Applying Urinary Biomarkers of 11-dehydrothromboxane B2 and 8-isoprostane to Understand the Health Effects of NO2 Exposure

https://github.com/Yang190809/data_repository_yang

Yang Wang

Abstract

Mounting evidence shows that exposure to NO2 generates oxidative stress. Oxidative stress can cause lipid damage, which plays an important role in various respiratory and cardiovascular diseases. Urinary 8-isoprostane is a product of lipid peroxidation which can reflect systemic lipid damage. I built linear mix models to analyze the relationship between the level of urinary 8-isoprostane with the level of NO2 exposure. I found that short term NO2 exposure was associated with lipid peroxidation, reflected by increased concentrations of urinary 8-isoprostane associated with increasing exposure. One IQR (7.41 µg/m3) incremental change of 12-h NO2 exposure was associated with an increase of 8-isoprostane level by 19.45% (95% Cl: 14.37%, 24.55%, p-value <0.001). One IQR (8.64 µg/m3) incremental change of 24-h NO2 exposure was associated an increase in 8-isoprostane level by 27.69% (95% Cl: 20.50%, 34.95% , p-value <0.001). One IQR (4.47 µg/m3) incremental change of 1-week NO2 exposure was associated with an increase in 8-isoprostane level by 25.15% (95% Cl: 14.14%,36.31%, p-value <0.05). One IQR (2.74 µg/m3) incremental change of 2-week NO2 exposure was associated with an increase in 8-isoprostane level by 15.68% (95% Cl: 8.87%,22.65%, p-value <0.05).

Contents

1 Rationale and Research Questions

Exposure to NO2 associated with cardiovascular disease, lung function impairment and asthma (Mölter, A., et. al.,2013, Collart, P., et. al.,2018, Takenoue, Y., et. al., 2012). Mounting evidence shows that exposure to NO2 generates oxidative stress (Hashemzadeh, B., et. al.,2019). Oxidative stress can cause lipid damage (Black, C. N., et. al., 2017), which plays an important role in various respiratory and cardiovascular diseases. (Zanolin, M. E., et. a.,2015, Lakshmi, S. V., et. al., 2009). I applied 8-isoprostane to investigate the health effects of short term NO2 exposure.

8-isoprostane is the product of the oxidized cell membrane after being attacked by reactive oxygen species such as peroxides. Therefore, 8-isoprostane can reflect lipid damage (Danielsen et al., 2009). 8-isoprostane in urine does not have a diurnal variation and has been proven to be a stable biomarker showing systemic lipid damage. It has been applied in studying diseases such as type 2 diabetes mellitus, obesity, coronary heart disease, asthma, and acute respiratory distress syndrome (Nuernberg, A. M., et al., 2008). Nevertheless, the use of this biomarker in the research of air pollution and its health effects is scarce. One study in Iran discovered a significant positive relationship between short-term NO2 exposure and 8-isoprostane in exhaled breath condensate (EBC) in healthy children aged 12-13 years old (Hashemzadeh, B., et. al., 2019). One study in New York City found that the increases in 1- to 5-day averages of nitrogen dioxide were significantly associated with increases in 8-isoprostane in EBC among 18-year old healthy and asthma affected individuals (Patel, M. M., et. al., 2013). My research question is: If the NO2 exposure is positively associated with urinary 8-isoprostane? Among periods of 12-hour,24-hour,1-week and 2-weeks, which period has the most significant association with urinary 8-isoprostane?

2 Dataset Information

All the data are from Jim Zhang's lab. Jim Zhang is a professor at the Nicholas School of Environment at Duke University. The file 8-iso is the data set of urinary biomarker 8-isoprostane. The file urine_info is the data set that has information about the subjects who provided the urine samples. The urine list is a data set with lists of information of sample ID, subject ID and visits. The meaning of the columns as well as units and class of the data in each foler is listed below. Column names without descriptors are irrelevant to this study.

2.1 Urine list raw data set

The urine list has the information about the sample ID, subject ID and visits.

Column Name	Meaning	units	class of the data
Sample ID	the identity number of the samples	NA	integer
Subject ID	the identity number of the subjects	NA	integer
visit	the number of the 4 visits (1,2,3 or 4)	NA	integer

Table 1

2.2 8-isoprostane data set

The 8-isoprostane data set has information about 8-isoprostane concentration tested in the mass spectrum and the concentration of 8-isoprostane in original urine samples. The limit of detection for 8-isoprostane was 0.016ng/ml. Any value which is below 0.016 in the column of calculated Conc in 8-is data set should be excluded as an error.

Column Name	Meaning	Units	Class of the data
Sample ID	the identity number of the samples	NA	integer
Calculated Conc	the concentration tested by the machine	ng/ml	numeric
Sample Conc	the concentration in the original urine	ng/ml	numeric

Table 2

2.3 Urine Info

The urine info is the data set that has information about subjects characterictics and average pollutants exposure (NO2, SO2, O3 and PM2.5)over the periods of 12 hours, 24 hours, one week and two weeks. It also includes average temperature and humidity over the periods of 12 hours, 24 hours, one week and two weeks.

Column name	Meaning	Units	Class of the data
ample_ID	the identity number of the samples	NA	integer
SubjectID	the identity number of the subjects	NA	integer
COLD	cold (represent respiratory infection)	NA	category
MNST	menstration during visit	NA	category
last.ate	the hours to the last meal	hours	integer
wkday.start	the day that the subject start their work	NA	category
dt_smoke	second-hand smoke exposure in hours	hours	numeric
USG urine	specific gravity	g/ml	numeric
o3exp.12h	the exposure of ozone in 12h	ug/m3	numeric
pmexp.12h	the exposure of PM2.5 in 12h	ug/m3	numeric
no2exp.12h	the exposure of NO2 in 12h	ug/m3	numeric
so2exp.12h	the exposure of SO2 in 12h	ug/m3	numeric
Temp.12h	temperature in 12h	ug/m3	numeric
RHx.12h	humidity in 12h	ug/m3	numeric

Table 3

2.4 Data Wrangling

My data wrangling started with loading the data set of 8-isoprostane and changing the column names. Then I merged this data set with the urine list to match the sample ID with subjects ID and visit. Then this merged data was merged with the urine info data set using subject ID and visit as the matching keys. After the second merging, the rows with NAs were removed. Then I calculated the normalized 8-isoprostane using the specific urine gravity. This normlization can adjust the dilution in urine. I obtained my final data by selecting columns that is revelant to my rearch question. Finally, the processed data was saved into a processed data folder.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
8-isoprostane	302	3.806	5.780	0.186	0.747	4.005	48.299
11dhTxB2	302	1.510	1.259	0.141	0.812	1.861	14.255

Table 4

1. Data wrangling

```
# Set working directory
getwd()
```

```
## [1] "C:/Users/26059/OneDrive/Desktop/ENV 872 R/Yang_ENV872/directory_yang"
```

```
# Load packages
library(tidyverse)
library(lmerTest)
library(MuMIn)
library(car)
library(tidyverse)
library(cowplot)
library(RColorBrewer)
# Set ggplot theme
mytheme <- theme_classic(base_size = 25) +
  theme(axis.text = element_text(color = "black"),
        legend.position = "top")
theme_set(mytheme)

# Load dataset 1
is <- read.csv("raw_data/8iso.csv")
is <- select(is, Sample.ID, Calculated.Conc, Sample.Conc)
colnames(is)[colnames(is) == "Calculated.Conc"] <- "is.origin"
colnames(is)[colnames(is) == "Sample.Conc"] <- "is.conc"

# Load dataset 2
list <- read.csv("raw_data/urine_list.csv")

#merge
merge1 <- merge(x=is,y=list,by="Sample.ID",all=TRUE)

urine.info <- read.csv("raw_data/urine_info.csv")
merge3 <- merge(x=merge1,y=urine.info,by=c("Subject.ID","visit"),all=TRUE)
dim(merge3)
```

```
## [1] 344  69
```

```
full.data <- na.omit(merge3)
dim(full.data)
```

```
## [1] 307  69
```

```r
mean.cre<- mean(full.data$USG)
mean.cre
```

```
## [1] 1.016117
```

```r
#calculate normalized biomarker concentration
full.data$is.ad <- (full.data$is.conc/(1-full.data$USG))*(1-mean.cre)

#select useful colums
final.dat<- select(full.data, Sample.ID, Subject.ID, is.ad, is.conc, is.origin, group,

#save processed data set
write.csv(final.dat, row.names = FALSE, file = "processed_data/final.dat.csv")
```

2.Explore the data

```r
library(ggplot2)
#load processed data set
dat <- read.csv("processed_data/final.dat.csv")
#remove outliers
summary(dat$is.origin)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.508   6.220  13.421  45.703  38.281 605.614
```

```r
dat1<- dat[-c(4,7,89,147,167),]

#test normality
shapiro.test(dat1$is.ad)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  dat1$is.ad
## W = 0.58916, p-value < 2.2e-16
```

```r
shapiro.test(log(dat1$is.ad))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  log(dat1$is.ad)
## W = 0.98127, p-value = 0.0005511
```

```
#explore the data
is.plot1 <- boxplot(dat1$is.ad,colors="darkblue",ylab="8-isoprostane ng/ml")
```

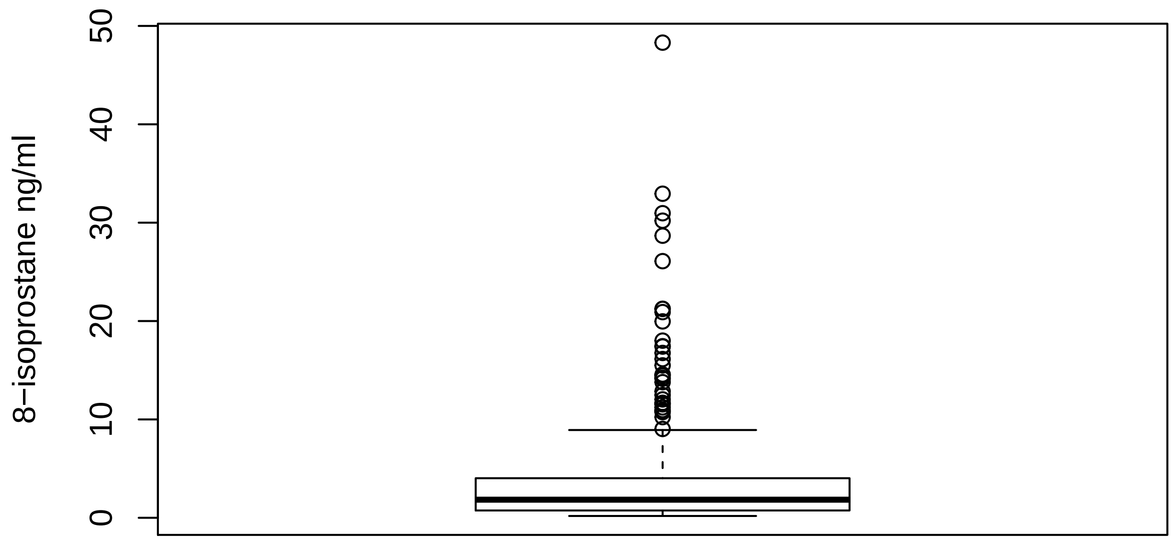

Figure 1: Plot of 8-isoprostane

Figure 1 shows that the range of concentraion of 8-isoprostane is huge and it is not normlized

```
is.plot1 <- boxplot(log(dat1$is.ad),color="darkblue",ylab="log(8-isoprostane ng/ml)")
```

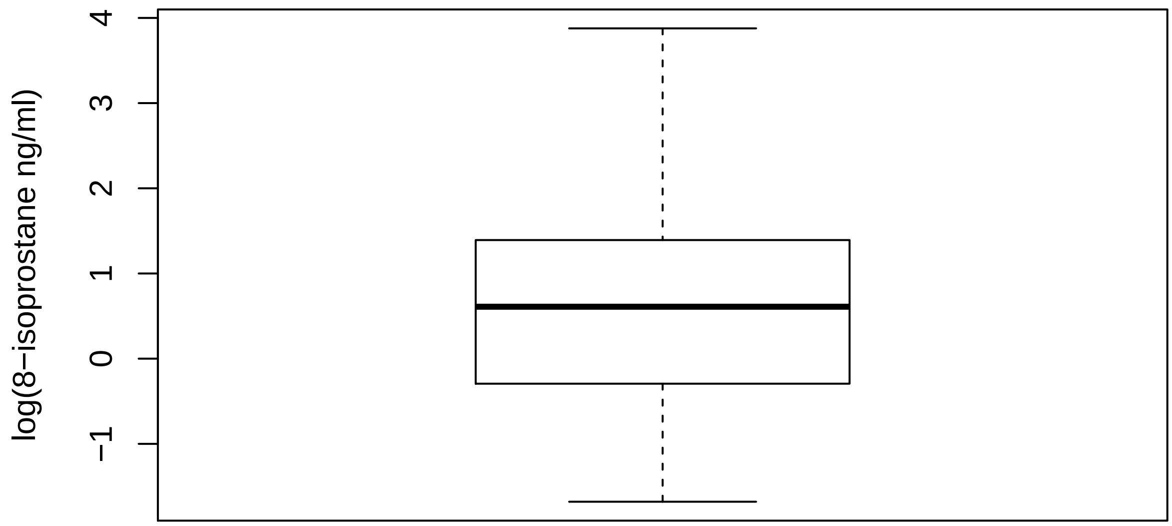

Figure 2: Plot of log(8-isoprostane)

Figure 2 shows that the log(8-isoprostane) looks more normlized than original scale of 8-isoprostane.

```
#calculate IQR for each period
sum3<-summary(dat$no2exp.12h)
no2.12 <-sum3[5]-sum3[2]
no2.12
```

```
##  3rd Qu.
## 7.409404
```

```
sum4<-summary(dat$no2exp.24h)
no2.24 <-sum4[5]-sum4[2]
no2.24
```

```
##  3rd Qu.
## 8.638027
```

```
sum5<-summary(dat$no2exp.1w)
no2.1w <-sum5[5]-sum5[2]
no2.1w
```

```
##  3rd Qu.
## 8.617326
```

```
sum6<-summary(dat$no2exp.2w)
no2.2w <-sum6[5]-sum6[2]
no2.2w
```

```
##  3rd Qu.
## 2.736119
```

3.Build models

3.1 Build models for 12-hour NO2 exposure

```
mo1<- lmer(log(is.ad) ~no2exp.12h+o3exp.12h+ pmexp.12h+so2exp.12h+Temp.12h+RHx.12h + las
#find the best model with lowest AIC
step(mo1)
```

```
## Backward reduced random-effect table:
##
##                  Eliminated npar  logLik     AIC    LRT Df Pr(>Chisq)
## <none>                         16 -474.63  981.26
## (1 | Subject.ID)          0    15 -493.94 1017.88 38.616  1  5.159e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##             Eliminated  Sum Sq Mean Sq NumDF   DenDF F value   Pr(>F)
## Temp.12h             1  0.0261  0.0261     1 276.301  0.0314 0.859411
## COLD                 2  0.0677  0.0677     1 266.912  0.0819 0.774976
## pmexp.12h            3  0.0956  0.0956     1 271.141  0.1159 0.733779
## go.home              4  0.6585  0.3293     2  83.839  0.4012 0.670812
## wkday.start          5  0.3986  0.3986     1 291.709  0.4848 0.486792
## so2exp.12h           6  0.4401  0.4401     1 293.944  0.5343 0.465386
## MNST                 7  0.8942  0.8942     1 272.520  1.0861 0.298261
## last.ate             8  0.8829  0.8829     1 276.178  1.0773 0.300216
## dt_smoke             9  0.9658  0.9658     1 293.061  1.1845 0.277332
```

```
## RHx.12h             10  2.0108  2.0108     1 251.786  2.4885 0.115939
## no2exp.12h           0 12.2057 12.2057     1 228.742 14.9910 0.000141 ***
## o3exp.12h            0  4.0334  4.0334     1 225.577  4.9538 0.027023 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Model found:
## log(is.ad) ~ no2exp.12h + o3exp.12h + (1 | Subject.ID)
```

```
#best model for 12-hour NO2 exposure
mo1.1<- lmer(log(is.ad) ~ no2exp.12h +o3exp.12h + (1 | Subject.ID),data=dat1)
summary(mo1.1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: log(is.ad) ~ no2exp.12h + o3exp.12h + (1 | Subject.ID)
##    Data: dat1
##
## REML criterion at convergence: 910.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.99963 -0.62729 -0.03595  0.60429  2.58701
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Subject.ID (Intercept) 0.5187   0.7202
##  Residual               0.8142   0.9023
## Number of obs: 302, groups:  Subject.ID, 89
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  -0.187878   0.212372 292.646414  -0.885 0.377064
## no2exp.12h    0.025912   0.006693 228.741989   3.872 0.000141 ***
## o3exp.12h     0.031066   0.013958 225.576673   2.226 0.027023 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) n2x.12
## no2exp.12h -0.812
## o3exp.12h  -0.489  0.132
```

```r
#check F value
anova(mo1.1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF  DenDF F value   Pr(>F)
## no2exp.12h 12.2057 12.2057     1 228.74 14.9910 0.000141 ***
## o3exp.12h   4.0334  4.0334     1 225.58  4.9538 0.027023 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#check collineariry
vif(mo1.1)
```

```
## no2exp.12h  o3exp.12h
##   1.017673   1.017673
```

```r
#get R^2
r.squaredGLMM (mo1.1)
```

```
##              R2m      R2c
## [1,] 0.03761012 0.412126
```

```r
#identify outliers
res1.1 <- resid(mo1, type = "pearson")
dat[which(abs(res1.1) > 2.5),]
```

```
##  [1] Sample.ID   Subject.ID  is.ad       is.conc     is.origin   group
##  [7] COLD        MNST        last.ate    wkday.start go.home     Smoker
## [13] dt_smoke    o3exp.12h   pmexp.12h   no2exp.12h  so2exp.12h  Temp.12h
## [19] RHx.12h     o3exp.24h   pmexp.24h   no2exp.24h  so2exp.24h  Temp.24h
## [25] RHx.24h     o3exp.1w    pmexp.1w    no2exp.1w   so2exp.1w   Temp.1w
## [31] RHx.1w      o3exp.2w    pmexp.2w    no2exp.2w   so2exp.2w   Temp.2w
## [37] RHx.2w
## <0 rows> (or 0-length row.names)
```

```r
#calculate IQR change
(exp(0.025912)-1)*no2.12
```

```
##   3rd Qu.
## 0.1945016
```

```
(exp(0.025912 +0.006693)-1)*no2.12
```

```
##   3rd Qu.
## 0.2455652
```

```
(exp(0.025912 -0.006693)-1)*no2.12
```

```
##   3rd Qu.
## 0.1437785
```

```
#plot residues
plot(mo1.1)
```

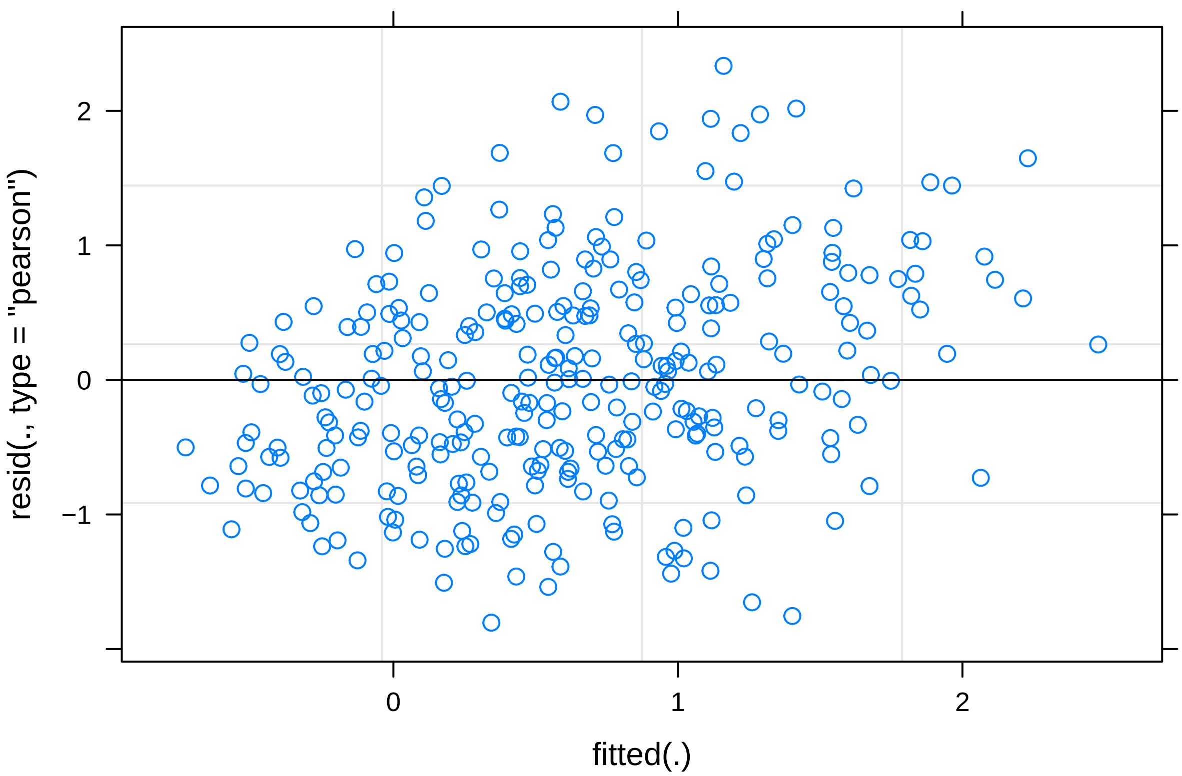

Figure 3: diagnostic plot of model 1.1

Figure 3 shows that the higher the fitted value, the higher the residue. However, in general it looks fine.

3.2 Build models for 24-hour NO2 exposure

```r
mo2<- lmer(log(is.ad) ~ no2exp.24h+o3exp.24h+ pmexp.24h+so2exp.24h+Temp.24h+RHx.24h + la
step(mo2)
```

```
## Backward reduced random-effect table:
##
##                    Eliminated npar  logLik     AIC    LRT Df Pr(>Chisq)
## <none>                          16 -475.04  982.07
## (1 | Subject.ID)           0    15 -493.92 1017.84 37.768  1  7.969e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##              Eliminated  Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## so2exp.24h            1  0.0016  0.0016     1 280.151  0.0019 0.9655516
## Temp.24h              2  0.0161  0.0161     1 277.135  0.0193 0.8895889
## wkday.start           3  0.0126  0.0126     1 279.050  0.0152 0.9021154
## COLD                  4  0.1555  0.1555     1 268.565  0.1873 0.6655198
## go.home               5  0.9235  0.4617     2  84.443  0.5575 0.5747522
## RHx.24h               6  0.4390  0.4390     1 258.415  0.5298 0.4673562
## dt_smoke              7  0.5832  0.5832     1 291.289  0.7037 0.4022322
## pmexp.24h             8  0.8868  0.8868     1 273.549  1.0799 0.2996466
## MNST                  9  0.8850  0.8850     1 274.024  1.0795 0.2997231
## last.ate             10  1.0377  1.0377     1 281.175  1.2708 0.2605860
## no2exp.24h            0 12.3002 12.3002     1 231.652 15.1608 0.0001292 ***
## o3exp.24h             0  5.0095  5.0095     1 225.520  6.1746 0.0136869 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Model found:
## log(is.ad) ~ no2exp.24h + o3exp.24h + (1 | Subject.ID)
```

```r
mo2.1<- lmer(log(is.ad) ~ no2exp.24h + o3exp.24h + (1 | Subject.ID),data=dat1)
summary(mo2.1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: log(is.ad) ~ no2exp.24h + o3exp.24h + (1 | Subject.ID)
##    Data: dat1
##
## REML criterion at convergence: 910.4
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.88882 -0.65529 -0.03642  0.58790  2.66141
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Subject.ID (Intercept) 0.5243   0.7241
##  Residual               0.8113   0.9007
## Number of obs: 302, groups:  Subject.ID, 89
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  -0.339832   0.243838 284.149326  -1.394 0.164503
## no2exp.24h    0.031555   0.008104 231.652052   3.894 0.000129 ***
## o3exp.24h     0.030798   0.012394 225.520144   2.485 0.013687 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) n2x.24
## no2exp.24h -0.857
## o3exp.24h  -0.505  0.196
```

```r
anova(mo2.1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##             Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## no2exp.24h 12.3002 12.3002     1 231.65 15.1608 0.0001292 ***
## o3exp.24h   5.0095  5.0095     1 225.52  6.1746 0.0136869 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
vif(mo2.1)
```

```
## no2exp.24h  o3exp.24h
##   1.039908   1.039908
```

```r
r.squaredGLMM (mo2.1)
```

```
##             R2m       R2c
## [1,] 0.03854593 0.4159847
```

```r
res2.1 <- resid(mo2.1, type = "pearson")
dat[which(abs(res2.1) > 2.5),]
```

```
##  [1] Sample.ID   Subject.ID  is.ad       is.conc     is.origin   group
##  [7] COLD        MNST        last.ate    wkday.start go.home     Smoker
## [13] dt_smoke    o3exp.12h   pmexp.12h   no2exp.12h  so2exp.12h  Temp.12h
## [19] RHx.12h     o3exp.24h   pmexp.24h   no2exp.24h  so2exp.24h  Temp.24h
## [25] RHx.24h     o3exp.1w    pmexp.1w    no2exp.1w   so2exp.1w   Temp.1w
## [31] RHx.1w      o3exp.2w    pmexp.2w    no2exp.2w   so2exp.2w   Temp.2w
## [37] RHx.2w
## <0 rows> (or 0-length row.names)
```

```r
#calculate IQR change
(exp(0.031555)-1)*no2.24
```

```
##  3rd Qu.
## 0.276919
```

```r
(exp(0.031555 +0.008104)-1)*no2.24
```

```
##   3rd Qu.
## 0.3494593
```

```r
(exp(0.031555 -0.008104)-1)*no2.24
```

```
##   3rd Qu.
## 0.2049643
```

```r
plot(mo2.1)
```

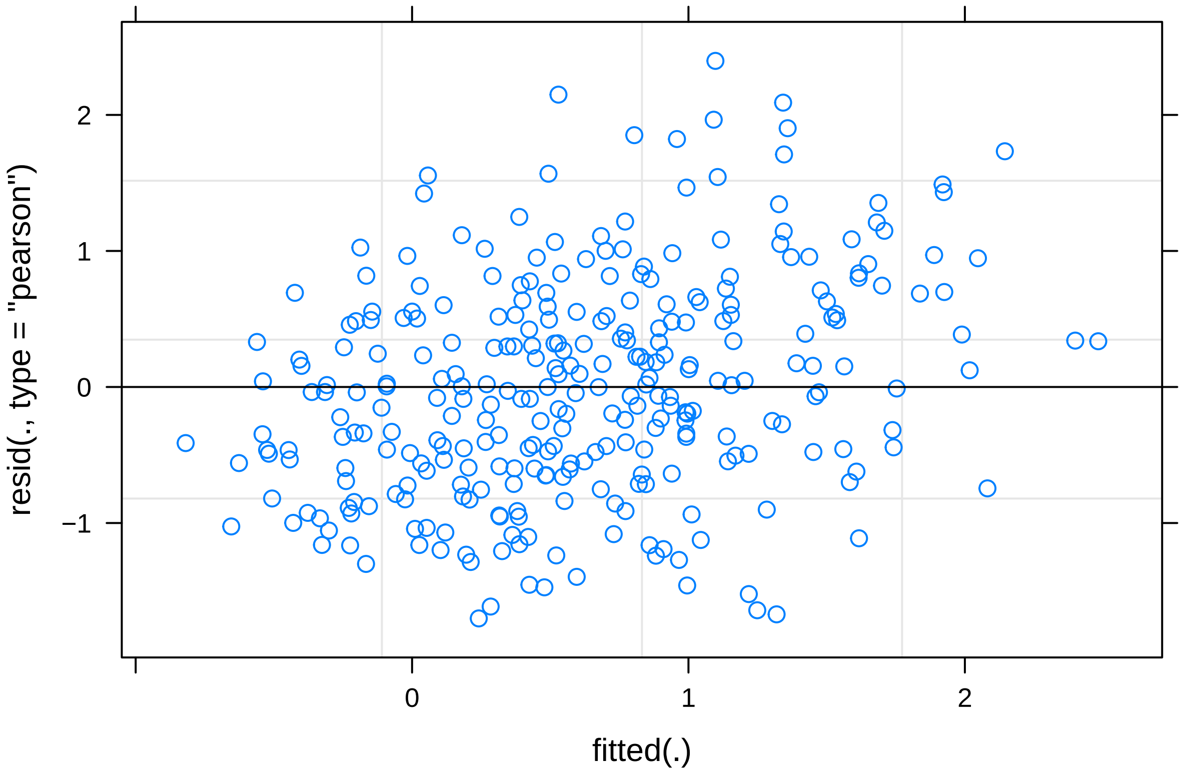

Figure 4: diagnostic plot of model 2.1

Figure 4 shows that the higher the fitted value, the higher the residue. However, in general it looks fine. 3.3 Build models for 1-week NO2 exposure

```
mo3<- lmer(log(is.ad) ~ no2exp.1w+o3exp.1w+ pmexp.1w+so2exp.1w+Temp.1w+RHx.1w + last.ate
step(mo3)
```

```
## Backward reduced random-effect table:
##
##                    Eliminated npar  logLik     AIC    LRT Df Pr(>Chisq)
## <none>                          16 -475.45  982.89
## (1 | Subject.ID)            0   15 -493.45 1016.90 36.001  1  1.972e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##               Eliminated Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Temp.1w                1 0.0000  0.0000     1 256.639  0.0000 0.9986105
```

```
## RHx.1w              2 0.0271  0.0271     1 288.772  0.0317 0.8588147
## COLD                3 0.0560  0.0560     1 271.043  0.0657 0.7979384
## wkday.start         4 0.1240  0.1240     1 258.909  0.1459 0.7027697
## go.home             5 0.8717  0.4359     2  84.614  0.5134 0.6002845
## pmexp.1w            6 0.2371  0.2371     1 293.316  0.2793 0.5975635
## dt_smoke            7 0.4455  0.4455     1 291.942  0.5270 0.4684347
## last.ate            8 0.6026  0.6026     1 277.597  0.7195 0.3970452
## MNST                9 0.7572  0.7572     1 273.657  0.9102 0.3409085
## o3exp.1w           10 1.4655  1.4655     1 229.000  1.7684 0.1849006
## no2exp.1w           0 4.4044  4.4044     1 229.706  5.2961 0.0222690 *
## so2exp.1w           0 9.8969  9.8969     1 230.979 11.9006 0.0006669 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Model found:
## log(is.ad) ~ no2exp.1w + so2exp.1w + (1 | Subject.ID)
```

```
mo3.1<- lmer(log(is.ad) ~ no2exp.1w + so2exp.1w + (1 | Subject.ID),data=dat1)
summary(mo3.1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: log(is.ad) ~ no2exp.1w + so2exp.1w + (1 | Subject.ID)
##    Data: dat1
##
## REML criterion at convergence: 913.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.95288 -0.61932 -0.07413  0.62697  2.63863
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Subject.ID (Intercept) 0.5199   0.7211
##  Residual               0.8316   0.9119
## Number of obs: 302, groups:  Subject.ID, 89
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   0.76276    0.29669 261.12832   2.571 0.010698 *
## no2exp.1w     0.02877    0.01250 229.70566   2.301 0.022269 *
## so2exp.1w    -0.12342    0.03578 230.97927  -3.450 0.000667 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation of Fixed Effects:
##           (Intr) n2xp.1
## no2exp.1w -0.624
## so2exp.1w -0.331 -0.476
```

```r
anova(mo3.1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## no2exp.1w 4.4044  4.4044     1 229.71  5.2961 0.0222690 *
## so2exp.1w 9.8969  9.8969     1 230.98 11.9006 0.0006669 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
vif(mo3.1)
```

```
## no2exp.1w so2exp.1w
##  1.293508  1.293508
```

```r
r.squaredGLMM (mo3.1)
```

```
##             R2m       R2c
## [1,] 0.02719595 0.4014215
```

```r
#calculate IQR change
(exp( 0.02877)-1)*no2.1w
```

```
##   3rd Qu.
## 0.2515213
```

```r
(exp( 0.02877 +0.01250)-1)*no2.1w
```

```
##   3rd Qu.
## 0.3630776
```

```r
(exp( 0.02877 -0.01250)-1)*no2.1w
```

```
##   3rd Qu.
## 0.1413507
```

```r
plot(mo3.1)
```

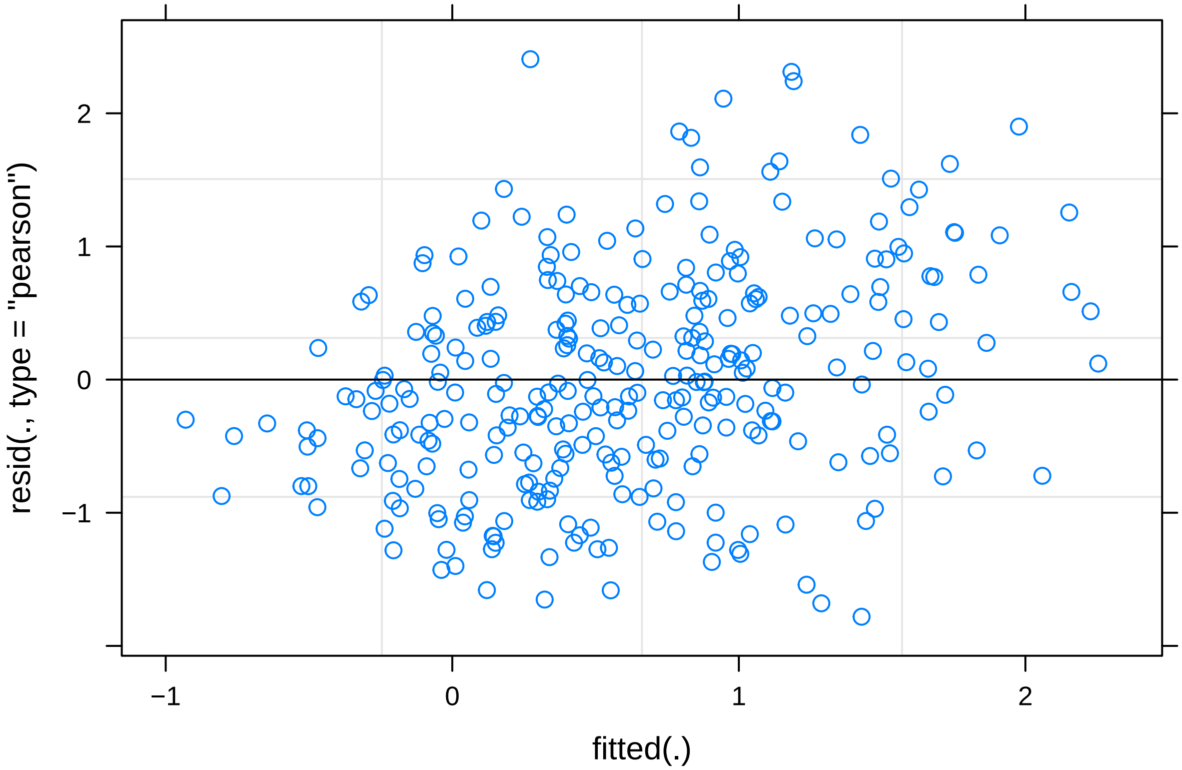

Figure 5: diagnostic plot of model 3.1

Figure 5 shows that the higher the fitted value, the higher the residue. However, in general it looks fine. 3.4 Build models for 2-week NO2 exposure

```r
mo4<- lmer(log(is.ad) ~ no2exp.2w+o3exp.2w+ pmexp.2w+so2exp.2w+Temp.2w+RHx.2w + last.ate
summary(mo4)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: log(is.ad) ~ no2exp.2w + o3exp.2w + pmexp.2w + so2exp.2w + Temp.2w +
##     RHx.2w + last.ate + wkday.start + go.home + COLD + MNST +
##     dt_smoke + (1 | Subject.ID)
##    Data: dat1
##
## REML criterion at convergence: 946.3
##
## Scaled residuals:
```

```
##      Min       1Q   Median       3Q      Max 
## -2.03734 -0.65274 -0.04204  0.61324  2.64864 
## 
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Subject.ID (Intercept) 0.4961   0.7043  
##  Residual               0.8617   0.9283  
## Number of obs: 302, groups:  Subject.ID, 89
## 
## Fixed effects:
##                  Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     -0.293745   2.831683 283.544175  -0.104    0.917
## no2exp.2w        0.066813   0.056690 283.012735   1.179    0.240
## o3exp.2w         0.040363   0.074186 278.563880   0.544    0.587
## pmexp.2w        -0.002124   0.005735 284.434919  -0.370    0.711
## so2exp.2w       -0.085834   0.104764 286.681042  -0.819    0.413
## Temp.2w          0.023248   0.120432 279.830238   0.193    0.847
## RHx.2w          -0.009879   0.036280 285.735687  -0.272    0.786
## last.ate         0.010671   0.013349 280.291066   0.799    0.425
## wkday.start      0.042621   0.084409 190.300335   0.505    0.614
## go.homewed      -0.252430   0.302338  79.849719  -0.835    0.406
## go.homeweekend  -0.130340   0.343381  80.299526  -0.380    0.705
## COLDY            0.058713   0.204809 270.397059   0.287    0.775
## MNSTY           -0.482831   0.490378 266.959455  -0.985    0.326
## dt_smoke        -0.022624   0.032845 284.361134  -0.689    0.492
```

```r
step(mo4)
```

```
## Backward reduced random-effect table:
## 
##                  Eliminated npar  logLik     AIC    LRT Df Pr(>Chisq)    
## <none>                        16 -473.16  978.33                         
## (1 | Subject.ID)          0   15 -490.95 1011.90 35.572  1  2.458e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite 
## 
##               Eliminated  Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)    
## Temp.2w                1  0.0321  0.0321     1 279.830  0.0373 0.8470709    
## COLD                   2  0.0678  0.0678     1 271.212  0.0790 0.7788488    
## pmexp.2w               3  0.0718  0.0718     1 280.831  0.0839 0.7722626    
## RHx.2w                 4  0.1283  0.1283     1 259.795  0.1506 0.6982381    
```

```
## go.home              5  0.5742  0.2871     2  83.645  0.3385 0.7138446
## o3exp.2w             6  0.1074  0.1074     1 270.074  0.1266 0.7222560
## dt_smoke             7  0.4478  0.4478     1 292.437  0.5298 0.4672622
## wkday.start          8  0.5832  0.5832     1 263.734  0.6961 0.4048516
## last.ate             9  0.7859  0.7859     1 278.303  0.9379 0.3336707
## MNST                10  0.8190  0.8190     1 274.894  0.9847 0.3219216
## no2exp.2w            0  4.5348  4.5348     1 221.626  5.4726 0.0202064 *
## so2exp.2w            0 10.7883 10.7883     1 222.380 13.0193 0.0003808 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Model found:
## log(is.ad) ~ no2exp.2w + so2exp.2w + (1 | Subject.ID)
```

```
mo4.1<- lmer(log(is.ad) ~ no2exp.2w + so2exp.2w + (1 | Subject.ID),data=dat1)
summary(mo4.1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: log(is.ad) ~ no2exp.2w + so2exp.2w + (1 | Subject.ID)
##    Data: dat1
##
## REML criterion at convergence: 910.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.07871 -0.63093 -0.03594  0.59131  2.66937
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Subject.ID (Intercept) 0.5184   0.7200
##  Residual               0.8286   0.9103
## Number of obs: 302, groups:  Subject.ID, 89
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   0.07802    0.54464 232.36291   0.143 0.886220
## no2exp.2w     0.05572    0.02382 221.62605   2.339 0.020206 *
## so2exp.2w    -0.13645    0.03782 222.37980  -3.608 0.000381 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) n2xp.2
```

```
## no2exp.2w -0.904
## so2exp.2w  0.079 -0.471
```

```r
anova(mo4.1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##            Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## no2exp.2w  4.5348  4.5348     1 221.63  5.4726 0.0202064 *
## so2exp.2w 10.7883 10.7883     1 222.38 13.0193 0.0003808 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
vif(mo4.1)
```

```
## no2exp.2w so2exp.2w
##  1.284372  1.284372
```

```r
r.squaredGLMM (mo4.1)
```

```
##             R2m       R2c
## [1,] 0.02811885 0.4021322
```

```r
#calculate IQR change
(exp(0.05572)-1)*no2.2w
```

```
##  3rd Qu.
## 0.156784
```

```r
(exp(0.05572 +0.02382)-1)*no2.2w
```

```
##   3rd Qu.
## 0.2265202
```

```r
(exp(0.05572 -0.02382)-1)*no2.2w
```

```
##    3rd Qu.
## 0.08868928
```

```r
plot(mo4.1)
```

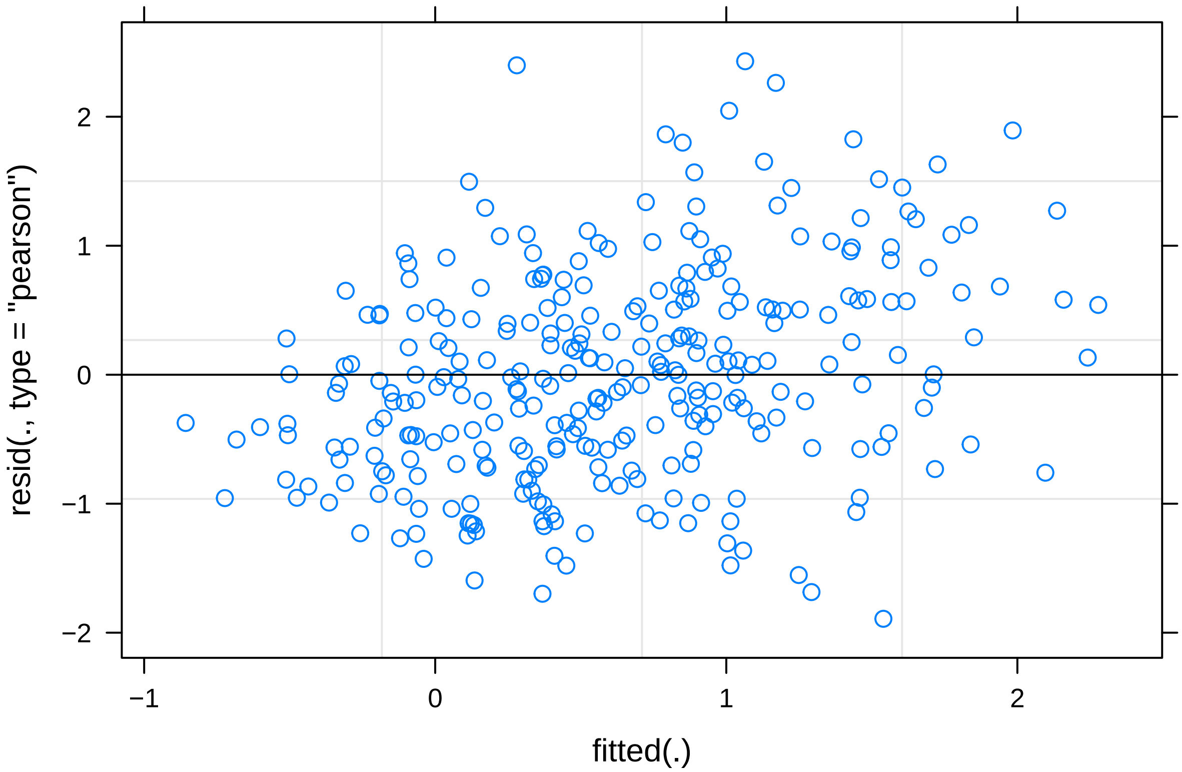

Figure 6: diagnostic plot of model 4.1

Figure 3 shows that the higher the fitted value, the higher the residue. However, in general it looks fine.

4.Ploting Results 4.1 Relationship between log(8-isoprostane) and 12-hour NO2 exposure

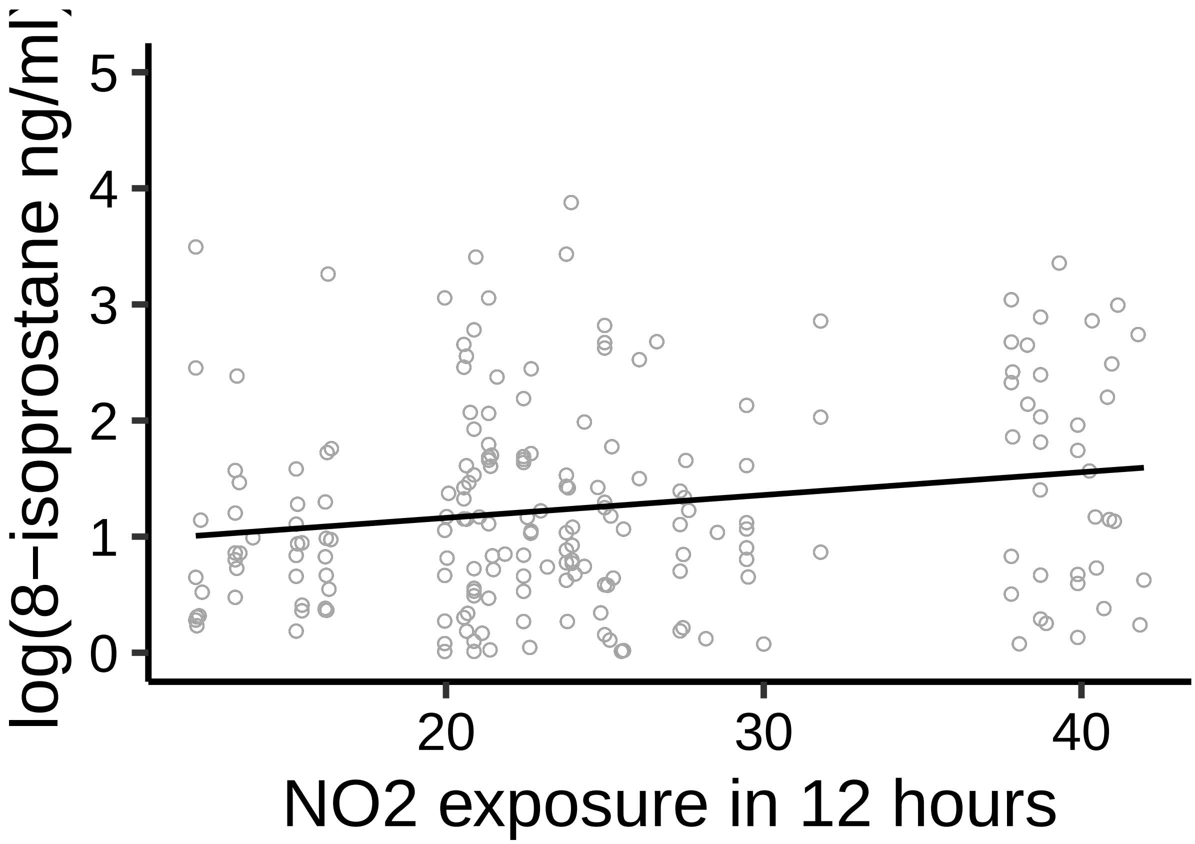

Figure 7:Relationship between log(8-isoprostane) and 12-hour NO2 exposure

Figure 7 shows that there is a positive relationship beteween 12-hour NO2 exposure and log(8-isoprostane).

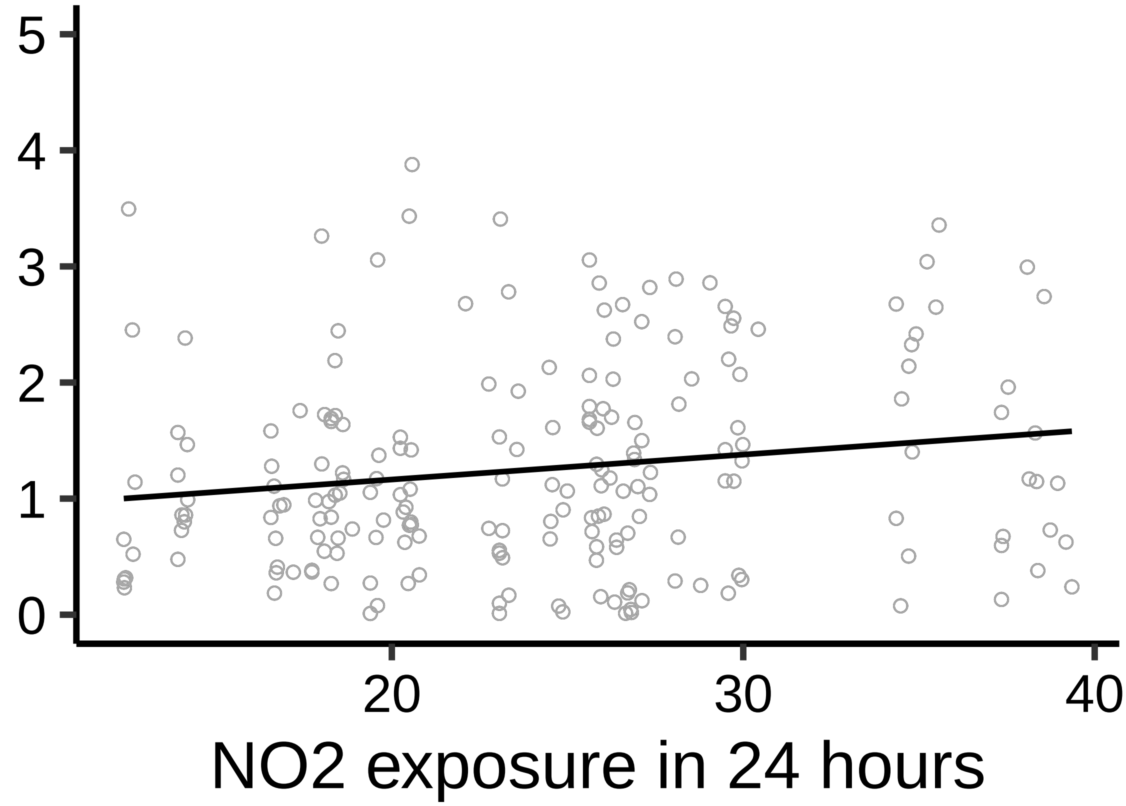

Figure 8:Relationship between log(8-isoprostane) and 24-hour NO2 exposure

Figure 8 shows that there is a positive relationship beteween 24-hour NO2 exposure and log(8-isoprostane).

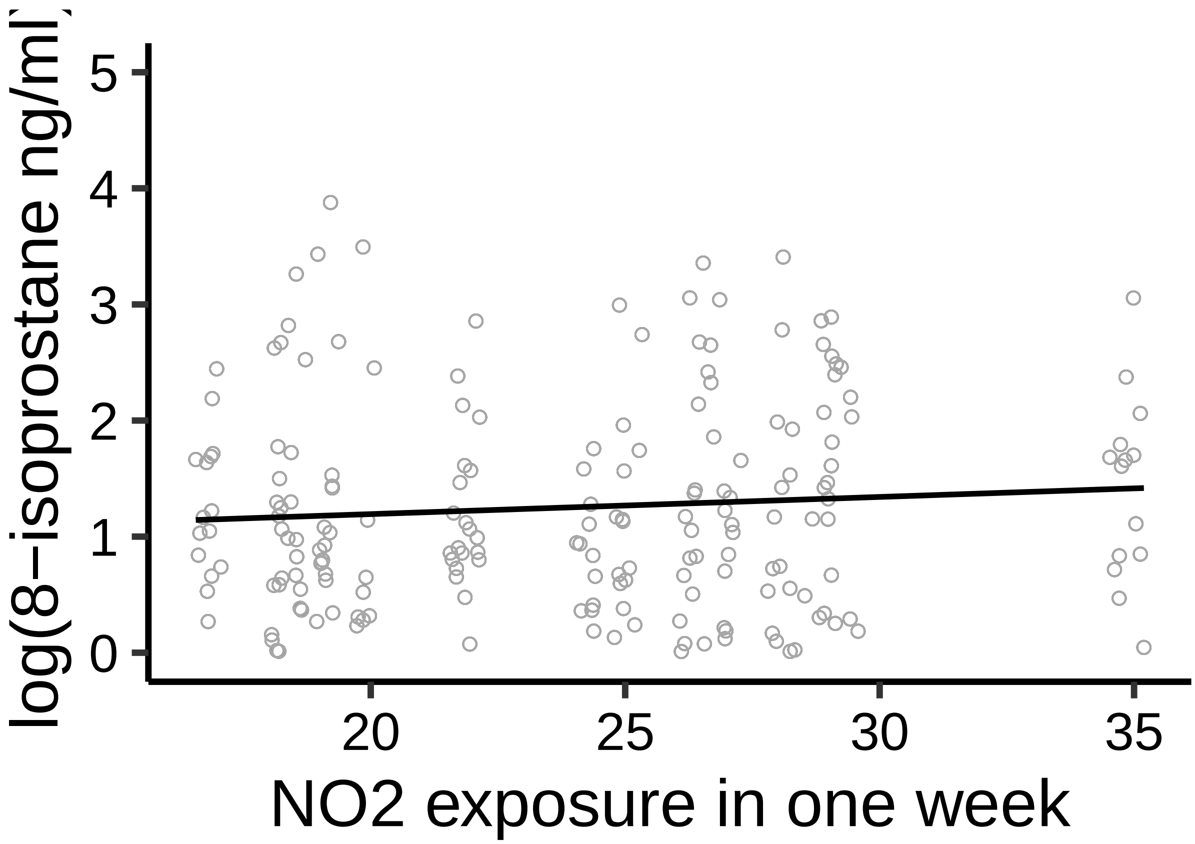

Figure 9:Relationship between log(8-isoprostane) and 1-week NO2 exposure

Figure 9 shows that there is a positive relationship beteween 1-week NO2 exposure and log(8-isoprostane).

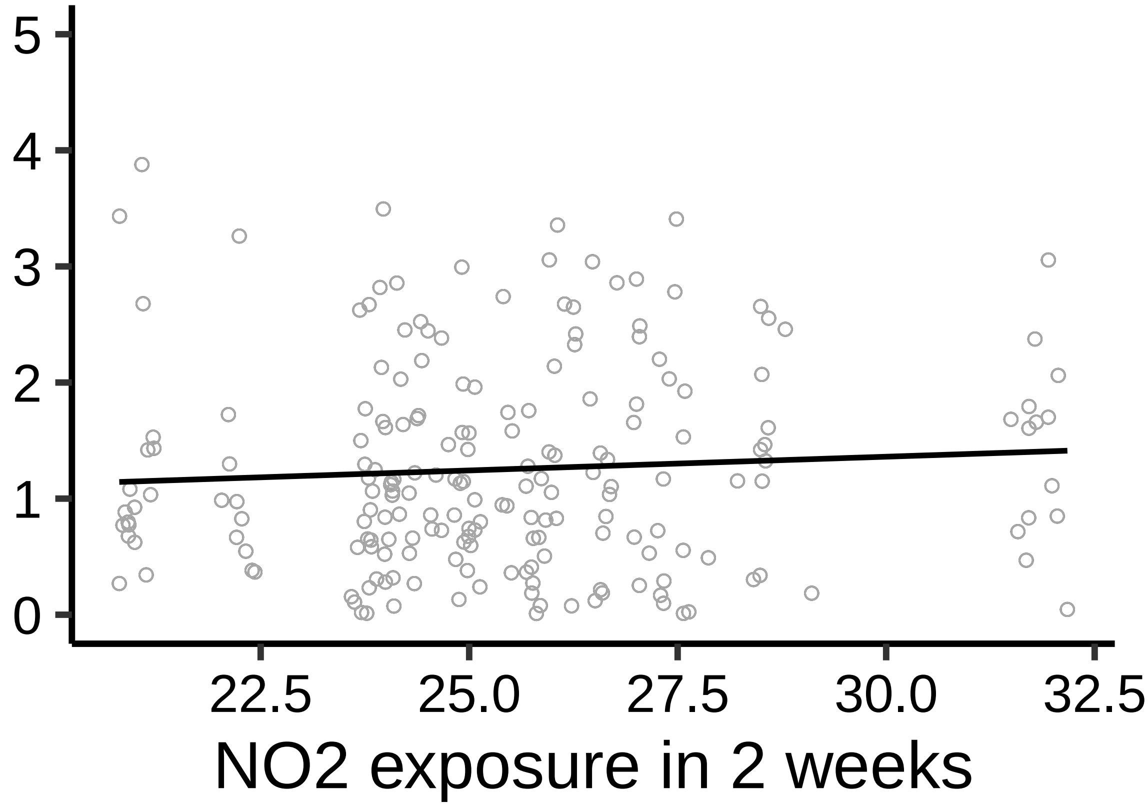

Figure 10:Relationship between log(8-isoprostane) and 2-week NO2 exposure

Figure 10 shows that there is a positive relationship beteween 2-week NO2 exposure and log(8-isoprostane).

3 Exploratory Analysis

The Shapiro tests showed that concentrations of urinary 8-isoprostane is not normalized (p-value<0.001). I log-transformed it to put it into my models as the respondent variables.

4 Analysis

I used linear mixed models with participant-specific intercepts. The inclusion of participants as a variable in mixed-effects models account for the correlation of repeated measurements from the same individuals and precludes the need to control for participant characteristics (e.g., age, gender, BMI, the identity of smoker or non-smoker) that do not change across the four longitudinal measurements.

I used NO2 exposure as predicting variables and log-transformed urinary biomarkers as dependent variables. For each biomarker, I built 4 models to the exposure over the periods of 12h, 24h, one week, and two weeks. I used a backward stepwise model selection method to select the co-variables for each model. The co-variables that I tested in the models were last meal, the start of the workday, respiratory infection status, menstruation, and the time of second-hand exposure. I also tested the SO2 exposure, ozone exposure and PM2.5 exposure, the average temperature and humidity during a corresponding period.

Both 12-hour, 24-hour, 1-week and 2-week NO2 exposure showed significant correlations with the level of urinary 8-isoprostane, especially the 2-week NO2 exposure. *One IQR (7.41 μg/m3) incremental change of 12-h NO2 exposure was associated with an increase of 8-isoprostane level by 19.45% (95% Cl: 14.37%, 24.55%, F value=14.99, p-value <0.001). The Adjusted R-squared =0.4121, which is the fraction of total variance explained by the model* One IQR (8.64 µg/m3) incremental change of 24-h NO2 exposure was associated an increase in 8-isoprostane level by 27.69% (95% Cl: 20.50%, 34.95% , F value=15.16, p-value <0.001). The Adjusted R-squared = 0.4160, which is the fraction of total variance explained by the model. *One IQR (4.47 μg/m3) incremental change of 1-week NO2 exposure was associated with an increase in 8-isoprostane level by 25.15% (95% Cl: 14.14%,36.31%, F value=5.30p-value <0.05). The Adjusted R-squared = 0.4014, which is the fraction of total variance explained by the model* One IQR (2.74 µg/m3) incremental change of 2-week NO2 exposure was associated with an increase in 8-isoprostane level by 15.68% (95% Cl: 8.87%,22.65%, F value=5.47p-value <0.05). The Adjusted R-squared = 0.4021, which is the fraction of total variance explained by the model.

5 Summary and Conclusions

Short term NO2 exposure was associated with urinary 8-isoprostane. This finding meet my hypothesis. Urinary 8-isoprostane indicates the lipid peroxidation from the whole body, reflecting the systemic oxidative stress. It indicates short term NO2 exposure can cause signifiant higher systemic oxidative stress.

6 References

Mölter, A., Agius, R. M., de Vocht, F., Lindley, S., Gerrard, W., Lowe, L., ... & Simpson, A. (2013). Long-term exposure to PM10 and NO2 in association with lung volume and airway resistance in the MAAS birth cohort. Environmental health perspectives, 121(10), 1232-1238. Collart, P., Dubourg, D., Levêque, A., Sierra, N. B., & Coppieters, Y. (2018). Short-term effects of nitrogen dioxide on hospital admissions for cardiovascular disease in Wallonia, Belgium. International journal of cardiology, 255, 231-236. Takenoue, Y., Kaneko, T., Miyamae, T., Mori, M., & Yokota, S. (2012). Influence of outdoor NO2 exposure on asthma in childhood: Meta-analysis. Pediatrics International, 54(6), 762-769. Hashemzadeh, B., Idani, E., Goudarzi, G., Ankali, K. A., Sakhvidi, M. J. Z., Babaei, A. A., ... & Neisi, A. (2019). Effects of PM2. 5 and NO2 on the 8-isoprostane and lung function indices of FVC and FEV1 in students of Ahvaz city, Iran. Saudi journal of biological sciences, 26(3), 473-480. Black, C. N., Bot, M., Révész, D., Scheffer, P. G., & Penninx, B. (2017). The association between three major physiological stress systems and oxidative DNA and lipid damage. Psychoneuroendocrinology, 80, 56-66. Lakshmi, S. V., Padmaja, G., Kuppusamy, P., & Kutala, V. K. (2009). Oxidative stress in cardiovascular disease. Zanolin, M. E., Chamitava, L., Degan, P., Pasini, A., Fratta-Pasini, A., Nicolis, M., ... & De Marco, R. (2015). Biomarkers of oxidative stress in chronic respiratory diseases. Danielsen, P. H., Loft, S., Kocbach, A., Schwarze, P. E., & Møller, P. (2009). Oxidative damage to DNA and repair induced by Norwegian wood smoke particles in human A549 and THP-1 cell lines. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 674(1-2), 116-122. Nuernberg, A. M., Boyce, P. D., Cavallari, J. M., Fang, S. C., Eisen, E. A., & Christiani, D. C. (2008). Urinary 8-isoprostane and 8-OHdG concentrations in boilermakers with welding exposure. Journal of occupational and environmental medicine, 50(2), 182-189. Hashemzadeh, B., Idani, E., Goudarzi, G., Ankali, K. A., Sakhvidi, M. J. Z., Babaei, A. A., ... & Neisi, A. (2019). Effects of PM2. 5 and NO2 on the 8-isoprostane and lung function indices of FVC and FEV1 in students of Ahvaz city, Iran. Saudi journal of biological sciences, 26(3), 473-480. Patel, M. M., Chillrud, S. N., Deepti, K. C., Ross, J. M., & Kinney, P. L. (2013). Traffic-related air pollutants and exhaled markers of airway inflammation and oxidative stress in New York City adolescents. Environmental research, 121, 71-78.