Kapraun et al. (submitted): Generic Human Gestational Model

Dustin F. Kapraun, Mark Sfeir, Robert Pearce, Sarah Davidson, Annie Lumen, Lesa Aylward, André Dallmann, Richard Judson, and John F. Wambaugh

2022-01-11

wambaugh.john@epa.gov

Abstract

Chemical risk assessment considers potentially susceptible populations including pregnant women and developing fetuses. Humans encounter thousands of chemicals in their environments, few of which have been fully characterized in terms of internal human dosimetry and mode-of-action. Toxicokinetic (TK) information is needed to relate chemical exposure to potentially bioactive tissue concentrations. Observational data describing human gestational exposures are unavailable for most chemicals, but physiologically based TK (PBTK) models provide a means for estimating such exposures. However, development of chemical-specific PBTK models requires considerable time and resources. As an alternative, generic PBTK approaches describe a standardized physiology and characterize chemicals with a set of standard physical and in vitro-measured descriptors. Here we report and evaluate a generic PBTK model of a human mother and developing fetus. We used a previously published set of formulas describing the major anatomical and physiological changes that occur during pregnancy to augment the High-Throughput Toxicokinetics (httk) software package. We then performed simulations using the httk pregnancy model to estimate the ratio of concentrations in maternal and fetal plasma and compared these estimates to literature in vivo measurements. We also evaluated the model with literature in vivo time-course measurements of maternal plasma concentrations in pregnant and non-pregnant women. Finally, , we demonstrated conceptual prioritization of chemicals measured in maternal [cord blood?] based on predicted fetal brain concentrations. This new chemical-independent model can be used with existing human-specific in vitro data for 957 chemicals as well as any new chemical data that become available. With sufficient evaluation, this gestational model may allow for in vitro to in vivo extrapolation of point of departure doses relevant to reproductive and developmental toxicity.

Prepare for session

```
knitr::opts_chunk$set(collapse = TRUE, comment = '#>')
options(rmarkdown.html_vignette.check_title = FALSE)
```

Load the relevant libraries

```
# Clear the memory:
rm(list=ls())
#
#
# Setup
#
#
#library(readxl)
library(ggplot2)
library(httk)
library(scales)
library(gridExtra)
```

Change working directory

```
setwd("c:/users/jwambaug/git/httk/work")
```

Prepare custom functions

```
scientific_10 <- function(x) {
  out <- gsub("1e", "10^", scientific_format()(x))
  out <- gsub("\\+","",out)
  out <- gsub("10\\^01","10",out)
  out <- parse(text=gsub("10\\^00","1",out))
}

RMSE <- function(x)
{
  mean(x$residuals^2)^(1/2)
}
```

Number of chemicals

The number of chemicals with in vitro TK data (Cl_{int} and f_{up}) that are also non-volatile and non-PFAS can be found using:

```
length(get_cheminfo(model="fetal_pbtk"))
#> [1] 952
```

Data sets were curated from the literature to allow evaluation of the gestational PBTK model. In all cases chemical identities from the original publications were mapped onto DTXSID's from the CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard) Williams et al., 2017. Statistical testing for correlation between predictions and observations was performed using R function "lm" and p-values were calculated according to an F-distribution.

Aylward cord-blood data

Aylward et al., 2014 compiled measurements of the ratio of maternal to fetal cord blood chemical concentrations at birth for a range of chemicals with environmental routes of exposure, including bromodiphenyl ethers, fluorinated compounds, organochlorine pesticides, polyaromatic hydrocarbons, tobacco smoke components, and vitamins. The PBTK model does not have an explicit cord blood compartment, so the ratio between maternal and fetal venous plasma concentrations was used as a surrogate when making comparisons of model predictions to these data.. For each chemical three daily oral doses (every eight hours) total 1 mg/kg/day were simulated starting from the 13th week of gestation until full term (40 weeks). Simulations were made both with and without the correction to f_{up}^f.

Load/format the data

```
#MFdata <- read_excel("Aylward-MatFet.xlsx")
MFdata <- httk::AylwardMatFet

cat(paste("summarized data from over 100 studies covering ",
  length(unique(MFdata$DTXSID)[!(unique(MFdata$DTXSID)%in%c("","-"))]),
  " unique chemicals structures\n",sep=""))
#> summarized data from over 100 studies covering 89 unique chemicals structures


MFdata.httk <- subset(MFdata,DTXSID %in% get_cheminfo(info="DTXSID",model="pbtk"))
MFdata.httk[MFdata.httk$Chemical.Category=="bromodiphenylethers",
  "Chemical.Category"] <- "Bromodiphenylethers"
MFdata.httk[MFdata.httk$Chemical.Category=="organochlorine Pesticides",
  "Chemical.Category"] <- "Organochlorine Pesticides"
  MFdata.httk[MFdata.httk$Chemical.Category=="polyaromatic hydrocarbons",
  "Chemical.Category"] <- "Polyaromatic Hydrocarbons"

colnames(MFdata.httk)[colnames(MFdata.httk) ==
  "infant.maternal.conc...Central.tendency..calculate.j.k..or.report.paired.result."] <-
  "MFratio"
colnames(MFdata.httk)[colnames(MFdata.httk) ==
  "PREFERRED_NAME"] <-
  "Chemical"
colnames(MFdata.httk)[colnames(MFdata.httk) ==
```

```
        "details.on.matrix.comparison...e.g...cord.blood.lipid..maternal.serum.lipid..or.cord.blood.wet.weight..maternal.whole.
         <-

  "Matrix"

# Format the columns:
MFdata.httk$MFratio <- as.numeric(MFdata.httk$MFratio)
MFdata.httk$Chemical <- as.factor(MFdata.httk$Chemical)
MFdata.httk$Matrix <- as.factor(MFdata.httk$Chemical)
MFdata.httk$Chemical.Category <- as.factor(MFdata.httk$Chemical.Category)

colnames(MFdata.httk)[15] <- "infant"
colnames(MFdata.httk)[16] <- "maternal"
colnames(MFdata.httk)[17] <- "obs.units"

MFdata.httk$infant <- suppressWarnings(as.numeric(MFdata.httk$infant))
MFdata.httk$maternal <- suppressWarnings(as.numeric(MFdata.httk$maternal))
MFdata.httk$AVERAGE_MASS <- as.numeric(MFdata.httk$AVERAGE_MASS)
```

Convert the units:

```
convert1.units <- c("ng/ml","ng/mL","ug/L","ug/l","ng/mL serum","ng/g",
  "ng/g wet wt.","ppb")

MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"infant"] <-
  MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"infant"] / # ng/ml = ug/L
  MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"AVERAGE_MASS"]  # ug/L -> uM
MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"maternal"] <-
  MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"maternal"] / # ng/ml = ug/L
  MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"AVERAGE_MASS"]  # ug/L -> uM
MFdata.httk[MFdata.httk$obs.units%in%convert1.units,"obs.units"] <- "uM"

convert2.units <- c("mg/L","ppm")

MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"infant"] <-
  MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"infant"] * 1000 / # mg/L = ug/L
  MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"AVERAGE_MASS"]  # ug/L -> uM
MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"maternal"] <-
  MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"maternal"]* 1000 / # mg/L = ug/L
  MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"AVERAGE_MASS"]  # ug/L -> uM
MFdata.httk[MFdata.httk$obs.units%in%convert2.units,"obs.units"] <- "uM"
```

Compare HTTK predictions with maternal:fetal observations from Aylward 2014

Note that we can't do an absolute scale comparison (for example, fetal:fetal or maternal:maternal) because we don't know the dose for the Aylward data but we assume that the maternal:fetal ratio is independent of dose and so we use the plasma ratio at full term (40 weeks) resulting from a 1 mg/kg/day exposure rate starting in week 13.

```
# Simulate starting from the 13th week of gestation until full term (40 weeks):
times <- sort(unique(c(seq(13 * 7, 40 * 7, 0.25),seq(278,280,.1))))

# For each chemical with maternal-fetal ratio data and HTTK in vitro data:
for (this.id in unique(MFdata.httk$DTXSID))
{
# Load the chemical-specifc paramaters:
  p <- suppressWarnings(parameterize_fetal_pbtk(dtxsid=this.id,
    fetal_fup_adjustment =TRUE,
    suppress.messages=TRUE))
# Skip chemicals where the 95% credible interval for Fup spans more than 0.1 to
# 0.9 (that is, Fup is effectively unknown):
    if (!is.na(p$Funbound.plasma.dist))
  {
```

```
    if (as.numeric(strsplit(p$Funbound.plasma.dist,",")[[1]][3])>0.9 &
      as.numeric(strsplit(p$Funbound.plasma.dist,",")[[1]][2])<0.11)
    {
      skip <- TRUE
    } else skip <- FALSE
  } else skip <- FALSE

  if (!skip)
  {
# Solve the PBTK equations for the full simulation time assuming 1 total daily
# dose of 1 mg/kg/day spread out over three evenly spaces exposures:
    out <- suppressWarnings(solve_fetal_pbtk(
      parameters=p,
      dose=0,
      times=times,
      daily.dose=1,
      doses.per.day=3,
      output.units = "uM",
      suppress.messages=TRUE))
# Identify the concentrations from the final (279th) day:
    last.row <- which(out[,"time"]>279)
    last.row <- last.row[!duplicated(out[last.row,"time"])]
# Average over the final day:
    MFdata.httk[MFdata.httk$DTXSID==this.id,"Mat.pred"] <- mean(out[last.row,"Cplasma"])
    MFdata.httk[MFdata.httk$DTXSID==this.id,"Fet.pred"] <- mean(out[last.row,"Cfplasma"])
    MFdata.httk[MFdata.httk$DTXSID==this.id,"MFratio.pred"] <-
      mean(out[last.row,"Cplasma"])/mean(out[last.row,"Cfplasma"])

# Repeat the analysis without the adjustment to fetal Fup:
    p <- suppressWarnings(parameterize_fetal_pbtk(dtxsid=this.id,
      fetal_fup_adjustment =FALSE,
      suppress.messages = TRUE))
    out <- suppressWarnings(solve_fetal_pbtk(
      parameters=p,
      dose=0,
      times=times,
      daily.dose=1,
      doses.per.day=3,
      output.units = "uM",
      maxsteps=1e7,
      suppress.messages = TRUE))

    last.row <- which(out[,"time"]>279) # The whole final day
    last.row <- last.row[!duplicated(out[last.row,"time"])]
    MFdata.httk[MFdata.httk$DTXSID==this.id,"Mat.pred.nofup"] <- mean(out[last.row,"Cplasma"])
    MFdata.httk[MFdata.httk$DTXSID==this.id,"Fet.pred.nofup"] <- mean(out[last.row,"Cfplasma"])
    MFdata.httk[MFdata.httk$DTXSID==this.id,"MFratio.pred.nofup"] <-
      mean(out[last.row,"Cplasma"])/mean(out[last.row,"Cfplasma"])
  }
}

# Something is wrong with cotinine:
MFdata.httk <- subset(MFdata.httk,Chemical!="Cotinine")

max.chem <- MFdata.httk[which(MFdata.httk$MFratio==max(MFdata.httk$MFratio)),]
min.chem <- MFdata.httk[which(MFdata.httk$MFratio==min(MFdata.httk$MFratio)),]
cat(paste("The minimum observed ratio was ",
  signif(min.chem[,"MFratio"],2),
  " for ",
  min.chem[,"Chemical"],
  " and the maximum was ",
  signif(max.chem[,"MFratio"],2),
  " for ",
  max.chem[,"Chemical"],
  ".\n",sep=""))
#> The minimum observed ratio was 0.11 for 16 and the maximum was 2.9 for 11.
```

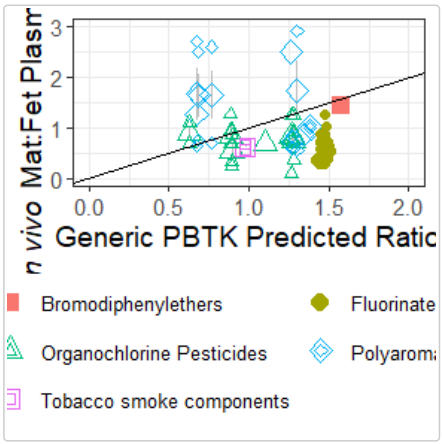
Aylward Ratio figures:

Clean up repeated observations:

```
MFdata.main <- NULL
MFdata.outliers <- NULL
for (this.id in unique(MFdata.httk$DTXSID))
{
  this.subset <- subset(MFdata.httk,DTXSID==this.id)
  this.row <- this.subset[1,]
  this.row$N.obs <- dim(this.subset)[1]
  this.row$MFratio <- median(this.subset$MFratio)
  this.row$MFratio.Q25 <- quantile(this.subset$MFratio,0.25)
  this.row$MFratio.Q75 <- quantile(this.subset$MFratio,0.75)
  MFdata.main <- rbind(MFdata.main,this.row)
  this.subset <- subset(this.subset,
    MFratio<this.row$MFratio.Q25 |
    MFratio>this.row$MFratio.Q75)
  MFdata.outliers <- rbind(MFdata.outliers,this.subset)
}
```

Cord Blood Ratio Predictions without Fup Adjustement:

```
FigCa  <- ggplot(data=MFdata.main) +
  geom_segment(color="grey",aes(
    x=MFratio.pred.nofup,
    y=MFratio.Q25,
    xend=MFratio.pred.nofup,
    yend=MFratio.Q75))+
  geom_point(aes(
    x=MFratio.pred.nofup,
    y=MFratio,
    shape=Chemical.Category,
    color=Chemical.Category),
    size=4)   +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  geom_point(data=MFdata.outliers,aes(
    x=MFratio.pred.nofup,
    y=MFratio,
    shape=Chemical.Category,
    color=Chemical.Category),
    size=2)   +
  xlim(0,2) +
  ylim(0,3) +
#   geom_text(aes(x=AUC,y=Critical.concentration,label=Compound.abbrev,color=Chemical)) +
#   scale_y_log10(label=scientific_10) +
#,limits=c(10^-7,100)) +
#   scale_x_log10(label=scientific_10) +
#   ,limits=c(10^-7,100)) +
#    annotation_logticks() +
  geom_abline(slope=1, intercept=0) +
#    geom_abline(slope=1, intercept=1,linetype="dashed") +
#    geom_abline(slope=1, intercept=-1,linetype="dashed") +
  ylab(expression(paste(italic("In vivo")," Mat:Fet Plasma Ratio"))) +
  xlab("Generic PBTK Predicted Ratio") +
#    scale_color_brewer(palette="Set2") +
  theme_bw()  +
  theme(legend.position="bottom")+
  theme( text  = element_text(size=14))+
  theme(legend.text=element_text(size=10))+
  guides(color=guide_legend(title="Class",nrow=3,byrow=TRUE))+
  guides(shape=guide_legend(title="Class",nrow=3,byrow=TRUE))

print(FigCa)
```



Generate text for results section:

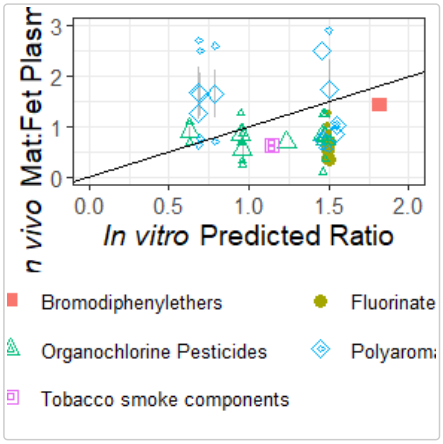
```
cat(paste("In Figure 4 we compare predictions made with our high throughput \
human gestational PBTK model with experimental observations on a per chemical \
basis wherever we had both in vitro HTTK data and in vivo observations (",
length(unique(MFdata.main$DTXSID)),
" chemicals).\n",sep=""))
#> In Figure 4 we compare predictions made with our high throughput
#> human gestational PBTK model with experimental observations on a per chemical
#> basis wherever we had both in vitro HTTK data and in vivo observations (24 chemicals).


repeats <- subset(MFdata.main,N.obs>1)

cat(paste("Multiple observations were available for ",
  dim(repeats)[1],
  " of the chemicals,\n",sep=""))
#> Multiple observations were available for 18 of the chemicals,


max.chem <- repeats[which(repeats$MFratio==max(repeats$MFratio)),]
min.chem <- repeats[which(repeats$MFratio==min(repeats$MFratio)),]

cat(paste("However, among the chemicals with repeated observations, the median \
observations only ranged from ",
  signif(min.chem[,"MFratio"],2),
  " for ",
  min.chem[,"Chemical"],
  " to ",
  signif(max.chem[,"MFratio"],2),
  " for ",
  max.chem[,"Chemical"],
  ".\n",sep=""))
#> However, among the chemicals with repeated observations, the median
#> observations only ranged from 0.36 for 21 to 1.7 for 11.

max.chem <- MFdata.main[which(MFdata.main$MFratio.pred==max(MFdata.main$MFratio.pred,na.rm=TRUE)),]
min.chem <- MFdata.main[which(MFdata.main$MFratio.pred==min(MFdata.main$MFratio.pred,na.rm=TRUE)),]

cat(paste("The predictions for all chemicals ranged from ",
  signif(min.chem[,"MFratio.pred"],2),
  " for ",
  min.chem[,"Chemical"],
  " to ",
  signif(max.chem[,"MFratio.pred"],2),
  " for ",
  max.chem[,"Chemical"],
  ".\n",sep=""))
#> The predictions for all chemicals ranged from 0.63 for 17 to 1.8 for 1.
```

```
fit1 <- lm(data=MFdata.main,MFratio~MFratio.pred.nofup)
summary(fit1)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred.nofup, data = MFdata.main)
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -0.5986 -0.3087 -0.1434  0.2120  1.5661
#>
#> Coefficients:
#>                   Estimate Std. Error t value Pr(>|t|)
#> (Intercept)          1.6512     0.4104   4.023  0.00057 ***
#> MFratio.pred.nofup  -0.5681     0.3391  -1.676  0.10799
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.5095 on 22 degrees of freedom
#> Multiple R-squared:  0.1132, Adjusted R-squared:  0.07286
#> F-statistic: 2.807 on 1 and 22 DF,  p-value: 0.108
RMSE(fit1)
#> [1] 0.4878525

fit1sub <- lm(data=subset(MFdata.main,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))),
  MFratio~MFratio.pred.nofup)
summary(fit1sub)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred.nofup, data = subset(MFdata.main,
#>     !(Chemical.Category %in% c("Fluorinated compounds", "Polyaromatic Hydrocarbons"))))
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -0.1803 -0.1409 -0.1071  0.1233  0.3677
#>
#> Coefficients:
#>                   Estimate Std. Error t value Pr(>|t|)
#> (Intercept)          0.2849     0.3165    0.90    0.398
#> MFratio.pred.nofup   0.4928     0.2801    1.76    0.122
#>
#> Residual standard error: 0.2247 on 7 degrees of freedom
#> Multiple R-squared:  0.3067, Adjusted R-squared:  0.2076
#> F-statistic: 3.096 on 1 and 7 DF,  p-value: 0.1219
RMSE(fit1sub)
#> [1] 0.198205
```

Cord Blood Ratio Predictions with Fup Adjustement:

```
FigCb  <- ggplot(data=MFdata.main) +
  geom_segment(color="grey",aes(
    x=MFratio.pred,
    y=MFratio.Q25,
    xend=MFratio.pred,
    yend=MFratio.Q75))+
  geom_point(aes(
    x=MFratio.pred,
    y=MFratio,
    shape=Chemical.Category,
    color=Chemical.Category),
    size=3)   +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  geom_point(data=MFdata.outliers,aes(
```

```
    x=MFratio.pred,
    y=MFratio,
    shape=Chemical.Category,
    color=Chemical.Category),
    size=1)   +
  xlim(0,2) +
  ylim(0,3) +
#   geom_text(aes(x=AUC,y=Critical.concentration,label=Compound.abbrev,color=Chemical)) +
#   scale_y_log10(label=scientific_10) +
#,limits=c(10^-7,100)) +
#   scale_x_log10(label=scientific_10) +
#   ,limits=c(10^-7,100)) +
#    annotation_logticks() +
  geom_abline(slope=1, intercept=0) +
#    geom_abline(slope=1, intercept=1,linetype="dashed") +
#    geom_abline(slope=1, intercept=-1,linetype="dashed") +
  ylab(expression(paste(italic("In vivo")," Mat:Fet Plasma Ratio"))) +
  xlab(expression(paste(italic("In vitro")," Predicted Ratio"))) +
#    scale_color_brewer(palette="Set2") +
  theme_bw()  +
  theme(legend.position="bottom")+
  theme( text  = element_text(size=14))+
  theme(legend.text=element_text(size=10))+
  guides(color=guide_legend(title="Class",nrow=3,byrow=TRUE))+
  guides(shape=guide_legend(title="Class",nrow=3,byrow=TRUE))

print(FigCb)
```



Examine performance when excluding certain chemical classes:

```
# Mean logHenry's law constant for PAH's:
mean(subset(chem.physical_and_invitro.data,DTXSID%in%subset(MFdata.main,Chemical.Category=="Polyaromatic
          Hydrocarbons")$DTXSID)$logHenry)
#> [1] -4.505444

nonvols <- subset(chem.physical_and_invitro.data,logHenry < -4.5)$DTXSID
fluoros <-
         chem.physical_and_invitro.data$DTXSID[regexpr("fluoro",tolower(chem.physical_and_invitro.data$Compound))!=-1]


fit2 <- lm(data=MFdata.main,MFratio~MFratio.pred)
summary(fit2)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred, data = MFdata.main)
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -0.5575 -0.3297 -0.1727  0.2149  1.5841
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
```

```
#> (Intercept)    1.4593     0.4061   3.594  0.00162 **
#> MFratio.pred  -0.3728     0.3085  -1.208  0.23978
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.524 on 22 degrees of freedom
#> Multiple R-squared:  0.06223,    Adjusted R-squared:  0.0196
#> F-statistic:  1.46 on 1 and 22 DF,  p-value: 0.2398
RMSE(fit2)
#> [1] 0.5016675

fit2sub <- lm(data=subset(MFdata.main,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))),
  MFratio~MFratio.pred)
summary(fit2sub)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred, data = subset(MFdata.main,
#>     !(Chemical.Category %in% c("Fluorinated compounds", "Polyaromatic Hydrocarbons"))))
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -0.1785 -0.1578 -0.1052  0.1247  0.3918
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
#> (Intercept)    0.3748     0.2897   1.294    0.237
#> MFratio.pred   0.3644     0.2256   1.615    0.150
#>
#> Residual standard error: 0.2304 on 7 degrees of freedom
#> Multiple R-squared:  0.2716, Adjusted R-squared:  0.1675
#> F-statistic:  2.61 on 1 and 7 DF,  p-value: 0.1502

RMSE(fit2sub)
#> [1] 0.2031591

cat(paste("When volatile and fluorinated chemicals are omitted only ",
  dim(subset(MFdata.main,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))))[1],
  " evaluation chemicals remain, but the R2 is ",
  signif(summary(fit1sub)$adj.r.squared,2),
  " and the RMSE is ",
  signif(RMSE(fit1sub),2),
  " without the correction. When the fetal plasma fraction unbound correction is used, the predictivity
        decreases, slightly: R2 is ",
  signif(summary(fit2sub)$adj.r.squared,2),
  " and the RMSE is ",
  signif(RMSE(fit2sub),2),
  " for the non-volatile, non-fluorinated chemicals.\n",
  sep=""))
#> When volatile and fluorinated chemicals are omitted only 9 evaluation chemicals remain, but the R2
        is 0.21 and the RMSE is 0.2 without the correction. When the fetal plasma fraction unbound
        correction is used, the predictivity decreases, slightly: R2 is 0.17 and the RMSE is 0.2 for
        the non-volatile, non-fluorinated chemicals.



cat(paste("We compare the RMSE for our predictions to the standard deviation \
of the observations ",
  signif(sd(MFdata.main$MFratio)[1],2),
  " (",
  signif(sd(subset(MFdata.main,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons")))$MFratio),2),
  " for non PAH or fluorinated compounds).\n",sep=""))
```

```r
#> We compare the RMSE for our predictions to the standard deviation
#> of the observations 0.53 (0.25 for non PAH or fluorinated compounds).

cat(paste("The average standard deviation for chemicals with repeated observations was ",
  signif(sd(subset(MFdata.main,N.obs>1)$MFratio)[1],2),
  " (",
  signif(sd(subset(MFdata.main,
  N.obs > 1 &
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons")))$MFratio),2),
  " for non PAH or fluorinated compounds).\n",sep=""))
#> The average standard deviation for chemicals with repeated observations was 0.45 (0.13 for non PAH
#>        or fluorinated compounds).

fit3 <- lm(data=repeats,MFratio~MFratio.pred.nofup)
summary(fit3)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred.nofup, data = repeats)
#>
#> Residuals:
#>      Min       1Q   Median       3Q      Max
#> -0.56671 -0.26929 -0.04221  0.23978  0.94562
#>
#> Coefficients:
#>                   Estimate Std. Error t value Pr(>|t|)
#> (Intercept)          1.8198     0.3530   5.156 9.57e-05 ***
#> MFratio.pred.nofup  -0.7925     0.2971  -2.667   0.0169 *
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.3865 on 16 degrees of freedom
#> Multiple R-squared:  0.3077, Adjusted R-squared:  0.2645
#> F-statistic: 7.112 on 1 and 16 DF,  p-value: 0.01688
RMSE(fit3)
#> [1] 0.3644247

fit3sub <- lm(data=subset(MFdata.main, N.obs > 1 &
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))),
  MFratio~MFratio.pred.nofup)
summary(fit3sub)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred.nofup, data = subset(MFdata.main,
#>     N.obs > 1 & !(Chemical.Category %in% c("Fluorinated compounds",
#>         "Polyaromatic Hydrocarbons"))))
#>
#> Residuals:
#>         1         2         3         4         5         6         7
#>  0.085881 -0.214067  0.075812 -0.001507  0.052055  0.135298 -0.133472
#>
#> Coefficients:
#>                     Estimate Std. Error t value Pr(>|t|)
#> (Intercept)         0.765526   0.238553   3.209   0.0238 *
#> MFratio.pred.nofup -0.007228   0.226018  -0.032   0.9757
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.1398 on 5 degrees of freedom
#> Multiple R-squared:  0.0002045,  Adjusted R-squared:  -0.1998
#> F-statistic: 0.001023 on 1 and 5 DF,  p-value: 0.9757


fit4 <- lm(data=subset(MFdata.main,N.obs > 1),MFratio~MFratio.pred)
summary(fit4)
#>
```

```
#> Call:
#> lm(formula = MFratio ~ MFratio.pred, data = subset(MFdata.main,
#>     N.obs > 1))
#>
#> Residuals:
#>      Min       1Q   Median       3Q      Max
#> -0.54893 -0.32402 -0.03503  0.25576  0.99811
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
#> (Intercept)    1.7238     0.3575   4.821 0.000188 ***
#> MFratio.pred  -0.6581     0.2790  -2.359 0.031392 *
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.4002 on 16 degrees of freedom
#> Multiple R-squared:  0.258,  Adjusted R-squared:  0.2116
#> F-statistic: 5.563 on 1 and 16 DF,  p-value: 0.03139
RMSE(fit4)
#> [1] 0.3772881

fit4sub <- lm(data=subset(MFdata.main, N.obs > 1 &
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))),
  MFratio~MFratio.pred)
summary(fit4sub)
#>
#> Call:
#> lm(formula = MFratio ~ MFratio.pred, data = subset(MFdata.main,
#>     N.obs > 1 & !(Chemical.Category %in% c("Fluorinated compounds",
#>         "Polyaromatic Hydrocarbons"))))
#>
#> Residuals:
#>         1         2         3         4         5         6         7
#>  0.084084 -0.215796  0.078259  0.001062  0.054644  0.131020 -0.133273
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
#> (Intercept)   0.77385    0.20585   3.759   0.0132 *
#> MFratio.pred -0.01363    0.17212  -0.079   0.9399
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.1398 on 5 degrees of freedom
#> Multiple R-squared:  0.001253,   Adjusted R-squared:  -0.1985
#> F-statistic: 0.006274 on 1 and 5 DF,  p-value: 0.9399

cat(paste("Overall, we observed relatively poor correlation (R2 = ",
  signif(summary(fit1)$adj.r.squared,2),
  ", RMSE = ",
  signif(RMSE(fit1),2),
  ") without our fetal fup correction that was unchanged with that assumption (R2 = ",
  signif(summary(fit2)$adj.r.squared,2),
  ", RMSE = ",
  signif(RMSE(fit2),2),
  ").\n",sep=""))
#> Overall, we observed relatively poor correlation (R2 = 0.073, RMSE = 0.49) without our fetal fup
        correction that was unchanged with that assumption (R2 = 0.02, RMSE = 0.5).

repeats <-subset(MFdata.main,N.obs > 1)
cat(paste("The RMSE of the predictions for the ",
  dim(subset(repeats,!(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons"))))[1],
  " non-PAH and non-fluorinated compounds with repeated observations is ",
  signif(RMSE(fit4sub),2),
  " with the fup correction and ",
  signif(RMSE(fit3sub),2),
```

```
  " without.\n",sep=""))
#> The RMSE of the predictions for the 7 non-PAH and non-fluorinated compounds with repeated
         observations is 0.12 with the fup correction and 0.12 without.

cat(paste("These values are close to the standard deviation of the mean but the p-value for the
         chemicals with repeated observations is ",
  signif(pf(
    summary(fit4sub)$fstatistic[1],
    summary(fit4sub)$fstatistic[2],
    summary(fit4sub)$fstatistic[3]),2),
" indicating some value for the predictive model, albeit for only ",
dim(subset(repeats,!(Chemical.Category %in% c(
  "Fluorinated compounds",
  "Polyaromatic Hydrocarbons"))))[1],
 " chemicals",sep=""))
#> These values are close to the standard deviation of the mean but the p-value for the chemicals with
         repeated observations is 0.06 indicating some value for the predictive model, albeit for only
         7 chemicals



Fig4.table <- data.frame()
Fig4.table["Number of Chemicals","All Fig 4"] <- length(unique(MFdata.httk$DTXSID))
Fig4.table["Number of Observations","All Fig 4"] <- dim(MFdata.httk)[1]
Fig4.table["Observed Mean (Min - Max)","All Fig 4"] <- paste(
  signif(mean(MFdata.httk$MFratio),2)," (",
  signif(min(MFdata.httk$MFratio),2)," - ",
  signif(max(MFdata.httk$MFratio),2),")",sep="")
Fig4.table["Observed Standard Deviation","All Fig 4"] <- signif(sd(MFdata.httk$MFratio),2)
Fig4.table["Predicted Mean (Min - Max)","All Fig 4"] <-  paste(
  signif(mean(MFdata.main$MFratio.pred),2)," (",
  signif(min(MFdata.main$MFratio.pred),2)," - ",
  signif(max(MFdata.main$MFratio.pred),2),")",sep="")
Fig4.table["RMSE","All Fig 4"] <- signif(RMSE(fit2),2)
Fig4.table["R-squared","All Fig 4"] <- signif(summary(fit2)[["adj.r.squared"]],2)
Fig4.table["p-value","All Fig 4"] <- signif(summary(fit2)[["coefficients"]]["MFratio.pred",4],2)

MFdata.sub1 <- subset(MFdata.httk,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons")))

Fig4.table["Number of Chemicals","No PFAS/PAH"] <- length(unique(MFdata.sub1$DTXSID))
Fig4.table["Number of Observations","No PFAS/PAH"] <- dim(MFdata.sub1)[1]
Fig4.table["Observed Mean (Min - Max)","No PFAS/PAH"] <- paste(
  signif(mean(MFdata.sub1$MFratio),2)," (",
  signif(min(MFdata.sub1$MFratio),2)," - ",
  signif(max(MFdata.sub1$MFratio),2),")",sep="")
Fig4.table["Observed Standard Deviation","No PFAS/PAH"] <- signif(sd(MFdata.sub1$MFratio),2)

MFdata.sub2 <- subset(MFdata.main,
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons")))

Fig4.table["Predicted Mean (Min - Max)","No PFAS/PAH"] <-  paste(
  signif(mean(MFdata.sub2$MFratio.pred),2)," (",
  signif(min(MFdata.sub2$MFratio.pred),2)," - ",
  signif(max(MFdata.sub2$MFratio.pred),2),")",sep="")
Fig4.table["RMSE","No PFAS/PAH"] <- signif(RMSE(fit2sub),2)
Fig4.table["R-squared","No PFAS/PAH"] <- signif(summary(fit2sub)[["adj.r.squared"]],2)
Fig4.table["p-value","No PFAS/PAH"] <- signif(summary(fit2sub)[["coefficients"]]["MFratio.pred",4],2)


MFdata.sub3 <- subset(MFdata.main,N.obs > 1 &
  !(Chemical.Category %in% c(
    "Fluorinated compounds",
    "Polyaromatic Hydrocarbons")))

Fig4.table["Number of Chemicals","Replicates Only"] <- length(unique(MFdata.sub3$DTXSID))
```

```
Fig4.table["Number of Observations","Replicates Only"] <- dim(MFdata.sub3)[1]
Fig4.table["Observed Mean (Min - Max)","Replicates Only"] <- paste(
  signif(mean(MFdata.sub3$MFratio),2)," (",
  signif(min(MFdata.sub3$MFratio),2)," - ",
  signif(max(MFdata.sub3$MFratio),2),")",sep="")
Fig4.table["Observed Standard Deviation","Replicates Only"] <- signif(sd(MFdata.sub3$MFratio),2)

Fig4.table["Predicted Mean (Min - Max)","Replicates Only"] <-  paste(
  signif(mean(MFdata.sub3$MFratio.pred),2)," (",
  signif(min(MFdata.sub3$MFratio.pred),2)," - ",
  signif(max(MFdata.sub3$MFratio.pred),2),")",sep="")
Fig4.table["RMSE","Replicates Only"] <- signif(RMSE(fit4sub),2)
Fig4.table["R-squared","Replicates Only"] <- signif(summary(fit4sub)[["adj.r.squared"]],2)
Fig4.table["p-value","Replicates Only"] <- signif(summary(fit4sub)[["coefficients"]]
         ["MFratio.pred",4],2)

write.csv(Fig4.table[1:6,],file="Fig4Table.csv")
```

Evaluation of Area Under the Curve Predictions

Dallmann et al., 2018 includes descriptions of toxicokinetics summary statistics, including time-integrated plasma concentrations (area under the curve or AUC) for six drugs administered to a sample of subjects including both pregnant and non-pregnant women. Additional data from X and Y for two chemicals among the httk library were collected from the literature.

```
#TKstats <- as.data.frame(read_excel("MaternalFetalAUCData.xlsx"))
TKstats <- httk::Dallmann2018Data

TKstats.Cmax <- subset(TKstats,DTXSID!="" & Parameter=="Cmax")
TKstats <- subset(TKstats,DTXSID!="" & Parameter=="AUCinf")

# Assumed body weight (kg) from Kapraun 2019
BW.nonpreg <- 61.103

#TKstats$Dose <- TKstats$Dose/BW
#TKstats$Dose.Units <- "mg/kg"
colnames(TKstats)[colnames(TKstats)=="Observed Pregnant"] <- "Observed.Pregnant"
colnames(TKstats)[colnames(TKstats)=="Observed Pregnant5"] <- "Observed.Pregnant5"
colnames(TKstats)[colnames(TKstats)=="Observed Non-Pregnant"] <- "Observed.Non.Pregnant"
colnames(TKstats)[colnames(TKstats)=="Observed Non-Pregnant2"] <- "Observed.Non.Pregnant2"
colnames(TKstats)[colnames(TKstats)=="Dose Route"] <- "Dose.Route"
```

Predict AUC

The circumstances of the dosing varied slightly between drugs and pregnancy status required variation in simulated dose regimen as summarized in Table 12. The function solve_fetal_pbtk was used to determine the time-integrated plasma concentration (area under the curve, or AUC) for the mothers both when pregnant and non-pregnant.

```
for (this.id in unique(TKstats$DTXSID))
{
  if (any(regexpr("ng",TKstats[TKstats$DTXSID==this.id,"Dose Units"])!=-1))
  {
  }
  if (this.id %in% get_cheminfo(info="DTXSID",model="fetal_pbtk"))
  {
    this.subset <- subset(TKstats,DTXSID==this.id)
    p <- suppressWarnings(parameterize_pbtk(
      dtxsid=this.id,
      suppress.messages=TRUE))
    p$hematocrit <- 0.39412 # Kapraun 2019 (unitless)
    p$Rblood2plasma <- suppressWarnings(calc_rblood2plasma(
      parameters=p,
      suppress.messages=TRUE))
```

```
    p$BW <- BW.nonpreg
    p$Qcardiacc <- 301.78 / p$BW^(3/4) # Kapraun 2019 (L/h/kg^3/4)
    for (this.route in unique(this.subset$Dose.Route))
    {
      this.route.subset <- subset(this.subset, Dose.Route==this.route)
      if (this.route.subset[1,"Gestational.Age.Weeks"] > 12)
      {
        this.dose <- this.route.subset$Dose
        # Non-pregnant PBPK:
        out.nonpreg <- suppressWarnings(solve_pbtk(
          parameters=p,
          times = seq(0, this.route.subset[1,"NonPreg.Duration.Days"],
                      this.route.subset[1,"NonPreg.Duration.Days"]/100),
          dose=this.dose/BW.nonpreg, # mg/kg
          daily.dose=NULL,
          iv.dose=(this.route=="iv"),
          output.units="uM",
          suppress.messages=TRUE))
        # Pregnant PBPK:
        BW.preg <- suppressWarnings(calc_maternal_bw(
          week=this.route.subset[1,"Gestational.Age.Weeks"]))
        out.preg <- suppressWarnings(solve_fetal_pbtk(
          dtxsid=this.id,
          times = seq(
            this.route.subset[1,"Gestational.Age.Weeks"]*7,
            this.route.subset[1,"Gestational.Age.Weeks"]*7 +
              this.route.subset[1,"Preg.Duration.Days"],
            this.route.subset[1,"Preg.Duration.Days"]/100),
          dose=this.dose/BW.preg, # mg/kg
          iv.dose=(this.route=="iv"),
          daily.dose=NULL,
          output.units="uM",
          suppress.messages=TRUE))
        if (any(regexpr("AUC",this.subset$Parameter)!=-1))
        {
          TKstats[TKstats$DTXSID==this.id &
                    TKstats$Dose.Route == this.route &
                    regexpr("AUC",TKstats$Parameter)!=-1,
                  "Predicted.Non.Pregnant.httk"] <- max(out.nonpreg[,"AUC"])*24 #uMdays->uMh
          TKstats[TKstats$DTXSID==this.id &
                    TKstats$Dose.Route == this.route &
                    regexpr("AUC",TKstats$Parameter)!=-1,
                  "Predicted.Pregnant.httk"] <- max(out.preg[,"AUC"])*24 # uM days -> uM h
        }
        if (any(regexpr("Cmax",this.subset$Parameter)!=-1))
        {
          TKstats[TKstats$DTXSID==this.id &
                    TKstats$Dose.Route == this.route &
                    regexpr("Cmax",TKstats$Parameter)!=-1,
                  "Predicted.Non.Pregnant.httk"] <- max(out.nonpreg[,"Cplasma"])
          TKstats[TKstats$DTXSID==this.id &
                    TKstats$Dose.Route == this.route &
                    regexpr("Cmax",TKstats$Parameter)!=-1,
                  "Predicted.Pregnant.httk"] <- max(out.preg[,"Cfplasma"])
        }
      }
    }
  }
}
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 273, R2 = 1.5093e-15
#>
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 273, R2 = 1.5093e-15
#>
```

```
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 216.3, R2 = 3.12567e-16
#>
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 216.3, R2 = 3.12567e-16
#>
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 245.7, R2 = 7.75163e-16
#>
#> DLSODA-  Warning..Internal T (=R1) and H (=R2) are
#>       such that in the machine, T + H = T on the next step
#>      (H = step size). Solver will continue anyway.
#> In above message, R1 = 245.7, R2 = 7.75163e-16
#>

TKstats$Ratio.obs <- TKstats$Observed.Pregnant / TKstats$Observed.Non.Pregnant
TKstats$Ratio.httk <- TKstats$Predicted.Pregnant.httk / TKstats$Predicted.Non.Pregnant.httk
TKstats <- subset(TKstats, !is.na(TKstats$Ratio.httk))

write.csv(TKstats,file="Table-Dallmann2018.csv",row.names=FALSE)
```

Generate AUC Predicted vs. Observed Figure

```
FigEa  <- ggplot(data=subset(TKstats,Parameter=="AUCinf")) +
  geom_point(aes(
    y=Observed.Non.Pregnant2,
    x=Predicted.Non.Pregnant.httk,
    shape=Dose.Route,
    color=Drug),
    size=3)   +
  geom_abline(slope=1, intercept=0) +
  geom_abline(slope=1, intercept=1, linetype=3) +
  geom_abline(slope=1, intercept=-1, linetype=3) +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  xlab("httk Predicted (uM*h)") +
  ylab("Observed") +
  scale_x_log10(#limits=c(10^-3,10^3),
    label=scientific_10) +
  scale_y_log10(#limits=c(10^-3,10^3),
    label=scientific_10) +
  ggtitle("Non-Pregnant") +
  theme_bw()  +
  theme(legend.position="none")+
  theme( text  = element_text(size=10))+
  theme(legend.text=element_text(size=8))

#print(FigEa)
cat(paste(
  "For the Dallman et al. (2018) data we observe an average-fold error (AFE)\n\ of",
  signif(mean(TKstats$Predicted.Non.Pregnant.httk/TKstats$Observed.Non.Pregnant2),2),
  "and a RMSE (log10-scale) of",
  signif((mean((log10(TKstats$Predicted.Non.Pregnant.httk)-
         log10(TKstats$Observed.Non.Pregnant2))^2))^(1/2),2),
  "for non-pregnant woman.\n"))
#> For the Dallman et al. (2018) data we observe an average-fold error (AFE)
#>  of 1.1 and a RMSE (log10-scale) of 1 for non-pregnant woman.

FigEb  <- ggplot(data=subset(TKstats,Parameter=="AUCinf")) +
  geom_point(aes(
    y=Observed.Pregnant5,
    x=Predicted.Pregnant.httk,
```

```
    shape=Dose.Route,
    color=Drug),
    size=3)   +
      geom_abline(slope=1, intercept=0) +
  geom_abline(slope=1, intercept=1, linetype=3) +
  geom_abline(slope=1, intercept=-1, linetype=3) +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  xlab("httk Predicted (uM*h)") +
  ylab("Observed") +
  scale_x_log10(#limits=c(10^-5,10^3),
                label=scientific_10) +
  scale_y_log10(#limits=c(10^-5,10^3),
                label=scientific_10) +
  ggtitle("Pregnant")+
  theme_bw()  +
  theme(legend.position="none")+
  theme( text  = element_text(size=10))+
  theme(legend.text=element_text(size=8))

#print(FigEb)
cat(paste(
  "We observe an AFE of",
  signif(mean(TKstats$Predicted.Pregnant.httk/TKstats$Observed.Pregnant5),2),
  "and a RMSE (log10-scale) of",
  signif((mean((log10(TKstats$Predicted.Pregnant.httk)-log10(TKstats$Observed.Pregnant5))^2))^(1/2),2),
  "for pregnant woman.\n"))
#> We observe an AFE of 1.2 and a RMSE (log10-scale) of 0.95 for pregnant woman.



FigEc  <- ggplot(data=subset(TKstats,Parameter=="AUCinf")) +
  geom_point(aes(
    x=Predicted.Non.Pregnant.httk,
    y=Predicted.Pregnant.httk,
    shape=Dose.Route,
    color=Drug),
    size=3)   +
      geom_abline(slope=1, intercept=0) +
  geom_abline(slope=1, intercept=1, linetype=3) +
  geom_abline(slope=1, intercept=-1, linetype=3) +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6),name="Route")+
  ylab("httk Pregnant (uM*h)") +
  xlab("httk Non-Pregnant (uM*h)") +
  scale_x_log10(#limits=c(10^-1,10^2),
                label=scientific_10) +
  scale_y_log10(#limits=c(10^-1,10^2),
                label=scientific_10) +
  ggtitle("Model Comparison") +
  theme_bw()  +
  theme(legend.position="left")+
  theme( text  = element_text(size=10))+
  theme(legend.text=element_text(size=8))

#print(FigEc)

FigEd  <- ggplot(data=subset(TKstats,Parameter=="AUCinf" &
  !is.na(Ratio.httk))) +
  geom_point(aes(
    y=Ratio.obs,
    x=Ratio.httk,
    shape=Dose.Route,
    color=Drug),
    size=3)   +
  scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  xlab("httk Predicted") +
  ylab("Observed") +
 # scale_x_continuous(limits=c(0.25,3)) +
 # scale_y_continuous(limits=c(0.25,3)) +
  geom_abline(slope=1, intercept=0) +
```

```
  geom_abline(slope=1, intercept=1.15, linetype=3) +
  geom_abline(slope=1, intercept=-1.15, linetype=3) +
  ggtitle("Ratio") +
  theme_bw()  +
  theme(legend.position="none")+
  theme( text  = element_text(size=10))+
  theme(legend.text=element_text(size=8))

dev.new()
grid.arrange(FigEa,FigEb,FigEc,FigEd,nrow=2)

write.csv(subset(TKstats,Parameter=="AUCinf" &
  !is.na(Ratio.httk)),
  file="DallmanTable.txt")
```

Evaluation of Partition Coefficient Predictions

For each tissue, the partition coefficient (which describes the ratio of the concentration in the tissue to concentration of unbound chemical in the plasma at equilibrium) is a constant. We calculate each partition coefficient using the method of Schmitt, 2008 as described by Pearce et al., 2017. The partition coefficient for any given type of tissue (for example, liver tissue) depends on fraction unbound in plasma (f_{up}^m or f_{up}^f), so in general these differ for mother and fetus.

Curley et al., 1969 compiled data on the concentration of a variety of pesticides in the cord blood of newborns and in the tissues of infants that were stillborn.

```
Curley <- as.data.frame(read_excel("Curley1969.xlsx"))
dim(Curley)
Curley.compounds <- Curley[1,4:13]
Curley <- Curley[4:47,]
colnames(Curley)[1] <- "Tissue"
colnames(Curley)[2] <- "N"
colnames(Curley)[3] <- "Stat"
```

The ratio of chemical concentration in tissue (C_y^f) vs. blood (C_{venb}^f) was related to the tissue-to-unbound-plasma concentration partition coefficients used in the PBTK model as

$$(C_y^f)/(C_{venb}^f) = (C_y^f)/(R_{b:p}^f \times C_{plas}^f) = (C_y^f \times f_{up}^f)/(R_{b:p}^f \times C_{up}^f) = (f_{up}^f)/(R_{b:p}^f) \times K_y^f$$

where C_{up}^f denotes the concentration of substance unbound in the fetal plasma.

```
Curley.pcs <- NULL
cord.blood <- subset(Curley, Tissue == "Cord Blood")
suppressWarnings(
for (this.tissue in unique(Curley$Tissue))
  if (this.tissue != "Cord Blood")
  {
    this.subset <- subset(Curley, Tissue == this.tissue)
    for (this.chemical in colnames(Curley)[4:13])
    {
      if (!is.na(as.numeric(subset(this.subset,Stat=="Mean")[,this.chemical])) &
        !is.na(as.numeric(subset(cord.blood,Stat=="Mean")[,this.chemical])))
      {
        this.row <- data.frame(
          Compound = Curley.compounds[,this.chemical],
          DTXSID = this.chemical,
          Tissue = this.tissue,
          PC = as.numeric(subset(this.subset,Stat=="Mean")[,this.chemical]) /
            as.numeric(subset(cord.blood,Stat=="Mean")[,this.chemical])
          )
        Curley.pcs <- rbind(Curley.pcs,this.row)
      } else if (!is.na((as.numeric(subset(this.subset,Stat=="Range")[,this.chemical]))) &
        !is.na((as.numeric(subset(cord.blood,Stat=="Mean")[,this.chemical]))))
      {
```

```
        this.row <- data.frame(
          Compound = Curley.compounds[,this.chemical],
          DTXSID = this.chemical,
          Tissue = this.tissue,
          PC = as.numeric(subset(this.subset,Stat=="Range")[,this.chemical]) /
            as.numeric(subset(cord.blood,Stat=="Mean")[,this.chemical])
          )
        Curley.pcs <- rbind(Curley.pcs,this.row)
      }
    }
  }
)
Curley.pcs[Curley.pcs$Tissue=="Lungs","Tissue"] <- "Lung"

pearce2017 <- read_excel("PC_Data.xlsx")
pearce2017 <- subset(pearce2017,DTXSID%in%Curley.pcs$DTXSID)
any(pearce2017$DTXSID%in%Curley.pcs$DTXSID)
print("None of the Curley chems were in the Pearce 2017 PC predictor evaluation.")
```

Partition coefficients were measured for tissues, including placenta, in vitro by Csanady et al., 2002 for Bisphenol A and Diadzen.

```
csanadybpa <- read_excel("Csanady2002.xlsx",sheet="Table 2")
csanadybpa$Compound <- "Bisphenol A"
csanadybpa$DTXSID <- "DTXSID7020182"
csanadybpa <- csanadybpa[,c("Compound","DTXSID","Tissue","Mean")]
colnames(csanadybpa) <- colnames(Curley.pcs)

csanadydaid <- read_excel("Csanady2002.xlsx",sheet="Table 3",skip=1)
csanadydaid$Compound <- "Daidzein"
csanadydaid$DTXSID <- "DTXSID9022310"
csanadydaid <- csanadydaid[,c("Compound","DTXSID","Tissue","Mean...6")]
colnames(csanadydaid) <- colnames(Curley.pcs)

Curley.pcs <- rbind(Curley.pcs,csanadybpa,csanadydaid)
Curley.pcs <- subset(Curley.pcs,Tissue!="Mammary gland")
```

Three of the chemicals studied by Curley et al., 1969 were modeled by Weijs et al., 2013 using the same partition coefficients for mother and fetus. The values used represented “prior knowledge” summarizing the available literature.

```
weijstab3 <- read_excel("Weijs2013.xlsx",sheet="Table3",skip=1)
weijstab3 <- subset(weijstab3, !is.na(Compound) & !is.na(Tissue))
weijstab4 <- read_excel("Weijs2013.xlsx",sheet="Table4",skip=1)
weijstab4 <- subset(weijstab4, !is.na(Compound) & !is.na(Tissue))

for (this.compound in unique(weijstab3$Compound))
  for (this.tissue in unique(weijstab3$Tissue))
  {
    Curley.pcs[
      Curley.pcs$DTXSID==this.compound & Curley.pcs$Tissue==this.tissue,
      "Weijs2013"] <- weijstab3[weijstab3$Compound==this.compound &
                                  weijstab3$Tissue==this.tissue,"value"]

  }

for (this.compound in unique(weijstab4$Compound))
  for (this.tissue in unique(weijstab4$Tissue))
  {
    Curley.pcs[
      Curley.pcs$DTXSID==this.compound & Curley.pcs$Tissue==this.tissue,
      "Weijs2013"] <- weijstab4[weijstab4$Compound==this.compound &
                                  weijstab4$Tissue==this.tissue,"value"]

  }

write.csv(Curley.pcs,"PCs-table.csv",row.names=FALSE)
```

Predict partition coefficients

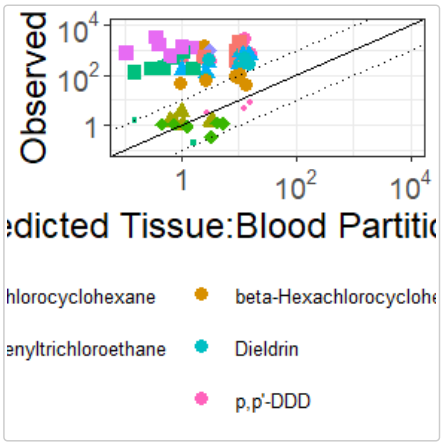
```
Curley.pcs <- httk::FetalPCData

suppressWarnings(load_sipes2017())
#> Loading predictions from Sipes et al. (2017) for 8758 chemicals.
#> Existing data are not being overwritten.
#> Please wait...
for (this.chemical in unique(Curley.pcs$DTXSID))
  if (this.chemical %in% get_cheminfo(info="DTXSID",model="fetal_pbtk"))
  {
    this.subset <- subset(Curley.pcs,DTXSID==this.chemical)
    p <- suppressWarnings(parameterize_fetal_pbtk(dtxsid=this.chemical,
      fetal_fup_adjustment = FALSE,
      suppress.messages=TRUE,
      minimum.Funbound.plasma = 1e-16))
    fetal.blood.pH <- 7.26
    Fup <- p$Fraction_unbound_plasma_fetus
    fetal_schmitt_parms <- suppressWarnings(
      parameterize_schmitt(dtxsid=this.chemical,
        suppress.messages=TRUE,
      minimum.Funbound.plasma = 1e-16))
    fetal_schmitt_parms$plasma.pH <- fetal.blood.pH
    fetal_schmitt_parms$Funbound.plasma <- Fup
    Curley.pcs[Curley.pcs$DTXSID==this.chemical,"Fup"] <- Fup
    # httk gives tissue:fup (unbound chemical in plasma) PC's:
    fetal_pcs <- suppressWarnings(
      predict_partitioning_schmitt(parameters = fetal_schmitt_parms,
        suppress.messages=TRUE,
        model="fetal_pbtk",
        minimum.Funbound.plasma = 1e-16))
    fetal_pcs.nocal <- suppressWarnings(predict_partitioning_schmitt(
      parameters = fetal_schmitt_parms,
      regression=FALSE,
      suppress.messages=TRUE,
      model="fetal_pbtk",
      minimum.Funbound.plasma = 1e-16))
    out <- suppressWarnings(solve_fetal_pbtk(
      dtxsid = this.chemical,
      fetal_fup_adjustment =FALSE,
      suppress.messages=TRUE,
      minimum.Funbound.plasma = 1e-16))
    Rb2p <- out[dim(out)[1],"Rfblood2plasma"]
    Curley.pcs[Curley.pcs$DTXSID==this.chemical,"Rb2p"] <- Rb2p
    # Convert to tissue:blood PC's:
    for (this.tissue in this.subset$Tissue)
      if (tolower(this.tissue) %in%
        unique(subset(tissue.data,Species=="Human")$Tissue))
      {
        Curley.pcs[Curley.pcs$DTXSID==this.chemical &
          Curley.pcs$Tissue == this.tissue, "HTTK.pred.t2up"] <-
          fetal_pcs[[paste("K",tolower(this.tissue),"2pu",sep="")]]
        Curley.pcs[Curley.pcs$DTXSID==this.chemical &
          Curley.pcs$Tissue == this.tissue, "HTTK.pred.nocal.t2up"] <-
          fetal_pcs.nocal[[paste("K",tolower(this.tissue),"2pu",sep="")]]
        Curley.pcs[Curley.pcs$DTXSID==this.chemical &
          Curley.pcs$Tissue == this.tissue, "HTTK.pred"] <-
          fetal_pcs[[paste("K",tolower(this.tissue),"2pu",sep="")]]*Fup/Rb2p
        Curley.pcs[Curley.pcs$DTXSID==this.chemical &
          Curley.pcs$Tissue == this.tissue, "HTTK.pred.nocal"] <-
          fetal_pcs.nocal[[paste("K",tolower(this.tissue),"2pu",sep="")]]*Fup/Rb2p
      } else {
       print(this.tissue)
      }
  } else print(this.chemical)
#> [1] "Spinal Cord"
#> [1] "Adrenals"
```

```
#> [1] "Pancreas"
#> [1] "Spinal Cord"
#> [1] "Adrenals"
#> [1] "DTXSID9020374"
#> [1] "Spinal Cord"
#> [1] "Adrenals"
#> [1] "Adrenals"
#> [1] "Spinal Cord"
#> [1] "Adrenals"
#> [1] "Adrenals"
#> [1] "Spinal Cord"
#> [1] "Adrenals"
#> [1] "Pancreas"
#> [1] "Adrenal Gland"
#> [1] "Adrenal gland"
reset_httk()

cat(paste(
  "For the two placental observations (",
  signif(subset(Curley.pcs,Compound=="Bisphenol A" & Tissue=="Placenta")[,"PC"],2),
  " for Bisphenol A and ",
  signif(subset(Curley.pcs,Compound=="Daidzein" & Tissue=="Placenta")[,"PC"],2),
  " for Diadzen) the predictions were ",
  signif(subset(Curley.pcs,Compound=="Bisphenol A" & Tissue=="Placenta")[,"HTTK.pred"],2),
  " and ",
  signif(subset(Curley.pcs,Compound=="Daidzein" & Tissue=="Placenta")[,"HTTK.pred"],2),
  ", respectively.\n",sep=""))
#> For the two placental observations (1.4 for Bisphenol A and 1.1 for Diadzen) the predictions were
         0.63 and 0.44, respectively.
```

Create Observed vs. Predicted Plot

First indicate compound:

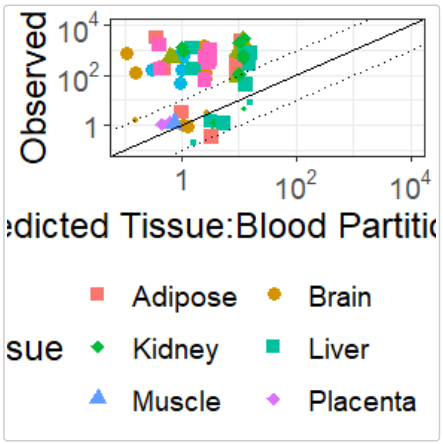
```
FigFa  <- ggplot(data=subset(Curley.pcs,!is.na(HTTK.pred.nocal))) +
  geom_point(size=1,aes(
    y=Weijs2013,
    x=HTTK.pred,
    shape=Compound,
    color=Compound)) +
  geom_point(size=3,aes(
    y=PC,
    x=HTTK.pred,
    shape=Compound,
    color=Compound)) +
  geom_abline(slope=1, intercept=0) +
  geom_abline(slope=1, intercept=1, linetype=3) +
  geom_abline(slope=1, intercept=-1, linetype=3) +
  scale_shape_manual(values=rep(c(15,16,17,18),4))+
  xlab("httk Predicted Tissue:Blood Partition Coefificent") +
  ylab("Observed") +
  scale_x_log10(label=scientific_10,limits=c(10^-1,10^4)) +
  scale_y_log10(label=scientific_10,limits=c(10^-1,10^4)) +
  theme_bw()  +
  theme(legend.position="bottom")+
  theme( text  = element_text(size=18))+
  theme(legend.text=element_text(size=10))+
  guides(shape=guide_legend(nrow=3,byrow=TRUE))

print(FigFa)
#> Warning: Removed 53 rows containing missing values (geom_point).
```




```
fitFa <- lm(data=Curley.pcs,log10(PC)~log10(HTTK.pred))
RMSE(fitFa)
#> [1] 1.081573
fitFb <- lm(data=Curley.pcs,log10(PC)~log10(HTTK.pred.nocal))
RMSE(fitFb)
#> [1] 1.101518
```

Then indicate tissues:

```
FigFb  <- ggplot(data=subset(Curley.pcs,!is.na(HTTK.pred))) +
  geom_point(size=1,aes(
    y=Weijs2013,
    x=HTTK.pred,
    shape=Tissue,
    color=Tissue)) +
  geom_point(size=3, aes(
    y=PC,
    x=HTTK.pred,
    shape=Tissue,
    color=Tissue))   +
  geom_abline(slope=1, intercept=0) +
  geom_abline(slope=1, intercept=1, linetype=3) +
  geom_abline(slope=1, intercept=-1, linetype=3) +
  scale_shape_manual(values=rep(c(15,16,17,18),4))+
  xlab("httk Predicted Tissue:Blood Partition Coefficent") +
  ylab("Observed") +
  scale_x_log10(label=scientific_10,limits=c(10^-1,10^4)) +
  scale_y_log10(label=scientific_10,limits=c(10^-1,10^4)) +
  theme_bw()  +
  theme(legend.position="bottom")+
  theme( text  = element_text(size=18))+
  theme(legend.text=element_text(size=14)) +
  guides(shape=guide_legend(nrow=3,byrow=TRUE))

print(FigFb)
#> Warning: Removed 53 rows containing missing values (geom_point).
```



Chemical Prioritization on the Basis of Fetal Brain Concentration

```
#Wangchems <- read_excel("Wang2018.xlsx",sheet=3,skip=2)
Wangchems <- httk::Wang2018Chems
maternal.list <- Wangchems$CASRN[Wangchems$CASRN%in%get_cheminfo(model="fetal_pbtk")]
nonvols <- subset(chem.physical_and_invitro.data,logHenry < -4.5)$CAS
nonfluoros <- chem.physical_and_invitro.data$CAS[
  regexpr("fluoro",tolower(chem.physical_and_invitro.data$Compound))==-1]
maternal.list <- maternal.list[maternal.list %in% intersect(nonvols,nonfluoros)]
```

Make predictions for three daily doses from week 13 to full term

```
pred.table1 <- subset(suppressWarnings(get_cheminfo(
  info=c("Compound","CAS","DTXSID","logP","pka_accept","pka_donor"),
  model="fetal_pbtk")),
  CAS %in% maternal.list)
pred.table1$Compound <- gsub("\"","",pred.table1$Compound)

for (this.cas in maternal.list)
{
  Fup <- suppressWarnings(subset(get_cheminfo(info="all"),CAS==this.cas)$Human.Funbound.plasma)
  if (regexpr(",",Fup)!=-1)
  {
    if (as.numeric(strsplit(Fup,",")[[1]][1])==0 |
      (as.numeric(strsplit(Fup,",")[[1]][3])>0.9 &
      as.numeric(strsplit(Fup,",")[[1]][2])<0.11))
    {
      skip <- TRUE
    } else skip <- FALSE
  } else if (Fup== 0)
  {
    skip <- TRUE
  } else skip <- FALSE

  if (!skip)
  {
    out <- suppressWarnings(solve_fetal_pbtk(
      chem.cas=this.cas,
      dose=0,
      daily.dose=1,
      doses.per.day=3,
      fetal_fup_adjustenment=FALSE,
      suppress.messages=TRUE))
    last.row <- which(out[,"time"]>279)
    last.row <- last.row[!duplicated(out[last.row,"time"])]
    pred.table1[pred.table1$CAS==this.cas,"Ratio.pred"] <-
```

```
      mean(out[last.row,"Cfplasma"])/mean(out[last.row,"Cplasma"])
  }
}
```

Create final table holding all predicitons for paper:

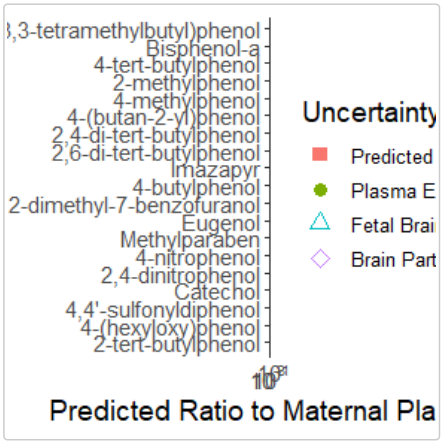
```
PC.table <- pred.table1
colnames(PC.table)[colnames(PC.table)=="Ratio.pred"] <- "R.plasma.FtoM"
pred.table1$Uncertainty <- "Predicted F:M Plasma Ratio"

pred.table3 <- pred.table1
pred.table3$Uncertainty <- "Plasma Error (Fig. 4b)"
empirical.error <- RMSE(fit2sub)
for (this.cas in maternal.list)
{

  pred.table3[pred.table3$CAS==this.cas,"Ratio.pred"] <-
#    pred.table2[pred.table2$CAS==this.cas,"MFratio.pred"]*(1+empirical.error)
    1/((1/pred.table1[pred.table3$CAS==this.cas,"Ratio.pred"])*(1-1.96*empirical.error))
}
# Update final table for paper:
PC.table$RMSE <- RMSE(fit2sub)
PC.table$R.plasma.FtoM.upper <- pred.table3$Ratio.pred

pred.table4 <- pred.table1
pred.table4$Uncertainty <- "Fetal Brain Partioning"
for (this.cas in maternal.list)
{
  p <- suppressWarnings(parameterize_fetal_pbtk(chem.cas=this.cas,
      fetal_fup_adjustment=FALSE,
      suppress.messages=TRUE))
  Kbrain2pu <- p$Kfbrain2pu
  fup <- p$Fraction_unbound_plasma_fetus
#  out <- solve_fetal_pbtk(
#    chem.cas=this.cas,
#    dose=0,
#    daily.dose=1,
#    doses.per.day=3,
#    fetal_fup_adjustenment=FALSE)
#  Rb2p <- out[dim(out)[1],"Rfblood2plasma"]
  pred.table4[pred.table4$CAS==this.cas,"Ratio.pred"] <-
    pred.table3[pred.table3$CAS==this.cas,"Ratio.pred"]*
    Kbrain2pu * fup
  PC.table[PC.table$CAS==this.cas,"Kbrain2pu"] <- Kbrain2pu
  PC.table[PC.table$CAS==this.cas,"fup"] <- fup
#  PC.table[PC.table$CAS==this.cas,"Rb2p"] <- Kbrain2pu
  PC.table[PC.table$CAS==this.cas,"R.brain.FtoM"] <-
    pred.table4[pred.table4$CAS==this.cas,"Ratio.pred"]
}

pred.table5 <- pred.table1
pred.table5$Uncertainty <- "Brain Partitioning Error"
for (this.cas in maternal.list)
{
  p <- suppressWarnings(parameterize_fetal_pbtk(chem.cas=this.cas,
      fetal_fup_adjustment=FALSE,
      suppress.messages=TRUE))
  Kbrain2pu <- p$Kfbrain2pu
  fup <- p$Fraction_unbound_plasma_fetus
#  out <- solve_fetal_pbtk(
#    chem.cas=this.cas,
#    dose=0,
#    daily.dose=1,
#    doses.per.day=3,
#    fetal_fup_adjustenment=FALSE)
#  Rb2p <- out[dim(out)[1],"Rfblood2plasma"]
# From Pearce et al. (2017) PC paper:
```

```
  Kbrain2pu.upper <- Kbrain2pu*10^(1.96*0.647)
  quad.error <- ((RMSE(fit2sub) /
    pred.table1[pred.table1$CAS==this.cas,"Ratio.pred"])^2 +
    (log(10)*10^0.647)^2)^(1/2)
  pred.table5[pred.table5$CAS==this.cas,"Ratio.pred"] <-
    pred.table1[pred.table1$CAS==this.cas,"Ratio.pred"]*Kbrain2pu*
    (1+1.96*quad.error)* fup
  PC.table[PC.table$CAS==this.cas,"Kbrain2pu.upper"] <- Kbrain2pu.upper
  PC.table[PC.table$CAS==this.cas,"Quad.error"] <- quad.error
  PC.table[PC.table$CAS==this.cas,"R.brain.FtoM.upper"] <-
    pred.table5[pred.table5$CAS==this.cas,"Ratio.pred"]
}


pred.levels <- pred.table5$Compound[order(pred.table5$Ratio.pred)]

pred.table <- rbind(
  pred.table1,
#  pred.table2,
  pred.table3,
  pred.table4,
  pred.table5)
pred.table$Compound <- factor(pred.table$Compound,
  levels = pred.levels)

pred.table$Uncertainty <- factor(pred.table$Uncertainty,
  levels = c(pred.table1[1,"Uncertainty"],
#    pred.table2[1,"Uncertainty"],
    pred.table3[1,"Uncertainty"],
    pred.table4[1,"Uncertainty"],
    pred.table5[1,"Uncertainty"]))
```

Make prioritization figure:

```
#Wang 2018 confirmed 6 chemical identities:
confirmed.chemicals <- c(
  "2,4-Di-tert-butylphenol",
  "2,4-Dinitrophenol",
  "Pyrocatechol",
  "2'-Hydroxyacetophenone",
  "3,5-Di-tert-butylsalicylic acid",
  "4-Hydroxycoumarin"
  )
confirmed.chemicals <- c(
  "96-76-4",
  "19715-19-6",
  "51-28-5",
  "120-80-9",
  "118-93-4",
  "1076-38-6")


FigG  <- ggplot(data=pred.table) +
  geom_point(aes(
    x=Ratio.pred,
    y=Compound,
    color = Uncertainty,
    shape = Uncertainty),
    size=3)   +
    scale_shape_manual(values=c(15, 16,2, 23, 0, 1, 17, 5, 6))+
  scale_x_log10(limits=c(10^-2,10^3),label=scientific_10)+
  ylab(expression(paste(
    "Chemicals Found in Maternal Plasma by Wang   ",italic("et al."),"  (2018)"))) +
  xlab("Predicted Ratio to Maternal Plasma") +
  theme_bw()  +
#  theme(legend.position="bottom")+
  theme( text  = element_text(size=14))+
```

```
  theme(legend.text=element_text(size=10))#+
 # guides(color=guide_legend(nrow=4,byrow=TRUE))+
  #guides(shape=guide_legend(nrow=4,byrow=TRUE))
  #+
 # theme(legend.justification = c(0, 0), legend.position = c(0, 0))

print(FigG)
```



Look at ioniation state:

```
for (this.col in 7:14)
  PC.table[,this.col] <- signif(PC.table[,this.col],3)

PC.table <- PC.table[order(PC.table$R.brain.FtoM.upper,decreasing=TRUE),]

for (this.row in 1:dim(PC.table)[1])
{
  out <- calc_ionization(
    pH=7.26,
    pKa_Donor=PC.table[this.row,"pKa_Donor"],
    pKa_Accept=PC.table[this.row,"pKa_Accept"])
  if (out$fraction_neutral>0.9) PC.table[this.row,"Charge_726"] <- "Neutral"
  else if (out$fraction_positive>0.1) PC.table[this.row,"Charge_726"] <-
    paste(signif(out$fraction_positive*100,2),"% Positive",sep="")
  else if (out$fraction_negative>0.1) PC.table[this.row,"Charge_726"] <-
    paste(signif(out$fraction_negative*100,2),"% Negative",sep="")
  else if (out$fraction_zwitter>0.1) PC.table[this.row,"Charge_726"] <-
    paste(signif(out$fraction_zwitter*100,2),"% Zwitterion",sep="")
}

PC.table <- PC.table[,c(
  "Compound",
  "CAS",
  "DTXSID",
  "logP",
  "Charge_726",
  "R.plasma.FtoM",
  "RMSE",
  "R.plasma.FtoM.upper",
  "Kbrain2pu",
  "fup",
  "R.brain.FtoM",
  "Kbrain2pu.upper",
  "R.brain.FtoM.upper")]

write.csv(PC.table,
  file="WangTable.txt",
  row.names=FALSE)
```

The next two figures take a long time to generate and so are disabled by default.

Protein Binding Figures

The fetal fraction unbound (f_{up}^f i) is calculated from the maternal fraction unbound and the serum protein concentration ratio in infants vs. mothers based on Equation 6 of McNamara et al., 2019,

$$f_{up}^f = (f_{up}^m)/(f_{up}^m + (P^f \ / \ P^m) \times (1 - f_{up}^m))$$

in which the maternal fraction unbound, f_{up}^m, is assumed to be equal to the in vitro measured value for fraction unbound in plasma, and the ratio of protein concentrations $P^f \ / \ P^m$ depends on the identity of the dominant binding protein for the chemical (assumed to be either albumin or AAG). Lacking data to model the gestational kinetics of albumin and AAG concentrations, we used the concentrations at birth (McNamara et al., 2019) to calculate a constant fetal f_{up}, using $P^f \ / \ P^m = 0.777$ for albumin and $P^f \ / \ P^m = 0.456$ for AAG (McNamara et al., 2019). We determine the charge state of a compound separately for maternal and fetal plasma as a function of plasma pH (7.38 for maternal and 7.28 for fetal (K.H. Lee, 1972) and chemical-specific predictions for ionization affinity (that is, pKa (Strope et al., 2018) using the “httk” function “calc_ionization” (Pearce et al., 2017). If the fraction of a chemical that is predicted to be in positive ionic form is greater than 50%, we treat the chemical as a base (which is in its conjugate acid form) and use only the maternal-to-infant ratio of AAG concentrations. Otherwise, we assume the chemical is neutral or an acid and use the ratio of albumin concentrations.

```
fup.table <- NULL
all.chems <- suppressWarnings(get_cheminfo(model="fetal_pbtk",info="all"))
# Get rid of median fup 0:
all.chems <- subset(all.chems,
  as.numeric(unlist(lapply(strsplit(
    all.chems$Human.Funbound.plasma,","),function(x) x[[1]])))!=0)
for (this.chem in all.chems[,"CAS"])
{
  temp <- suppressWarnings(
    parameterize_fetal_pbtk(chem.cas=this.chem,suppress.messages = TRUE))
  state <- calc_ionization(
      pH=7.26,
      pKa_Donor=temp$pKa_Donor,
      pKa_Accept=temp$pKa_Accept)
  if (state$fraction_positive > 0.5) this.charge <- "Positive"
  else if (state$fraction_negative > 0.5) this.charge <- "Negative"
  else this.charge <- "Neutral"
  this.row <- data.frame(DTXSID=all.chems[all.chems[,"CAS"]==this.chem,"DTXSID"],
    Compound=all.chems[all.chems[,"CAS"]==this.chem,"Compound"],
    CAS=this.chem,
    Fup.Mat.Pred = temp$Funbound.plasma,
    Fup.Neo.Pred = temp$Fraction_unbound_plasma_fetus,
    Charge = this.charge
    )
  fup.table <- rbind(fup.table,this.row)
}

fup.table[fup.table$Charge=="Positive","Charge"] <- paste("Positive (n=",
  sum(fup.table$Charge=="Positive"),
  ")",sep="")
fup.table[fup.table$Charge=="Negative","Charge"] <- paste("Negative (n=",
  sum(fup.table$Charge=="Negative"),
  ")",sep="")
fup.table[fup.table$Charge=="Neutral","Charge"] <- paste("Neutral (n=",
  sum(fup.table$Charge=="Neutral"),
  ")",sep="")
```

Plot the fetal protein binding changes predicted:

```
FigA  <- ggplot(data=fup.table) +
  geom_point(alpha=0.25, aes(
    x=Fup.Mat.Pred,
    y=Fup.Neo.Pred,
    shape=Charge,
```

```
    color=Charge),
    size=3)   +
  geom_abline(slope=1, intercept=0) +
  ylab(expression(paste("Predicted Neonate ",f[up]))) +
  xlab(expression(paste(italic("In vitro"),," Measured Adult ",f[up]))) +
   scale_x_log10(label=scientific_10) +
   scale_y_log10(label=scientific_10) +
  theme_bw()  +
  theme( text  = element_text(size=14))

print(FigA)
```

Maternal-Fetal Predictions across HTTK Chemical Library:

```
times <- sort(unique(c(seq(13 * 7, 40 * 7, 0.25),seq(278,280,.1))))

MFratio.pred <- NULL
all.chems <- get_cheminfo(model="fetal_pbtk", info=c("Compound","DTXSID","CAS","Funbound.plasma"))
for (this.cas in all.chems$CAS)
  if ((this.cas %in% nonvols) &
    !(this.cas %in% fluoros))
{
  this.id <- all.chems[all.chems$CAS==this.cas,"DTXSID"]
  Fup <- subset(all.chems,DTXSID==this.id)$Human.Funbound.plasma
  if (regexpr(",",Fup)!=-1)
  {
    if (as.numeric(strsplit(Fup,",")[[1]][1])==0 |
      (as.numeric(strsplit(Fup,",")[[1]][3])>0.9 &
      as.numeric(strsplit(Fup,",")[[1]][2])<0.1))
    {
      skip <- TRUE
    } else skip <- FALSE
  } else if (Fup== 0)
  {
    skip <- TRUE
  } else skip <- FALSE

  if (!skip)
  {
    p <- suppressWarnings(parameterize_fetal_pbtk(dtxsid=this.id,
    fetal_fup_adjustment =TRUE,
    suppress.messages=TRUE))
    out <- suppressWarnings(solve_fetal_pbtk(
      parameters=p,
      fetal_fup_adjustment =FALSE,
      dose=0,
      times=times,
      daily.dose=1,
      doses.per.day=3,
      output.units = "uM",
      suppress.messages=TRUE))
    last.row <- which(out[,"time"]>279)
    last.row <- last.row[!duplicated(out[last.row,"time"])]
    new.row <- data.frame(
      Chemical = all.chems[all.chems$DTXSID==this.id,"Compound"],
      DTXSID = this.id,
      Mat.pred = mean(out[last.row,"Cplasma"]),
      Fet.pred = mean(out[last.row,"Cfplasma"]),
      MFratio.pred = mean(out[last.row,"Cplasma"])/mean(out[last.row,"Cfplasma"])
      )
    MFratio.pred <- rbind(MFratio.pred,new.row)
  }
  }
```

Histogram of maternal-fetal ratio

```
FigD <- ggplot(data=MFratio.pred)+
   geom_histogram(binwidth = 0.05,fill="Red",aes(MFratio.pred))+
  xlab("Maternal:Fetal Plasma Concentration Ratio") +
  ylab("Number of chemicals")+
    theme_bw()+
    theme( text  = element_text(size=14))


print(FigD)
```

Statistics on maternal-fetal ratio for full HTTK library

```
max.chem <- MFratio.pred[which(
  MFratio.pred$MFratio.pred==max(MFratio.pred$MFratio.pred,na.rm=TRUE)),]
min.chem <- MFratio.pred[which(
  MFratio.pred$MFratio.pred==min(MFratio.pred$MFratio.pred,na.rm=TRUE)),]
cat(paste("In Figure X we examine the ratios predicted for the ",
  dim(MFratio.pred)[1],
  " appropriate (non-volatile or PFAS) chemicals with measured HTTK data.\n",
  sep=""))


cat(paste("We observe a median value of ",
  signif(median(MFratio.pred$MFratio.pred,na.rm=TRUE),3),
  " ranging from ",
  signif(min.chem[,"MFratio.pred"],3),
  " for ",
  min.chem[,"DTXSID"],
  " to ",
  signif(max.chem[,"MFratio.pred"],3),
  " for ",
  max.chem[,"DTXSID"],
  ".\n",sep=""))

# Check out phys-chem > 1.6, < 1:
highratio <-
         subset(chem.physical_and_invitro.data,DTXSID%in%subset(MFratio.pred,MFratio.pred>1.6)$DTXSID)
# all highly bound
highratio$Compound
suppressWarnings(apply(highratio,2,function(x) mean(as.numeric(x),na.rm=2)))


lowratio <-
         subset(chem.physical_and_invitro.data,DTXSID%in%subset(MFratio.pred,MFratio.pred<0.9)$DTXSID)
# No obvious pattern
```

Code used to create data distributed with vignette

```
AylwardMatFet <-MFdata
Dallmann2018Data <- TKstats
FetalPCData <- Curley.pcs
Wang2018Chems <- Wangchems

save(AylwardMatFet,Dallmann2018Data,FetalPCData,Wang2018Chems,
     file="Kapraun2022Vignette.RData",version=2)
```