SELECTING NANOPARTICLE PROPERTIES TO MITIGATE RISKS TO WORKERS AND THE PUBLIC – A MACHINE LEARNING MODELING FRAMEWORK TO COMPARE PULMONARY TOXICITY RISKS OF NANOMATERIALS

Jeremy M. Gernand
Carnegie Mellon University
Pittsburgh, PA, USA

Elizabeth A. Casman
Carnegie Mellon University
Pittsburgh, PA, USA

ABSTRACT

Due to their size and unique chemical properties, nanomaterials have the potential to interact with living organisms in novel ways, leading to a spectrum of negative consequences. Though a relatively new materials science, already nanomaterial variants in the process of becoming too numerous to be screened for toxicity individually by traditional and expensive animal testing. As with conventional pollutants, the resulting backlog of untested new materials means that interim industry and regulatory risk management measures may be mismatched to the actual risk. The ability to minimize toxicity risk from a nanomaterial during the product or system design phase would simplify the risk assessment process and contribute to increased worker and consumer safety.

Some attempts to address this problem have been made, primarily analyzing data from *in vitro* experiments, which are of limited predictive value for the effects on whole organisms. The existing data on the toxicity of inhaled nanomaterials in animal models is sparse in comparison to the number of potential factors that may contribute to or aggravate nanomaterial toxicity, limiting the power of conventional statistical analysis to detect property/toxicity relationships. This situation is exacerbated by the fact that exhaustive chemical and physical characterization of all nanomaterial attributes in these studies is rare, due to resource or equipment constraints and dissimilar investigator priorities.

This paper presents risk assessment models developed through a meta-analysis of *in vivo* nanomaterial rodent-inhalational toxicity studies. We apply machine learning techniques including regression trees and the related ensemble method, random forests in order to determine the relative contribution of different physical and chemical attributes on observed toxicity. These methods permit the use of data records with missing information without substituting presumed values and can reveal complex data relationships even in nonlinear contexts or conditional situations.

Based on this analysis, we present a predictive risk model for the severity of inhaled nanomaterial toxicity based on a given set of nanomaterial attributes. This model reveals the anticipated change in the expected toxic response to choices of nanomaterial design (such as physical dimensions or chemical makeup). This methodology is intended to aid nanomaterial designers in identifying nanomaterial attributes that contribute to toxicity, giving them the opportunity to substitute safer variants while continuing to meet functional objectives.

Findings from this analysis indicate that carbon nanotube (CNT) impurities explain at most 30% of the variance pulmonary toxicity as measured by polymorphonuclear neutrophils (PMN) count. Titanium dioxide nanoparticle size and aggregation affected the observed toxic response by less than ±10%. Difference in observed effects for a group of metal oxide nanoparticle associated with differences in Gibbs Free Energy on lactate dehydrogenase (LDH) concentrations amount to only 4% to the total variance. Other chemical descriptors of metal oxides were unimportant.

INTRODUCTION

Nanoparticles or ultrafine particles, sometimes designated PM0.1 (i.e. particulate matter 0.1 micrometers or less), are solid particles where at least one dimension is in the range of 1-100 nanometers. A nanometer is one billionth of a meter, and is just slightly longer than 3 molecules of H_2O would stretch if lined up end to end. Since many properties of solids including magnetism, electrical conductivity, and tensile strength arise only after sufficient numbers of atoms have aggregated together, the properties exhibited by nanoparticles can either be in gentle transition from solute to solid behavior or reveal steep threshold transitions to some combination of

characteristics as particle size increases. Conversely, at small sizes, particles can exhibit an entirely new property like semi-conductivity that is unexpected based on studies of the material bulk properties.

These peculiar aspects of nanoparticles have led to the conjecture that some particular substance in nano-form could prove to be anomalously toxic or more toxic [1]. Since many biological structures and molecules are nano-sized themselves (e.g., like the protein immunoglobulin-G that measures 33 nm across), there is a potential for unique interactions between these small solids and biological processes.

That exposure to nanomaterials in the workplace may result in dangerous pathologies associated with their unique properties is reflected in recent National Institutes of Occupational Safety and Health (NIOSH) recommendations. Existing regulations by the Occupational Safety and Health Administration (OSHA) limit titanium dioxide particulate to a concentration in the workplace of 15 mg/m^3, and limit carbon particulate to a concentration of 5 mg/m^3 [2]. But, upon review of the available research, NIOSH published a recommendation that titanium dioxide nanoparticles be limited to no more than 0.3 mg/m^3 [3], and also proposed that carbon nanotubes (CNTs) be limited to concentrations no greater than 7 $\mu g/m^3$ [4]. The recommended maximum allowable quantities of nano-particulate in the workplace are 50 and 700 times smaller than for carbon soot and micron and larger diameter titanium dioxide respectively. No new OSHA regulations for nanoparticles have yet been formally adopted.

The new proposed exposure limits apply to all particles with a primary size smaller than 100 nm equally, even though substantial differences in toxicity may exist between 20 nm and 40 nm particles, between thinner or thicker nanotubes, or between lightly aggregated or significantly aggregated nanoparticles. Small differences in chemistry including coatings, functionalizations, or contaminants may produce divergent toxic responses. Thus, a limit based on all shapes and varieties for a certain chemical compound sized on average below a single threshold may be inadequate.

Developing specific standards for nanoparticle variants would require a lot more data than the size-based standard and, animal studies, while the most applicable to human risks, are expensive. Since, there are potentially dozens of variable characteristics between different nanoparticles, such studies can only offer limited conclusions on the importance of specific characteristics (for example, that multi-walled carbon nanotubes may be more toxic than single-walled carbon nanotubes [5]).

Reaching conclusions on the interactions of specific characteristics and dose on observed toxic effects is further complicated by the inconsistent or incomplete measurement of nanoparticle properties among the published studies. So, in many published nanotoxicology studies there is uncertainty about the properties of the substance that was tested.

Since, there is limited capability of cellular *in vitro* studies to predict the outcome of *in vivo* mammalian studies [6], and individual animal studies have not adequately studied the variations in nanoparticle properties, understanding the contributions of several nanoparticle properties at once must rely on a virtual experiment assembled from combinations of individual studies across the literature, a meta-analysis [7].

The identification of properties responsible for toxicity could conceivable allow product and process designers to design safer nanoparticles while achieving the same functional objectives. For example, if the length of a carbon nanotube was critical to a design's functionality, but its diameter was not, and diameter proved to be a critical determinant of the carbon nanotube's toxicity risk, careful selection of the CNT's diameter could mitigate that risk without compromising functionality.

There has been limited research to date on predictively modeling the toxicity of nanomaterials, and those studies have focused entirely on cell culture toxicity *in vitro* [8–10]. Summarizing the knowledge gained to date from *in vivo* mammalian pulmonary nanoparticle toxicity studies of relevance to workplace safety is the intent of this paper.

This work seeks to quantify and visualize the degree to which changes in certain nanoparticle characteristics change the overall magnitude of the toxic response to a given dose of nanoparticles. To accomplish this, our study utilizes a machine learning algorithm called the random forest (RF), which has unique capabilities for quantitatively learning from data with a high proportion of missing values, revealing relationships that may be conditional or only applicable after a certain threshold has been passed, and without having to assume statistical independence between each of the inputs. These strengths make RF models especially suited to risk assessment activities in the early stages of implementing a new technology.

METHODS

We perform a meta-analysis of pulmonary nanoparticle toxicity studies in order to determine the degree to which design variables such as chemical composition, dimensions, shape, and surface treatments affect the magnitude of the toxic dose response.

Data Sources

We collected data from published peer-reviewed literature describing experiments where rodents were exposed to nanoparticles through inhalation (dry aerosols), aspiration (a small volume of saline fluid with suspended nanoparticles positioned just beyond the trachea and naturally inhaled by the animal), or instillation (a small volume of saline fluid with suspended nanoparticles injected into the bronchial tubes).

All of the included studies reported quantitative toxicity measures for either the concentration of lactate dehydrogenase (LDH) or the number of polymorphonuclear neutrophils (PMN) in bronchoalveolar lavage (BAL) fluid. BAL is a

procedure where saline fluid is used to rinse out the lungs of the rodent some time following exposure. The fluid is collected along with dislodged cells, particles, and biomolecules and analyzed for indicators of toxicity, like PMN, LDH, and Total Protein.

LDH is a cellular protein and as such is an indicator of cytotoxicity or cell membrane damage in the lungs. This is typically measured as a concentration on a picogram per milliliter basis, but we translate all data for this analysis as a multiple change from control basis (i.e. fold of control). The mean and standard deviation of LDH concentration for a group of animals exposed to nanoparticulate are normalized to the mean measurement for the control group. Control animals are either exposed to only air for inhalation experiments, or an instillation of saline fluid for instillation experiments.

The PMN cell count is an indicator of inflammation and the early stages of an immune response. PMNs are measured in terms of absolute cell counts per milliliter of BAL fluid. We translate these values into "fold of control" in the same way as the LDH values to reflect the change from the control group response.

Studies included in this meta-analysis had to include characterization of the nanoparticles used in the experiment as well as control groups and a quantitatively measured output for PMN, LDH, or Total Protein including reported uncertainty. A complete listing of all data sources utilized in this study is provided in Annex C.

Data Preparation

The input for the data analysis is a matrix with rows representing data specific to individual animals in the selected studies and columns containing experimental and output variables.

Before analysis, the output data for this study were expanded using a Monte Carlo resampling technique. At the rate of 100 samples per animal subject, a given set of experimental inputs including dose levels and nanoparticle characteristics were associated with 100 discrete realizations of the reported distribution of measured output responses. Distributions for measured endpoints were assumed to be normal with values deemed to be impossible (cell counts less than zero, for example) excluded. For example, if the PMN average of 6 animals with a given exposure to nanoparticles was measured to be 10 ± 2, the data set would contain 600 rows (6×100) with identical input characteristics, but each row having a discrete sample from a normal distribution with a mean of 10 and a standard deviation of 2. We found that higher sampling rates (e.g., 500 rows per animal or 1,000 rows per animal) did not alter the results.

This procedure accomplishes two important tasks for this analysis: (1.) it reduces the likelihood of overfitting, since the model must measure its error against the entire range of experimental outcomes and not just the mean values; and (2.) it permits the uncertainty of the multi-dimensional model with respect to the actual measured results to be traced through and evaluated at any desired point or sub-region of interest.

The final data set contained one column for each experimental variable including measured nanoparticle characteristics and attribute of exposure (e.g., total dose, length of recovery, mode of exposure, etc.), and 100 rows for each animal subject utilized in the experiments.

Random Forest Models

We employ random forest (RF) models [11]—an unsupervised machine learning method—to discover and quantify the relationships in the existing data set. RF models are made up of ensembles of regression trees (RT) [12], which are hierarchical structures of decision rules that divide observations into two groups on the basis of a specific criterion (see Figure 1). The decision rules are automatically selected by the algorithm on the basis of those with produce the greatest possible information gain. Random forests extend this process by creating large numbers of regression trees using subsets of the data randomly omitting a fraction of the experimental variables. The results are averaged and considered to be more robust than a single regression tree constructed from the complete data set.

We have implemented Breiman's random forest algorithm via the MATLAB™ function `treebagger`, creating RF models of 1,000 trees each, each branch being established from a randomly selected sub-set of one third of the available input variables. The learning progression diagrams shown in Annex B indicate that the models have already reached their maximum performance with several hundred fewer trees. One may download the final model objects for this study and instructions on implementation here: http://nanohub.org/resources/17539. For further detail on the model implementation, internal structure and validation results, see Annex B.

Visualizing Interactions

In order to visualize the interactions between multiple nanoparticle and exposure characteristics at the same time, we record the RF predicted output while changing 2 or 3 of the input parameters of interest from their minimum to maximum values in 20 steps. For 2 variables, this creates a matrix of 400 values, which we represent as a filled contour plot (see Figure 2). Changes in 3 variables are represented by multiples of two dimensional contour plots, for example showing the relationship between toxicity and changes in length and diameter for different doses of carbon nanotubes as shown in Figure 3.

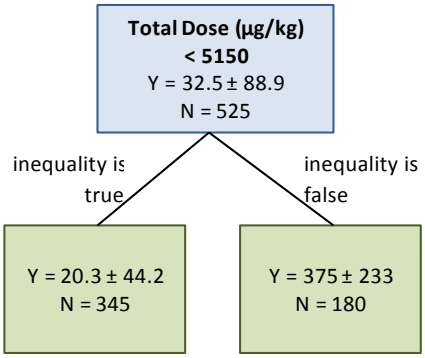


Figure 1: Example of a single-branch regression tree with two leaf nodes displaying the predicted mean output (Y) and standard deviation at each node along with the sample size (N). The branch decision rule in this case is an inequality based on the total dose of nanoparticles received.

RESULTS

Examining the dose-response effects of exposure to nanomaterials (see Figure 2), one can see that for carbon nanotubes the total dose dominates the effects from the length of recovery time in their influence on PMNs, while the recovery from exposure to titanium dioxide nanoparticles dominates the expected LDH and total protein concentrations, while total dose explains most of the LDH observations for metal oxides (Figure A9).

To consider the effects of nanoparticle design tradeoffs, one can see, for example, how length and diameter of CNTs appear to affect toxicity across a range of dose levels (Figure 3). With the highest observed responses in PMN and LDH occuring when the diameter of the CNTs is large and the length of the CNTs is short. These effects are consistent proportionally across several dose levels, even as the total magnitude of the observed response increases.

For titanium dioxide we display the effects of chemical purity and aggregation (MMAD, mass mode aerodynamic diameter—a metric for the average size of aggregated particles) as total dose changes (Figure 4). Aggregation appears to have a limited effect on pulmonary inflammation as compared to changes in purity.

For a broader set of metal oxides, we find that the total dose for an animal subject is a much more important predictor of measured BAL LDH than any other physical or chemical attribute of the nanoparticles (Figure 5). The total dose of metal oxide nanoparticles appears to explain almost all of the variation (Figure 5 and A9)

Additional contour plots generated by these models are included in Annex A. These include the relationship between carbon nanotube dose, cobalt impurity dose and PMN and LDH (Figures A1 and A2); the relationship between titanium dioxide dose and aggregate diameter for LDH and total protein (Figures A3 and A4); the relationship between aggregate diameter and recovery time for titanium dioxide nanoparticles and LDH and total protein (Figures A5 and A6); and the relationship between aggregate diameter and purity (Figure A7) and aggregate diameter and Gibbs Free Energy for BAL LDH following exposure to metal oxide nanoparticles.

DISCUSSION

In terms of an individual designing or specifying a nanomaterial for a particular application and wanting to minimize risks from toxicity at the same time, certain factors including particle size, shape, and chemical makeup would be at the forefront of easily manipulatable design characteristics. If a designer can reduce the toxic risk through careful selection of these factors while continuing to meet functional objectives, they would likely do so.

However, the effects of changes in particle size on toxicity are still a matter of some debate. Aggregation of nanoparticles into larger particles is also debated as whether it may exacerbate toxicity or alternatively to not have a significant effect [13]. It is also unclear from the published literature whether impurities should be considered important or unimportant contributors to toxicity [14], or further, to what actual extent differences in chemical makeup account for differences in toxicity between different nanoparticles.

Effects of Particle Size and Aggregation

Particle size and aggregation were thought to be important determinants of toxicity for nanomaterials, especially the idea that as the particles became smaller and potentially more highly reactive, their toxicity could markedly increase [15]. Although the experimental data including cellular-level in vitro experiments is mixed, larger particle sizes and aggregate do at least sometimes increase the resulting toxicity [16].

For carbon nanotubes, we see an overall effect of both length and diameter. Increasing diameter, which may also be an indicator of carbon nanotube stiffness, was associated with increasing toxicity (Figure 3), with two thresholds of 5nm and 30nm. CNTs with a diameter less than 5nm are single-walled nanotubes, while those with larger diameters are multi-walled. The greatest toxicity is exhibited by CNTs with large diameters and short lengths. By way of comparison, asbestos fibers are on average longer (by 2-5 times) and have much larger diameters (by about two orders of magnitude) than the typical carbon nanotube.

For titanium dioxide, very small particles do in fact seem to generally produce higher dose-response effects than larger titanium dioxide nanoparticles (Figure A10). This only occurs for very small particles, and fewer data points are available in this size range making this conclusion more uncertain than the general observation that dose-response is not influenced much by particle size over most of the wide range of tested sizes.

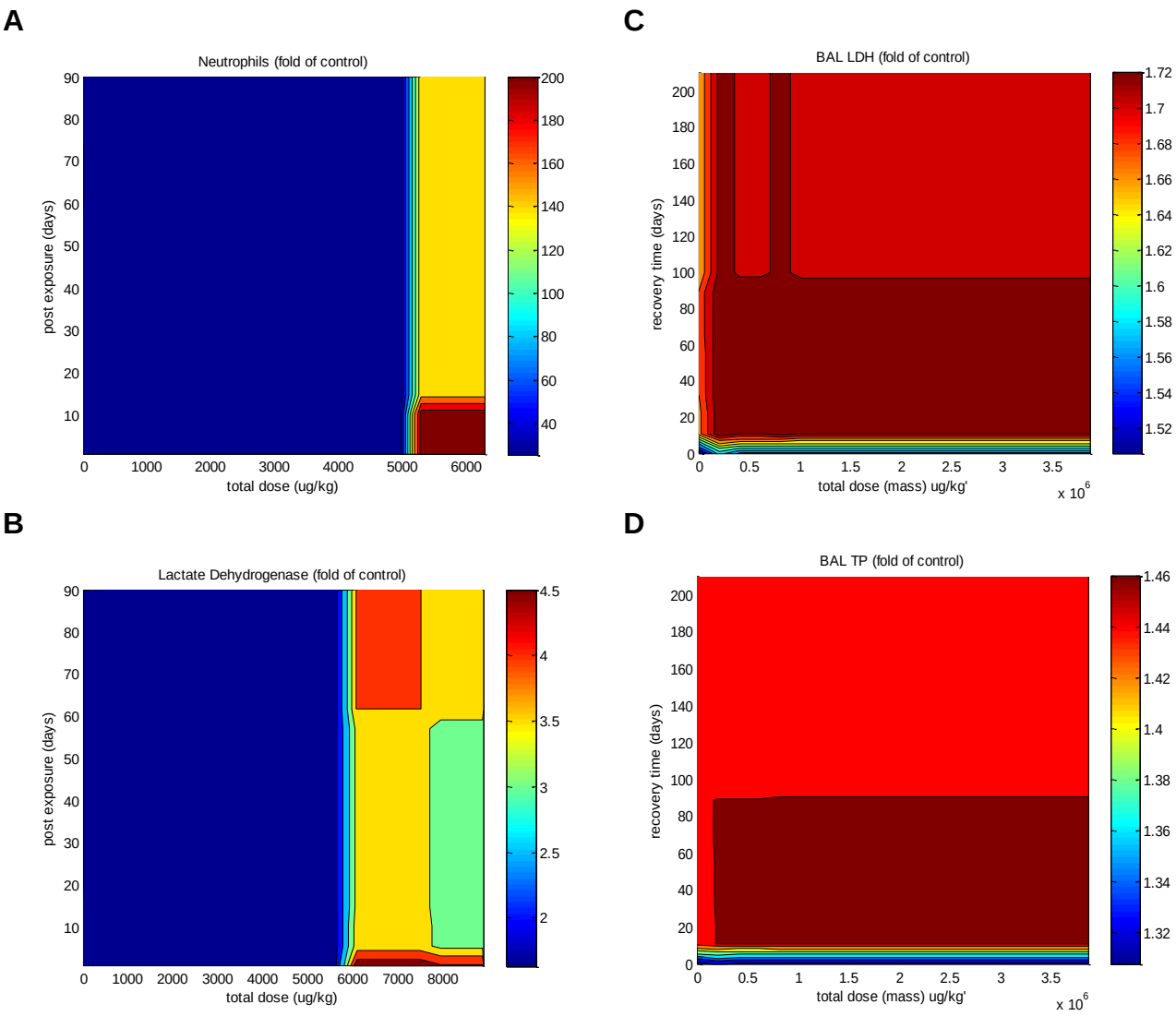
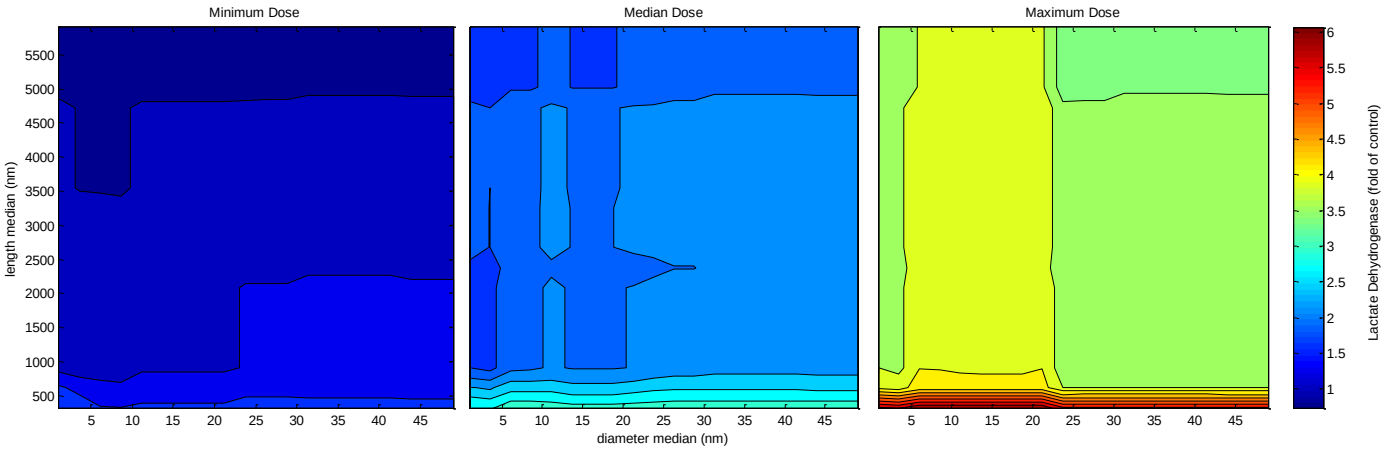


Figure 2: [A] Change in Neutrophil count in BAL fluid following pulmonary exposure to carbon nanotubes. [B] Change in lactate dehydrogenase in BAL fluid following pulmonary exposure to carbon nanotubes. [C] Change in lactate dehydrogenase in BAL fluid following exposure to titanium dioxide nanoparticles. [D] Change in total protein in BAL fluid following exposure to titanium dioxide nanoparticles.

A



B

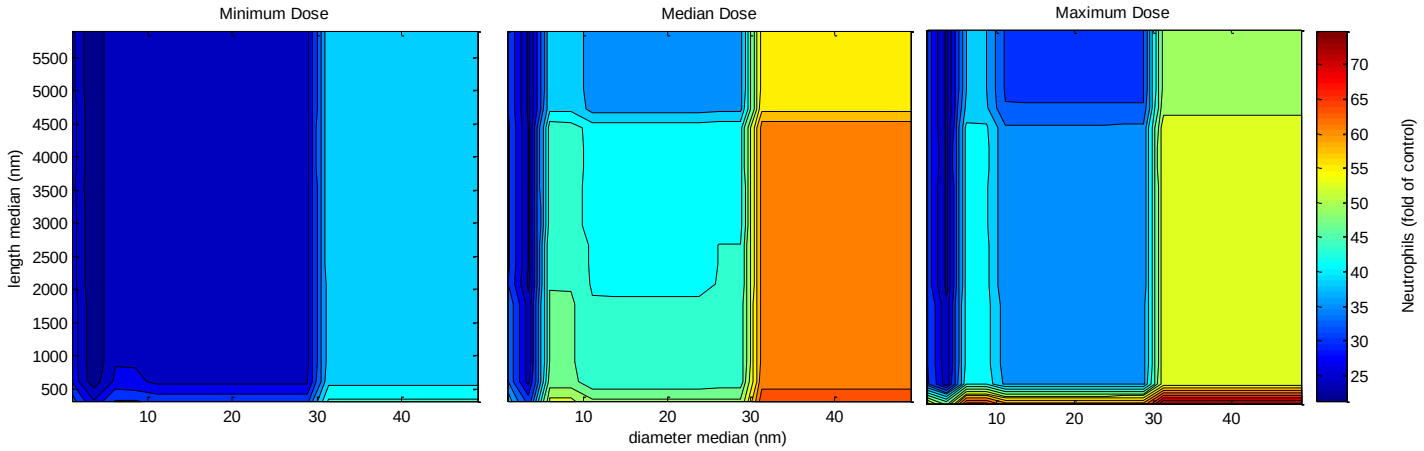
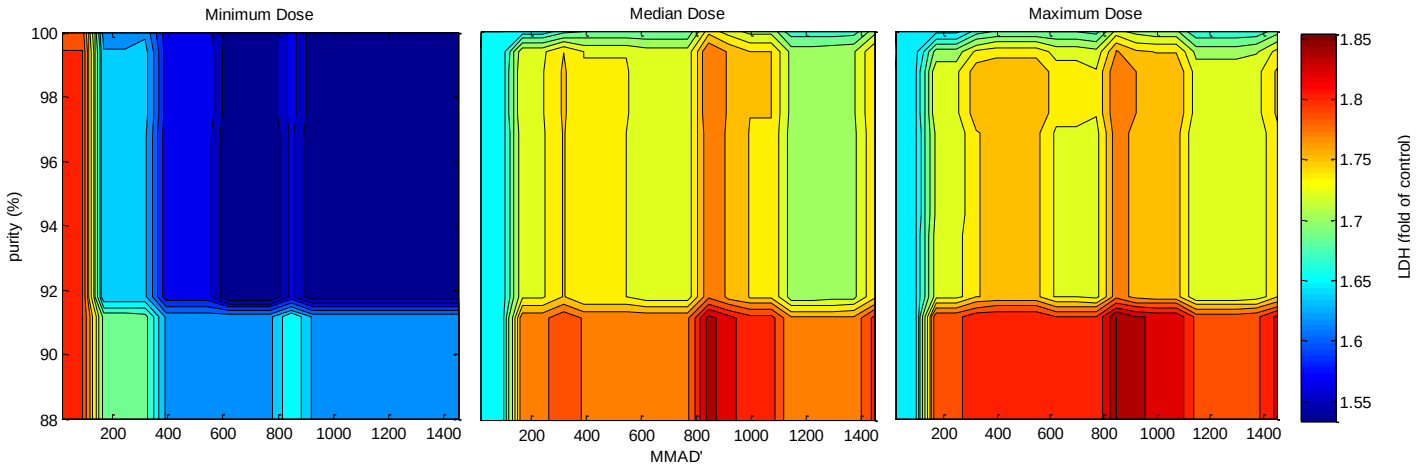


Figure 3: Effects of pulmonary exposure to carbon nanotubes at three dose levels, and all values of nanotube length and diameter: minimum dose is 2 µg/kg; median dose is 3250 µg/kg; maximum dose is 6500 µg/kg. [A] Change in Lactate dehydrogenase (LDH) in BAL fluid following exposure. [B] Change in Neutrophils count in BAL fluid following exposure. Values other than dose, length, and diameter, such as recovery period, and % cobalt impurity are held constant at their median reported values. These results suggest that larger diameter CNTs (multi-walled CNTs) produce a significantly increased immune response (PMN counts), but only a mildly increased LDH concentration.

A



B

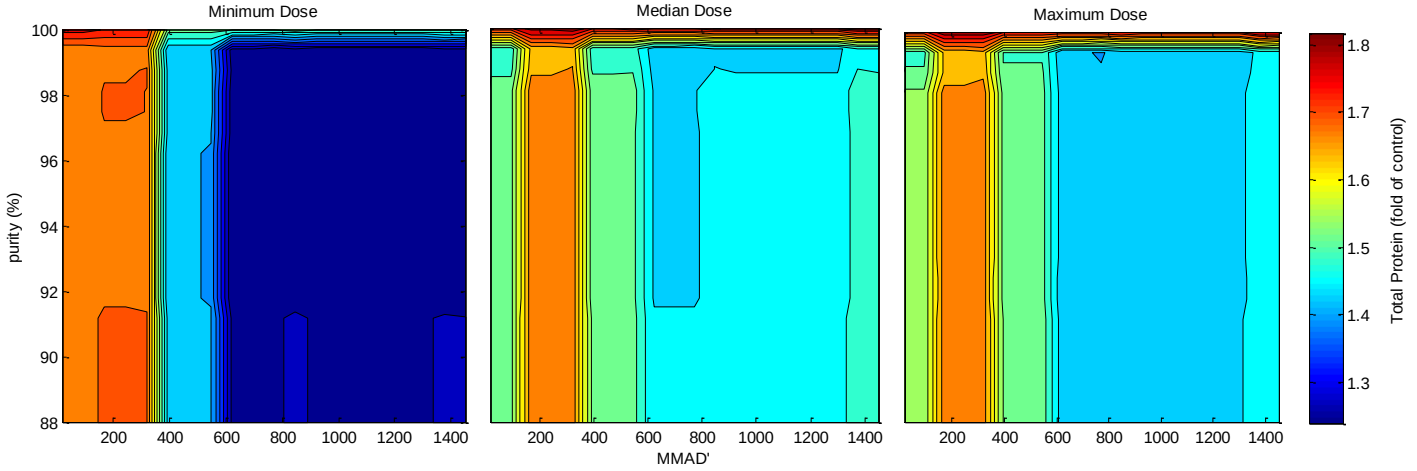
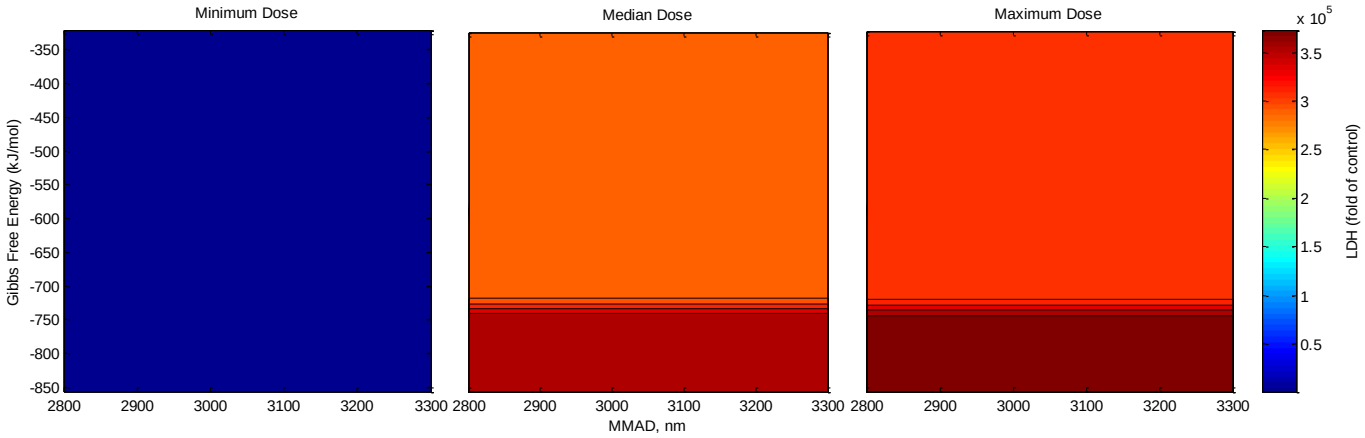


Figure 4: Effects of pulmonary exposure to titanium dioxide nanoparticles based on changes in dose, aggregate diameter (MMAD), and purity. [A] Changes in lactate dehydrogenase (LDH) in BAL fluid [B] Changes in total protein concentration in BAL fluid. Other variables in the model are held constant at their median values. The minimum dose is 35 µg/kg. The median dose is 1.8×10^{6} µg/kg. The maximum dose is 3.5×10^{6} µg/kg. These results indicate that increasing purity is associated with a mildly decreasing LDH concentration, but has little impact on total protein concentration. The size of particle aggregates appears to have negligible effect for either measure.

A



B

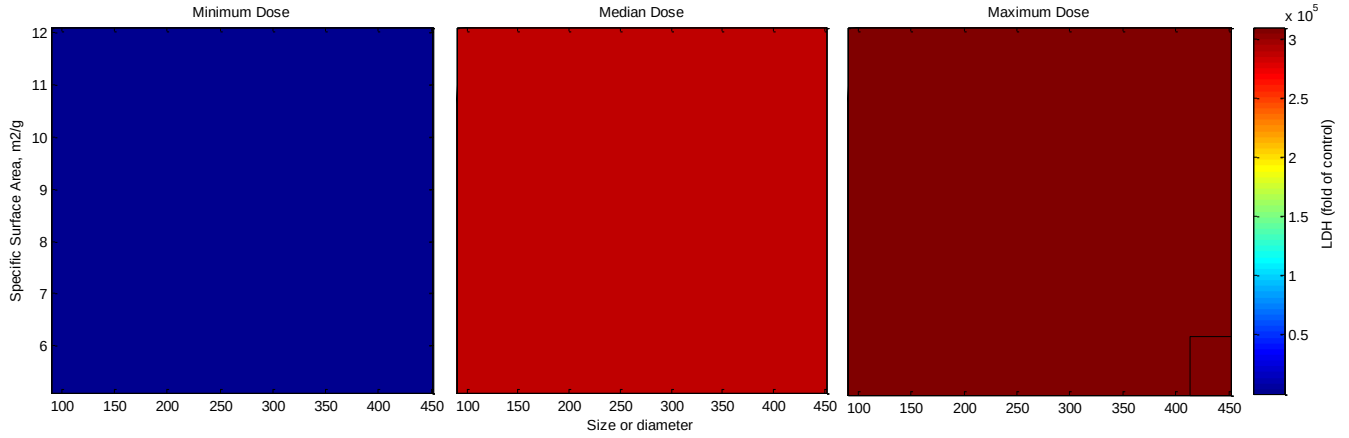


Figure 5: Effects of pulmonary exposure to metal oxide nanoparticles including titanium dioxide, zinc oxide, magnesium oxide, and silicon dioxide, based on changes in [A] aggregation (MMAD) and the Gibbs free energy, and primary particle size and specific surface area [B]. The minimum dose is 300 µg/kg. The median dose is 8,000 µg/kg. And, the maximum dose is 16,000 µg/kg. These plots indicate that changes in total dose by mass affect the observed toxicity to a much greater degree than any effects from size or chemical factors.

When considering the metal oxides as a group, particle size does affect the BAL concentration of LDH to some extent, with larger particles causing higher LDH concentrations. Particle size contributes to model variance reduction (Figure B5), but the magnitude of the difference in LDH is dwarfed by the change associated with higher doses (Figure 5).

This is opposite of the effect observed for titanium dioxide data analyzed alone, where titanium dioxide nanoparticles smaller than 40 nm caused at most a 2-fold increase in LDH. It must be noted that the entire metal oxide data set did not include any particles that small except for titanium dioxide nanoparticles, so the effect of very small diameter metal oxide nanoparticles should not be considered to be well defined.

Effects of Impurities

The importance of impurities in explaining the toxicity of nanoparticles has long been debated for carbon nanotubes, and these data appear to clearly indicate that the cobalt content of CNTs (see Figure A1) has the effect of increasing the immune response, whether sensitizing the system to the effects of CNTs exposure, or causing such an effect independently. The metallic impurities that exist together with the CNTs are remants of the metallic catalysts used in the manufacturing process.

For titanium dioxide nanoparticles, on the other hand, the impurities (or purity) of the particles is not a significant contributor (see Figure 4) to inhalational toxicity. This is most likely due to the fact that impurities in titanium dioxide nanoparticle manufacturing include much more inert materials than the metals associated with carbon nanotubes. The impurities were not often characterized in the nano-TiO2 toxicology studies.

Effects of Chemical Differences

A variety of different quantitative chemical descriptors have been proposed and tested with models to predict *in vitro* toxicity of metal oxide nanoparticles [9,17]. But, as seen in Figure 5, the results of pulmonary exposure studies on rodents appear to indicate that the total mass of metal oxide nanoparticles is a much more important predictor than any chemical or physical descriptors. In fact, if all chemical descriptors were excluded from the model, the fraction of explained variance (or R^2 value) only decreased from 0.97 to 0.93.

The magnitude of the change in LDH as shown in Figure 5 due to particle size, aggregation, or Gibb's Free Energy are dwarfed by the magnitude of change due to simply increasing the total mass of metal oxide nanoparticles the animals are exposed to. While this analysis only includes a few different metal oxides, these oxides do differ significantly in terms of solubility, thermodynamic stability, and reactivity. While other quantitative chemical descriptors were tested including metal group or period from the periodic table, the mean isoelectric point, the surface charge, the enthalpy of formation, and crystalline structure, the Gibb's Free Energy proved to have the greatest apparent effect, but only a slight one.

CONCLUSIONS

Random forest models even when trained on an incomplete data set can provide useful risk assessment of the benefits or costs of possible design tradeoffs in the area of nanoparticle toxicity. Using these models to quantitatively summarize the current knowledge and visualize the relationships between particle design parameters contributes to understanding the risks of a new technology. This is especially true during the early stages of implementation when the science may not have developed mechanistic explanations for why one material may pose a higher risk than another.

The pulmonary toxicity measured by LDH release of metal oxide nanoparticles as a group including titanium dioxide, magnesium oxide, silicon dioxide, and zinc oxide does not appear to be highly dependent on physical characteristics of the particles, and depends only slightly on chemical characteristics, at least within the ranges that have been tested to date in animals. This leads to the conclusion that for these materials, the first and best risk mitigation may be only to minimize exposure.

Design characteristics for carbon nanotubes are much more important, relatively, to pulmonary toxicity, at least for the relatively short term exposures that have been examined so far. These characteristics include the proportion of metallic impurities like cobalt, and the nanotube length and diameter. CNT diameter is important over a wide range of doses and combinations of other variables and should be minimized to mitigate toxicity.

Meta-analysis of toxicity studies such as this one have the ability to quantitatively compare the claims of single studies against the larger field of study and to quantify the relative contributions of a large number of factors. Those the individual studies form the basis for this analysis, their conclusions are re-evaluated in light of other findings and minor effects can be distinguished from major ones. Such information could be taken into account in future product and process design decisions that utilize nanoparticles in order to mitigate risks to workers, consumers, and businesses. Meta-analyses could also play a role in determining future regulatory decisions regarding these materials, by helping distinguish significant from insignificant effects on toxicity.

ACKNOWLEDGMENTS

We would like to acknowledge the contributions of Amy Wang and Samantha Frady of the Environmental Protection Agency (EPA), who compiled the published pulmonary toxicity data on titanium dioxide nanoparticles.

We also acknowledge the assistance of Sophie Grodsinsky, who collected and compiled the published pulmonary toxicity data for the silicon dioxide, zinc oxide, and magnesium oxide nanoparticles.

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, the Bertucci Fellowship in Engineering, 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] Auffan M., Rose J., Bottero J.-Y., Lowry G. V, Jolivet J.-P., and Wiesner M. R., 2009, "Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective.," Nature nanotechnology, **4**(10), pp. 634–41.

[2] Limits for Air Contaminants. Code of Federal Regulations. Title 29. 1910.1000.

[3] NIOSH, 2011, Current Intelligence Bulletin 63: Occupational Exposure to Titanium Dioxide, DHHS (NIOSH) Publication No. 2011-160.

[4] NIOSH, 2010, NIOSH Current Intelligence Bulletin: Occupational Exposure to Carbon Nanotubes and Nanofibers, DHHS (NIOSH) Publication No. 2010-XXX.

[5] Nygaard U. C., Hansen J. S., Samuelsen M., Alberg T., Marioara C. D., and Løvik M., 2009, "Single-walled and multi-walled carbon nanotubes promote allergic immune responses in mice.," Toxicol. Sci., **109**(1), pp. 113–123.

[6] Warheit D. B., Sayes C. M., and Reed K. L., 2009, "Nanoscale and fine zinc oxide particles: can in vitro assays accurately forecast lung hazards following inhalation exposures?," Environmental science & technology, **43**(20), pp. 7939–45.

[7] Gernand J., and Casman E., 2013, "A meta-analysis of carbon nanotube pulmonary toxicity studies – How physical dimensions and impurities affect the toxicity of carbon nanotubes," Risk Analysis, **XX**, p. XXX–XXX [Submitted].

[8] Sayes C., and Ivanov I., 2010, "Comparative study of predictive computational models for nanoparticle-induced cytotoxicity.," Risk analysis : an official publication of the Society for Risk Analysis, **30**(11), pp. 1723–34.

[9] Puzyn T., Rasulev B., Gajewicz A., Hu X., Dasari T. P., Michalkova A., Hwang H.-M., Toropov A., Leszczynska D., and Leszczynski J., 2011, "Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles.," Nature nanotechnology, **6**(February), pp. 175–178.

[10] Fourches D., Pu D., and Tropsha A., 2011, "Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles," Comb. Chem. High T. Scr., **14**, pp. 217–225.

[11] Breiman L., 2001, "Random Forests," Machine Learning, **45**, pp. 5–32.

[12] Breiman L., Friedman J., Stone C., and Olshen R. A., 1984, Classification and Regression Trees, Chapman and Hall/CRC, New York.

[13] Gosens I., Post J. A., De la Fonteyne L. J. J., Jansen E. H. J. M., Geus J. W., Cassee F. R., and De Jong W. H., 2010, "Impact of agglomeration state of nano- and submicron sized gold particles on pulmonary inflammation.," Particle and fibre toxicology, **7**(1), p. 37.

[14] Pauluhn J., 2010, "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., **113**(1), pp. 226–242.

[15] Yu L. E., Yung L.-Y. L., Ong C., Tan Y.-L., Balasubramaniam K. S., Hartono D., Shui G., Wenk M. R., and Ong W., 2007, "Translocation and effects of gold nanoparticles after inhalation exposure in rats," Nanotoxicology, **1**(1-4), pp. 234–241.

[16] Sayes C., and Ivanov I., 2010, "Comparative study of predictive computational models for nanoparticle-induced cytotoxicity.," Risk Analysis, **30**(11), pp. 1723–1734.

[17] Fourches D., Pu D., Tassa C., Weissleder R., Shaw S. Y., Mumper R. J., and Tropsha A., 2010, "Quantitative nanostructure-activity relationship modeling.," ACS nano, **4**(10), pp. 5703–5712.

[18] Pauluhn J., 2010, "Multi-walled carbon nanotubes (Baytubes): approach for derivation of occupational exposure limit.," Regul. Toxicol. Pharm., **57**(1), pp. 78–89.

[19] Muller J., Huaux F., Moreau N., Misson P., Heilier J.-F., Delos M., Arras M., Fonseca A., Nagy J. B., and Lison D., 2005, "Respiratory toxicity of multi-wall carbon nanotubes.," Tox. App. Pharma., **207**(3), pp. 221–231.

[20] Shvedova A. A., Kisin E., Murray A. R., Johnson V. J., Gorelik O., Arepalli S., Hubbs A. F., Mercer R. R., Keohavong P., Sussman N., Jin J., Yin J., Stone S., Chen B. T., Deye G., Maynard A., Castranova V., Baron P. A., and Kagan V. E., 2008, "Inhalation vs. aspiration of single-walled carbon nanotubes in C57BL/6 mice: inflammation, fibrosis, oxidative stress, and mutagenesis.," Am J. Physiol.-Lung C., **295**(4), pp. 552–565.

[21] Porter D. W., Hubbs A. F., Mercer R. R., Wu N., Wolfarth M. G., Sriram K., Leonard S., Battelli L., Schwegler-Berry D., Friend S., Andrew M., Chen B. T., Tsuruoka S., Endo M., and Castranova V., 2010, "Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes," Toxicology, **269**, pp. 136–147.

[22] Inoue K.-I., Takano H., Koike E., Yanagisawa R., Sakurai M., Tasaka S., Ishizaka A., and Shimada A., 2008, "Effects of pulmonary exposure to carbon nanotubes on lung and systemic inflammation with coagulatory disturbance induced by lipopolysaccharide in mice.," Exp. Biol. Med., **233**(12), pp. 1583–1590.

[23] Shvedova A. A., Kisin E. R., Mercer R., Murray A. R., Johnson V. J., Potapovich A. I., Tyurina Y. Y., Gorelik O., Arepalli S., Schwegler-Berry D., Hubbs A. F., Antonini J., Evans D. E., Ku B.-K., Ramsey D., Maynard A., Kagan V. E., Castranova V., and Baron P., 2005, "Unusual inflammatory and fibrogenic pulmonary responses to single-walled carbon nanotubes in mice.," Am J. Physiol.-Lung C., **289**(5), pp. 698–708.

[24] Shvedova A. A., Kisin E. R., Murray A. R., Gorelik O., Arepalli S., Castranova V., Young S., Gao F., Tyurina Y. Y., Oury T. D., and Kagan V. E., 2007, "Vitamin E deficiency enhances pulmonary inflammatory response and oxidative stress induced by single walled

carbon nanotubes in C57BL/6 mice," Toxicol. Appl. Pharmacol., **221**(3), pp. 339–348.

[25] Warheit D. B., Laurence B. R., Reed K. L., Roach D. H., Reynolds G. A. M., and Webb T. R., 2004, "Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats," Toxicol. Sci., **77**, pp. 117–125.

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

[27] Ellinger-Ziegelbauer H., and Pauluhn J., 2009, "Pulmonary toxicity of multi-walled carbon nanotubes (Baytubes) relative to a-quartz following a single 6 h inhalation exposure of rats and a 3 months post-exposure period," Toxicology, **266**, pp. 16–29.

[28] Bermudez E., Mangum J. B., Asgharian B., Wong B. a, Reverdy E. E., Janszen D. B., Hext P. M., Warheit D. B., and Everitt J. I., 2002, "Long-term pulmonary responses of three laboratory rodent species to subchronic inhalation of pigmentary titanium dioxide particles.," Toxicological sciences : an official journal of the Society of Toxicology, **70**(1), pp. 86–97.

[29] Grassian V. H., O'shaughnessy P. T., Adamcakova-Dodd A., Pettibone J. M., and Thorne P. S., 2007, "Inhalation exposure study of titanium dioxide nanoparticles with a primary particle size of 2 to 5 nm.," Environmental health perspectives, **115**(3), pp. 397–402.

[30] Nemmar A., and Melghit K. et al., 2008, "The acute proinflammatory and prothrombotic effects of pulmonary exposure to rutile TiO2 nanorods in rats," Exp. Biol. Med., **233**(5), pp. 610–619.

[31] Oberdorster G., Ferin J., and Lehnert B. E., 1994, "Correlation between Particle Size, In Vivo Particle Persistence, and Lung Injury," Environmental Health Perspectives, **102**, p. 173.

[32] Warheit D. B., Webb T. R., Sayes C. M., Colvin V. L., and Reed K. L., 2006, "Pulmonary instillation studies with nanoscale TiO2 rods and dots in rats: toxicity is not dependent upon particle size and surface area.," Toxicological sciences : an official journal of the Society of Toxicology, **91**(1), pp. 227–236.

[33] Renwick L. C., 2004, "Increased inflammation and altered macrophage chemotactic responses caused by two ultrafine particle types," Occupational and Environmental Medicine, **61**(5), pp. 442–447.

[34] Rehn B., and Selier F. et al, 2003, "Investigations on the inflammatory and genotoxic lung effects of two types of titanium dioxide: untreated and surface treated," Tox. App. Pharma., **189**(2), pp. 84–95.

[35] Osier M., and Baggs R. et al., 1997, "Intratracheal instillation versus intratracheal inhalation: influence of cytokines on inflammatory response," Environ. Health Perspect., **105**(S5), pp. 1265–1271.

[36] Warheit D. B., Hoke R. A., Finlay C., Donner E. M., Reed K. L., and Sayes C. M., 2007, "Development of a base set of toxicity tests using ultrafine TiO2 particles as a component of nanoparticle risk management," Toxicology letters, **171**, pp. 99–110.

[37] Warheit D. B., Webb T. R., Reed K. L., Frerichs S., and Sayes C. M., 2007, "Pulmonary toxicity study in rats with three forms of ultrafine-TiO2 particles: differential responses related to surface properties," Toxicology, **230**, pp. 90–104.

[38] Warheit D. B., Sayes C. M., Frame S. R., and Reed K. L., 2010, "Pulmonary exposures to Sepiolite nanoclay particulates in rats: resolution following multinucleate giant cell formation.," Toxicology letters, **192**(3), pp. 286–93.

[39] Kobayashi N., Naya M., Endoh S., Maru J., Yamamoto K., and Nakanishi J., 2009, "Comparative pulmonary toxicity study of nano-TiO2 particles of different sizes and agglomerations in rats: different short- and long-term post-instillation results," Toxicology, **264**, pp. 110–118.

[40] Sayes C. M., Reed K. L., and Warheit D. B., 2007, "Assessing toxicity of fine and nanoparticles: comparing in vitro measurements to in vivo pulmonary toxicity profiles.," Toxicol. Sci., **97**(1), pp. 163–80.

[41] Gernand J., 2012, "Carbon Nanotube (CNT) Pulmonary Toxicity Data Set" [Online]. Available: http://nanohub.org/resources/13515.

ANNEX A

NANOPARTICULATE TOXICITY RISK CONTOUR PLOTS

This annex contains additional toxicity risk contour plots generated by the random forest models.

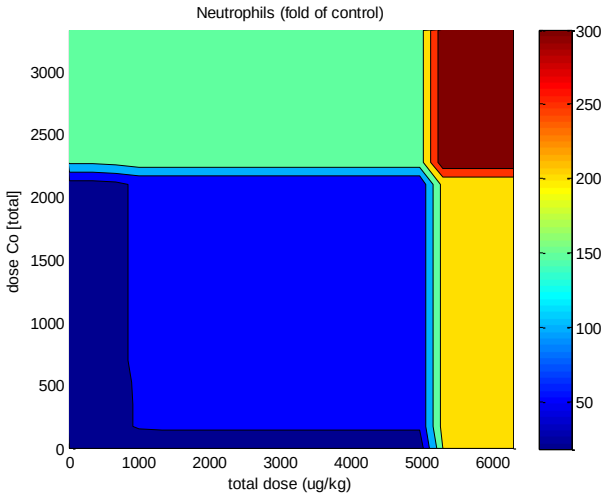


Figure A1: Changes in RF model predicted BAL neutrophils count following exposure to carbon nanotubes as a function of changes in total dose and the dose of cobalt, a common toxic impurity (up to 0.53% by weight of total CNTs). This suggests that Co and total CNTs both independently contribute to higher neutrophils count.

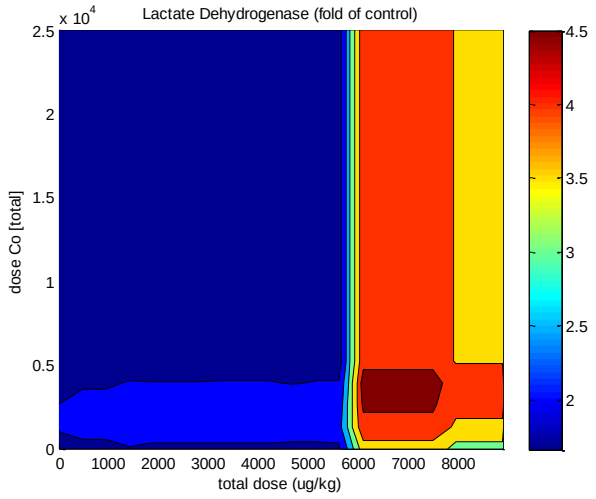


Figure A2: Changes in RF model predicted BAL LDH following exposure to carbon nanotubes as a function of changes in total dose and the dose of cobalt, a common toxic impurity (up to 0.53% by weight of total CNTs). This suggests that total dose is much more important than Co content for increasing LDH concentration.

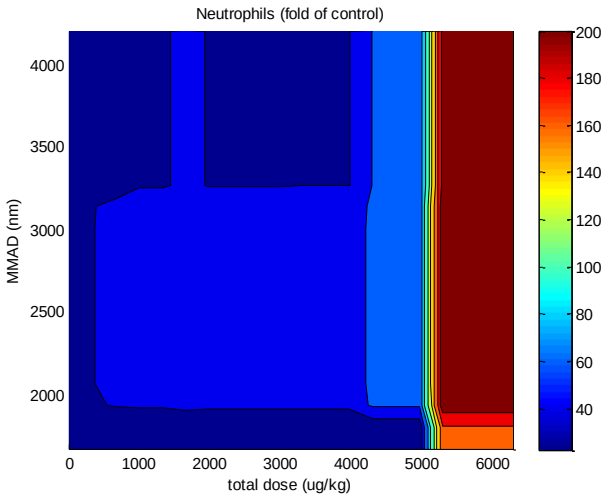


Figure A3: Changes in RF model predicted BAL neutrophils count following exposure to carbon nanotubes as a function of changes in total dose and aggregation (MMAD, mass mode aerodynamic diameter). This suggests that aggregation only has a small effect on neutrophils count as compared to total dose, and also that low to moderate doses are relatively similar in response.

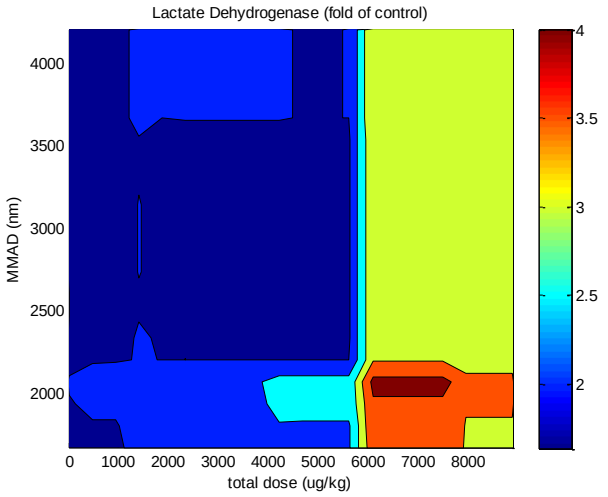


Figure A4: Changes in RF model predicted BAL LDH following exposure to carbon nanotubes as a function of changes in total dose and aggregation (MMAD, mass mode aerodynamic diameter). This suggests both that total dose is a more important predictor of LDH than aggregation, but also that less aggregation can increase LDH as well.

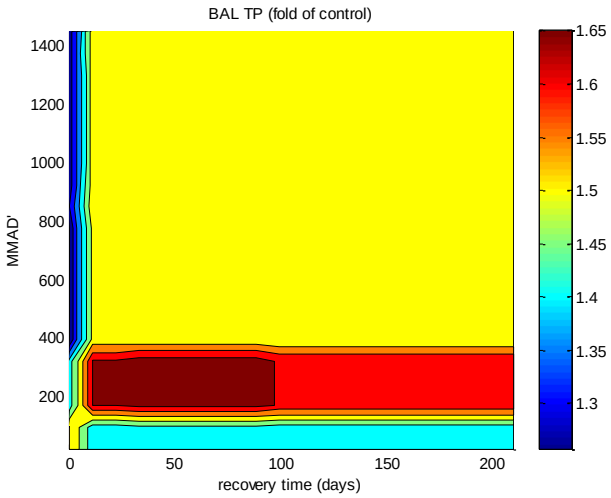


Figure A5: Changes in RF model predicted BAL total protein following exposure to titanium dioxide nanoparticles as a function of changes in total dose and aggregation (MMAD, mass mode aerodynamic diameter). This suggests that low aggregation levels are more toxic than higher ones, and that recover time is not an important factor, but the scale of these differences in small overall.

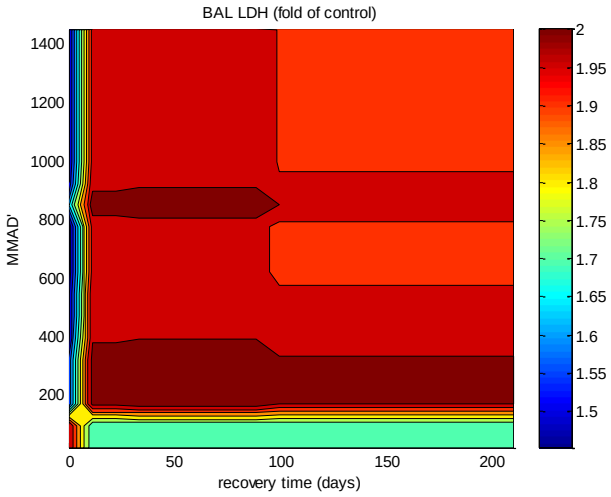


Figure A6: Changes in RF model predicted BAL LDH following exposure to titanium dioxide nanoparticles as a function of changes in total dose and aggregation (MMAD, mass mode aerodynamic diameter). This suggests that in terms of predicting LDH response, neither recovery time nor aggregation is consistently detrimental or beneficial, so more data would be required.

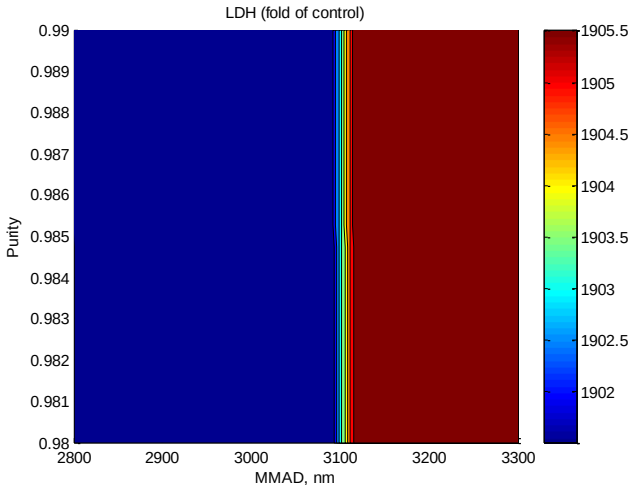


Figure A7: Changes in RF model predicted BAL LDH following exposure to metal oxide nanoparticles (TiO_2, MgO, ZnO, SiO_2) as a function of changes in aggregation (MMAD, mass mode aerodynamic diameter), and purity. Based on the scale, there is little difference in toxicity across this range of variables indicating these factors play little role in toxicity.

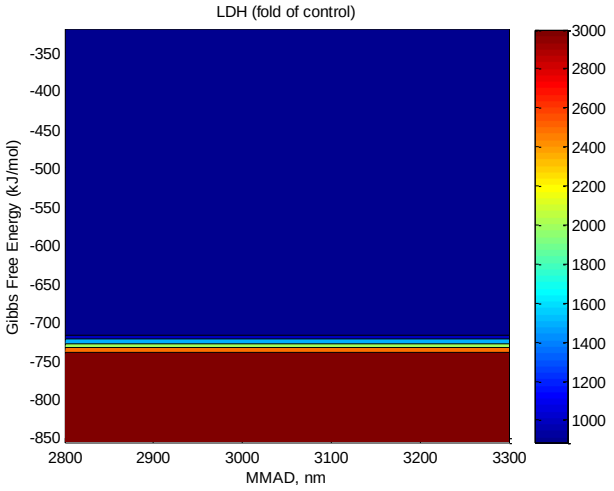


Figure A8: Changes in RF model predicted BAL LDH following exposure to metal oxide nanoparticles (TiO_2, MgO, ZnO, SiO_2) as a function of changes in aggregation (MMAD, mass mode aerodynamic diameter), and Gibbs Free Energy, a descriptor of the chemical energy available in the metal oxide compound. This suggests that aggregation is much less important than differences in chemical makeup [other results, see Figure 5, indicate that Gibbs Free Energy is much less important than total dose].

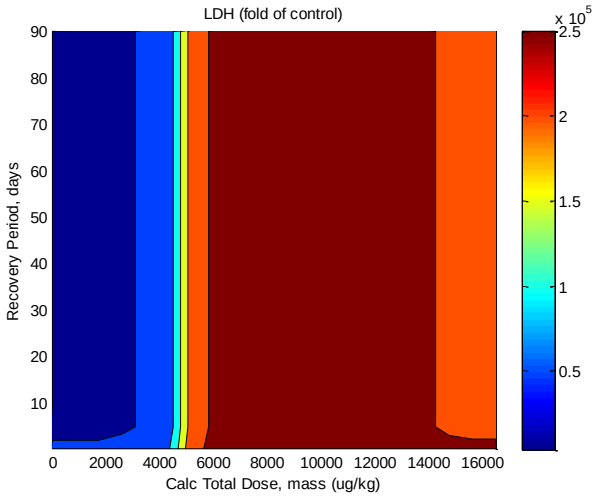


Figure A9: Changes in RF model predicted BAL LDH following exposure to metal oxide nanoparticles (TiO_2, MgO, ZnO, SiO_2) as a function of total dose and recovery period. This indicates that total dose dominates the change in LDH due to longer recovery periods.

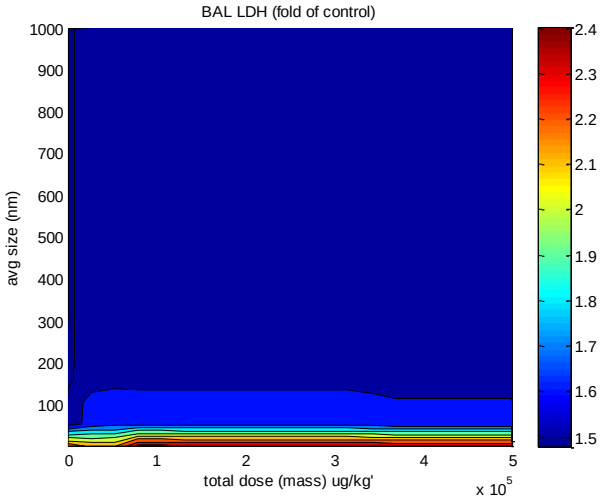


Figure A10: Changes in RF model predicted BAL LDH following exposure to titanium dioxide nanoparticles as a function of total dose and average particle size. This indicates that very small TiO_2 nanoparticles are more toxic than those in most of the possible range of sizes.

ANNEX B

DETAILS ON RANDOM FOREST MODEL STRUCTURE AND LEARNING

This annex contains details on the random forest model structures employed in this analysis, learning statistics, error, and goodness-of-fit metrics.

Random Forest Model Variable Importance

There are many different ways to represent the importance of variables in a random forest models. All of these methods consider the information gain achieved by each branch node summed by input variable and averaged across all of the regression trees in the forest. Some methods of calculating information gain include entropy, standardized mean difference, Gini coefficient, and variance reduction. Generally the results as calculated by these methods are very similar. We have chosen to utilize variance reduction as the primary information gain metric primarily due to comparability to other methods of evaluating different kinds of models.

The following figures (B1 through B6) display the internal structure of the RF models and their relative reliance on different input variables to reduce the error of the models. The height of the columns is not directly analogous to magnitude of the change in outcome associated with a unit change in input—the magnitude of change in toxicity is better observed in the contour plots.

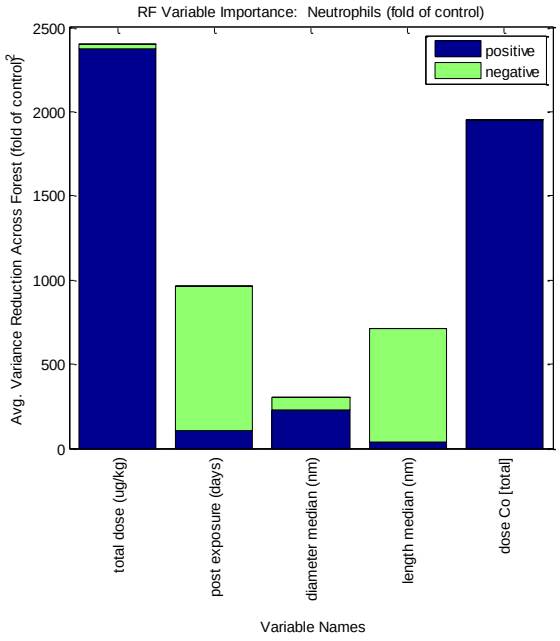


Figure B1: RF model variable importance as measured by variance reduction attributable to each variable for the prediction of BAL neutrophils count following exposure to carbon nanotubes.

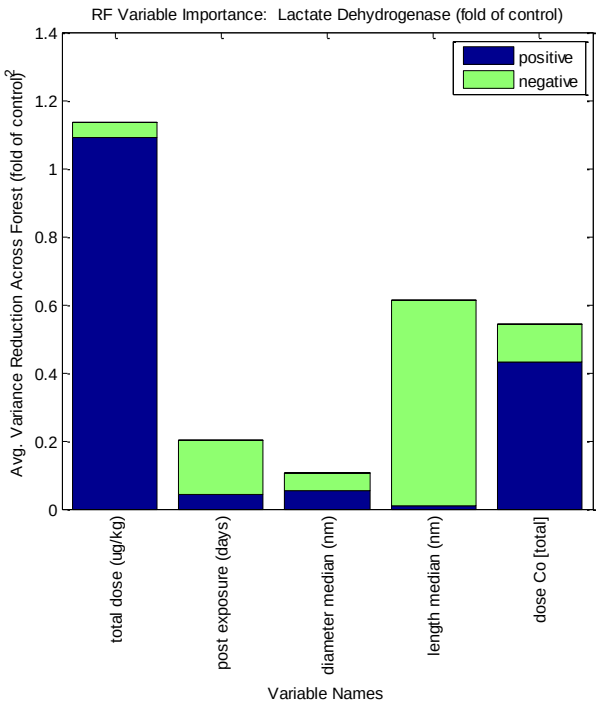


Figure B2: RF model variable importance as measured by variance reduction attributable to each variable for the prediction of BAL LDH count following exposure to carbon nanotubes.

Column colors reflect whether changes in a given variable when applied to a branch split in the RF model were associated with a positive or negative change in the model output. For example, increases total dose is usually associated with increasing toxic responses, and increasing recovery time is usually associated with decreasing recovery time. Sometimes, these relationships can be complex or non-linear and result in a variable having different effects in different parts of the variable space.

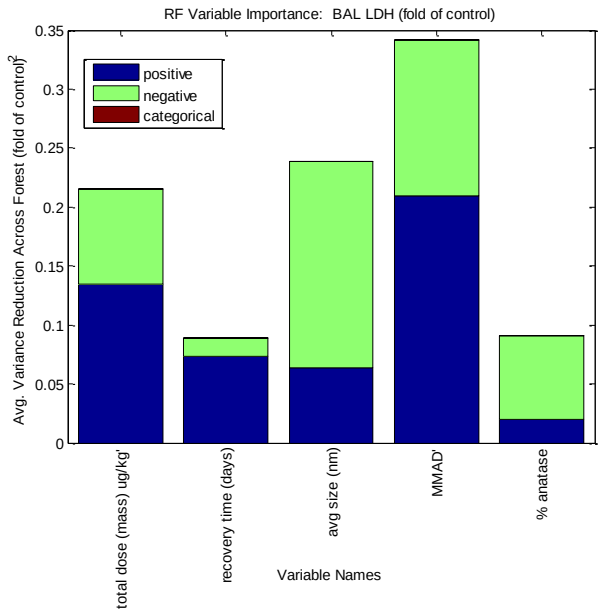


Figure B3: RF model variable importance as measured by variance reduction attributable to each variable for the prediction of BAL LDH count following exposure to titanium dioxide nanoparticles.

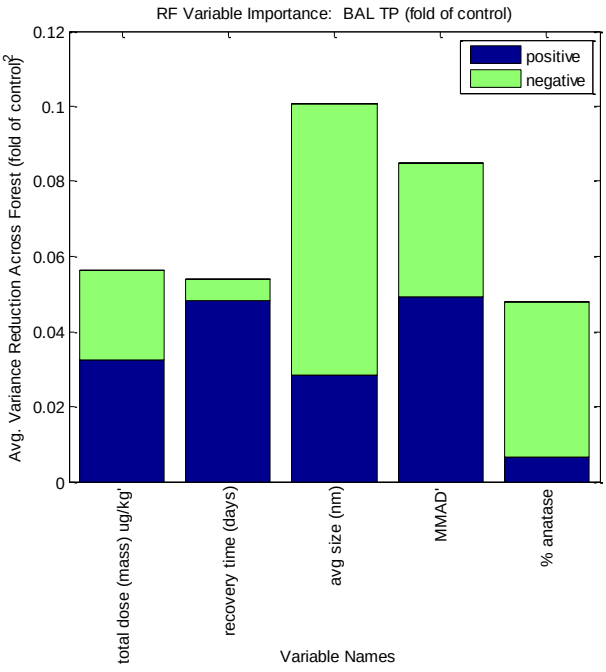


Figure B4: RF model variable importance as measured by variance reduction attributable to each variable for the prediction of BAL total protein count following exposure to titanium dioxide nanoparticles.

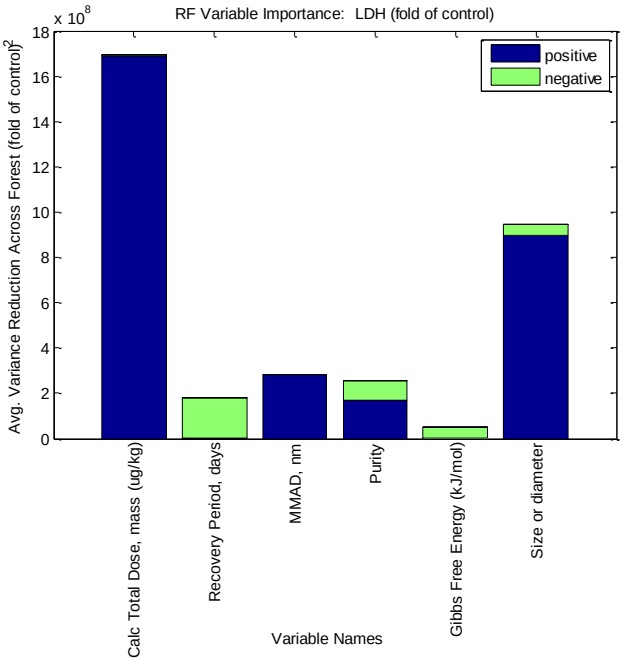


Figure B5: RF model variable importance as measured by variance reduction attributable to each variable for the prediction of BAL LDH count following exposure to metal oxide nanoparticles include titanium dioxide, magnesium oxide, silicon dioxide,

Random Forest Model Learning Progression

These figures display the error of the RF model in predicting the actual observed values as additional regression trees are added to the forest. Each RF model contains 1,000 trees. As these graphs show, the minimal error state is usually realized by the model once the model has achieved a size of 100-200 trees.

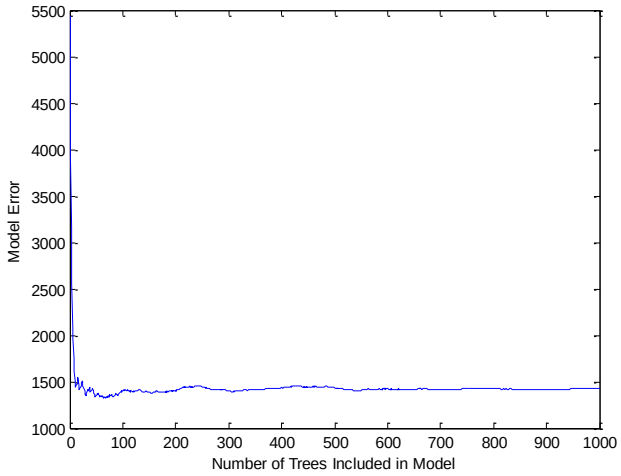


Figure B6: RF model error as a function of trees included in model for prediction of BAL neutrophils following pulmonary exposure to carbon nanotubes.

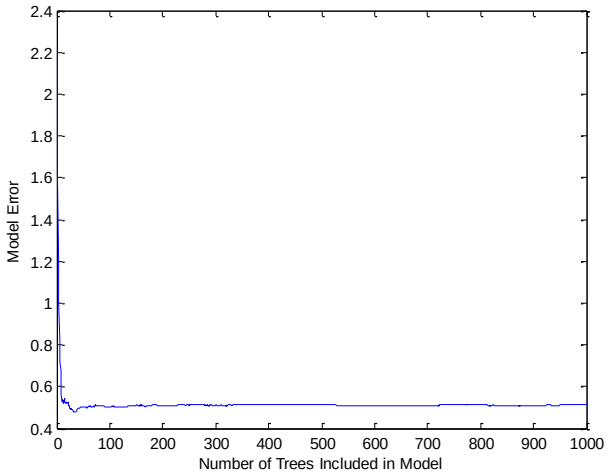


Figure B7: RF model error as a function of trees included in model for prediction of BAL LDH following pulmonary exposure to carbon nanotubes.

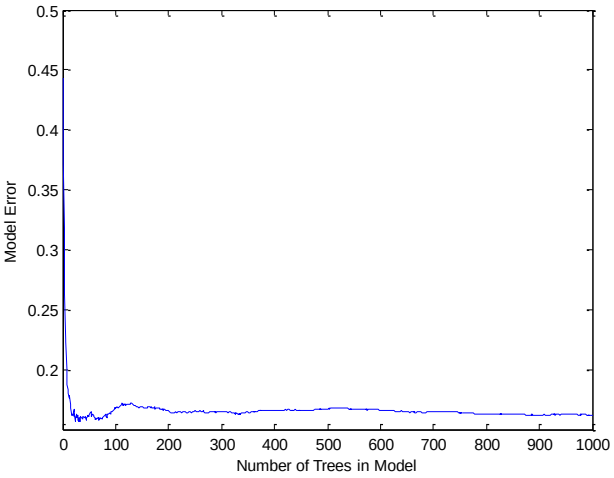


Figure B8: RF model error as a function of trees included in model for prediction of BAL total protein following pulmonary exposure to titanium dioxide nanoparticles.

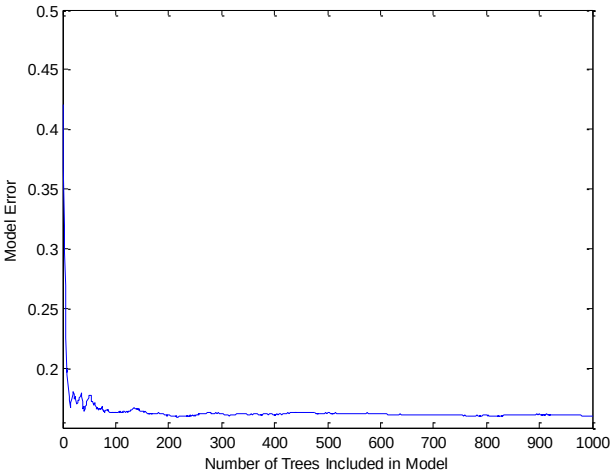


Figure B9: RF model error as a function of trees included in model for prediction of BAL LDH following pulmonary exposure to titanium dioxide nanoparticles.

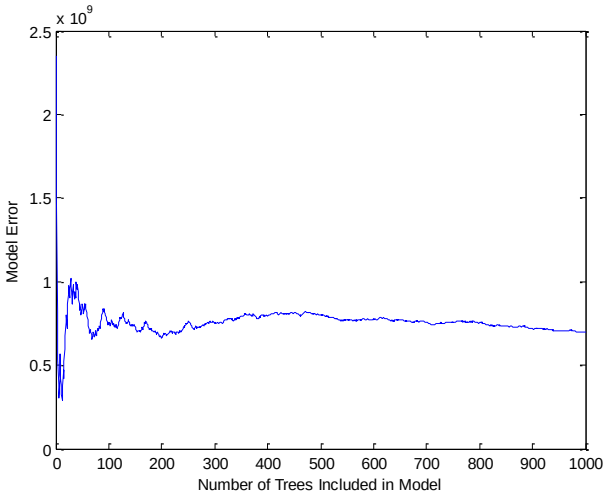


Figure B10: RF model error as a function of trees included in model for prediction of BAL LDH following pulmonary exposure to metal oxide nanoparticles including titanium dioxide, magnesium oxide, silicon dioxide, and zinc oxide.

ANNEX C

TABLE OF DATA SOURCES

Table C-1: Listing of all rodent pulmonary toxicity studies included in this meta-analysis.

Reference	First Author	Year	Nanoparticle	Total Dose (μg/kg)	Recovery Period (days)	Exposure Mode	Endpoint(s) Measured
[18]	Pauluhn J.	2010	CNT*	105 – 6,290	1 – 90	Inhalation	PMN, LDH
[19]	Muller J.	2005	CNT*	2 – 8	3	Instillation	PMN, LDH
[20]	Shvedova A.	2008	CNT*	250 – 1,000	1 – 7	Instillation	PMN, LDH
[21]	Porter	2010	CNT*	435 – 1,740	1 – 28	Aspiration	PMN, LDH
[22]	Inoue	2008	CNT*	4,000	1	Instillation	PMN
[23]	Shvedova A.	2005	CNT*	490 – 1,970	1 – 60	Aspiration	PMN
[24]	Shvedova A.	2007	CNT*	1,851	1	Aspiration	PMN, LDH
[25]	Warheit D.	2004	CNT*	1,000 – 5,000	1 – 30	Instillation	LDH
[26]	Muller J.	2008	CNT*	8,890	3	Instillation	LDH
[27]	Ellinger-Ziegelbauer H.	2009	CNT*	180 – 3,900	7 – 90	Inhalation	LDH
[28]	Bermudez E.	2002	TiO_2	10,000 – 90,000	0 – 365	Inhalation	PMN
[29]	Grassian V.	2007	TiO_2	35 – 3,300	0 – 14	Inhalation	PMN, LDH
[30]	Nemmar A.	2007	TiO_2	1,000 – 5,000	1	Instillation	PMN
[31]	Oberdorster G.	1994	TiO_2	295 – 2300	0	Instillation	PMN
[32]	Warheit D.	2006	TiO_2	300 – 10,000	0 – 84	Instillation	PMN, LDH
[33]	Renwick L.	2004	TiO_2	300 – 1200	1	Instillation	PMN
[34]	Rehn B.	2003	TiO_2	750 – 6,000	0 – 90	Instillation	PMN
[35]	Osier M.	1997	TiO_2	750 – 3,750	0 – 7	Inhalation	PMN
[36,37]	Warheit D.	2007	TiO_2	1,000 – 5,000	0 – 84	Instillation	PMN, LDH
[38]	Warheit D.	2010	TiO_2	1,000 – 5,000	0 – 30	Instillation	LDH
[39]	Kobayashi N.	2009	TiO_2	5,000	0 – 10	Instillation	LDH
[40]	Sayes C.	2007	SiO_2, ZnO	1,000 – 5,000	1 – 90	Instillation	LDH
[13]	Gosens I.	2010	SiO_2	1,600	1	Instillation	LDH
[6]	Warheit D.	2009	ZnO, MgO	1,000 – 5,000	1 – 90	Instillation, Inhalation	LDH

*The carbon nanotube (CNT) portion of this pulmonary toxicity data set is available for download at http://nanohub.org/resources/13515 [41]