# EFSA (2020): Bi-linear dose-response modelling of PFAS induced immunotoxicity

**National Institute for Public Health and the Environment (RIVM)**

## Conclusion

Just like other types of regression models (constant response model, linear response model, quadratic response model) a bi-linear response model gives a valid description of the immunotoxicity data as presented in EFSA (2020). In the case of a Diph vaccination response a bi-linear response model was a statistically significant improvement compared to the other models. However, as the dose-response calibration was mainly determined by the small number of data for high doses, more data are necessary to conclude more definitively that the bi-linear model is *the* relevant type of regression model for analyzing PFAS induced immunotoxicity.

## Summary

EFSA (2020) and Abraham *et al.* (2020) present a dose-response analysis of the vaccination response to influenza (Hib), tetanus (Tet) and diphteria (Diph) on serum PFOA and serum sumPFAS, i.e. the sum of PFOA, PFNA, PFOS and PFHxS, in one-year old children. This analysis did not address various aspects of (model) uncertainty. As the latter is a prerequisite in quantitative risk assessment RIVM has (re)analyzed the available data using various simple regression models. These applied models were of increasing complexity, the simplest being a constant model (absence of a dose-response relationship), followed by a linear model, a quadratic model and a bi-linear model (absence of a response up to a certain dose ("knee"), followed by a linear response thereafter). Note that the bi-linear model was used in the mentioned by EFSA/Abraham analysis . The chosen models guaranteed model uncertainty to be addressed, i.e. to scale different models on their performance to describe the available data. The performances of the models were compared on the basis of model fits taking account of the differences in the number of model parameters used (non-linear model: 1; linear and quadratic model: 2; bi-linear model: 3). The performances were measured using the Akaike Information Criterium (AIC) and the Bayesian Information Criterium (BIC). The AIC and the BIC differ slightly in the way they measure the penalty of increasing the number of model parameters. Note that the AIC is the default approach in EFSA's and RIVM's current BMD modelling to compare the performance of different models.

The current analysis revealed the following:

- The high variability of the data, resulting in marginal dose-response information in the data, led to small differences between the performances of the different types of regression models.
For example, this resulted in cases where a particular model performed better according to the AIC criterium function, but not according to the BIC criterium function.

- The trend of the calibrated dose-response relations were mainly determined by the response data for the highest dose values.

- The analysis showed that for the Diph vaccination response the bi-linear model performed better than other models (non-linear, linear and quadratic) models in terms of AIC.

- For Diph the serum concentration from which the a linear response visible was found to be 35.8 µg sumPFAS/L [95% confidence interval: 27.4 – 46.7]. In the case of Hib this concentration was found to be 32.5 [18.3 – 57.6], and for Tet 35.3 [23.7 – 52.6] µg sumPFAS/L, to be compared with 12.0 (Hib), 37.1 (Tet) and 37.1 (Diph) (EFSA, 2020, Figures K.1-K.3).

## Rationale

The EFSA Draft Opinion on PFAS addresses the derivation of a HBGV for an (assumed) equipotent PFAS mixture consisting of PFOA, PFNA, PFOS and PFHxS. This derivation was based on PFAS induced immunotoxicity, i.e. a decreased vaccination response, as described by Abraham *et al.* in 1-year olds (2020, in press, the accepted manuscript kindly provided to RIVM by the authors). Abraham *et al.* (2020) present a data base consisting of elicited influenza Hib, tetanus (Tet) and diphteria (Diph) antibody titers upon vaccination, and corresponding PFOA, PFNA, PFOS and PFHxS concentrations in the serum of one-year old children. In this context EFSA presents the results of a bi-linear dose-response modeling analysis (also referred to as "knee" type modeling or a "piece wise" linear modeling, for details see EFSA (2020), Figures K.1 – K.3) on the Abraham data. However, the uncertainty in the model calibration, i.e. the accuracy with which parameters were estimated, was not provided, neither was the issue of model uncertainty addressed. For this reason RIVM (re)produced the mentioned bi-linear dose-response analysis, incorporating the mentioned uncertainty issues.

## Methodology

The performances of different types of regression models to fit the data were compared. These comparisons were made along two ways. The first way was fitting several non-linear regression models on the data using a generic non-linear model fitting routine (nls of R). These regression models contained the crossing-point as a model parameter and thus were non-linear in the model parameters. The performances of these models were defined in terms of the AIC and BIC criterium functions that measure the goodness of fit with a penalty based on the number of parameters included. The second way was fitting linear regression models with linear and/or quadratic terms with respect to the dose using the lm-routine of R. The latter regression models can be interpreted as re-parametrizations of the former regression models. That means, the latter regression models did not include the crossing-points as a model parameter, but only included polynomial functions of the dose. Again the performances of these models were defined in terms of AIC and BIC criterium functions. However, since the latter regression models can be interpreted as nested models, we also used the Likelihood-Ratio test to check whether the better fit of a more complex model was statistically significant. By comparing the results of the different ways of model comparisons the results of the nls analyses were validated.
Note that the bi-linear regression model is essentially non-linear in the dose. Therefore this regression model could not be fitted with the lm-routine. Hence for this model only the nls-routine was used to fit this model. Also note that because of the non-linearities the calculated distributions of the parameters after fitting the non-linear models were clearly skewed, all parameters were log-transformed.

As results the predictions of the non-linear regression and related linear regression models and the AIC criterium function values that resulted from fitting the non-linear regression model using the nls-routine and fitting the related linear regression models using the lm-routine are presented.

## Dose –response models

1 $y(x) = a$

2 $y(x) = a ( x_0 – x )$

3 $y(x) = a ( x_0^2 – x^2 )$

4 $y(x) = \min( a ( x_0 – x ), a ( x_0 – x_1 ) )$

with: a constant regression parameter

$x_0$ y-axis crossing-point

$x_1$ “knee” of the bi-linear model

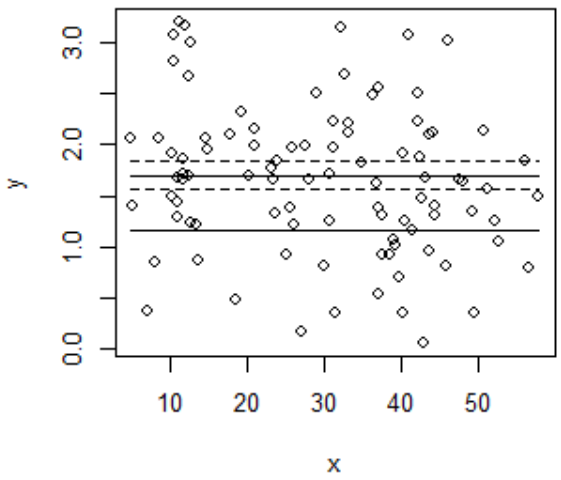
## Available data

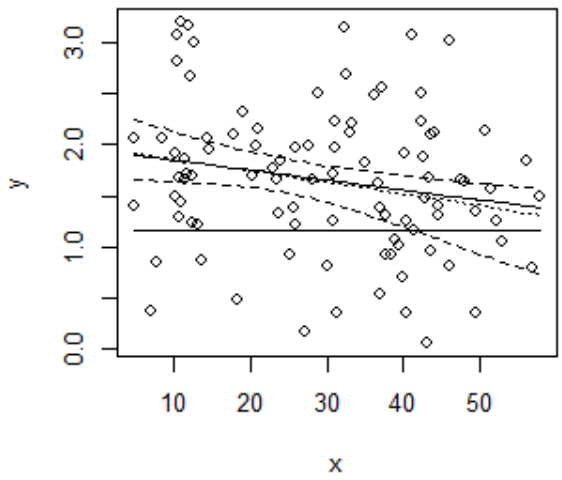
EFSA (2020) present dose-response data on $^{10}$log transformed Hib, Tet and Diph antibody titers and corresponding sum PFAS serum concentrations in 98 (Hib) resp. 100 (Tet; Diph) 100 healthy one-year old children. As the primary data were not available to RIVM, relevant data were electronically retrieved from the original publication, i.e. EFSA (2020), Appendix K, Figures K.1-K.3. In this way 95, 97 and 97 of the Hib, Tet and Diph data points could be retrieved with an accuracy of > 99 % .
The log-transformed antibody titers were used unmodified in the dose-response analysis.

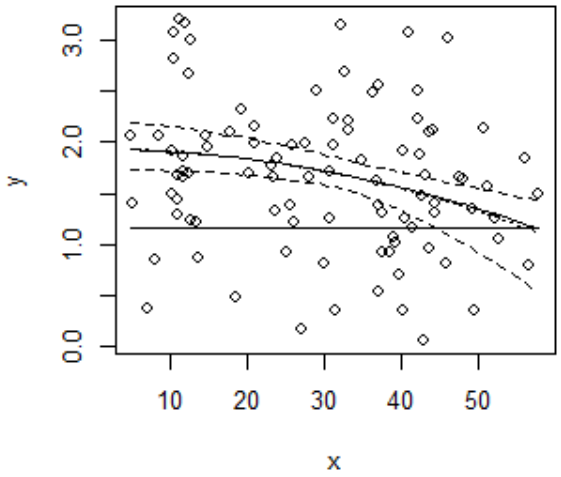
## Results

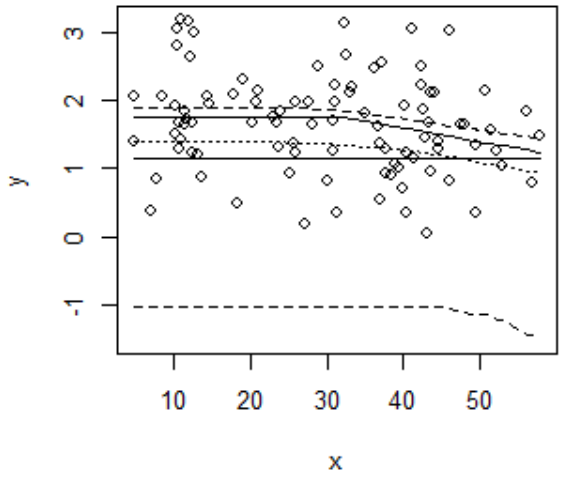
### 1. Graphical output

**Hib**

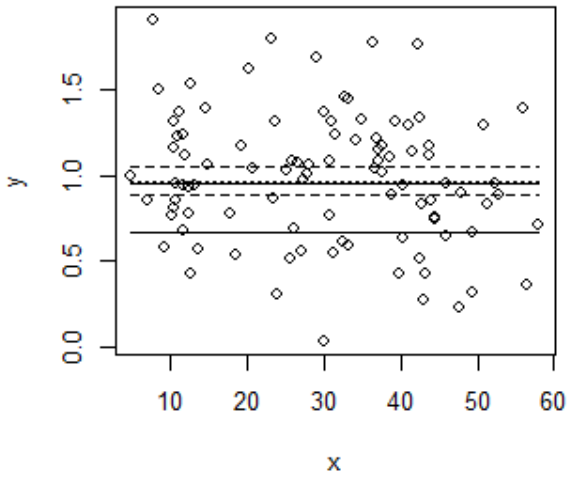


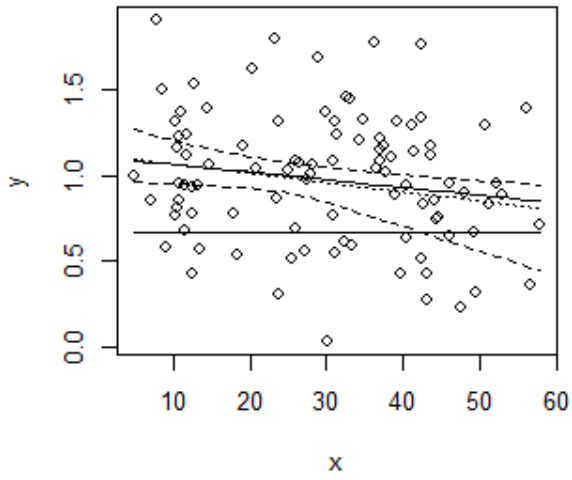



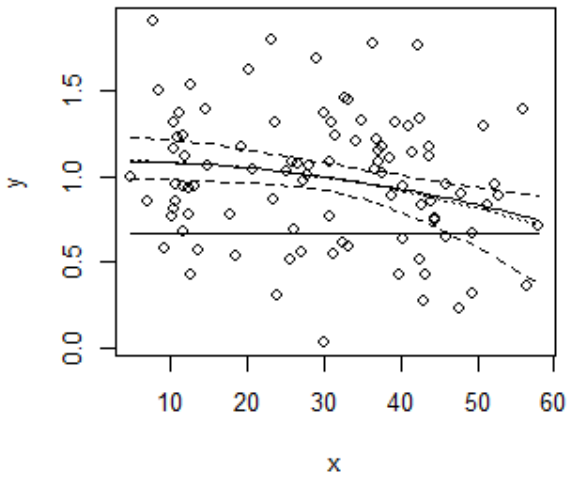



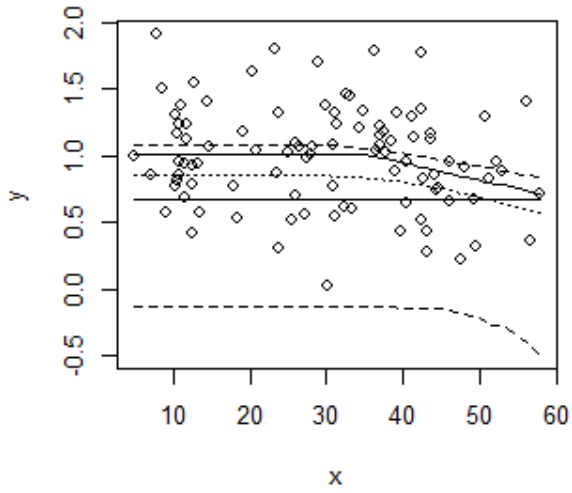



**Tet**

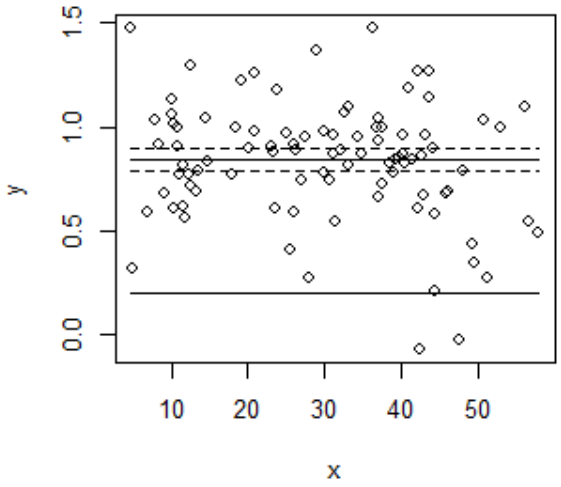

y
x
0.0
0.5
1.0
1.5
10
20
30
40
50
60

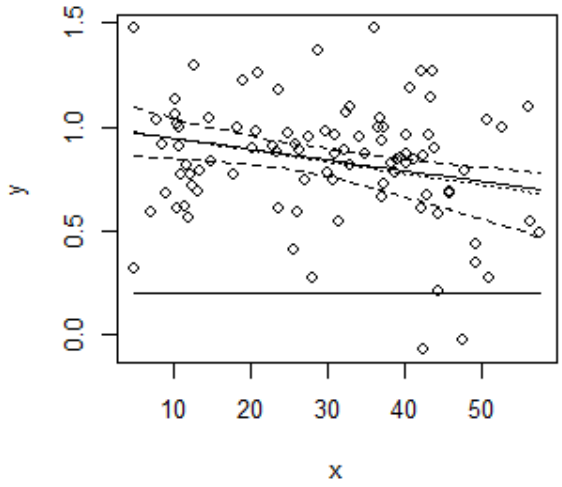

y
x
0.0
0.5
1.0
1.5
10
20
30
40
50
60

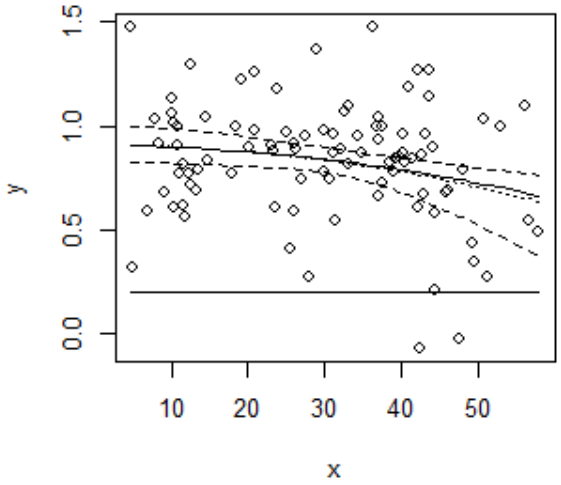

y
x
0.0
0.5
1.0
1.5
10
20
30
40
50
60

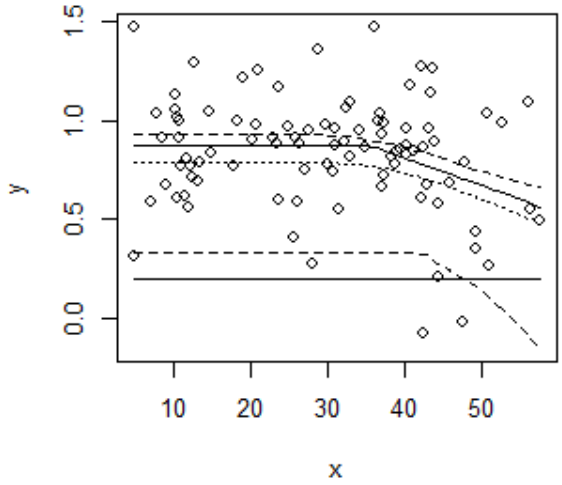

y
x
-0.5
0.0
0.5
1.0
1.5
2.0
10
20
30
40
50
60

**Diph**


y
x
0.0
0.5
1.0
1.5
10
20
30
40
50


y
x
0.0
0.5
1.0
1.5
10
20
30
40
50


y
x
0.0
0.5
1.0
1.5
10
20
30
40
50


y
x
0.0
0.5
1.0
1.5
10
20
30
40
50

## 2. Parameter estimations (after back-transformation) including crossing-points

Note: model 1 (constant only) does not include crossing-point; i.e. it assumes no crossing-point

**Hib, model1**

```
[1] "parameter distribution"
        [,1]      [,2]      [,3]
x0 0.5359735 0.4048006 0.7096521
```

Hib, model2

```
[1] "parameter distribution"
              [,1]         [,2]         [,3]
x0   81.368482362 50.555559099 130.96146180
alfa  0.009845037  0.003922343   0.02471094
```

Hib, model 3

```
[1] "parameter distribution"
              [,1]         [,2]         [,3]
x0   5.792883e+01 4.928631e+01 6.808683e+01
alfa 2.327116e-04 1.151707e-04 4.702123e-04
```

Hib, model 4

```
[1] "parameter distribution"
             [,1]         [,2]        [,3]
x0   62.35268071 46.724258609 83.20852824
alfa  0.02026656  0.004162802  0.09866759
beta 32.46646198 18.312849679 57.55910041
```

Tet, model 1

```
[1] "parameter distribution"
        [,1]      [,2]      [,3]
x0 0.2892714 0.2210161 0.3786056
```

Tet, model 2

```
[1] "parameter distribution"
              [,1]         [,2]        [,3]
x0   98.294359885 54.657162707 176.7706318
alfa  0.004504479  0.001600726   0.0126757
```

Tet, model 3

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 6.431482e+01 | 5.327659e+01 | 7.764004e+01 |
| alfa | 1.022862e-04 | 4.452362e-05 | 2.349868e-04 |

Tet, model 4

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 61.03476839 | 48.455439404 | 76.87976827 |
| alfa | 0.01321233 | 0.003024166 | 0.05772358 |
| beta | 35.31872078 | 23.721639776 | 52.58540512 |

Diph, model 1

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 0.640946 | 0.5848962 | 0.702367 |

Diph, model 2

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 1.533232e+02 | 98.687783624 | 2.382059e+02 |
| alfa | 5.199198e-03 | 0.002736567 | 9.877944e-03 |

Diph, model 3

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 9.846515e+01 | 8.056617e+01 | 1.203407e+02 |
| alfa | 7.279413e-05 | 3.071853e-05 | 1.725013e-04 |

Diph, model 4

[1] "parameter distribution"

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| x0 | 83.49882668 | 63.617033996 | 109.59413886 |
| alfa | 0.01411822 | 0.005268863 | 0.03783058 |
| beta | 35.80000846 | 27.428943881 | 46.72584592 |

**AIC's en BIC's**

Rows: 1: Hib, 2: Tet, 3: Diph; Colums: models 1-4.

AIC's based on fitting log-linear models with nls

```
[1,] 210.40835 208.52194 209.23959 210.88616
[2,]  90.35735  89.43853  89.14392  89.63997
[3,]  34.65290  34.59059  32.37776  31.43283
```

AIC's based on fitting linear models with lm

```
[1,] 210.18650 208.51501 208.46582
[2,]  89.99872  89.42306  88.86811
[3,]  34.51769  33.69055  32.35491
```

BIC's based on fitting log-linear models with nls

```
[1,] 215.49494 216.15182 216.86947 221.05934
[2,]  95.48604  97.13158  96.83697  99.89736
[3,]  39.76065  42.25222  40.03939  41.64834
```

BIC's based on fitting linear models with lm

```
[1,] 215.27309 216.14489 216.09571
[2,]  95.12741  97.11611  96.56116
[3,]  39.62544  41.35218  40.01654
```