Mathematics   Undergraduate

2016

Benchmark Dose Modeling with Covariates for Nanomaterials

Sarah E. Davidson
Xavier University - Cincinnati, davidsons1@xavier.edu

Follow this and additional works at: http://www.exhibit.xavier.edu/undergrad_mathematics

Part of the Materials Science and Engineering Commons, and the Mathematics Commons

Benchmark Dose Modeling with Covariates for Nanomaterials

Davidson, Sarah E.[1], Advisors: Buot, Max PhD [1], Kuempel,
Eileen PhD[2], Smith, Randall MA[2], and Drew, Nathan MS[2]

[1]Department of Mathematics, Xavier University
[2]National Institute for Occupational Safety and Health (NIOSH)

5 April 2016

Abstract

In the last decade, the use of engineered nanomaterials (ENMs) such
as titanium dioxide (TiO2), carbon nanotubes (CNTs), carbon nanofibers
(CNFs), as well as a variety of other materials have become increasingly
popular in commerce because of their many beneficial properties (e.g.
ability to manufacture products that are lighter, stronger, and/or more
compact). However, according to the National Institute of Occupational
Safety and Health, with the development of new nanotechnology it is pru-
dent to ensure the health and safety of workers who are producing or using
these materials at the forefront. For many ENMs, occupational exposure
limits (OELs) are not available and the OELs developed for microscale
materials may not be adequate for ENMs. In the absence of human data,
rodent assays are often used to find a dose estimate which can then be used
as a point of departure (POD) to extrapolate to humans. Some bioassays
report summary statistics, which can be used to determine benchmark
dose (BMD) estimates – the dose that corresponds to a specified level
of increased response called a benchmark response or BMR [4]. Pooling
data across studies with a small number of dose groups (as in many of the
studies in this dataset) provides a more robust dataset by increasing the
sample size, although also adding variability across different experimental
designs (i.e. species, strain, gender). Thus, the aim of this project was to
examine the influence of material type on the dose-response relationship
using statistical regression modeling in R (statistical software) since the
EPA's Benchmark Dose Software (BMDS) does not allow for covariates,
and building upon these regression models by adding covariates to account
for experimental design features which add variability that may obscure
these relationships.

Disclaimer: The findings and conclusions in this report are those of the
author(s) and do not necessarily represent the views of the National Institute
for Occupational Safety and Health.

Background

Engineered nanomaterials (ENMs or nanomaterials) are particles with a diameter less ranging from 1-100 nanometers which have become increasingly more popular in industry because of their chemical properties that make it possible to manufacture products that are lighter, stronger, and/or more compact. However, individuals working with these materials in the production of goods that are produced using ENMs are susceptible to inhaling any excess nanomaterials that escapes into the air. These nanoparticles when inhaled can be deposited into various parts of the respiratory tract (lungs and other air pathways) which can lead to the development of adverse health effects. In an effort to protect individuals from developing chronic illnesses agencies like the National Institute of Occupational Safety and Health (NIOSH) work to set occupational exposure limits (OELs) to improve the working conditions of individuals. For many ENMs, OELs are not available and the OELs for the microscale materials may not be adequate for nanomaterials. In absence of human data, rodent bioassays are often used to develop OELs. Rodent doses associated with no or low level of adverse effect is used as a point of departure (POD) to extrapolate to humans. These dose estimates also provide a measure of the relative potency of nanoscale particles, which could then be used for developing OELs for ENMs.

Dose estimates provide a measure of the relative potency of nanoscale particles. These include NOAELs (no observed adverse effect level), LOAELs (lowest observed adverse effect level), BMD (benchmark dose), and BMDL (lower one sided 95% confidence limit on the BMD). All of these dose estimates are used as a POD for developing OELs, but there are advantages and disadvantages to the various methods and there is a strong case of moving towards working with BMD/BMDL estimates.

NOAEL is the highest experimental dose group that is not statistically different from the control group. LOAEL is the lowest experimental dose group that is statistically different from the control group. The use of NOAELs/LOAELs are advantageous for setting recommended exposure limits because they are conceptually they are relatively simple to use [7]. They are also nice to use since the dose estimates come from the experimental dose groups in the toxicology studies. Also, they can be used to compare the potency of various materials in the absence of common endpoints [4]. However, this simplicity also comes at a price and presents a number of disadvantages:

1. Larger studies tend to produce smaller NOAELs which discourages larger experiments to be conducted [4].

2. NOAELs do not lend themselves to cost-benefit with experimental studies (i.e. failed studies may force a requirement of additional studies) [4].

3. Since the NOAEL/LOAEL come from the dose groups of the experiment then the actual NOAEL may be lower than the observed NOAEL if the response if not statistically significant (e.g. due to small sample size) (Gezondheidsraad, 25). Thus, when extrapolating to a recommended OEL

it will produce a more conservative exposure limit than is necessary. On the surface this seems good, but the feasibility of managing that conservative level of exposure may not be reasonable for financial and practical reasons.

4. The NOAEL corresponds to a single dose group in a toxicology study. So, results at other dose groups are not taken into account eliminating information about the variability of the data, it does not giving a clearer picture of the precision of the NOAEL, and it also does not allow for information about the shape of the dose-response curve [7].

5. The experimental design of the toxicology study (such as number of animals, number of dose groups, and intervals between the level dose groups) greatly influences the NOAEL, for example in general a lower number of animals per study group may result in a higher value for the NOAEL [7].

6. Since the NOAEL/LOAEL method is dependent upon the design of the experiment and the dose groups, then that also means that if an experiment is not designed well then it is possible that the results do not produce a NOAEL, which would result in more testing which can be expensive and time consuming.

Although using the NOAEL/LOAEL method is commonly used practice, its lack in statistical robustness as well as its other disadvantages (mentioned above) promotes the use of the BMD method when it it feasible to do so.

The BMD method, like the NOAEL/LOAEL method, has its advantages and disadvantages. Some of the disadvantages that come about with the BMD method are, firstly one must take care to fit a model which is not only statistically significant but biologically significant as well. Another disadvantage to using the BMD method is that when there is not a common endpoint between two studies, then it is not possible to compare the resulting BMD estimates [4]. On the other hand there are many advantages of using BMD estimation:

1. In the absence of a NOAEL a BMD estimate can still be calculated for a specific dose-response study [4].

2. Less likely to involve difficult decisions about particular experimental groups defining a NOAEL [4].

3. The full dose-response relationship is used to estimate the BMD and BMDL [4],

4. The BMD method also reflects sample size, and makes better use of the dose-response relationship [4].

Thus, the BMD methods is statistically more robust, makes use of more of the data in each of the studies, and the health-based exposure limits that result from the BMD method have an inherently lower uncertainty than those generated from the NOAEL method [7].

The EPA's benchmark dose software was developed to model dose-response relationships and calculate BMD/BMDLs. The BMDL is the lower one sided confidence limit on the BMD, with a 95% confidence that the true BMD estimate is not lower, and is the preferred POD because it accounts for variability in the data used to estimate the BMD. However, BMDS does not allow for the addition of covariates and thereby does not allow data from multiple studies (which are not identically designed) to be pooled together and modeled. The motivation for this analysis was to be able to pool summary data together from multiple inhalation toxicology studies for ENMs to evaluate the hazard potency across materials using a more flexible software for dose-response modeling. Being able to pool together data and writing a model with covariates could be helpful in evaluating experimental variables as well as estimate exposure limits to prevent adverse health effects in workers.

Methods

Data Collection. The data in this analysis was extracted from a pre-exisiting dataset used in a previous research project, and the studies included from the original dataset were chosen based on the criterion that each of the studies were adequately fit with a dose-response model in BMDS.

All of the data from the pre-existing dataset came from one of two published compiled datasets (OECD and Gernand) or from 8 recently published toxicology studies that were added in an effort to supplement the compiled datasets with more current up to date studies. In order to be included in the dataset the studies had to fit the following criteria:

1. The rodent bioassay was focused on pulmonary (lungs) health effects (inhalation, intratracheal instillation, intranasal instillation, intraperitoneal, or pharyngeal aspiration).

2. An original published study could be found, so that data could be collected.

3. Eliminating and consolidating duplicates between the three main sources of data.

4. The individual studies had available quantitative dose-response data for the biological endpoint of interest (polymorphonuclear leukocytes/neutrophils - PMNs; which is a biological endpoint for measuring inflammation) including dose, number of animals per group, proportions/count of PMNs, and standard deviation or standard error of PMNs.

The studies that fit these criteria were modeled in BMDS to determine BMD estimates for the various materials. In all we started with 82 studies and ran 28 studies in through BMDS looking for model fits with reasonable BMD estimates.

The dataset that was used for our study resulted from the studies that when modeled in BMDS had a model that passed through the model fit criteria

and an acceptable BMD estimate (based on the criteria for model fit to the data) was produced. As a result, in our dataset there were 12 studies that were included. Then, after further review of the quantitative data that was provided by each one of these studies, two of the studies did not provide sufficient summary statistics about variance of the percent of PMNs, and estimating those values based of other information that are already rough estimates would have introduced uninterpretable error into our dataset. Thus, it was necessary to remove those two studies from our dataset, and we were left with 10 studies that would provide summary data for us to produce our model in R.

Since each one of the individual studies followed a different experimental design, in particular differences in the dosage rates and exposure times, it was necessary to normalize the dose metrics and make sure all of the response point estimates were of the same type (i.e. percent PMNs). Instead of using the dose concentration in the chamber at a particular time, the cumulative exposure dose ($Cumulative.Exposure = [concentration] * exposure.time$) was used in order to compare all of the dose groups from the various studies, and take into account how much a subject was exposed to overall and the PMN response in accordance to that given dose.

Exploring the Data. Since the response at various post-exposures will differ based on clearance of the material from the lungs and/or reduction of inflammation caused by the material. Then it was first necessary to stratify the dataset into subgroups based dose groups from four different post-exposure periods, either short-term, moderate, moderately long, or long term. Short-term post-exposure were any studies that had a post-exposure period from 0,1, or 3 days, moderate post-exposure were any studies with 16 days post-exposures, moderately long post-exposure were any studies with 45 days, and long term post-exposure were any studies with 90 days or more post-exposure. Then looked at plots of each of the subgroups to see if there was a distinct trend across data, and to see if there was a particular subgroup that would be more beneficial to use for this study.

Modeling in R

Replication of Basic Dose-Response Model from BMDS to R. Since BMDS does not allow for the addition of covariates into model fits, then it was necessary to replicate the model fits in some other software that would allow for covariate additions (i.e. R statistical software). For this analysis the focus was placed on the following dose-response models (assuming constant variance) from BMDS:

- Linear (Polynomial 1)

  ○ $y = b_0 + b_1(dose)$

- Quadratic (Polynomial 2)

○ $y = b_0 + b_1(dose) + b_2(dose)^2$

- Cubic (Polynomial 3)

    ○ $y = b_0 + b_1(dose) + b_2(dose)^2 + b_3(dose)^3$

The following models were chosen as the first set of models to be replicated from BMDS to R because of their ability to be replicated quite simply into another software.

In R there is a function called `glm` (or generalized linear model), which is a statistical regression modeling function that was used to replicate the model fitting process that BMDS implements. The function `glm` was chosen on the criterion that it is able to incorporate more arguments when fitting a regression model (if necessary) than the function `lm` (linear model), which is another statistical regression modeling function in R. Two individual bioassays were chosen to replicate the polynomial models form BMDS into R, and the studies were chosen based on the following criteria:

1. The inhalation studies were included in the pooled dataset for this analysis

2. The inhalation studies had multiple experimental dose groups (i.e. had a control group, and multiple dose groups exposed to some dose of the aerosalized ENM)

The first criterion was set to ensure that the studies used were adequately fit with a model in BMDS - does not necessarily mean one of the polynomial models was the best fit models but it was a sufficient dataset to be able to fit a dose-response model. The second criterion was set to ensure that there was enough data in the individual study dataset to be able to adequately replicate a cubic (polynomial 3) model from BMDS to R. Also, it was necessary to use two individual inhalation datasets to ensure that model fits being output in R were in fact replicas of the model fits in BMDS. Thus, the initial study was meant to see if the `gglm` function does in fact use essentially the same statistical modeling methods as BMDS. The second dataset was meant to verify - in the case of replication of model fits for the initial study - that the replication was not just a matter of circumstances and R does in fact use essentially the same modeling methods as BMDS.

Now, when fitting the 2 individual datasets we decided the following criteria would be sufficient for ruling BMDS and R model fits to be approximately the same and thus a replication of BMDS models into the R statistical software:

1. The intercept estimate (b_0) in the dose-response model in R is the same as intercept from the model output in BMDS

2. All of the parameter estimates (b_n ; $n \in N$) in the dose-response model in R are the same as the parameter estimates from the model output in BMDS

These two criteria are sufficient to determine a successful replication of the model fit because if the intercept and all of the parameter estimates are essentially the same, then the two model fits would also be essentially the same (i.e. one could lay the dose-response model fit from BMDS on top of the model fit from R and they would be relatively indistinguishable).

Fitting the Pooled data with Dose-Response Models. The next step was to fit the pooled dataset with each one of the basic polynomial models. However, since the data that is being fit now with a dose-response curve is a collection of multiple studies and all of the data points are mean responses for each dose group, then in order for our model to be valid it is necessary to weight the model accounting for some variability in the data within each dose group (from the individual studies). Thus, weighting the model takes away the assumption that each mean response contributes the same amount of information to a dose-response model. Two approaches were explored in the analysis:

1. Weighting the model using precision - which is derived using the standard deviation from each dose group (i.e. the sufficient summary statistic of variability for a dose groups) and the number of subjects per dose group

2. Weighting the model using the number of animals per dose group

The first method, which is the more favorable of the two methods because of its use of the sufficient statistic of variability for each of the groups, utilizes something called precision. Precision is defined in the following manner:

$$Precision = \frac{1}{SEM^2}$$

SEM -Standard error of the mean (or $(\frac{sd}{\sqrt{n}})$ - standard deviation of the dose group divided by the square root of the number of animals in the dose group)

Weighting the model based on the precision is telling our model fitting function that mean responses with a lot of variability should contribute less to the overall fitting of the dose-response model. In turn, that also means that the mean responses with low variability should contribute a lot of information to the dose-response model fit. Also, weighting by precision is optimal if the SEM values are correctly estimated. However, with the small sample sizes in the dose groups, the estimates of the SEM may be poor; thus, a reason to weight by the number of animals per dose group, which gives higher weight to responses based on larger sample size.

The second method, which is also valid but less statistically robust, would be finding the number of animals for each mean response and weighting the model based off of the number of animals that produced the mean response for a specified dose. One advantage of using this method is that it is relatively simple as well as it accounts for some of the variability of the study designs.

Secondly, this method responses that are based on larger sample sizes contribute more weight to the model fit. If the sample sizes for dose groups are relatively small then the standard deviations could be very imprecise, which could cause the model to be inadequately fit when weighting with precisions.

Both weighting methods were explored in this analysis and the one that was most feasible to fit a dose-response model was implemented in our later model.

Adding Covariates to the Pooled Data Dose-Response Linear Model. After fitting each of the basic polynomial models (Linear, Poly 2, and Poly 3) with the respective weighting method, then it was possible to begin adding covariates to our model. For this portion of the analysis it was best to again begin with the simpliest model to add covariates to. Therefore, the model that was focused on for adding covariates was the linear model. All of the covaritates that were included in our dataset were categorical varibles and the main goal for adding these covariates was to explain the variability in the data as well as explain their effect on the observed response in PMNs. When adding the covariates to the linear model in R, the covariates were coded to observe slope effects to determine any explanation in a significant increase in the PMN response.

The model development method implemented to add covariates to the linear model was a step-wise method, i.e. adding one variable at a time to see if it would help to develop a better model fit. The order of adding covariates was as follows in order of importance:

1. Material Type

2. Species/Strain

3. Gender

The purpose for this order of importance stemmed from the main objectives of the analysis. The main purpose for this analysis was to evaluate the hazard potency of the various ENM types, and thus was placed as the most important covariate to consider for developing the model fit. Following that would be species and strain, since species was the more general case of categorizing animal subjects it made sense to add it as a covariate first. Furthermore, strain being a more specific categorizations of the animal subjects, that was also considered a reasonable next step in the model development. However, the addition of species as well as strain in the model is not necessary and may make unreasonable parameter estimates if both are included in the model fit. For example, calculating a parameter estimate for the species: hamster and strain: Wistar would not make sense considering there is not such a combination, because hamsters correspond with the Syrian strain in our dataset and Wistar is a strain of rat. Finally, gender effects are always important to consider in biological studies, but not necessarily the primary concern for explaining variability in the dose-response data.

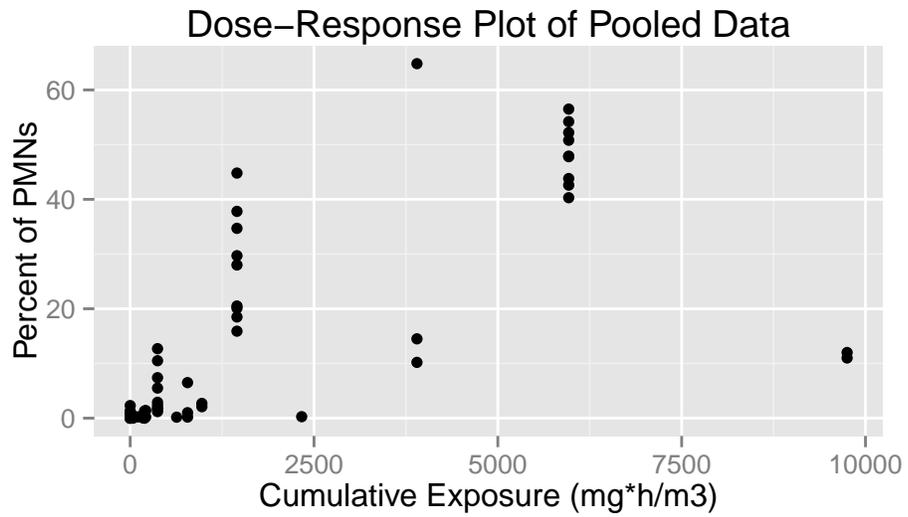The following criteria were used to help determine our final model fit:

1. The addition of a covariate to the linear model must lower the AIC (Akaike information criterion) of the model fit

2. There was no direct correlation between covariates

3. The addition of the covariate not only had to be statistically significant in developing a best fit model but it also needed to be biologically significant

The AIC is a goodness-of-fit test for a proposed statistical regression model, which is assessed by the log-likelihood function and includes a penalty for increasing the number of parameter to the proposed model [5]. The idea of the AIC is to discourage over-fitting a model. If an addition of a parameter lowers the overall AIC of the regression model, then the variability in the data being explained by the addition of that particular covariate outweighs the penalty assigned for the addition of the covariate; the lower the AIC then the better the model fit.

The second criterion was set in place to ensure that the model fit does not attempt to estimate separate parameters that provide the same information about the data and to avoid over-parameterizing the model. Finally, the third criterion was set in place in order to ensure statistically adequate model fits did not make unreasonable biological assumptions to achieve a "better model fit", developing models that are not biologically significant can result inaccurate BMD/BMDL estimates which would be used as a POD to extrapolate to humans thereby developing an OEL that was not sufficient to protect workers adequately.

Results

Exploring the Data.

Figure 1: Pooled dataset of all of the published inhalation toxicology studies contained within the dataset for this analysis. This plot contains studies varying in material type being tested, species/strain of subjects, gender, as well as post-exposure time.
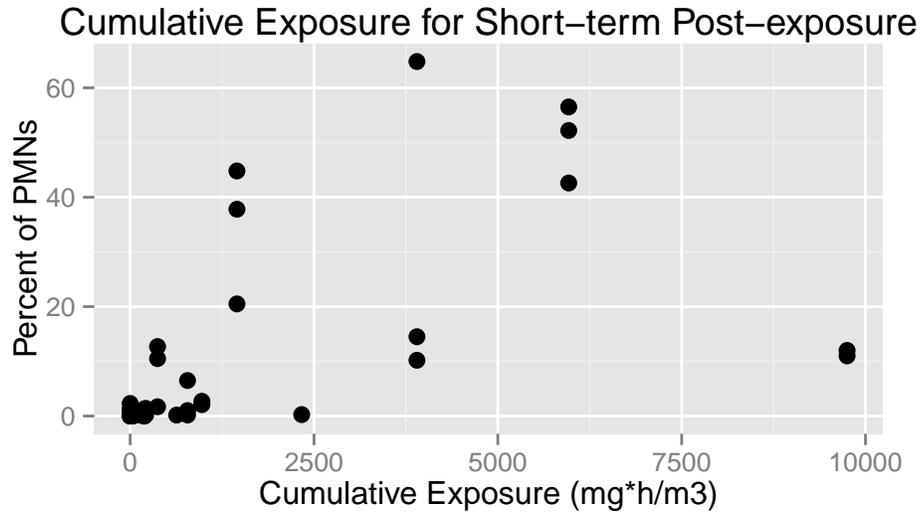
Figure 2: Pooled dataset of only the short-term post-exposure (0-3 days) inhalation toxicology studies contained within the dataset for this analysis.

In both of the plots the pooled data looks as if it follows a typical dose-response trend with much of the data is contained in the lower dose regions. However, the dose-response trend that is seen in these plots does not account that the mean responses come from a mix of various material types.

Replication of Model Fits

Plot of Test Dataset No.1
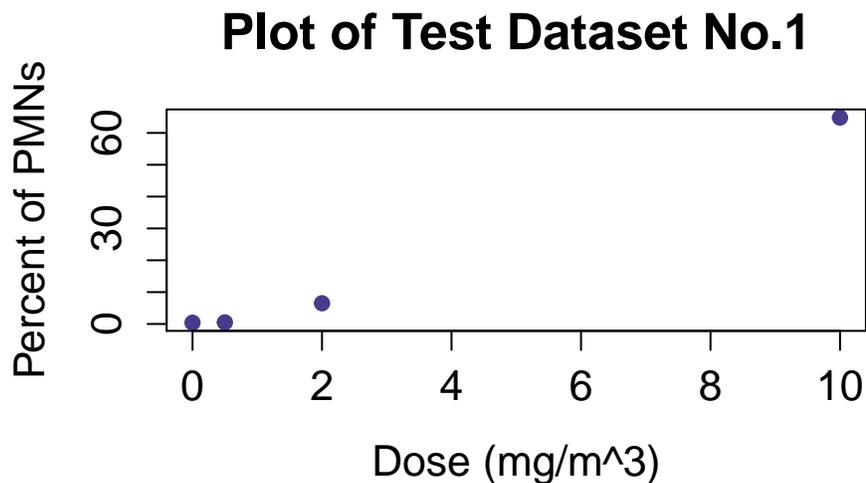
Figure 3: Plot of the individual study that was used to replicate the polynomial models from BMDS into R. - Test Dataset No.1 corresponds to Study 85 (Bermudez, E. et al.) in the original dataset.

Linear:

Table 1: Test Dataset No. 1, Replication Results for the Linear (Poly 1) Model. BMDS output (top) ; R output (bottom).

F: Linear_84_ped_0_TiO2_linear_d_BMR-10				
Variable	**Estimate**	**Std. Err.**	**Lower Conf. Limit**	**Upper Conf. Limit**
alpha	16.2034	5.12398	6.16062	26.2462
beta_0	-2.83327	1.13827	-5.06423	-0.602306
beta_1	6.68265	0.222965	6.24564	7.11965

```
> lin.reg.a<-glm(Percent.of.PMNs~Dose,data=study.no.85.acute)
> lin.reg.a

Call:  glm(formula = Percent.of.PMNs ~ Dose, data = study.no.85.acute)

Coefficients:
(Intercept)         Dose
     -2.833        6.683

Degrees of Freedom: 3 Total (i.e. Null);  2 Residual
Null Deviance:      2938
Residual Deviance: 27.36  AIC: 25.04
```

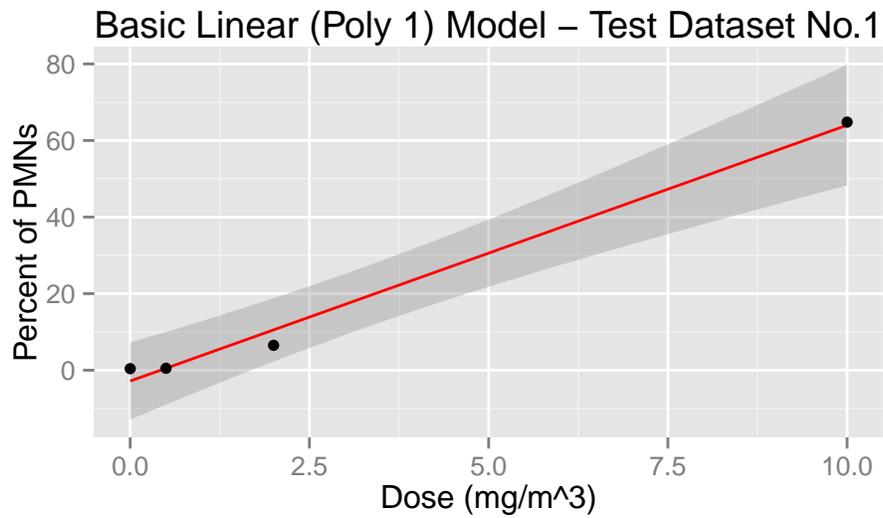Figure 4: Test Dataset No. 1, BMDS Linear (Poly 1) Model plot.

Figure 5: Test Dataset No. 1, R Linear (Poly 1) Model plot.

Quadratic (Polynomial 2):

Table 2: Test Dataset No. 1, Replication Results for the Quadratic (Poly 2) Model. BMDS output (top) ; R output (bottom).

G: Polynomial_84_ped_0_TiO2_poly2_d_BMR-10				
Variable	**Estimate**	**Std. Err.**	**Lower Conf. Limit**	**Upper Conf. Limit**
alpha	9.5549	3.02152	3.63282	15.477
beta_0	-0.0943176	1.14153	-2.33167	2.14304
beta_1	2.3702	1.16861	0.0797578	4.66064
beta_2	0.411988	0.110438	0.195533	0.628444

```
> poly2.reg.a<-glm(Percent.of.PMNs~Dose+I(Dose^2),data=study.no.85.acute)
> poly2.reg.a

Call:  glm(formula = Percent.of.PMNs ~ Dose + I(Dose^2), data = study.no.85.acute)

Coefficients:
(Intercept)         Dose     I(Dose^2)
   -0.09432      2.37020       0.41199

Degrees of Freedom: 3 Total (i.e. Null);  1 Residual
Null Deviance:      2938
Residual Deviance: 0.7681  AIC: 12.75
```
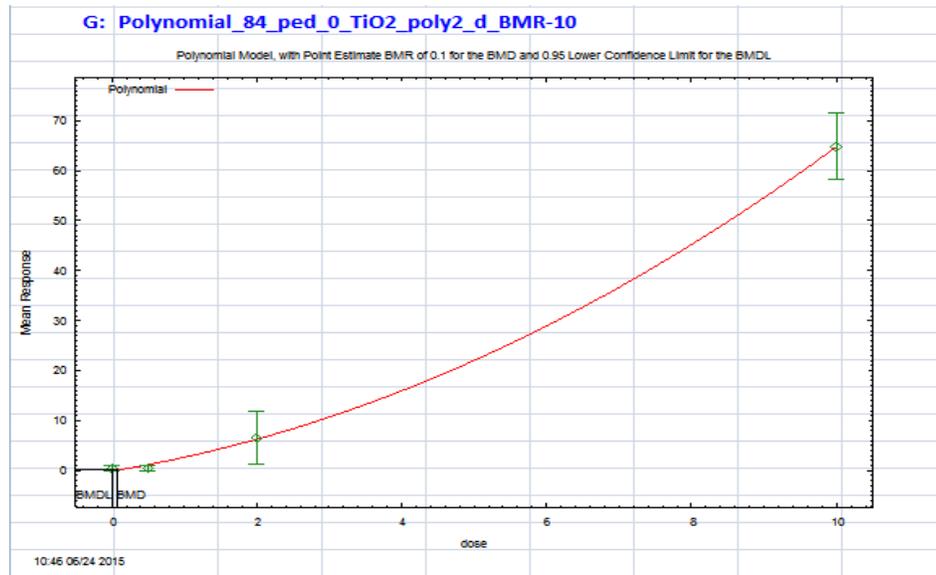

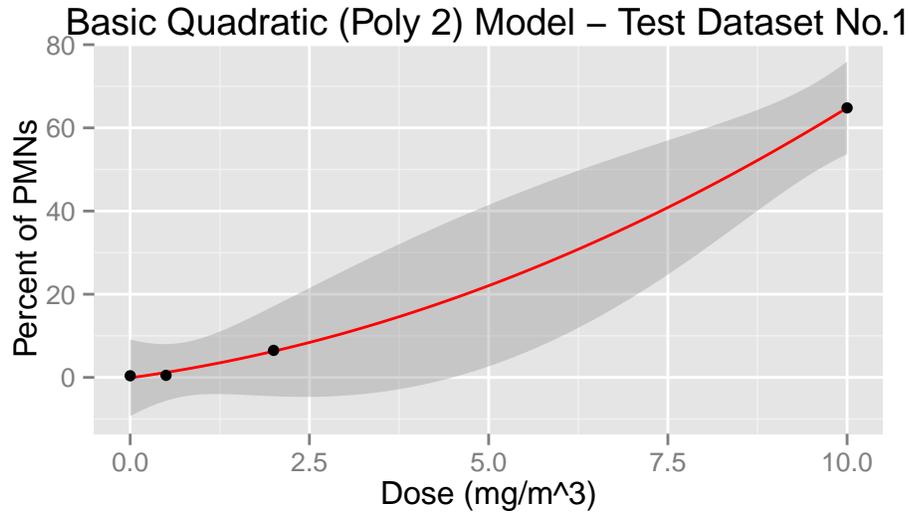Figure 6: Test Dataset No. 1, BMDS Quadratic (Poly 2) Model plot.

14

Figure 7: Test Dataset No. 1, R Quadratic (Poly 2) Model plot.

Cubic (Polynomial 3):

Table 3: Test Dataset No. 1, Replication Results for the Cubic (Poly 3) Model.
BMDS output (top) ; R output (bottom).

H:	Polynomial_84_ped_0_TiO2_poly3_d_BMR-10			
Variable	**Estimate**	**Std. Err.**	**Lower Conf. Limit**	**Upper Conf. Limit**
alpha	9.36286	2.9608	3.55981	15.1659
beta_0	0.4	1.36842	-2.28205	3.08205
beta_1	-0.905395	5.24352	-11.1825	9.37172
beta_2	2.28849	2.9319	-3.45792	8.0349
beta_3	-0.155395	0.242624	-0.63093	0.32014

```
> poly3.reg.a<-glm(Percent.of.PMNs~Dose+I(Dose^2)+I(Dose^3),data=study.no.85.acute)
> poly3.reg.a

Call:  glm(formula = Percent.of.PMNs ~ Dose + I(Dose^2) + I(Dose^3),
    data = study.no.85.acute)

Coefficients:
(Intercept)         Dose      I(Dose^2)      I(Dose^3)
     0.4000      -0.9054         2.2885        -0.1554

Degrees of Freedom: 3 Total (i.e. Null);  0 Residual
Null Deviance:      2938
Residual Deviance: 1.201e-27  AIC: -232.1
```
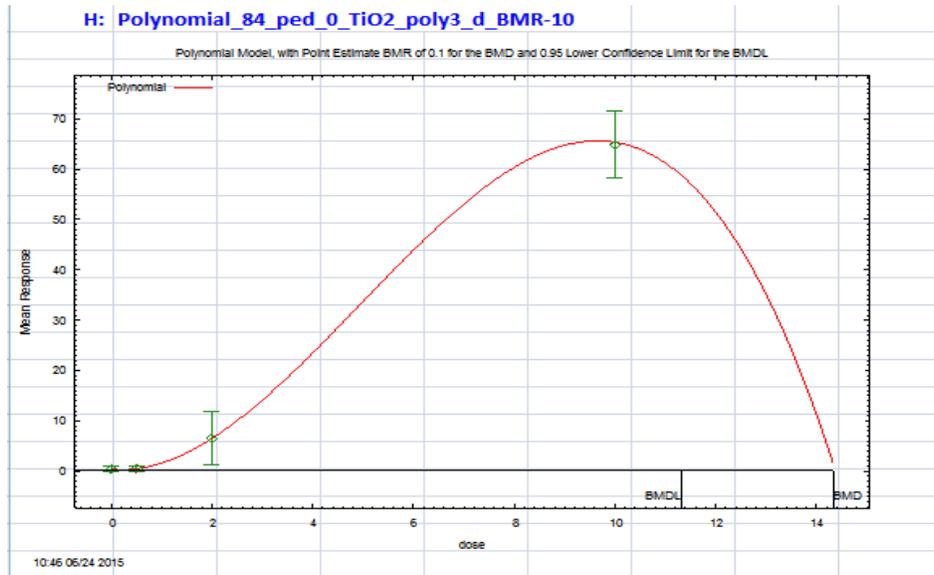
15

Figure 8: Test Dataset No. 1, BMDS Cubic (Poly 3) Model plot.
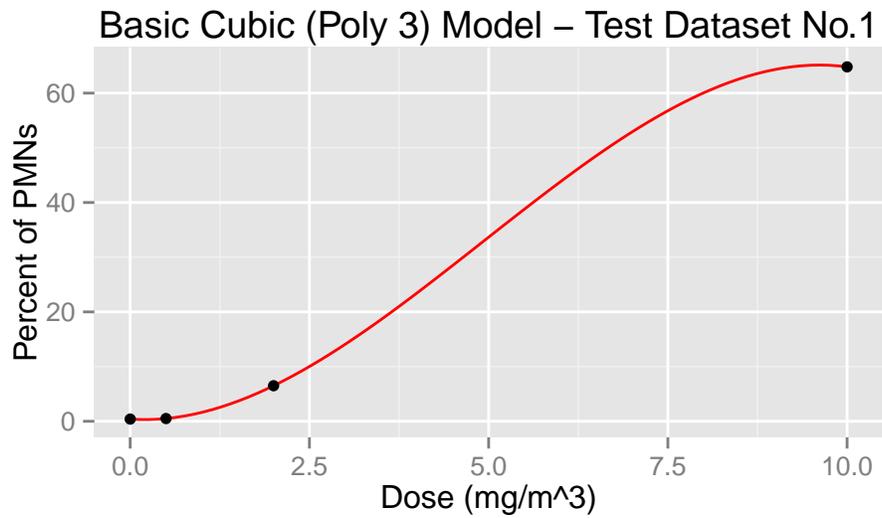

Figure 9: Test Dataset No. 1, R Cubic (Poly 3) Model plot.

All of the replication model fits (Linear, Poly2, and Poly 3) for test dataset No.1 in R had essentially the same parameter estimates - down to the thousandths

16

decimal place - as model fits from BMDS. Similar results for the test dataset No. 2 were found, all of the model fits were matched to the thousanths place. Since test dataset No. 2 was meant to be a cross-check of model fitting method, see the R code in the Appendix for these replication results for test dataset No.2. Thus, the replication of BMDS modeling methods (for models assuming constant variance) were successful.

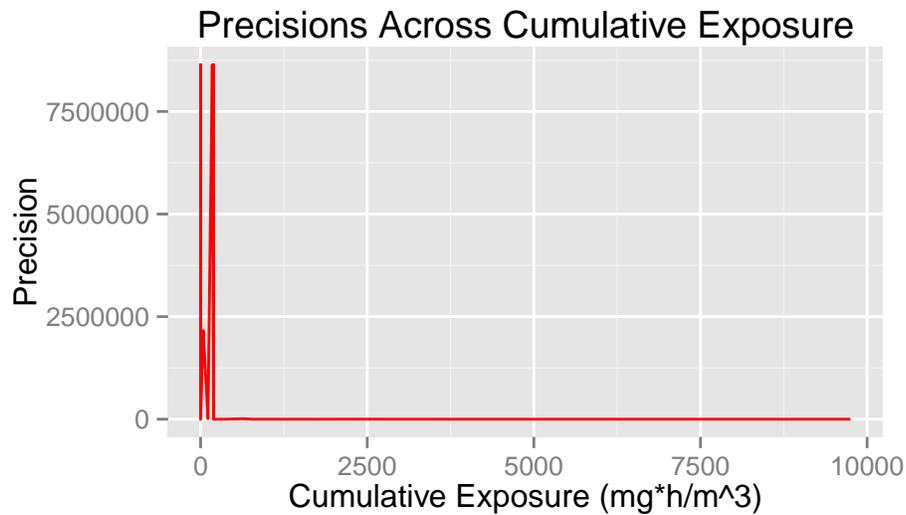## Model fitting the Pooled Data

Weighting Assessment

Figure 10: Comparison plot of precisions ($\frac{1}{SEM^2}$) across Cumulative Exposure groups. Large variability in the precisions at the low exposure end and almost no precision seen in the the higher exposure groups.

The estimates of SEM may be poor given the small number of animals per group. In addition, the true SEM may be increasing with dose due to greater variability in responses from a distribution of individual animal sensitivities to the exposure (i.e. nonconstant variance). The high exposure regions have very low precision esitimates relative to those in the low exposure region where most of the data are (about 8 orders of magnitude difference between precisions).

Consider the number of animals per exposure group as a method of weighting the model with variability. This approach will provide more weight to the responses that are based on larger sample sizes.
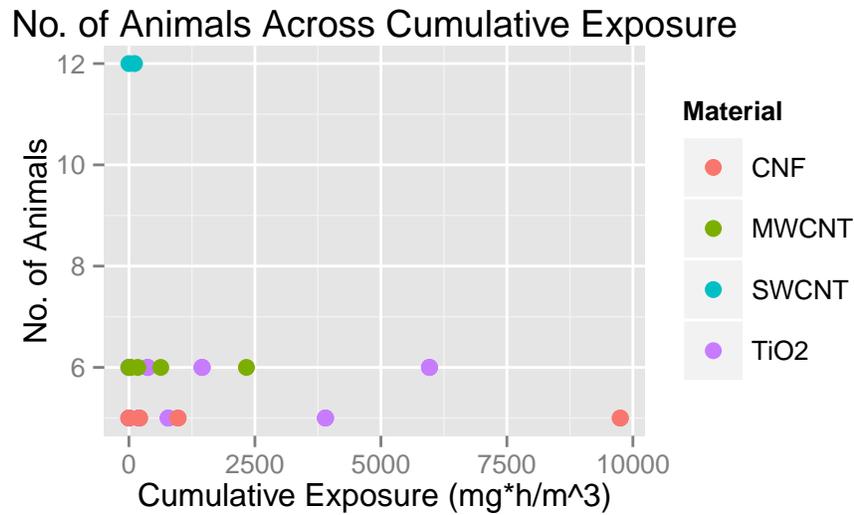
Figure 11: Comparison plot of No. of Animals per Cumulative Exposure group. The number of animals per Cumulative Exposure group ranges from 5 to 12 animals per group.

Since the range number of animals per exposure (dose) group is more evenly distributed across all of the exposure groups and there are issues with the variance structure of the data, then it is sufficient to weight the pooled model with the number of animals per dose group (as mentioned in the methods section).

Model Fits to the Short-term Post-exposure Pooled Dataset (weighting based on number of animals per dose group)

Linear (Polynomial 1):

Table 4: Linear (Poly 1) Model.

```
> my.lin.reg.3<-glm(Percent.of.PMNs~Cumulative.Dose..mg.h.m3.,weights = No.of.Subjects.N.,data =
ds.1.acute)
> summary(my.lin.reg.3)#with weights=No. of Animals

Call:
glm(formula = Percent.of.PMNs ~ Cumulative.Dose..mg.h.m3., data = ds.1.acute,
    weights = No.of.Subjects.N.)

Deviance Residuals:
    Min      1Q   Median      3Q      Max
 -74.144  -13.275  -9.532  -3.262   99.682

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                4.2628386  2.7872785   1.529  0.13467
Cumulative.Dose..mg.h.m3. 0.0040918  0.0009577   4.273  0.00013 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 1278.189)

    Null deviance: 70628  on 38  degrees of freedom
Residual deviance: 47293  on 37  degrees of freedom
AIC: 325.98

Number of Fisher Scoring iterations: 2
```
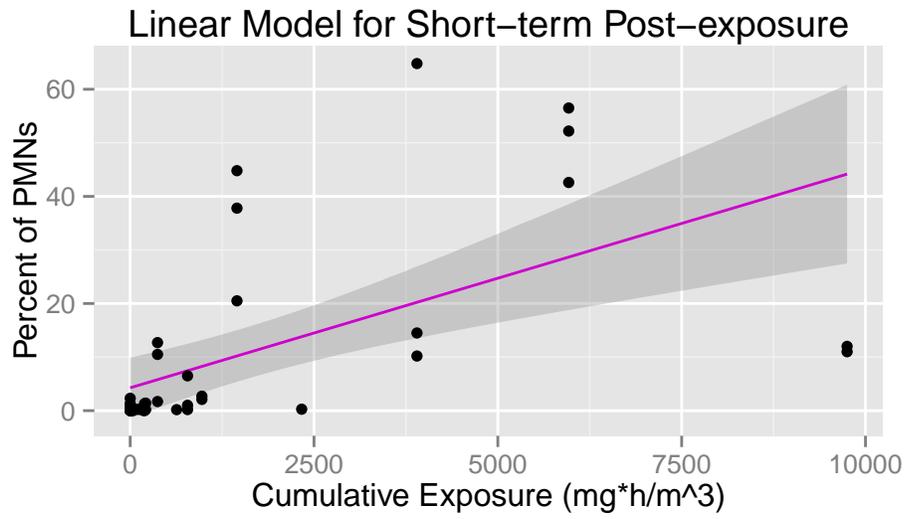
Thus, the linear model that follows is $y = b_0 + b_1(dose)$; $b_0 = 4.2628386$, $b_1 = 0.0040918$; and is shown in the plot below:

Figure 12: Linear (Poly 1) Model fit plot for the short-term post-exposure pooled dataset.

Quadratic (Polynomial 2):

Table 5: Quadratic (Poly 2) Model fit for the short-term post-exposure pooled dataset.

```
> my.poly2.reg.3<-glm(Percent.of.PMNs~Cumulative.Dose..mg.h.m3.+I(Cumulative.Dose..mg.h.m3.^2),
weights = No.of.Subjects.N., data=ds.1.acute)
> summary(my.poly2.reg.3)#with weights=No. of Animals

Call:
glm(formula = Percent.of.PMNs ~ Cumulative.Dose..mg.h.m3. + I(Cumulative.Dose..mg.h.m3.^2),
    data = ds.1.acute, weights = No.of.Subjects.N.)

Deviance Residuals:
    Min       1Q    Median       3Q       Max
-66.264   -8.032     1.058    5.202    64.605

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                      -1.024e+00  2.352e+00  -0.435    0.666
Cumulative.Dose..mg.h.m3.         1.544e-02  2.281e-03   6.770 6.58e-08 ***
I(Cumulative.Dose..mg.h.m3.^2)   -1.410e-06  2.684e-07  -5.253 6.92e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 743.6777)

    Null deviance: 70628  on 38  degrees of freedom
Residual deviance: 26772  on 36  degrees of freedom
AIC: 305.79

Number of Fisher Scoring iterations: 2
```

Thus, the quadratic model that follows is $y = b_0 + b_1(dose) + b_2(dose)^2$; $b_0 = $ -1.024e+00, b_1=1.544e-02, b_2= -1.410e-06 ; and is shown in the plot below:
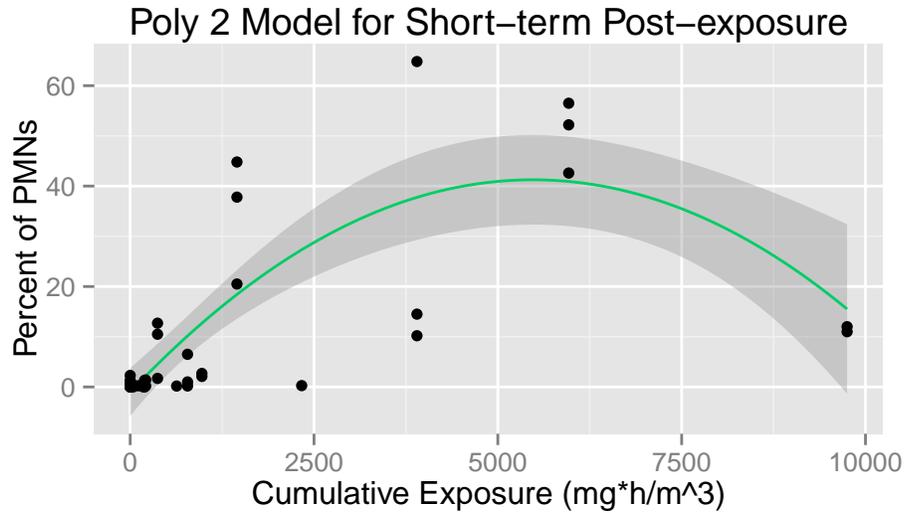
Figure 13: Quadratic (Poly 2) Model fit plot for the short-term post-exposure pooled dataset.

Cubic (Polynomial 3):

Table 6: Cubic (Poly 3) Model fit for the short-term post-exposure pooled dataset.

```
> my.poly3.reg.3<-glm(Percent.of.PMNs~Cumulative.Dose..mg.h.m3.+I(Cumulative.Dose..mg.h.m3.^2)+
I(Cumulative.Dose..mg.h.m3.^3), weights = No.of.Subjects.N.,data=ds.1.acute)
> summary(my.poly3.reg.3)#with weights=No. of Animals

Call:
glm(formula = Percent.of.PMNs ~ Cumulative.Dose..mg.h.m3. + I(Cumulative.Dose..mg.h.m3.^2) +
    I(Cumulative.Dose..mg.h.m3.^3), data = ds.1.acute, weights = No.of.Subjects.N.)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-59.507   -5.371   -2.358    1.129   74.804

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                    6.659e-01  2.595e+00   0.257    0.799
Cumulative.Dose..mg.h.m3.      8.505e-03  5.293e-03   1.607    0.117
I(Cumulative.Dose..mg.h.m3.^2) 8.256e-07  1.567e-06   0.527    0.602
I(Cumulative.Dose..mg.h.m3.^3) -1.615e-10 1.116e-10  -1.447    0.157

(Dispersion parameter for gaussian family taken to be 721.7406)

    Null deviance: 70628  on 38  degrees of freedom
Residual deviance: 25261  on 35  degrees of freedom
AIC: 305.52

Number of Fisher Scoring iterations: 2
```

22

Thus, the cubic model that follows is $y = b_0 + b_1(dose) + b_2(dose)^2 + b_3(dose)^3$; $b_0 = 6.659\text{e-}01$, $b_1 = 8.505\text{e-}03$, $b_2 = 8.256\text{e-}07$, $b_3 = -1.615\text{e-}10$; and is shown in the plot below:
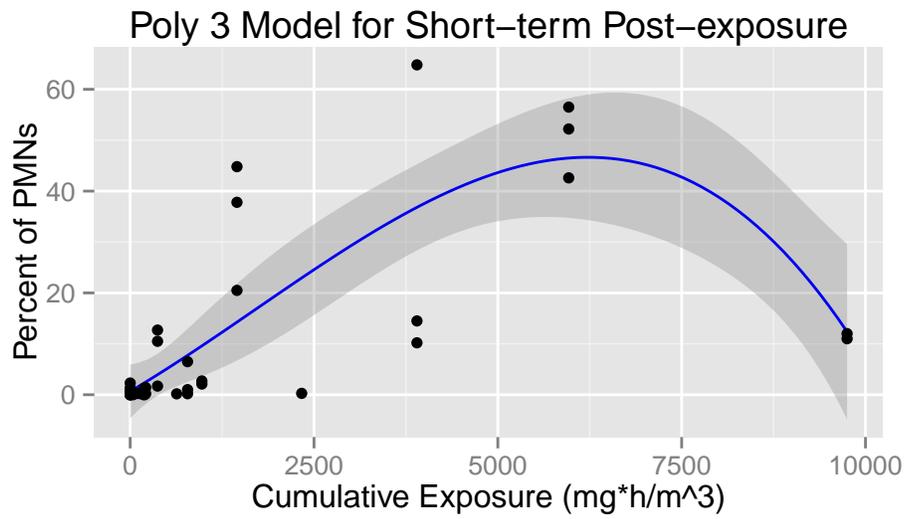

Figure 14: Cubic (Poly 3) Model fit for short-term post-exposure pooled dataset.
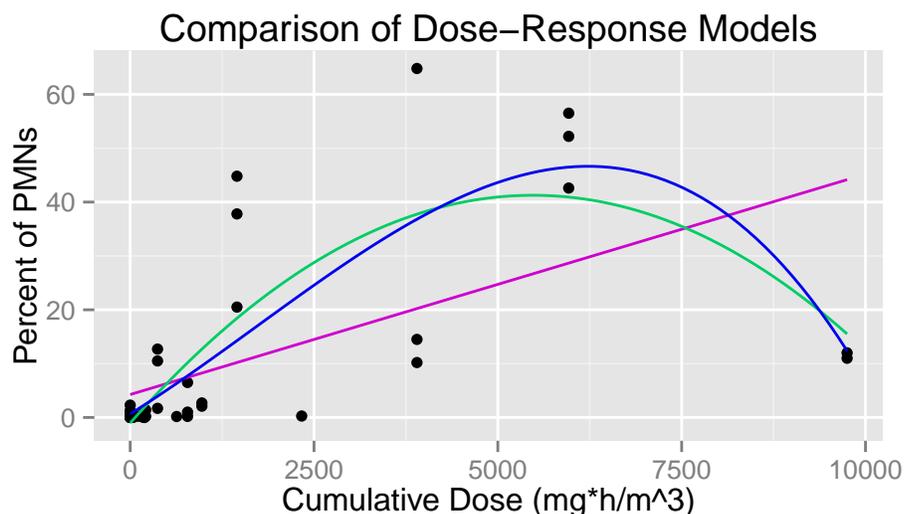
Comparison of Polynomial Models

Figure 15: Comparison of the weighted polynomial dose-response models. Pink - Linear (Poly 1); Green - Quadratic (Poly 2); Blue - Cubic (Poly 3).

Adding Covariates to the Basic Linear Model of the Pooled Data

Table 7: Possible covariates to consider for explaining variability among our mean responses from various individual toxicology studies.

	TiO₂	MWCNT	SWCNT	CNF
Species/Strain	Rat Fischer , Wistar Mouse B6C3F1/CrlBr Hamster Syrian	Rat Wistar	Mouse C57BL/6	Rat Sprague-dawley
Gender	Female / Male	Male	Female	Female / Male

The table of possible covariates shows that there are direct correlations between material with species/strain as well as material and gender, because for all of the materials (except TiO2) the mean responses only come from one species/strain and the same for gender. Thus, planning the modeling development it is only sufficient to model the linear model adding material as a covariate, to evaluate experimental factors that may be significant for explaining the variability in

the data. Adding material as a covariate helps to evaluate the relative hazard potencies across the ENMs as well.

The linear model with material a covariate is shown as follows:

Table 8: Linear Model fit for the short-term post-exposure pooled dataset with Material as a Covariate.

```
> my.lin.reg.2.se.a<-glm(Percent.of.PMNs~Material*Cumulative.Dose..mg.h.m3.,weights =
No.of.Subjects.N.,data = ds.1.acute)
> summary(my.lin.reg.2.se.a)#Covariate(s): Material #weighted=No. of Animals

Call:
glm(formula = Percent.of.PMNs ~ Material * Cumulative.Dose..mg.h.m3.,
    data = ds.1.acute, weights = No.of.Subjects.N.)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-55.321   -8.175   -0.067    0.788   72.044

Coefficients:
                                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                                0.9653231  4.9133176   0.196    0.846
MaterialMWCNT                             -0.9339423  7.6121695  -0.123    0.903
MaterialSWCNT                             -0.9653231  8.9198976  -0.108    0.915
MaterialTiO2                               2.8256788  5.6750233   0.498    0.622
Cumulative.Dose..mg.h.m3.                  0.0010835  0.0010026   1.081    0.288
MaterialMWCNT:Cumulative.Dose..mg.h.m3.  -0.0009728  0.0054586  -0.178    0.860
MaterialSWCNT:Cumulative.Dose..mg.h.m3.   0.0010977  0.0953715   0.012    0.991
MaterialTiO2:Cumulative.Dose..mg.h.m3.    0.0069034  0.0014712   4.693 5.17e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 665.0866)

    Null deviance: 70628  on 38  degrees of freedom
Residual deviance: 20618  on 31  degrees of freedom
AIC: 305.6

Number of Fisher Scoring iterations: 2
```

The linear model can be written in the following form:

$$y = b_0 + b_1(dose) + b_2(I_{MWCNT}) + b_3(I_{SWCNT}) + b_4(I_{TiO_2}) + b_5(I_{MWCNT})(dose) + b_6(I_{SWCNT})(dose) + b_1 = 0.001$$

where b_0= 0.9653231; b_1= 0.0010835; b_2= -0.9339423; b_3= -0.9653231; b_4= 2.8256788; b_5= -0.0009728; b_6= 0.0010977; b_7= 0.0069034 ; and I is an indicator parameter (0 or 1 depending on if the material is present).
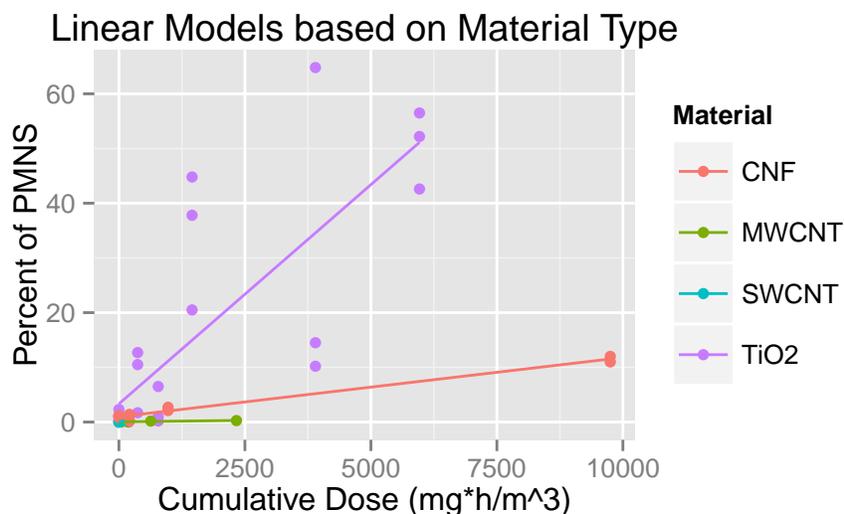
25

Figure 16: Comparison of Linear Models for the various Material Types. This plot shows the effect of various nano-material types on PMN response, TiO2 (titatnium dioxide) looks to be significant in increasing the PMN response over CNF, MWCNT, and SWCNT for this biological endpoint.

From the slope estimates of the model output and the corresponding visual plot of the various linear models based on material, then for the PMN endpoint, TiO2 is significantly more potent than CNFs, MWCNTs, and SWCNTs. TiO2 is followed in potency by CNF, but dose not seem to be as significant in the low dose regions above MWCNTs and SWCNTs, and the multi-walled and single-walled carbon nanotubes do not seem to have much of an effect over the PMN response.

Discussion

Model Fitting and Adding Covariates. The polynomial models (assuming constant variance) from BMDS were successfully replicated in R with a function that uses similar best fit optimization methods as implemented in BMDS. This provided us the opportunity to fit multiple regression models to the pooled dataset.

Before fitting the pooled data with the polynomial models it was necessary to considered the underlying variance structure so that the model could be weighted appropriately. The first method that was evaluated was weighting by precision, but further exploration into the precisions across the data it was apparent that there are some nuances in the variance structure. Across the board there were vast differences in precisions across the data, in the low dose

region were higher in precision as opposed to those in the higher dose regions. Between the maximum and minimum precisions there were about 8 orders of magnitude difference, and the highly precise data points were all contained within that low dose region, and even the precisions within that low dose region were quite varied. The large magnitude of difference between precisions makes the higher dose groups have essentially zero precision, which then leads to a model fit that is determined mostly by data points in the low dose region and does not give a good description of the data. Thus, the next method of weighting that was explored was to weight the model using the number of animals per dose group. The number of animals per dose groups was a more reasonable representation, for this analysis, of variability in the data than weighting by the precisions because of the smaller range in the number of animals per dose group (5-12 animals) across all of the dose groups.

The short-term post-exposure pooled data was modeled with all of the polynomial models (i.e. Linear, Polynomial 2 and Polynomial 3), seen in Fig 12-14. The AIC decreased from the linear model to the quadratic, but the cubic AIC did not have a very significant drop from the quadratic. Thus, the quadratic and cubic are about the same when describing the trend in the data (for the dose-response trend for general nanomaterials) and is a "better model fit" over the linear model. However, when adding the covariates the simplest model to consider was the linear model.

Adding covariates to the linear model, material was the main variable of interest so was added first to the model output seen in Table 8. Other variables of interest were species/strain, and gender, but there were direct correlations between material and species/strain as well as material and gender. If these variables were added to the model fit then there are two consequences. Firstly, adding directly correlated variables to a model might over-parameterize the model, fitting a model that is inaccurate. Secondly, might fit a model that estimates parameters that would rely upon making assumptions about biological processes that may not be feasible to make. Thus, it was only valid to add material as a covariate to the linear model.

Comparing the basic linear model and the linear model with with material as a covariate, there was drop in the AIC from 325.98 to 305.6, thus material is significant variable. Thus, this means that there is a difference in the potencies across the various ENMs and we can evaluate their potencies using the model. Looking at the slope estimates and Figure 16 we can see that for the PMN endpoint that TiO2 is significantly more potent that CNFs, MWCTs, and SWCNTs. The statistical tests for significance of the model parameters (and associated p-values) are not considered reliable due to the nonconstant variance observed in the data (i.e. increasing variance with increasing dose, as seen in the TiO2 data). A model that adequately accounts for the variance structure was not feasible in these data due to the large differences in the precision estimates based on small sample sizes.

Data Limitations.   Although the replication of models from BMDS to R was successful and it was possible to write a model with covariates to evaluate the potencies of various ENMs, there were some limitations in the data that forced some assumptions to be made. Some of the limitations and their consequences are as follows:

- Sparse data for several variables and correlations in the variables.

A small amount of data and sparse data for several of the variables limited the number of variables that could be evaluated for significance on their effect of evaluating the potecies of various nanomaterials. Thus, may have limited our ability to describe some of the variability in the data.

- Working with Summary Statistics.

Working with summary statistics limited our ability to see how the individual data was distributed, which affects our ability to model variability.

- Relatively few material types represented in the dataset.

The number of nanomaterials that exist and are used in industry are numerous, and the dataset for this analysis only covers a small snippet of a broad spectrum of materials. This limits the ability of researchers to be able to evaluate potencies of the numerous nanomaterials and set OELs to protect workers.

- Much of the data was in the low dose region.

Data in the low dose region had a huge impact on the overall model fit and made the information in the higher dose region "less" important even though all of the data in the higher dose regions are important in the grand scheme of developing an adequate OEL.

- Heterogeneity in variability in PMN response.

Heterogeneous variance issues cannot be ignored in these data, and more exploration into the underlying variance structure is necessary to be able to write an adequate dose-response model from which a BMD and BMDL can be estimated from. Furthermore, for this analysis since the model was weighted using an alternative method than precision and the uncertainty in the variance structure, then there were also limitations in the validity of statistical tests for the model fit of the data (i.e. cannot rely upon the significance tests for the model parameters when adding covariates due to the nonconstant variance observed in the data).

   All of these data limitations would be important to explore and account for in further explorations of dose-response modeling problems.

Conclusion

This analysis was crucial in demonstrating that data from multiple toxicology studies can be pooled together and evaluated to estimate potency based on

material type (and other factors if sufficient data are available for modeling). Importantly, BMDS models could be replicated into R, which gives more flexibility to dose-response modeling and allows one to be able to evaluate different variables across a number of toxicology studies. Finally, this analysis showed that material is in fact an important factor for estimating BMD/BMDLs, which means that potencies across ENMs are not the same and therefore OELs is not sufficient for all ENMs.

References

[1] Bermudez, E. et al. (2014). Pulmonary responses of Mice, Rats, and Hamsters to Subchronic Inhalation of Ultrafine Titanium Dioxide Particles, Toxicol Sci. 77(2): 347-357.

[2] Creutzenberg, O. (2013). Final Report of the Project: "Toxic Effects of Various Modifications of a Nanoparticle Following Inhalation" – Project F 2246 – on behalf of the Federal Institute for Occupational Safety and Health (BAuA), Germany.

[3] Crump KS [1984]. A new method for determining allowable daily intakes. Fund Appl Toxicol 4(5):854–871.

[4] Crump KS [1995]. Calculation of benchmark doses from continuous data. Risk Anal 15(1):79–89.

[5] EPA [2012]. Benchmark Dose Technical Guidance. Risk Assessment Forum. U.S. Environmental Protection Agency Washington, DC. EPA/100/R-12/001.

[6] Gernand JM, Casman EA [2014]. A meta-analysis of carbon nanotube pulmonary toxicity studies–how physical dimensions and impurities affect the toxicity of carbon nanotubes. Risk Anal 34(3):583-597.

[7] Health Council of the Netherlands. Benchmark dose method: health-based recommended exposure limits in new perspective. The Hague: Health Council of the Netherlands, 2003; publication no. 2003/06E.

[8] Kalberlah F, Föst U, Schneider K [2002]. Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Ann Occup Hyg 46(2):175-85.

[9] Kalberlah F, Schneider K, Schuhmacher-Wolz U [2003]. Uncertainty in toxicological risk assessment for non-carcinogenic health effects. Regul Toxicol Pharmacol 37(1):92-104.

[10] NAS [1983]. Risk assessment in the federal government: Managing the process. Washington DC: National Academy of Sciences.

[11] NAS [2009]. Science and Decisions: Advancing Risk Assessment. Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies, National Research Council of the National Academies of Sciences. Washington, DC: The National Academies Press. ISBN-10: 0-309-12046-2.

[12] NIOSH (2011). Current Intelligence Bulletin 63: Occupational Exposure to Titanium Dioxide. Cincinnati, Ohio: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health. NIOSH (DHHS) Publication No. 2011-160

[13] NTP [2015]. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Office of Health Assessment and Translation (OHAT), Division of the National Toxicology Program, National Institute of Environmental Health Sciences. January 9.

[14] OECD [2014]. Guidance Notes on Nanomaterials: Interspecies Variability Factors in Human Health Risk Assessment. Working Party on Manufactured Nanomaterials. Organisation for Economic Co-operation and Development 18-Dec-2014. ENV/CHEM/NANO(2015)4.

[15] Stahlmann R, Dabrowski A [2013]. Expert Opinion: Interspecies Variability of the Toxicity of Nanomaterials. Prepared in Support of the OECD Working Party for Manufactured Nanomaterials (WPMN) and REACH Adaption on Nanomaterials. [Unpublished report].

Appendix

The R Code for this analysis can be found at the following website http://bit.ly/1peenBa.