + BCIPHIPP: 3-6 years"]]),describe(BCIPP_matrix[["BCIPP + BCIPHIPP: 20-39 years"]]),
      describe(BCIPP_matrix[["BCIPHIPP: 20-60+ years"]]))


summary_stat_BCIPP= summary_stat_BCIPP[,-1]
rownames(summary_stat_BCIPP) = age_groups;


writeData(wb, sheet = "Sim_BCIPP+BCIPHIPP", x =summary_stat_BCIPP, rowNames = T, startCol = 14, startRow = 2)
writeData(wb, sheet = "Sim_BCIPP+BCIPHIPP", x =BCIPP_lnorm , startCol = 6, startRow = 22)




sheet_name = "BCIPP+BCIPHIPP_Estimates"


addWorksheet(wb,paste0(sheet_name))


writeData(wb, sheet = paste0(sheet_name), x ="Table. Daily intake - Estimated parameters based on probabilistic models using LitStream data", borders = "columns", startCol = 1, startRow = 1)


## Table Caption
writeData(wb, sheet = sheet_name, x = "Unit: ug/kg-day", borders = "columns", startCol = 1, startRow = 2)


tbl <- BasicTable$new()
tbl$addData(NH_result_matrix, firstColumnAsRowHeaders=T,
            explicitColumnHeaders=c("Study", "Biomarker", "Age Group", "Arithmetic mean", "Arithmetic SD", "Geometric Mean", "Geometric SD", "25th percentile", "50th percentile", "75th percentile", "90th percentile", "95th percentile"));


# the arguments are (rFrom, cFrom, rTo, cTo, declarations)
```

```
tbl$setStyling(2, 3, 16, 12, declarations=list("xl-value-format"="##0.0000"))


tbl$writeToExcelWorksheet(wb=wb, wsName=sheet_name,
                      topRowNumber=4, leftMostColumnNumber=1,
                      applyStyles=T, mapStylesFromCSS=F, outputValuesAs="rawValue")



saveWorkbook(wb, file=paste0(path_results,exceldata_name), overwrite = TRUE);



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


#################################
#########  THE END   ############
#################################
```