A meta-analysis of carbon nanotube pulmonary toxicity studies – How physical dimensions and impurities affect the toxicity of carbon nanotubes

Jeremy Gernand[†] **and Elizabeth Casman**[†]

[†]Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA

This document is a currently unpublished chapter of the doctoral dissertation of J.M. Gernand. It has been submitted for publication to a peer reviewed journal and is currently under review.

ABSTRACT

This paper presents a regression-tree-based meta-analysis of rodent pulmonary toxicity studies of uncoated, non-functionalized carbon nanotube (CNT) exposure. The resulting analysis provides quantitative estimates of the contribution of CNT attributes (impurities, physical dimensions, aggregation) to pulmonary toxicity indicators in bronchoalveolar lavage fluid: neutrophil and macrophage count, and lactate dehydrogenase and total protein concentrations. The method employs classification and regression tree (CART) models, techniques which are relatively insensitive to data defects that impair other types of regression analysis: high-dimensionality, non-linearity, correlated variables, and significant quantities of missing values. Three types of analysis are presented, the regression tree, the random forest, and a random forest based dose-response model. The regression tree shows the best single model supported by all the data and typically contains a small number of variables. The random forest shows how much variance reduction is associated with every variable in the data set. The dose-response model is used to isolate the effects of CNT attributes from the CNT dose, showing the shift in the dose-response caused by the attribute across the measured range of CNT doses. It was found that the CNT attributes that contribute the most to pulmonary toxicity were metallic impurities, CNT length and diameter, and aggregate size. Increasing CNT specific surface area decreased toxicity indicators.

Keywords: Carbon nanotubes, inhalation, toxicity, regression tree, meta-analysis

1. INTRODUCTION

At one time the fibrous structure of carbon nanotubes (CNTs) led some to suspect that inhaled CNTs might cause pathologies similar to those associated with inhaled asbestos[(1)]. While CNT pulmonary exposure has not been observed to cause mesotheliomas[1], it does lead to the formation of granulomas[(2)], lung inflammation, and fibrotic responses[(3)].

The expected proliferation of products containing engineered carbon nanotubes in many configurations will mean that efforts to test toxicity of these variants through conventional animal experimentation will be burdensome. If the attributes of exposure and nanotube characteristics responsible for variations in CNT toxicity could be identified, this could potentially reduce the testing burden and provide guidance for the design and production of safer CNTs. This is the motivation of various attempts to identify property/toxicity relationships for nanomaterials, including this one.

The development of quantitative structure activity relationships (QSARs) for nanomaterials is also motivated by the desire to transition out of the animal testing paradigm. The s goal of QSARs is to predict the activity—in this case, toxicity—of a material from its chemical and physical structure. QSAR studies and toxicity modeling studies for nanomaterials have focused predominately on *in vitro* toxicity indicators, especially those amenable to high throughput screening[(4–9)].The only QSAR toxicity study on CNTs[(10)] to date concluded that the cytotoxicity of multi-walled carbon nanotubes (MWCNT) to bacteria is enhanced by chemical treatments that increase MWCNT-cell contact opportunities (uncapping, debundling, length shortening, and dispersion).

---

[1]Mesothelioma has been induced in p53+/- mice by injecting CNTs into their peritoneal cavities[(47)] and in Fischer 344 rats via intra-scrotal injection[(48)], but not by pulmonary routes of exposure[(49)].

Our work differs from these studies in its focus on quantitative *in vivo* pulmonary toxicity indicators, and those factors varying between different batches of carbon nanotubes, specifically physical dimensions and impurity content. For nearly all of the other traditional QSAR descriptors, such as polarizability, hydrophobicity, or surface charge, there are currently no of studies of the *in vivo* toxicity of functionalized or otherwise chemically modified CNTs that could be used to determine the effects of these properties. The major differences that do exist between batches of CNTs used in the literature on pulmonary toxicity studies are variations in dimensions, dispersion or aggregation, impurity content, and CNT configuration (multi-walled versus single-walled).

Only one effort has been made to date towards meta-analysis of CNT toxicity studies, and it did not address physical or chemical variation in the CNT materials themselves[11], focusing instead on the question of whether penetration of cell membranes is likely to occur in *in vitro* experiments. The study concluded that penetration of human cells is "somewhat possible" and could result in cellular damage, recommending the minimization of human exposure.

One challenge of our meta-analysis was that in the absence of standardized nanomaterial characterization protocols. Investigators typically used similar but different procedures. Researchers often only reported the properties of the nanomaterial believed to be important for their research question, leaving a substantial amount of uncertainty regarding the actual form of the material. For example, CNT inhalation studies differ significantly in measurement and reporting of the distribution of aggregate sizes[12, 13], the amount and type of metallic impurities[14, 15], and the presence of carbon impurities[16, 17]. These practices can make employing traditional meta-analysis techniques difficult, because they result in a data set with many missing values.

This paper presents a meta-analysis based on machine-learning-algorithms, specifically regression trees (RT)[18] and the associated ensemble method, random forests (RF)[19]. These methods experience less degradation from missing data than ordinary least squares regression, enabling meaning to be extracted from the relatively thinly populated data set.

The objectives of this research are three-fold, (1) to rank CNT properties and experimental conditions by their information content, (2) to quantify the influence of different CNT properties on dose-response relationships, and (3) to lay the computational foundation RT and RF based meta-analysis for toxicology studies.

2. DATA AND METHODS

2.1. Data selection

As of October 2012, the archival literature contained 17 carbon nanotube toxicity studies using rodent (rat and mouse) models exposed through inhalation, aspiration, or intratracheal instillation that met our screening criteria. To be included, studies had to report at least minimal CNT characterization and quantitative toxicity output measures, at least one of which also occurred in another published study (Table I). We did not include studies whose endpoints were the presence or absence of gross pathologies because limiting the data set to studies with continuous toxicity endpoints permitted greater contrast to be made between the effects of the different input variables.

(Table 1 near here)

Four individual pulmonary toxicity endpoints covering a range of effects were reported in the identified studies: polymorphonuclear neutrophils (PMN), macrophages (MAC), lactate dehydrogenase (LDH) and total protein (TP). These endpoints reflect several dimensions of

immune response and cell membrane damage and death. These indicators were all measured in bronchoalveolar lavage (BAL) fluid extracted from the lungs of the mice or rats, and were reported as a counts per subject or fold of control measurements (the indicator count or concentration in test subjects divided by the count or concentration in control animals).

We converted all toxicity measures reported as absolute counts or concentration to fold of control measurements, the form in which many of the studies already reported. We found by experimentation that this step was key for comparing test subjects across species and across modes of exposure. This data set is available at http://nanohub.org/resources/13515[20].

Table I. Summary of published data on *in vivo* pulmonary exposures of carbon nanotubes included in the meta-analysis.

Publication Reference	Author(s)	Year	CNT Configuration	Exposure Mode	Animal	Average Aggregate Size (nm)	Purity (%)	Total Dose (μg/kg)	Post Exposure Period (days)
(21)	Warheit, D. et al.	2004	SWCNT	instillation	rats		90	1000 - 5000	1 - 90
(14)	Muller, J. et al.	2005	MWCNT	instillation	rats		97.8	2 - 8	3 - 60
(22)	Shvedova, A. et al.	2005	SWCNT	aspiration	mice		99.7	490 - 1970	1 - 60
(23)	Shvedova, A. et al.	2007	SWCNT	aspiration	mice		99.7	1851	1
(24)	Shvedova, A. et al.	2008	SWCNT/MWCNT	instillation	mice	4200	82	250 - 1000	1 - 7
(25)	Muller, J. et al.	2008	MWCNT	instillation	rats		95 - 99	8890	3
(26)	Elgrabli D. et al.	2008	MWCNT	instillation	rats			5 - 500	1 - 180
(27)	Mercer R. et al.	2008	SWCNT	aspiration	mice	690	98	303	1 - 30
(28)	Inoue K. et al.	2008	SWCNT/MWCNT	instillation	mice			4000	1
(29)	Ryman-Rasmussen J. et al.	2008	MWCNT	inhalation	mice	714	94	12000	1 - 14
(17)	Ma-Hock, L. et al.	2009	MWCNT	inhalation	rats	400	90	190 - 2400	3 - 21
(12)	Nygaard, U.	2009	SWCNT/MWCNT	aspiration	mice		95	5000 - 20000	26
(30)	Ellinger-Ziegelbauer H. et al.	2009	MWCNT	inhalation	rats	2900		180 - 3900	7 - 90
(13)	Pauluhn, J.	2010	MWCNT	inhalation	rats	1670 - 2190	98.6	105 - 6290	1 - 90
(31)	Porter D. et al.	2010	MWCNT	aspiration	mice		99.5	435 - 1740	1 - 56
(32)	Park, E-J et al.	2011	SWCNT	instillation	mice		90	100	1 - 28
(33)	Teeguarden, J.G. et al.	2011	SWCNT	aspiration	mice		99.7	12000	1

2.2. Data preparation

We treated all nanoparticle properties and experimental conditions as independent variables. In Table II there are four categorical variables, CNT configuration, exposure mode, sex, and animal species. The inclusion of these variables is a key feature of the meta-analysis, allowing the dissimilar experimental systems to be included in a single analysis. Another important variable was the Post-Exposure Period, the time after exposure that elapsed before the BAL fluid was examined for toxicity indicators. Table I only shows dose measured in μg CNT per kg of animal. This is just one of five ways dose was reported in the input data (Table II), and it was the dose unit that resulted in the regression trees with the highest explanatory power.

(Table II near here)

The input for RT and RF analysis is a matrix in which the rows represent individual experimental animals. The 41 columns of the matrix contained 20 experimental conditions, 17 nanoparticle properties (such as CNT length, diameter, specific surface area, and average aggregate size), and the 4 experimental endpoints (toxicity measures). For diagnostic purposes, a series of 2 to 6 uniformly distributed random variables between 0 and 1 were added to the data array as columns at different stages of the analysis. The random variables—by definition information-free and not correlated with any toxicity endpoint values—supplied an independent measure with which to compare the discriminatory performance between information-rich and information-poor attributes of the RT and RF model generation and pruning algorithms. Missing values for any particular nanomaterial property or experimental condition in any published study were left blank.

Several variables utilized in the analysis were calculated or inferred from the data when not reported directly in the individual studies. For example, values for median CNT length and

diameter were derived from reported maximum and minimum values. When not reported, total and 24-hour doses for each experiment were calculated from reported concentrations. Specific surface area and total mass dose were combined to produce a value for surface area dose. In no cases were nanomaterial properties or experimental values inserted based on typical values or model-predicted quantities. In the individual studies, the toxicity endpoint results were all reported as a mean and standard deviation. We assumed the endpoint values were normally distributed.

The number of matrix rows per experiment was 100 times the number of animals in each experimental group. For example, if an experiment involved 12 test animals at one dose and 12 test animals at a different dose, the experiment would occupy 1200 + 1200 = 2400 rows. Each of these rows would contain identical nanomaterial and experimental attributes, the appropriate exposure metrics, but for the experimental endpoints (PMN, MAC, LDH, and TP), each row would have a unique, discrete realization of a normal distribution with mean and standard deviation reported for its group of animals. (Endpoint values deemed impossible—negative protein counts, for instance—were discarded and replaced with the nearest plausible value, usually zero.) Higher rates of sampling at 500 or 1000 samples per experimental animal did not alter the ordinal value of model variables or goodness of fit measures in the resulting RTs.

The probabilistic representation of the experimental results preserves the inherent variability in the measured responses and reduces the likelihood of the models reporting noise or of their oversensitivity to small differences in experimental exposure conditions. The assumption of a normal distribution for the toxicity endpoint measures is conservative, a choice that likely overestimates the amount of uncertainty present. This procedure also has the effect of weighting the influence of each study on the RT and RF models in relation to the number of animal subjects

involved. Furthermore, incorporating experimental uncertainty of the data in this manner enables the models to output the corresponding range of likely values rather than being limited to a single value.

Table II. List of the 37 input attributes included in models and their definitions.

Category	Variable	Units	Definition
Size and Shape	configuration	MWCNT or SWCNT	a categorical variable indicating whether the carbon nanotubes are multi-walled (MWCNTs) or single walled (SWCNTs)
	minimum CNT length	nm	the minimum reported length of the free individual CNT fibers either measured or stated by manufacturer's specifications
	median CNT length	nm	the median length of the free individual CNT fibers either measured or stated by manufacturer's specifications
	maximum CNT length	nm	the maximum reported length of the free individual CNT fibers either measured or stated by manufacturer's specifications
	minimum CNT diameter	nm	the minimum reported diameter of the free individual CNT fibers either measured or stated by manufacturer's specifications
	median CNT diameter	nm	the median diameter of the free individual CNT fibers either measured or stated by manufacturer's specifications
	maximum CNT diameter	nm	the maximum reported diameter of the free individual CNT fibers either measured or stated by manufacturer's specifications
	aggregate diameter (MMAD)	nm	mass mode aerodynamic diameter – the most frequent size of the particle aggregates by the mode of the distribution by mass
	specific surface area	m^2/g	specific surface area as measured by the N2-BET (Nitrogen, Brunauer-Emmett-Teller) gas adsorption method
	purity	%	the fraction by percent mass of the amount of the CNT sample composed of carbon atoms
Impurities	dose cobalt [total]	pg/kg	the total dose received by the animal subject of cobalt impurities present in the CNT particulate
	dose cobalt [24 hr average]	pg/kg	the average daily dose received by the animal subject of cobalt impurities present in the CNT particulate
	dose aluminum [total]	pg/kg	the total dose received by the animal subject of aluminum impurities present in the CNT particulate
	dose aluminum [24 hr average]	pg/kg	the average daily dose received by the animal subject of aluminum impurities present in the CNT particulate
	dose iron [total]	pg/kg	the total dose received by the animal subject of iron impurities present in the CNT particulate
	dose iron [24 hr average]	pg/kg	the average daily dose received by the animal subject of iron impurities present in the CNT particulate
	dose copper [total]	pg/kg	the total dose received by the animal subject of copper impurities present in the CNT particulate
	dose copper [24 hr average]	pg/kg	the average daily dose received by the animal subject of copper impurities present in the CNT particulate
	dose chromium [total]	pg/kg	the total dose received by the animal subject of chromium impurities present in the CNT particulate

Category	Variable	Units	Definition
	dose chromium [24 hr average]	pg/kg	the average daily dose received by the animal subject of chromium impurities present in the CNT particulate
	dose nickel [total]	pg/kg	the total dose received by the animal subject of nickel impurities present in the CNT particulate
	dose nickel [24 hr average]	pg/kg	the average daily dose received by the animal subject of nickel impurities present in the CNT particulate
	dose oxidized carbon [total]	pg/kg	the total dose received by the animal subject of oxidized carbon impurities present in the CNT particulate
	dose oxidized carbon [24-hr average]	pg/kg	the average daily dose received by the animal subject of oxidized carbon impurities present in the CNT particulate
Exposure characteristics	exposure hours	hours	the number of hours that the animal subject was exposed to the CNTs
	first to last exposure period	hours	the time period in hours between the first hour of exposure and the last hour of exposure by the animal subject to CNTs
	animal species	rat or mouse	a categorical variable indicating the type of the animal subject
	animal breed or strain	Sprague-Dawley, Wistar, C57BL/6, ICR, Crl:CD(SD)IGS BR,BALB/cAnNCrl	a categorical variable indicating the specific breed or strain of the animal subject
	sex	male or female	A categorical variable indicating the sex of the animal subject
	mean animal mass	g	the mean mass of the animal subjects in a given experiment
	post exposure	days	the number of days between the final exposure to CNTs and the sacrifice and measurement of the toxicity status of the subject, also referred to as recovery period
	exposure mode	inhalation, instillation, aspiration	a categorical variable indicating the mode of exposure
	mass concentration	mg/m^3	the mass concentration of CNTs in the air of the animal subject inhalation chamber (inhalation exposures only)
	total mass dose	µg/kg	the total mass dose of CNTs received over the course of the experiment by the animal subject
	average mass dose	µg/kg/24-hr	the average daily mass dose of CNTs received over the course of the experiment (first to last exposure period) by the animal subject
	total surface area dose	m^2/kg	the highest peak hourly surface area dose of CNTs received over the course of the experiment by the animal subject
	average surface area dose	$m^2/kg/24$-hr	the average daily surface area dose of CNTs received over the course of the experiment (first to last exposure period) by the animal subject

2.3. Regression trees

For each measured toxicity endpoint we created a RT model using the MATLAB™ function "classregtree." The RT algorithm successively divides the population of observations (toxicity endpoint values, also called output variables) into binary groups based on an inequality for quantitative input variables or a categorical grouping for categorical variables. At each branching point, the input variable and the split criterion (a value of the input variable) are chosen to produce the greatest possible information gain between the resulting two populations of observations. In this study we developed and applied a calculation of variance reduction as a measure of information gain. While entropy increase may also be used as a measure of information gain, it does not result in substantially different results, and variance reduction is more consistent with the statistics produced by alternative regression models, facilitating comparisons.

To prevent over-fitting, each RT model is pruned through a process similar to backwards stepwise elimination for linear models. In order to prune the model, the error of the model is calculated through ten-fold cross-validation—a process by which the data set is divided randomly into 10 subsets of rows, the RT model branches and split criteria are set based on a matrix containing 9 of those subsets, and the error is measured against the model predictions for the 10^{th} subset. This process is repeated until each of the 10 subsets has been withheld from the model and used for error calculation, and then the total sum-squared error for the model is calculated. The RT model is then pruned by removing the branches providing the least error reduction until the model reaches the smallest size possible within one standard error of the minimum error. The RT model for neutrophils is shown in Figure 1. The other three RTs can be found at http://nanohub.org/resources/15901.

The importance of each variable included in the RT model is calculated based on the total variance reduction achieved by each branch in the final model. For variables appearing on multiple branches, the variance reduction is summed across the entire model.

One reads a tree from the top down, following each branch to its terminal leaf. For example, in Figure 1, if the total CNT dose is less than 5150 μg/kg and the dose of Cr is more than 3.13 μg/kg and neutrophils were measured less than 17.5 days after exposure, and the dose of Fe was less than 2600 μg/kg, and the neutrophils were measured more than 4 days after exposure, and the total CNT dose was more than 1300 μg/kg, then the neutrophil counts in the 600 matrix rows representing animals satisfying these conditions would be on average 293 times higher than neutrophils in control animals.

While linear models and other algebraic models are undefined whenever any of the input variables are missing, this is not true of a RT model. Since overall 19% of the characterization or exposure attributes in the data matrix are missing and 74% of the variables have at least one missing value, complete utilization of all available data is a distinct advantage for RTs over multiple linear regression (MLR) models. Further, this RT property of producing a prediction regardless of the completeness of the data record prevents the underestimation of error that is common when regression models are applied to data sets with missing values.

While other machine learning methods such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) can also handle nonlinear relationships and may be more computationally efficient, they become undefined for cases where any of the input data is missing. So, while imputation of missing values would be required for MLR, ANN, or SVM models to take full advantage of the entire experimental data set, this step and potential source of artificially introduced uncertainty is unnecessary for RTs.

RTs show the best model for a given data set in the sense of parsimony, but there are ways to interrogate the data that are even less sensitive to their biases, such as the Random Forest (RF).

2.4. Random Forests

RF models have the advantage of more fully exploring the value of each variable in the data set, compared to RT models, which typically include a small subset of them. They also are less sensitive to weaknesses in the data. With RTs, variables can appear influential due to their over-representation in the data. The randomized variable selection process for the RF model generation procedure ensures that all variables are evaluated.

We generated a RF model for each toxicity endpoint implementing Breiman's algorithm with the MATLAB™ function "treebagger." The RF model is composed of an aggregate collection of RT models, each created from a data set where each branch is selected from a random subset of one third of the available variables (matrix columns). Each RF model contains at least 1000 RT models, whose elements are averaged to produce the RF output. The RFs were grown (i.e. trees added) until the error decreased less than 0.1% with the addition of 100 trees to the RF model. (See http://nanohub.org/resources/15901 for full model documentation.)

The variable importance results from the RF model are calculated in the same way as for the RT model; however, the final variance reduction for each variable is calculated from the average across all of the trees in the RF model. This is accomplished by calculating the variance reduction caused by each branch variable in each individual tree in the forest, summing the variance reductions by variable across all of the trees in the RF, and then dividing the final values by the total number of trees in the RF model. These error reduction values form the

measure of variable information content used in ranking the variables by their importance in Figure 2.

The RF models used for ranking variables and in the dose-response curve generation (in the next section) were trimmed by first removing all variables with less calculated importance than the average of the uniform random variables. Then, model refinement proceeded in a forward stepwise fashion evaluated on the basis of the R^2 value. In each step, the additional variable producing the highest possible R^2 value was selected.

To investigate the contribution of individual variables to a baseline CNT dose-response curve, we use a reduced version of the RF model containing only the variables accounting for >99% of the variance reduction plus the variable of interest. We ran this model at minimum and maximum levels of the variable of interest in the input matrix, holding all other variables constant at their median values while varying the dose variable. The two resulting curves show the shift in the dose-response associated with the variable of interest (Figure 3). Calculating the mean shift across the range of applied CNT doses enables a quantitative evaluation of the relative effect of different variables on toxicity (See Results).

In order to estimate the uncertainty in each of these model results, a RT model was trained to predict the observed experimental variance rather than the observed mean output value. Mean dose response curves plus and minus one estimated standard deviation are plotted in Figure 3. The RF standard deviations reflect an expectation of the population standard deviation.

The pixelated appearance of RF dose-response curves is due to the number of unique experimental data points available in a given sub-section of the attribute space. If more data were available, the shape of the response curve would be smoothed. Locations of abrupt steps in the RF curves are reflective of fewer data points being available under those conditions.

Of the 37 input variables, the RT and RF models contain only between 3 and 13 variables (Table III), yet the information these models explain approaches the maximum possible information value a model could contain if all differences between experimental groups[2] are fully captured. (See the 'Maximum Model Performance' column in Table III.) When the number of variables in a RF model is much smaller than the number of exposure groups, this gives confidence that the RF model is not over-fit. That the performance of the RF models approaches the limit of explaining all but the inherent experimental variability means that the most significant controllable factors affecting CNT toxicity are explicitly or implicitly included.

(Table III near here)

Table III. Model goodness-of-fit performance and the number of model input attributes included.

Output Variable	**Regression Tree**		**Random Forest**		**Maximum Model Performance**	
	R^2	# of variables	R^2	# of variables	R^2	# of exposure groups
Neutrophils	0.89	5	0.83	5	0.90	103
Macrophages	0.62	3	0.84	3	0.88	62
Lactate Dehydrogenase	0.84	7	0.89	5	0.90	84
Total Protein	0.92	13	0.95	5	0.96	67

2.4. Model Validation

Table IV displays a study-by-study validation of the RF model in terms of mean squared error (MSE). The best RF model was generated after withholding the target study from the training data set. Then, the specifics of the withheld study were fed into the model as test data inputs. The output of the model is what would be predicted if that experiment had yet to be performed. The

[2] An exposure group is defined as a group of animals that experienced the same level of exposure to the same batch of CNTs and the same length of recovery period.

experimental results and model predictions (and residuals) display the overall good performance of this modeling technique for anticipating the results of future experiments. The RF model cannot extrapolate based on trends, however, and can only be used in this manner for combinations of inputs that lie within the limits of the training data.

(Table IV near here.)

As seen in Table IV, the mean squared error of the overall model to all observations is similar to that resulting when each individual study is withheld and used as a test case. The one outlier, the study by Pauluhn et al., 2010, involved an inhalation exposure with a much longer exposure period than the other studies. The MSE also reflects the relative similarity in results between the selected study and the information contained in the others, as illustrated in particular by the 90-day inhalation study (Pauluhn, 2010), and the LDH results for the Shvedova et al., 2007 study. The overall impression from these results indicates that the RF models are generally reliable when predicting the results of a new study so long as no significant extrapolation of any of the input variables of the training data is required.

A complete model description and validation document including code with model training results is available at http://nanohub.org/resources/15901.

Table IV. The RF model validation results in terms of mean squared error (MSE) for each of the studies included in this analysis and the observed change in neutrophil count when that particular study was excluded from the training data.

First Author, Publication Date	Total Dose (μg/kg)	Recovery Period (days)	Exposure Mode	RF Model MSE (fold of control squared)
Neutrophils				
Pauluhn, 2010	105 - 6290	1 - 90	inhalation	38,500
Muller, 2005	2 - 8	3	instillation	6,300
Shvedova, 2008	250 - 1000	1	instillation	6,700
Porter, 2010	435 - 1740	1 - 28	aspiration	1,600
Inoue, 2008	4000	1	instillation	2,100
Shvedova, 2005	490 - 1970	1 - 60	aspiration	260
Shvedova, 2007	1851	1	aspiration	2,900
MSE for all points in full model	--	--	--	2,700
Macrophages				
Pauluhn, 2010	105 - 6,290	1 - 90	inhalation	46,700
Shvedova, 2008	250 - 1,000	1 - 7	instillation	7,000
Nygaard, 2009	5,000 - 20,000	26	aspiration	41
Mercer, 2008	303	1 - 30	aspiration	1,200
Shvedova, 2005	490 - 1,970	1 - 60	aspiration	740
MSE for all points in full model	--	--	--	7,900
Lactate Dehydrogenase				
Pauluhn, 2010	105 - 6,290	1 - 90	inhalation	46,700
Muller, 2005	2 - 8	3	instillation	6,200
Shvedova, 2008	250 - 1,000	1 - 7	instillation	4,100
Shvedova, 2008	250 - 1,000	1	instillation	4,900
Warheit, 2004	1,000 - 5,000	1 - 30	Instillation	10,000
Muller, 2008	8,890	3	Instillation	2,300
Porter, 2010	435 - 1,740	1 - 28	aspiration	300
Ellinger-Ziegelbauer, 2009	180 - 3,900	7 - 90	inhalation	1,500
Shvedova, 2007	1,851	1	aspiration	28,800
MSE for all points in full model	--	--	--	8,400
Total Protein				
Pauluhn, 2010	105 - 6,290	1 - 90	inhalation	46,000
Ma-Hock, 2009	190 - 2,400	3 - 21	inhalation	8,100
Muller, 2005	2 - 8	3	instillation	37
Muller, 2008	8,890	3	instillation	1,000
Inoue, 2008	4,000	1	instillation	2,400
Ellinger-Ziegelbauer, 2009	180 - 3,900	7 - 90	inhalation	2,700
Shvedova, 2005	490 - 1,970	1 - 60	aspiration	22
Shvedova, 2007	1,851	1	aspiration	2,800
MSE for all points in full model	--	--	--	8,500

3. RESULTS

The RT model for PMN displayed in Figure 1 demonstrates the structure and grouping of PMN observations based on differences between the various experiments and exposure conditions in the literature. Five unique variables are included in this model, some appearing in multiple branches. The model variables are a mix of experimental conditions and CNT properties, with the total dose by mass (positively correlated with PMN count) providing the greatest variance reduction, followed by post-exposure time (mostly negatively correlated with PMN count), the dose of chromium (positively correlated with PMN count), the dose of iron (negatively correlated with PMN count), and the dose of cobalt (positively correlated with PMN count).

(Figure 1 near here)

The random forest model permits us to put a more comprehensive quantitative value on the information gain or variance reduction provided by each possible model variable, rather than just the few that might appear in the best RT model. Figure 2 shows that the total dose by mass, the dose of cobalt, the post exposure time, and mass concentration (an exposure attribute for inhalation experiments, measured in $\mu g/m^3$) contribute most to the RF PMN model's ability to explain variance, a similar but more robust result than the best RT. It would be imprecise to say that the variables with the longest bars in Figure 2 are the most important for CNT toxicity. It is better to say that those variables are responsible for most of the performance of the best RF model. If those variables were deleted from the data set, it is likely that other variables that are correlated with them would take their place in the new RF without totally undermining the RF model's ability to explain the variance in the data.

(Figure 2 near here)

We also used RF models to contrast the effects of different levels of input variables on the output variables controlling for CNT dose. This is analogous to the kind of experiment that would contrast two batches of CNTs made with different catalysts at the same CNT dose, except that it does this for the full range of reported CNT doses. Since RF models are not for extrapolation, we could only vary the tested variables over the range represented in the data set. For the output variable PMN, several variables could increase the dose response relationship by more than 25% (Figure 3 and Table V). These were median diameter, mass mode aerodynamic diameter (MMAD), and cobalt content. Shorter median lengths and smaller specific surface areas also increased the dose response relationship. The negative signs in Table V indicate inverse relationships. Thus, for example, CNTs with the highest specific surface area (1,040 m^2/g) decreased the neutrophil response by 90%. For Macrophages, the same five variables were also important. For Lactate Dehydrogenase, the most influential variable was short median length, followed by cobalt content; and for Total Protein, the geometric variables, median length and median diameter, and MMAD were salient.

(Figure 3 near here)

(Table V near here)

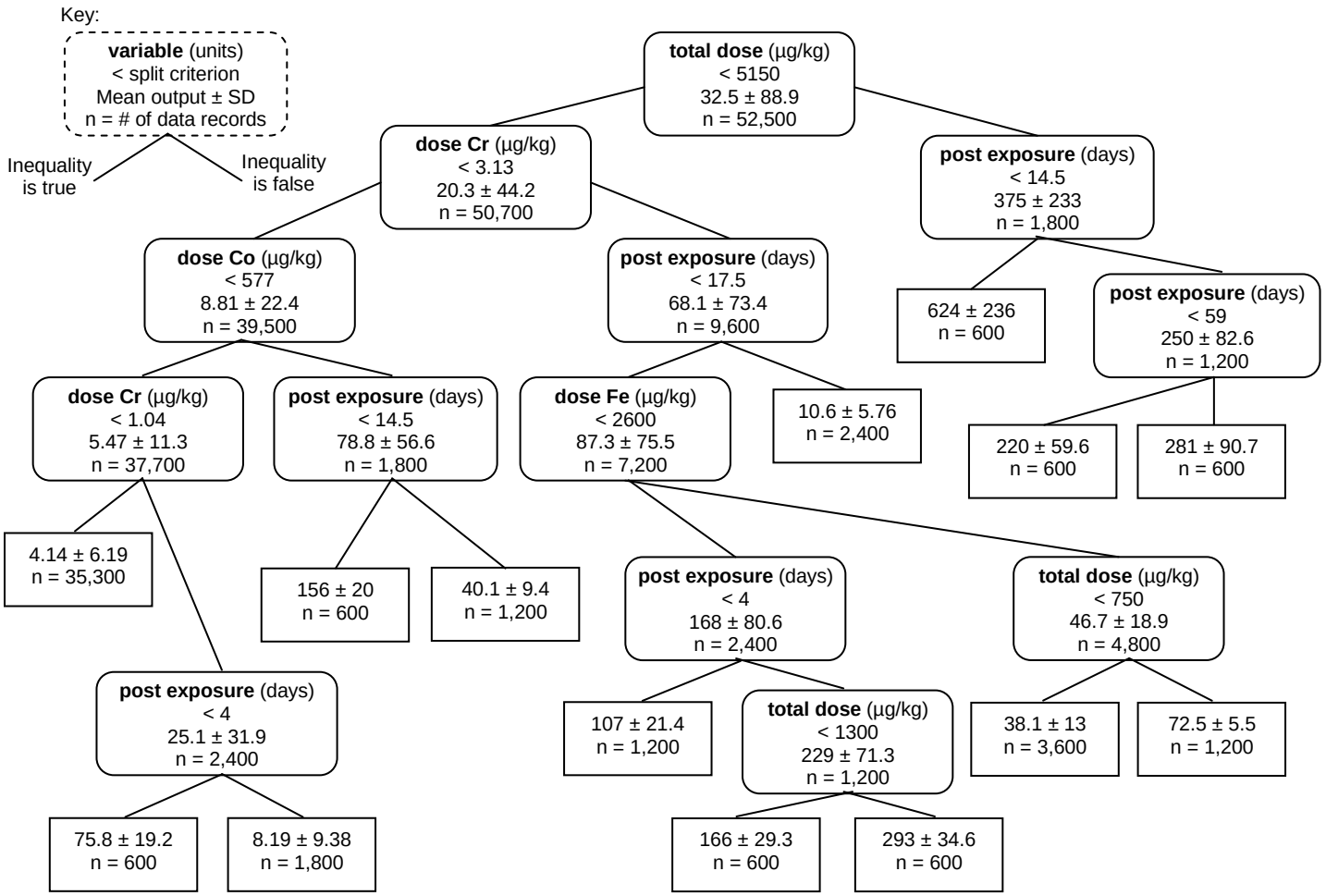

Figure 1. Regression tree model for the effect of CNT exposure on BAL neutrophils. The mean output values in the leaf nodes (rectangular terminal nodes) are the model's predictions.

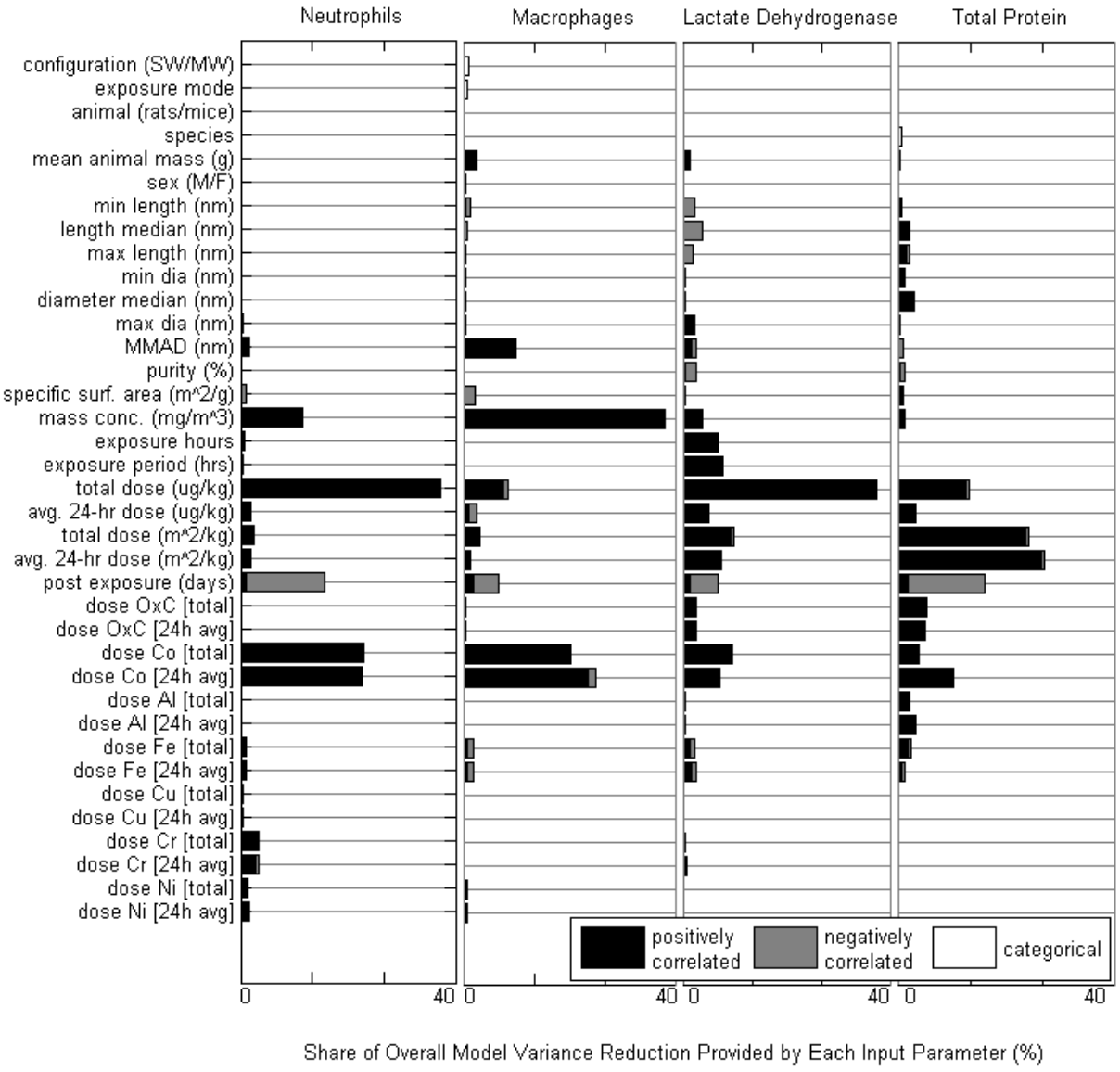

Figure 2. Random forest model variable importance comparison as measured by mean variance reduction (shown as share of total model variance reduction) for all four toxicity indicators. Bar shading indicates whether the variance reduction occurred with a positive or negative correlation between the input and output, or with a branch based on a non-numeric categorical variable.

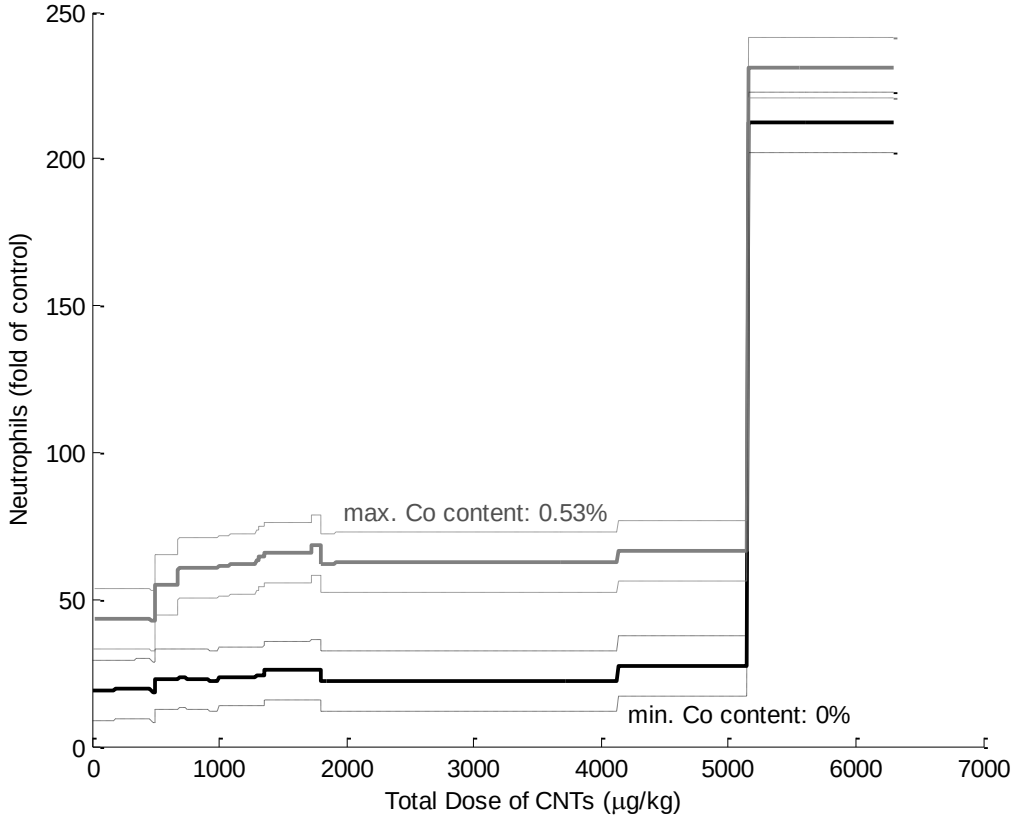

Figure 3. RF dose-response graphs of the effects of CNT exposure on PMNs at 0% and 0.53% cobalt impurity by mass. At CNT doses below 5000 μg/kg, CNTs with 0.5% cobalt double the neutrophil response. Dashed and dotted lines indicate the response ± the experimental standard deviation. All other model input variables are held at their median values.

Table V. Mean relative shift (%) observed in CNT total mass dose-response curves in four BAL inflammation indicators from changes in select CNT attributes based on RF model results. In each case the reported shift reflects the effect of changing the input variable from its reported minimum numerical value to its maximum in the input data set.

		Output Variables			
Input Variable	**Minimum/Maximum Value (units)**	**Neutrophils (%)**	**Macrophages (%)**	**Lactate Dehydrogenase (%)**	**Total Protein (%)**
Median Length	320/100,000 (nm)	−51	−36	−60	+30
Median Diameter	0.8/49 (nm)	+46	+32	+13	+72
MMAD	400/4,200 (nm)	+62	+29	−23	−64
Specific Surface Area	109/1,040 (m^2/g)	−90	−28	−19	+2
Aluminum Content	0/9.6(%)	+0	+16	+15	+10
Cobalt Content	0/0.53 (%)	+57	+58	+28	+1
Iron Content	0/17.7 (%)	+14	−20	+0	+16
Copper Content	0/0.16 (%)	−4	−8	−1	−16
Nickel Content	0/5.53 (%)	+3	+0	−1	+0
Chromium Content	0/0.05 (%)	+21	−8	+3	−25

4. DISCUSSION

While the literature conveys different conclusions with respect to whether normal levels of impurities dominate[34] or insignificantly affect[35] CNT toxicity, whether increasing aggregation is mitigating[36] or aggravating[37], whether thinner SWCNTs are more toxic[38] or less toxic[12], and whether aspect ratio[39] or surface area[40] could account for differences in biological interactions with CNTs, the results of this study reveal that some of these effects are non-linear (aggregation), others apparently insignificant as currently measured (aspect ratio), and some confirmed aggravators (impurities and increasing CNT diameter).

4.1. Effects of CNT dimensions and aggregation

Larger aggregate sizes increased PMN and MAC and decreased Total Protein and LDH. Larger aggregate sizes increased the post-exposure time required to recover to levels approaching normal. The largest aggregates resulted in cell counts as much as 30% higher than those observed with the smallest aggregates, for the same post-exposure times. These results are consistent with the idea that larger aggregates are more difficult to clear from the lungs, but cause less damage to cell membranes on an equivalent mass basis. While this result is inconsistent with the findings of some *in vitro* studies that increasing aggregate size leads to greater cell membrane damage[41], it does confirm the observation from one *in vivo* study where larger aggregates increased neutrophil count and reduced LDH release[14].

Increases in the median length of CNTs produced a marked decrease in the observed toxicity (−30% to −60%), except for in BAL total protein (+30%), contrary to expectations that increasing aspect ratio would increase toxicity. Aspect ratio (length divided by diameter) was tested as a predictive factor in the models, but provided negligible predictive value compared to either length or diameter alone. While CNT lengths varied by several orders of magnitude (from

a few hundred nanometers up to 100 micrometers), diameters for all CNTs varied within a much smaller range, of between 1 and 35 nm.[3] This significantly reduces the utility of aspect ratio as a CNT toxicity predictor, relative to length and stiffness.

While the categorical attribute, single-walled *vs.* multi-walled CNTs, produced only a limited shift in the dose-response curves (<2%), the diameter of the CNTs (a plausible proxy for the difference between single-walled and multi-walled CNTs) proved to consistently increase toxicity (causing shifts in dose response curves between +13 and +72%). These findings are consistent with the hypothesis that stiffer nanofibers (MWCNTs) produce greater cell damage in the lungs and present more resistance to the body's natural particle breakdown and clearance systems[(42, 43)].

The model preference of the continuous variable, diameter, over the categorical variable suggests that not all multi-walled CNTs are equally toxic. It is important, however, to realize that this could also reflect other properties correlated in certain parts of the variable space, such as different catalyst recipes for different CNT structures.

Contrary to observations of non-fibrous nanomaterials, increasing the specific surface area of the CNTs had a generally negative effect on toxicity, reducing the responses of most toxicity indicators or having very little effect[(44)]. CNT surface area measurements were all made by the N_2-BET method, which relates the amount of N_2 adsorbed to a sample to its surface area. Surface area values in our data generally increased with or were unaffected by increasing aggregate sizes. Micrographs of CNT aggregates in the source papers typically showed loose,

[3] It should be noted then that aspect ratio may be comparable between a long, but relatively thin and flexible SWCNT and a much stiffer MWCNT with an even greater length, while their toxicity mechanisms may differ.

tangled structures, relatively open and more consistently 3-dimensional from the perspective of N_2 molecules than disaggregated CNTs. Thus the interiors of the tangled aggregates would be shielded from interactions with biological surfaces, even while contributing significantly to the surface area measurement. The N_2-BET measurement methodology appears to act as a proxy for the degree of CNT aggregation as it is positively correlated with a coefficient of 0.88. This interpretation is consistent with the experimental finding that CNTs manipulated to be smaller and better dispersed had a smaller measured specific surface area and induced a more severe toxic response [(14)].

4.2. Effects of impurities

CNT pulmonary toxicity is significantly increased but not dominated by the amounts of metallic impurities. As these metals are removed from the CNTs, toxicity is reduced, but not eliminated. Of all the metal contaminants, toxicity measures were most negatively impacted by the content of cobalt. Other metals including chromium, nickel, aluminum, iron, and copper had relatively weak or inconsistent effects on observed toxicity.

Impurity contents are not truly independent variables but reflect a few specific catalyst constituent combinations, so the relative toxicity of one metal over another cannot be definitively addressed with the current data. One of the principal metals co-occurrence groupings is composed of cobalt and nickel (correlation coefficient of 0.992) with another grouping composed of iron, copper, and chromium (correlation coefficient of 0.988 between Fe and Cu or Cr). However, given the greatest RF response to impurities results from cobalt exposure, and the fact that cobalt is a known sensitizer [(45)], and has been associated together with cobalt carbides in acute lung toxicity [(46)], it is plausible that the dominant effect of cobalt is real. Though the cobalt

content of the CNTs in these experiments only varied between 0% and 0.53%, this was sufficient to produce a large change in the CNT dose response curve.

4.3. Exposure Mode

Contrary to expectations, the exposure mode of the experiment did not have a significant value in explaining the difference in study outcomes, only accounting for a relative variance reduction (in relation to the variance reduction of the total model) of between 0.01% and 0.5%.We expected that the differences in CNT deposition patterns between instillation exposures and inhalation exposures and the uncertainty in translating ambient particle concentration to total dose received would have produced a significant difference in the observed response for a given dose. The fact that the models did not find this factor significant means that either there is a non-obvious proxy for exposure mode contained within the combinations of other experimental variables, or that for these conditions (relatively high dose and short duration experiments) the mode of exposure is minimally influential to experimental outcomes.

5. CONCLUSIONS

We have developed a suite of machine-learning-based methods for meta-analysis of nanotoxicity studies that has enabled the extraction of information across studies that was not observable in the individual studies. This work has identified a short list of CNT attributes that contribute to uncoated unfunctionalized CNT pulmonary toxicity, metallic impurities, CNT length and diameter, surface area, and aggregate size (the variables listed in Table V), because they alter the CNT dose-response relationship. Being the result of a meta-analysis, our results are limited by the experimental choices made in the archival literature and do not involve a theory of toxicity. Nevertheless, this work has demonstrated the utility of RT and RF methods for identifying and

ranking attributes of nanoparticles that contribute to their toxicity and quantifying those contributions.

ACKNOWLEDGEMENTS

Funding for this research was provided by the National Science Foundation (NSF) and the Environmental Protection Agency (EPA) under NSF Cooperative Agreement EF-0830093, Center for the Environmental Implications of NanoTechnology (CEINT), the Carnegie Institute of Technology (CIT) Dean's Fellowship, the Prem Narain Srivastava Legacy Fellowship, the Neil and Jo Bushnell Fellowship, and the Steinbrenner Institute for Environmental Education and Research (SEER). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or the EPA. This work has not been subjected to EPA review and no official endorsement should be inferred.

REFERENCES

1. Kane B, Hurt RH. The asbestos analogy revisited. Nature Nanotechnology 2008;3(July):378–379.

2. Varga C, Szendi K. Carbon nanotubes induce granulomas but not mesotheliomas. In Vivo 2010;24(2):153–156.

3. Muller J, Delos M, Panin N, Rabolli V, Huaux F, Lison D. Absence of carcinogenic response to multiwall carbon nanotubes in a 2-year bioassay in the peritoneal cavity of the rat. Toxicol. Sci. 2009 Aug;110(2):442–448.

4. Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A, Hwang H-M, Toropov A, Leszczynska D, Leszczynski J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nature nanotechnology 2011 Feb;6(February):175–178.

5. Sayes C, Ivanov I. Comparative study of predictive computational models for nanoparticle-induced cytotoxicity. Risk Analysis 2010 Nov;30(11):1723–1734.

6. Liu R, Rallo R, George S, Ji Z, Nair S, Nel A, Cohen Y. Classification nanoSAR development for cytotoxicity of metal oxide nanoparticles. Small 2011;7(8):1118–1126.

7. Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ, Tropsha A. Quantitative nanostructure-activity relationship modeling. ACS nano 2010 Oct;4(10):5703–5712.

8. Fourches D, Pu D, Tropsha A. Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles. Comb. Chem. High T. Scr. 2011;14:217–225.

9. Kang S, Mauter MS, Elimelech M. Physicochemical Determinants of Multiwalled Carbon Nanotube Bacterial Cytotoxicity. Environmental Science & Technology 2008 Oct;42(19):7528–7534.

10. Kang S, Mauter MS. Physicochemical Determinants of Multiwalled Carbon Nanotube Bacterial Cytotoxicity. Environmental Science & Technology 2008;42(19):7528–7534.

11. Genaidy A, Tolaymat T, Sequeira R, Rinder M, Dionysiou D. Health effects of exposure to carbon nanofibers: systematic review, critical appraisal, meta analysis and research to practice perspectives. The Science of the total environment 2009 Jun;407(12):3686–701.

12. Nygaard UC, Hansen JS, Samuelsen M, Alberg T, Marioara CD, Løvik M. Single-walled and multi-walled carbon nanotubes promote allergic immune responses in mice. Toxicol. Sci. 2009 May;109(1):113–123.

13. Pauluhn J. Subchronic 13-week inhalation exposure of rats to multiwalled carbon nanotubes: toxic effects are determined by density of agglomerate structures, not fibrillar structures. Toxicol. Sci. 2010;113(1):226–242.

14. Muller J, Huaux F, Moreau N, Misson P, Heilier J-F, Delos M, Arras M, Fonseca A, Nagy JB, Lison D. Respiratory toxicity of multi-wall carbon nanotubes. Tox. App. Pharma. 2005;207(3):221–231.

15. Warheit DB, Webb TR, Colvin VL, Sayes CM. Pulmonary bioassay studies with nanoscale and fine-quartz particles in rats: toxicity is not dependent upon particle size but on surface characteristics. Toxicol. Sci. 2007;95(1):270–280.

16. Mitchell LA, Gao J, Wal R Vander, Gigliotti A, Burchiel SW, McDonald JD. Pulmonary and systemic immune response to inhaled multiwalled carbon nanotubes. Toxicol. Sci. 2007 Nov;100(1):203–214.

17. Ma-Hock L, Treumann S, Strauss V, Brill S, Luizi F, Mertler M, Wiench K, Gamer AO, Van Ravenzwaay B, Landsiedel R. Inhalation toxicity of multiwall carbon nanotubes in rats exposed for 3 months. Toxicol. Sci. 2009;112(2):468–481.

18. Breiman L, Friedman J, Stone C, Olshen RA. Classification and Regression Trees. New York: Chapman and Hall/CRC; 1984.

19. Breiman L. Random Forests. Machine Learning 2001 May;45:5–32.

20. Gernand J. Carbon Nanotube (CNT) Pulmonary Toxicity Data Set [Internet]. 2012;Available from: http://nanohub.org/resources/13515

21. Warheit DB, Laurence BR, Reed KL, Roach DH, Reynolds GAM, Webb TR. Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats. Toxicol. Sci. 2004;77:117–125.

22. Shvedova AA, Kisin ER, Mercer R, Murray AR, Johnson VJ, Potapovich AI, Tyurina YY, Gorelik O, Arepalli S, Schwegler-Berry D, Hubbs AF, Antonini J, Evans DE, Ku B-K, Ramsey D, Maynard A, Kagan VE, Castranova V, Baron P. Unusual inflammatory and fibrogenic pulmonary responses to single-walled carbon nanotubes in mice. Am J. Physiol.-Lung C. 2005;289(5):698–708.

23. Shvedova AA, Kisin ER, Murray AR, Gorelik O, Arepalli S, Castranova V, Young S, Gao F, Tyurina YY, Oury TD, Kagan VE. Vitamin E deficiency enhances pulmonary inflammatory response and oxidative stress induced by single walled carbon nanotubes in C57BL/6 mice. Toxicol. Appl. Pharmacol. 2007;221(3):339–348.

24. Shvedova AA, Kisin E, Murray AR, Johnson VJ, Gorelik O, Arepalli S, Hubbs AF, Mercer RR, Keohavong P, Sussman N, Jin J, Yin J, Stone S, Chen BT, Deye G, Maynard A, Castranova V, Baron PA, Kagan VE. Inhalation vs. aspiration of single-walled carbon nanotubes in C57BL/6 mice: inflammation, fibrosis, oxidative stress, and mutagenesis. Am J. Physiol.-Lung C. 2008;295(4):552–565.

25. Muller J, Huaux F, Fonseca A, Nagy JB, Moreau N, Delos M, Raymundo-Pinero E, Beguin F, Kirsch-Volders M, Fenoglio I, Fubini B, Lison D. Structural Defects Play a Major Role in the Acute Lung Toxicity of Multiwall Carbon Nanotubes: Toxicological Aspects. Chem. Res. Toxicol. 2008;21:1698–1705.

26. Elgrabli D, Abella-Gallart S, Robidel F, Rogerieux R, Boczkowski J, Lacroix G. Induction of apoptosis and absence of inflammation in rat lung after intratracheal instillation of multiwalled carbon nanotubes. Toxicology 2008;253:131–136.

27. Mercer RR, Scabilloni J, Wang L, Kisin E, Murray a R, Schwegler-Berry D, Shvedova a a, Castranova V. Alteration of deposition pattern and pulmonary response as a result of improved dispersion of aspirated single-walled carbon nanotubes in a mouse model. Am J. Physiol.-Lung C. 2008 Jan;294(1):L87–97.

28. Inoue K-I, Takano H, Koike E, Yanagisawa R, Sakurai M, Tasaka S, Ishizaka A, Shimada A. Effects of pulmonary exposure to carbon nanotubes on lung and systemic inflammation

with coagulatory disturbance induced by lipopolysaccharide in mice. Exp. Biol. Med. 2008 Dec;233(12):1583–1590.

29. Ryman-Rasmussen JP, Tewksbury EW, Moss OR, Cesta MF, Wong B a, Bonner JC. Inhaled multiwalled carbon nanotubes potentiate airway fibrosis in murine allergic asthma. Am. J. Respir. Cell Mol. Biol. 2009 Mar;40(3):349–358.

30. Ellinger-Ziegelbauer H, Pauluhn J. Pulmonary toxicity of multi-walled carbon nanotubes (Baytubes) relative to alpha-quartz following a single 6h inhalation exposure of rats and a 3 months post-exposure period. Toxicology 2009 Dec;266(1-3):16–29.

31. Porter DW, Hubbs AF, Mercer RR, Wu N, Wolfarth MG, Sriram K, Leonard S, Battelli L, Schwegler-Berry D, Friend S, Andrew M, Chen BT, Tsuruoka S, Endo M, Castranova V. Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes. Toxicology 2010;269:136–147.

32. Park E-J, Roh J, Kim S-N, Kang M-S, Han Y-A, Kim Y, Hong JT, Choi K. A single intratracheal instillation of single-walled carbon nanotubes induced early lung fibrosis and subchronic tissue damage in mice. Archives of toxicology 2011 Sep;85(9):1121–31.

33. Teeguarden J, Webb-Robertson B, Waters K, Murray A, Kisin E, Varnum S, Jacobs J, Pounds J, Zanger R, Shvedova A. Comparative proteomics and pulmonary toxicity of instilled single-walled carbon nanotubes, crocidolite adbestos, and ultrafine carbon black in mice. Toxicol. Sci. 2011;120(1):123–135.

34. Hull MS, Kennedy AJ, Steevens J a, Bednar AJ, Weiss C a, Vikesland PJ. Release of metal impurities from carbon nanomaterials influences aquatic toxicity. Environ. Sci. Technol. 2009 Jun;43(11):4169–4174.

35. Pauluhn J. Multi-walled carbon nanotubes (Baytubes): approach for derivation of occupational exposure limit. Regul. Toxicol. Pharm. 2010 Jun;57(1):78–89.

36. Oberdörster G, Oberdörster E, Oberdörster J. Nanotoxicology: An Emerging Discipline Evolving from Studies of Ultrafine Particles. Environ. Health 2005;113(7):823–839.

37. Helland A, Wick P, Koehler A, Schmid K, Som C. Reviewing the environmental and human health knowledge base of carbon nanotubes. Environ. Health Perspect. 2007;115(8):1125–1131.

38. Kang S, Herzberg M, Rodrigues DF, Elimelech M. Antibacterial effects of carbon nanotubes: size does matter! Langmuir 2008 Jun;24(13):6409–6413.

39. Kang S, Mauter MS, Elimelech M. Physicochemical determinants of multiwalled carbon nanotube bacterial cytotoxicity. Environ. Sci. Technol. 2008 Oct;42(19):7528–7534.

40. Hubbs AF, Mercer RR, Benkovic SA, Harkema J, Sriram K, Schwegler-Berry D, Goravanahally MP, Nurkiewicz TR, Castranova V, Sargent LM. Nanotoxicology--A Pathologist's Perspective. Toxicol. Path. 2011 Jan;39(2):301–324.

41. Wick P, Manser P, Limbach LK, Dettlaff-Weglikowska U, Krumeich F, Roth S, Stark WJ, Bruinink A. The degree and kind of agglomeration affect carbon nanotube cytotoxicity. Toxicol. Lett. 2007 Jan;168(2):121–131.

42. Kostarelos K. The long and short of carbon nanotube toxicity. Nature biotechnology 2008 Jul;26(7):774–776.

43. Donaldson K, Aitken R, Tran L, Stone V, Duffin R, Forrest G, Alexander A. Carbon nanotubes: a review of their properties in relation to pulmonary toxicology and workplace safety. Toxicol. Sci. 2006;92(1):5–22.

44. Watari F, Takashi N, Yokoyama A, Uo M, Akasaka T, Sato Y, Abe S, Totsuka Y, Tohji K. Material nanosizing effect on living organisms: non-specific, biointeractive, physical size effects. Journal of the Royal Society 2009 Jun;6 Suppl 3(April):S371–88.

45. Lison D. Human toxicity of cobalt-containing dust and experimental studies on the mechanism of interstitial lung disease (hard metal disease). Crit. Rev. Toxicol. 1996;26(6):585–616.

46. Lison D, Boeck M De, Verougstraete V, Kirsch-Volders M. Update on the genotoxicity and carcinogenicity of cobalt compounds. J. Occup. Environ. Med. 2001;58(10):619–625.

47. Takagi A, Hirose A, Nishimura T, Fukumori N, Ogata A, Ohashi N, Kitajima S, Kanno J. Induction of mesothelioma in p53+/- mouse by intraperitoneal application of multi-wall carbon nanotube. J. of Tox. Sci. 2008 Feb;33(1):105–116.

48. Sakamoto Y, Nakae D, Fukumori N, Tayama K, Maekawa A, Imai K, Hirose A, Nishimura T, Ohashi N, Ogata A. Induction of mesothelioma by a single intrascrotal administration of multi-wall carbon nanotube in intact Fischer 344 rats. J. of Tox. Sci. 2009;34:65–76.

49. Hubbs AF, Mercer RR, Benkovic SA, Harkema J, Sriram K, Schwegler-Berry D, Goravanahally MP, Nurkiewicz TR, Castranova V, Sargent LM. Nanotoxicology--A Pathologist's Perspective. Toxicol. Path. 2011 Jan;39(2):301–324.