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

##############################################################
##### IMPORTANT!!!!!!!!!!!!!!!!!!!                 ######
##### Please first run the R code of               ######
#####  "4.Probablistic results_litstream_1.BCEP.R" ######
#####  and                                  ######
#####  "4.Probablistic results_litstream_2.BCIPP.R ######
##############################################################
#####                                   #######
#####     This code is only for BDCIPP             #######
#####                                   #######
##############################################################


######################################################
### Author: Dr. Wei-Wen Hsu          #############
### University of Cincinnati         #############
### Division of Biostatistics and Bioinformatics ###
### hsuwe@uc.edu                     #############
### Latest modification: 2/27/2024     #####
###########################################
#### 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)



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

#### 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 from LitStream ##
## By age groups                     ##
########################################

study        = c("29-Phillips-2018","144-Hoffman-2017","379-Percy et al-2022_age3","379-Percy et al-2022_age5","382-Yang et al-2023","39-Hoffman-2021")
age_groups   = c("3-6 years"      ,"20-39 years"    ,"3 years"        ,"5 years"        ,"20-39 years"      ,"20-60+ years")

GM_conc_BDCIPP  = c(5.63,      1.8      , 2.85 , 3.19, 0.6 , 0.82)
GSD_conc_BDCIPP = c(2.520302665,2.54899379, 3.78 , 3.21, 3.29, 1.547527939)

#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

## BDCIPP

BDCIPP_matrix=matrix(,1000,6); ## for 6 groups
```

```
yr=0

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

yr= yr+1

FUE_BDCIPP_rand = rtri(n, min=FUE_BDCIPP[[3]], max=FUE_BDCIPP[[2]], mode=FUE_BDCIPP[[1]])

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

GM_conc_BDCIPP_rand = rlnorm(n, meanlog = log(GM_conc_BDCIPP[[yr]]), sdlog = log(GSD_conc_BDCIPP[[yr]]))

DI_BDCIPP = (GM_conc_BDCIPP_rand * UFRBW_BDCIPP_rand) / FUE_BDCIPP_rand; ## pay attention to the unit

BDCIPP_matrix[,yr]=DI_BDCIPP
}

BDCIPP_matrix = as.data.frame(BDCIPP_matrix)
names(BDCIPP_matrix)= study

BDCIPP_summary = BDCIPP_matrix %>%   summarise(
    mean_grp1 = (mean(BDCIPP_matrix[[1]], na.rm = TRUE)),
    mean_grp2 = (mean(BDCIPP_matrix[[2]], na.rm = TRUE)),
```

```
mean_grp3 = (mean(BDCIPP_matrix[[3]], na.rm = TRUE)),
mean_grp4 = (mean(BDCIPP_matrix[[4]], na.rm = TRUE)),
mean_grp5 = (mean(BDCIPP_matrix[[5]], na.rm = TRUE)),
mean_grp6 = (mean(BDCIPP_matrix[[6]], na.rm = TRUE)),

SD_grp1 = (sd(BDCIPP_matrix[[1]], na.rm = TRUE)),
SD_grp2 = (sd(BDCIPP_matrix[[2]], na.rm = TRUE)),
SD_grp3 = (sd(BDCIPP_matrix[[3]], na.rm = TRUE)),
SD_grp4 = (sd(BDCIPP_matrix[[4]], na.rm = TRUE)),
SD_grp5 = (sd(BDCIPP_matrix[[5]], na.rm = TRUE)),
SD_grp6 = (sd(BDCIPP_matrix[[6]], na.rm = TRUE)),

Gmean_grp1 = exp(mean(log(BDCIPP_matrix[[1]]), na.rm = TRUE)),
Gmean_grp2 = exp(mean(log(BDCIPP_matrix[[2]]), na.rm = TRUE)),
Gmean_grp3 = exp(mean(log(BDCIPP_matrix[[3]]), na.rm = TRUE)),
Gmean_grp4 = exp(mean(log(BDCIPP_matrix[[4]]), na.rm = TRUE)),
Gmean_grp5 = exp(mean(log(BDCIPP_matrix[[5]]), na.rm = TRUE)),
Gmean_grp6 = exp(mean(log(BDCIPP_matrix[[6]]), na.rm = TRUE)),

GSD_grp1 = exp(sd(log(BDCIPP_matrix[[1]]), na.rm = TRUE)),
GSD_grp2 = exp(sd(log(BDCIPP_matrix[[2]]), na.rm = TRUE)),
GSD_grp3 = exp(sd(log(BDCIPP_matrix[[3]]), na.rm = TRUE)),
GSD_grp4 = exp(sd(log(BDCIPP_matrix[[4]]), na.rm = TRUE)),
GSD_grp5 = exp(sd(log(BDCIPP_matrix[[5]]), na.rm = TRUE)),
GSD_grp6 = exp(sd(log(BDCIPP_matrix[[6]]), na.rm = TRUE)),

p25_grp1 = quantile(BDCIPP_matrix[[1]],0.25),
p25_grp2 = quantile(BDCIPP_matrix[[2]],0.25),
p25_grp3 = quantile(BDCIPP_matrix[[3]],0.25),
```

```
p25_grp4 = quantile(BDCIPP_matrix[[4]],0.25),
p25_grp5 = quantile(BDCIPP_matrix[[5]],0.25),
p25_grp6 = quantile(BDCIPP_matrix[[6]],0.25),

p50_grp1 = quantile(BDCIPP_matrix[[1]],0.5),
p50_grp2 = quantile(BDCIPP_matrix[[2]],0.5),
p50_grp3 = quantile(BDCIPP_matrix[[3]],0.5),
p50_grp4 = quantile(BDCIPP_matrix[[4]],0.5),
p50_grp5 = quantile(BDCIPP_matrix[[5]],0.5),
p50_grp6 = quantile(BDCIPP_matrix[[6]],0.5),

p75_grp1 = quantile(BDCIPP_matrix[[1]],0.75),
p75_grp2 = quantile(BDCIPP_matrix[[2]],0.75),
p75_grp3 = quantile(BDCIPP_matrix[[3]],0.75),
p75_grp4 = quantile(BDCIPP_matrix[[4]],0.75),
p75_grp5 = quantile(BDCIPP_matrix[[5]],0.75),
p75_grp6 = quantile(BDCIPP_matrix[[6]],0.75),

p90_grp1 = quantile(BDCIPP_matrix[[1]],0.9),
p90_grp2 = quantile(BDCIPP_matrix[[2]],0.9),
p90_grp3 = quantile(BDCIPP_matrix[[3]],0.9),
p90_grp4 = quantile(BDCIPP_matrix[[4]],0.9),
p90_grp5 = quantile(BDCIPP_matrix[[5]],0.9),
p90_grp6 = quantile(BDCIPP_matrix[[6]],0.9),

p95_grp1 = quantile(BDCIPP_matrix[[1]],0.95),
p95_grp2 = quantile(BDCIPP_matrix[[2]],0.95),
p95_grp3 = quantile(BDCIPP_matrix[[3]],0.95),
p95_grp4 = quantile(BDCIPP_matrix[[4]],0.95),
```

```
        p95_grp5 = quantile(BDCIPP_matrix[[5]],0.95),

        p95_grp6 = quantile(BDCIPP_matrix[[6]],0.95)


)


names(BDCIPP_summary) =
c(paste0("Mean_",age_groups),paste0("SD_",age_groups),paste0("GM_",age_groups),paste0("GSD_",ag
e_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     ########

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


Age_grp = age_groups

Mean= c(BDCIPP_summary[1:6])

SD  = c(BDCIPP_summary[7:12])

GM  = c(BDCIPP_summary[13:18])

GSD = c(BDCIPP_summary[19:24])

p25 = c(BDCIPP_summary[25:30])

p50 = c(BDCIPP_summary[31:36])

p75 = c(BDCIPP_summary[37:42])

p90 = c(BDCIPP_summary[43:48])

p95 = c(BDCIPP_summary[49:54])


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

```
names(NH_result_matrix) = c("Study", "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,"Simulated_BDCIPP")


writeData(wb, sheet = "Simulated_BDCIPP", x =BDCIPP_matrix, borders = "columns", startCol = 1, startRow = 1)



### graphs

BDCIPP_data_long <- gather(BDCIPP_matrix, age_groups, BDCIPP, study, factor_key=TRUE)

g2.s = ggplot(BDCIPP_data_long, aes(x=BDCIPP, color=age_groups, fill=age_groups)) +

geom_histogram(alpha=0.6, show.legend = FALSE) +

  scale_x_continuous(name = "BDCIPP 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, "Simulated_BDCIPP", startCol = 8, startRow = 2)

### log-normal distribution test
BDCIPP_lnorm_test_results1 = lnorm_test(BDCIPP_matrix[["29-Phillips-2018"]])
BDCIPP_lnorm_test_results2 = lnorm_test(BDCIPP_matrix[["144-Hoffman-2017"]])
BDCIPP_lnorm_test_results3 = lnorm_test(BDCIPP_matrix[["379-Percy et al-2022_age3"]])
BDCIPP_lnorm_test_results4 = lnorm_test(BDCIPP_matrix[["379-Percy et al-2022_age5"]])
BDCIPP_lnorm_test_results5 = lnorm_test(BDCIPP_matrix[["382-Yang et al-2023"]])
BDCIPP_lnorm_test_results6 = lnorm_test(BDCIPP_matrix[["39-Hoffman-2021"]])


BDCIPP_lnorm = rbind(

cbind(BDCIPP_lnorm_test_results1$method,BDCIPP_lnorm_test_results1$data.name,BDCIPP_lnorm_test_results1$p.value),

cbind(BDCIPP_lnorm_test_results2$method,BDCIPP_lnorm_test_results2$data.name,BDCIPP_lnorm_test_results2$p.value),

cbind(BDCIPP_lnorm_test_results3$method,BDCIPP_lnorm_test_results3$data.name,BDCIPP_lnorm_test_results3$p.value),

cbind(BDCIPP_lnorm_test_results4$method,BDCIPP_lnorm_test_results4$data.name,BDCIPP_lnorm_test_results4$p.value),
```

```
cbind(BDCIPP_lnorm_test_results5$method,BDCIPP_lnorm_test_results5$data.name,BDCIPP_lnorm_test_results5$p.value),

cbind(BDCIPP_lnorm_test_results6$method,BDCIPP_lnorm_test_results6$data.name,BDCIPP_lnorm_test_results6$p.value) )


colnames(BDCIPP_lnorm) = c("Method:", "Data:","p-value=")

BDCIPP_lnorm=transform(BDCIPP_lnorm)

BDCIPP_lnorm[[3]]=as.numeric(BDCIPP_lnorm[[3]])


summary_stat_BDCIPP = rbind(describe(BDCIPP_matrix[["29-Phillips-2018"]]),describe(BDCIPP_matrix[["144-Hoffman-2017"]]),

    describe(BDCIPP_matrix[["379-Percy et al-2022_age3"]]),describe(BDCIPP_matrix[["379-Percy et al-2022_age5"]]),

    describe(BDCIPP_matrix[["382-Yang et al-2023"]]),describe(BDCIPP_matrix[["39-Hoffman-2021"]]) )


summary_stat_BDCIPP = summary_stat_BDCIPP[,-1]

rownames(summary_stat_BDCIPP) = study;


writeData(wb, sheet = "Simulated_BDCIPP", x =summary_stat_BDCIPP, rowNames = T , startCol = 14, startRow = 2)

writeData(wb, sheet = "Simulated_BDCIPP", x =BDCIPP_lnorm , startCol = 8, startRow = 22)


sheet_name = "BDCIPP_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)

row_headers = paste0(rownames(NH_result_matrix))

tbl <- BasicTable$new()

tbl$addData(NH_result_matrix, firstColumnAsRowHeaders=T,

	explicitColumnHeaders=c("Study", "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, 11, 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 ############

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