The Pennsylvania State University

The Graduate School

TOXICITY BASED CLASSIFICATION OF ENGINEERED NANO-POLLUTANTS: A META -ANALYSIS BASED APPROACH

A Dissertation in

Energy and Mineral Engineering

by

Vignesh Ramchandran

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2023

The dissertation of Vignesh Ramchandran was reviewed and approved by the following:

Jeremy M. Gernand
Assistant Professor of Industrial Health and Safety
Dissertation Advisor
Chair of Committee
Associate Department Head for Graduate Education

William Groves
Associate Professor of Industrial Health and Safety

Eugene Morgan
Assistant Professor of Petroleum and Natural Gas Engineering

Ephraim Hanks
Associate Professor of Statistics

ABSTRACT

Experimental toxicology testing studies are time intensive and financially taxing. The rapid growth of the nanotechnology sector necessitates the advancement of methodology that is quick to identify nanomaterials with high toxic potential. Toxicity assays for the purposes of setting occupational exposure limits for nanomaterials have drawbacks which could be overcome or limited by the development of computational approaches. Quantitative relationships between the physicochemical characteristics of emerging nanomaterials and their corresponding toxicity are desired to better assist in the subsequent mitigation of toxicity by design.

This dissertation presents a hierarchical clustering methodology that is capable of identifying and categorizing the nanomaterials by the similarity of their *in-vivo* dose-response and recovery relationship. The potency of the newly generated clusters is examined thoroughly to generate hypotheses regarding what factors influence the observed toxicity of the particles. The clustering methodology was developed and deployed on 2 separate nanomaterial datasets using 5 toxicity endpoints. The first dataset is comprised of peer-reviewed *in-vivo* pulmonary toxicity data collected from exposure to Carbon nanotubes (CNTs). The second dataset is a collection of peer-reviewed *in-vivo* pulmonary toxicity data from exposure to several Metal oxide nanoparticles (MONPs).

The clustering results indicate that both sets of nanomaterials (CNTs and MONPs) can be categorized into 4 toxicologically unique clusters. The potency of the clusters can be extracted from their respective dose-response and recovery relationships. In the case of CNTs, the cluster with the largest potency was found to be 80 to 400 times more potent than the lowest which is indicative of a large spread across the responses. The MONP clusters showed a much larger spread with the estimated potency being 400000 times higher.

Investigation of physical differences between the CNT and MONP clusters revealed a pattern but the lack of sufficient data and high degree of data overlap across the clusters precludes the possibility of assigning statistically meaningful labels based on their physical properties.

The standardized potency of the clusters for the 5 endpoints was compared in both cases. 2 clusters (C and D) showed increased levels of immune response activity, 1 cluster (A) only showed increased cell damage indicators and, 1 cluster (B) displayed both elevated immune response and cell damage indicators based on the short-term endpoints measured in the test subjects. Amongst the 4 MONP clusters compared, 3 clusters (I, II and IV) showed signs of elevated immune system activity and cell membrane damage, these clusters were primarily composed of Iron oxide nanoparticles (cluster I), Silica (cluster II) and Titanium dioxide (cluster IV). 1 cluster (III) showed below average activity across all 5 responses tested, this cluster primarily featured Zinc oxide and Cerium oxide nanoparticles.

The final part of the dissertation discusses the derivation of model-based no-observed-adverse-effect-level (NOAEL) predictions for the clusters, referred to as MP-NOAELs. The MP-NOAELs for the provisionally titled “Long and Thin” variety of CNTs were found to be the lowest, indicating that those CNTs showed adverse effects at the lowest doses. A cluster comprised of Cerium oxide, Nickel oxide and Zinc oxide nanoparticles was estimated to have the lowest MP-NOAEL amongst the MONPs. A sensitivity analysis of the MP-NOAEL derivation highlighted the dependency of the process on the shape and type of the fitted dose response model, its parameters, dose selection and spacing, and the sample size analyzed.

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES ix

ACKNOWLEDGEMENTS xi

Chapter 1 Introduction 1

Chapter 2 A Dose-Response-Recovery Clustering Algorithm for Categorizing Carbon Nanotube Variants into Toxicologically Distinct Groups 6

1. Introduction 8
2. Methods 11
   2.1 Data Selection 11
   2.2 Model Definition 14
   2.3 Clustering Algorithm 15
3. Results and Discussion 16
4. Conclusions 24

Author Information 26
Author Contributions 26
Funding Sources 26
Acknowledgement 26
References 27

Chapter 3 Examining the In Vivo Pulmonary Toxicity of Engineered Metal Oxide Nanomaterials using a Genetic Algorithm-based Dose-Response-Recovery Clustering Model 30

1. Introduction 31
2. Methods 35
   2.1 Data Selection 35
   2.2 Model Definition 36
   2.3 Clustering Algorithm 37
   2.4 Genetic Algorithm 38
3. Results and Discussion 44
   3.1 Composition-Blind Clustering Results 44
   3.2 Results of Clustering by Patricle Composition 47
      3.2.1 Observed Differences in Toxicity 52
      3.2.2 Possible Policy Implications 56
4. Conclusions 57

Author Information 59
Author Contributions 59
Funding Sources 59

Acknowledgement 59
References 60

Chapter 4 Evaluation of Risk and Uncertainty for Model-Predicted NOAELs of Engineered Nanomaterials Based on Dose-Response-Recovery Clusters 67

1. Introduction 68
2. Methods 71
   2.1 Data Selection 71
   2.2 Model and Algorithm Definition 76
   2.3 Model-Predicted NOAELs 77
3. Results and Discussion 79
   3.1 Clustering Results 79
   3.2 Characterization Evaluation 80
   3.3 Model-Predicted NOAELs 83
   3.4 Sensitivity Analysis 84
4. Conclusions 91
Acknowledgement 92
References 93

Chapter 5 Conclusions 100

LIST OF FIGURES

Figure 2-1: Clustering tree for CNT pulmonary toxicity using LDH as response. The tree illustrates the progress of the algorithm from a single cluster containing all the CNT variants into 4 new toxicologically distinct clusters along with the distribution of particles across them 17

Figure 2-2: Contour plot of cluster model for LDH. The yellow region indicates the high dose and low post exposure response (elevated end-points). The bluer regions correspond to lower dosage and increasing post exposure periods (decreasing response/potential return to homeostasis) 18

Figure 2-3: Corrected Akaike Information Criterion (AICc) as a function of the number of clusters in the mode. The figure illustrates the optimum number of clusters is 4 based on the minimum achieved for all endpoints indicating there are 4 distinct groups of CNTs 18

Figure 2-4: The cluster averaged potency for each response. The values are indicative of the relative increase in measured response for the same given dosage. 22

Figure 2-5: The variation between natural logs of the potency of the responses in each cluster provides insight towards the observed effects each cluster of particles has on the response. A single CNT particle constituent of Cluster B possessed low potency as derived from the dose-response-recovery model 23

Figure 2-6: The variation of the standardized potency of the responses across the clusters allows for easier interpretation of the observed effects of exposure to the particle clusters. Cluster A shows signs of increased cell damage. Cluster B shows signs of both increased immune response and cell damage while clusters C & D only displayed signs of increased immune response activity 23

Figure 3-1: Variation between the "Total no. of configurations" analyzed vs "No. of Nanoparticles". The plot demonstrates the increasing computational load as more particles are added to the analysis 39

Figure 3-2: Penalized Akaike Information Criterion (AICc) as a function of the number of clusters. The figure illustrates the variation in the number of particles analyzed for each response and hence the lack of a consensus optimum number across the 5 responses 45

Figure 3-3: Corrected Akaike Information Criterion (AICc) as a function of the number of clusters in the mode. The figure illustrates the optimum number of clusters is 4 based on the minimum achieved for all endpoints indicating there are 4 distinct groups of MONPs 49

Figure 3-4: Corrected Akaike Information Criterion (AICc) comparison between models with a linear recovery (L) versus models with an exponential recovery (E). The lack

of any discernable difference between the AICc values is evidence to back our assumption of a linear recovery 50

Figure 3-5: The variation between natural logs of the potency of the responses in each cluster provides insight towards the observed effects each cluster of particles has on the response. LDH response levels were highly elevated for clusters 'II', 'III' and 'IV'.....55

Figure 3-6: The variation of the standardized potency of the responses across the clusters allows for easier interpretation of the observed effects of exposure to the particle clusters. Cluster 'I', 'II" and 'IV' shows signs of increased cell damage and elevated immune response. Cluster 'III' shows below average toxic potency across all 5 measured endpoints.....56

Figure 4-1: Confidence spread for the 4 clusters across the 5 responses for the CNT particles. Tighter spreads correlate to lower mean squared error computations. Larger spreads are associated with increased error in parameter estimation 85

Figure 4-2: Dose-response including the 95% confidence interval of the response for the 4 clusters across the 5 responses for the MONP particles. Tighter spreads correlate to lower mean squared error computations. Larger spreads are associated with increased error in parameter estimation. The color of shade is used to present the type of model used in evaluate the dose-response variation. Red is used to represent the traditional exponential dose-response model discussed in this chapter. Blue is used for a linear representation of the dose-response variation. Green is used to represent a "simple" exponential model was substituted for the traditional dose-response model.....87

LIST OF TABLES

Table 2-1: Compilation of the data (with sources) used for the CNT clustering process. (NR designation indicates “Not Reported”). 13

Table 2-2: Clusters generated across the 5 endpoints analyzed. The highlighted clusters are to illustrate the consistency of the clustering process. Certain clusters (1021) are found to exist independently or in conjunction with one another (1001, 1002, 1009 & 1003, 1007). 21

Table 2-3: Potency values for response across cluster categories (b x 10^{-10}). The disparity between the highest and lowest values for potency can be observed. The highest value is more than 400 times larger than the lowest across all the responses which further stresses the need for sub-groups. 21

Table 3-1: Compilation of the data (with sources) used for the MONP clustering process. (NR designation indicates “Not Reported”) 42

Table 3-2: Mean "typical particle size" for MONP clusters along with constituent cluster composition for PMN response. The composition of the clusters indicates that multiple metal oxide particles can be categorized together by virtue of their dose-response-recovery relationships 47

Table 3-3: Fractional overlap of exposure group data between the measured endpoints from the BAL fluid of the animals 50

Table 3-4: Clusters generated across the 5 endpoints for the various metal oxides (data sources are indicated in parentheses). The bolded elements illustrate particle composition “anchors” across the clustering results 51

Table 3-5: Potency values for response across cluster categories (b x 10^{-14}). The disparity between the highest and lowest values for potency can be observed. The highest value is more than 400,000 times larger than the lowest across all the responses which further stresses the need for sub-groups 52

Table 4-1: Compilation of the data (with sources) used for the analysis presented in this publication 73

Table 4-2: Median physical attributes of the CNT clusters for PMN response. The attributes can be used to group the clusters based on their physical properties. Provisional thresholds can be set at a median length of 2200 nm and median diameter of 10 nm to generate labels for the clusters 82

Table 4-3: Mean "typical particle size" for MONP clusters along with constituent cluster composition for PMN response. The composition of the clusters indicates that multiple metal oxide particles can be categorized together by virtue of their dose-response-recovery relationships 82

Table 4-4: Mean particle size and predicted NOAELs for the newly formed clusters based on chemical isolation of the MONPs. U indicates that a range of values for the cluster NOAEL is undefined.....84

Table 4-5: Model based Predicted NOAEL (MP-NOAEL) ranges (95% level of confidence) for Carbon Nanotube particles. The negative range of values indicate the source publications analyzed lack test results at low dose levels. U indicates that the evaluated NOAELs are undefined. The bolded range of values are examples where the confidence spread appears rational, yet the derived NOAEL range includes irrational values (e.g., negative values). All values presented below are in (μg/kg).....89

Table 4-6: Model based Predicted NOAEL (MP-NOAEL) ranges (95% level of confidence) for Metal Oxide Nanoparticles. The negative range of values indicate the source publications analyzed lack test results at low dose levels. U indicates that the evaluated NOAELs cannot be defined. The bolded range of values are examples where the confidence spread appears statistically valid, yet the derived NOAEL range contains irrational limits (e.g., negative values). All values presented below are in (μg/kg).....89

ACKNOWLEDGEMENTS

I would like to sincerely thank Dr. Jeremy Gernand for his invaluable guidance and immense patience throughout my time here as his student. I have been extremely lucky to have an advisor who cared so much about my work as well as my personal and professional development. He was always available to respond to my queries, and most importantly helped me realize my potential during this journey.

I want to sincerely thank NIOSH, NSF, EPA, and CEINT for their support and funding for the duration of my research. This dissertation would not have been possible without them. I would like to emphasize that the findings and conclusions of the research presented in this dissertation do not reflect the views of the funding agencies.

I want to give my wholehearted appreciation and thanks to the Department of Energy and Mineral Engineering at Penn State that provided with me the platform, the resources, and avenues for me to perform up to my abilities and strengths. I would like to extend my gratitude towards all the professors and staff, especially to Jaime Harter and Sue Hyde who work tirelessly to keep the department vibrant and welcoming.

I want to thank Dr. William Groves, Dr. Ephraim Hanks, and Dr. Eugene Morgan for serving on my committee and helping me navigate through the various obstacles and guide me in the right direction with regards to the quality of my research. Their valuable suggestions have made a significant difference.

A special note of thanks to Dr. Jonathan Mathews for his patience, guidance, and mentorship during the first two years of my stay at Penn State as a graduate student. I cannot imagine making it to Penn State without his help, it was a real pleasure working with him and learning from him.

Last but not in any way the least, I want to thank my parents, sister, and grandmother for their unending love and encouragement. They continue to be my unshakeable pillars of support. Special shoutout and thanks to all my friends and colleagues at Penn State, without your presence and support this journey would never have been as much fun or eventful.

Chapter 1

INTRODUCTION

Nanomaterials and nanostructures have developed at a tremendous rate in the last 10 years and projections surrounding their future growth are highly positive. Reports have indicated that the global market for nanotechnology could eclipse $90 billion by 2021. The sky is literally the limit because NASA is exploring the viability of carbon-based nanoparticles for use in outer space. Nanoparticles are highly valuable and coveted for the special properties exhibited by the particles at the nanoscale (<100 nm). They have found extensive usage in a multitude of areas, food and consumer products, fuel and automobiles, and pharmaceutical drug field are just a few prominent examples.[1] While nanomaterials provide a significant advantage due to their size, their size also makes them susceptible to be regarded as potentially toxic and harmful to upon exposure.

The study of the acute and chronic effects of exposure to nanomaterials and their associated pathology has been extensively documented over the years. The effects of "*in-vivo*" and "*in-vitro"* nanomaterial exposure have been studied via multiple modes of exposure (inhalation, instillation, aspiration etc.) to investigate and assess the potential damage.[2-4] In-vitro studies modeling nanoparticle toxicity use a wide variety of assays to study the effects of exposure, they can be scaled for large scale repeatability and are generally more affordable to conduct than their counterpart. A shortcoming of in-vitro testing is the inability to simulate the identical environment and conditions found in living organisms. In-vivo testing employs a wide variety of animals but primarily rats and mice, the mammals are exposed to varying concentrations of multiple nanomaterials via inhalation or instillation and subsequently their bronchoalveolar lavage fluid (BALF) is extracted.

The measures derived from the BALF are indicators of the degree of response exhibited by the subjects tested. The measures include indicators of inflammation such as Polymorphonuclear neutrophils (PMN), cell damage indicators such as Lactate dehydrogenase (LDH), immune response activity indicator Macrophages (MAC), Total Protein (TP) content in the BALF, and the overall Total Cell Count (TCC) in the BALF. The major shortcomings of in-vivo testing lie in the degree of variability associated with the experimentation and the lack of reliable repeatability of results obtained, they are also time-consuming and expensive to be conducted at a large-scale. The time and cost factor associated with in-vivo toxicity precludes the possibility of testing every newly synthesized nanomaterial, a pre-screening process that is capable of identifying the particles exhibiting signs of elevated toxic potential would be helpful in prioritizing the order of nanoparticle testing.[5]

The development of computational and quantitative approaches towards analyzing and assessing the toxicity of nanomaterials would assist manufacturers and regulators in implementing early-stage controls, these controls would aid the design and synthesis of new nanomaterials that are less toxic. Quantitative structure activity relationships (QSARs) are an example of a modeling tool that is able to express the biological activity of a nanostructure as a mathematical function of its physiochemical properties. QSARs have been extensively used to study the cytotoxicity of multiple nanomaterial types, but they have been limited in their scope.[6-15] Meta-analyses offer an alternate route to analyze the toxicity by curating published data across the complete timeline of study of nanoparticle toxicology. Meta-analysis of various nanomaterial types has been published over the years.[10, 12, 15-20] The benefits and limitation of the process are also well documented.[21, 22] The plethora of computational approaches combined with access to an ever expanding knowledge base of in-vitro and in-vivo testing studies makes it possible to compare nanomaterials, identify their relative toxicity, and identify the impactful properties behind it.

This dissertation presents 3 papers that describe the development of an analytics driven approach towards categorizing engineered nanomaterials by their toxicity, identifying the properties responsible for their potency, and deriving a dose-response model based no observed adverse effect level (NOAEL) for the nanomaterials. The first two papers presented in Chapters 2 and 3, discuss the development of a hierarchical clustering methodology that is capable of categorizing nanomaterials on the basis of the dose-response and recovery similarity. Chapter 2 is a case study on the meta-analysis of Carbon nanotube (CNT) in-vivo pulmonary toxicity. Chapter 3 presents a case study on the meta-analysis of Metal oxide nanoparticles (MONP). The last paper in Chapter 4 presents the derivation of cluster-based model predicted NOAEL referred to as the MP-NOAEL for the CNTs and MONPs. Chapter 5 presents the conclusions drawn across the 3 papers and direction for future research.

REFERENCES

1. Smijs, T.G. and S. Pavel, *Titanium dioxide and zinc oxide nanoparticles in sunscreens: focus on their safety and effectiveness.* Nanotechnol Sci Appl, 2011. **4**: p. 95-112.

2. Sayes, C.M., et al., *Comparative pulmonary toxicity assessments of C-60 water suspensions in rats: Few differences in fullerene toxicity in vivo in contrast to in vitro profiles.* Nano Letters, 2007. **7**(8): p. 2399-2406.

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

4. Warheit, D.B., C.M. Sayes, and K.L. Reed, *Nanoscale and Fine Zinc Oxide Particles: Can in Vitro Assays Accurately Forecast Lung Hazards following Inhalation Exposures?* Environmental Science & Technology, 2009. **43**(20): p. 7939-7945.

5. Choi, J.Y., G. Ramachandran, and M. Kandlikar, *The Impact of Toxicity Testing Costs on Nanomaterial Regulation.* Environmental Science & Technology, 2009. **43**(9): p. 3030-3034.

6. Puzyn, T., D. Leszczynska, and J. Leszczynski, *Toward the Development of "Nano-QSARs": Advances and Challenges.* Small, 2009. **5**(22): p. 2494-2509.

7. Puzyn, T., et al., *Nanomaterials - the Next Great Challenge for Qsar Modelers.* Recent Advances in Qsar Studies: Methods and Applications, 2010. **8**: p. 383-409.

8. Puzyn, T., et al., *Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles.* Nature Nanotechnology, 2011. **6**(3): p. 175-178.

9. Shao, C.Y., et al., *Dependence of QSAR Models on the Selection of Trial Descriptor Sets: A Demonstration Using Nanotoxicity Endpoints of Decorated Nanotubes.* Journal of Chemical Information and Modeling, 2013. **53**(1): p. 142-158.

10. Bakhtyari, N.G., et al., *Prediction of Genotoxicity of Nano Metal Oxides by Computational Methods: A New Decision Tree QSAR Model.* Environmental and Molecular Mutagenesis, 2014. **55**: p. S43-S43.

11. Odziomek, K., et al., *Toward quantitative structure-activity relationship (QSAR) models for nanoparticles.* Abstracts of Papers of the American Chemical Society, 2014. **248**.

12. Pathakoti, K., et al., *Using experimental data of Escherichia coli to develop a QSAR model for predicting the photo-induced cytotoxicity of metal oxide nanoparticles.* Journal of Photochemistry and Photobiology B-Biology, 2014. **130**: p. 234-240.

13. Singh, K.P. and S. Gupta, *Nano-QSAR modeling for predicting biological activity of diverse nanomaterials.* Rsc Advances, 2014. **4**(26): p. 13215-13230.

14. Esposito, E.X., et al., *Exploring possible mechanisms of action for the nanotoxicity and protein binding of decorated nanotubes: interpretation of physicochemical properties from optimal QSAR models.* Toxicology and Applied Pharmacology, 2015. **288**(1): p. 52-62.

15. Pan, Y., et al., *Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors.* Rsc Advances, 2016. **6**(31): p. 25766-25775.

16. Chang, X.H., et al., *Health effects of exposure to nano-TiO2: a meta-analysis of experimental studies.* Nanoscale Research Letters, 2013. **8**: p. 1-10.

17. Cho, W.S., et al., *Metal Oxide Nanoparticles Induce Unique Inflammatory Footprints in the Lung: Important Implications for Nanoparticle Testing.* Environmental Health Perspectives, 2010. **118**(12): p. 1699-1706.

18. Genaidy, A., et al., *Health effects of exposure to carbon nanofibers: Systematic review, critical appraisal, meta analysis and research to practice perspectives.* Science of the Total Environment, 2009. **407**(12): p. 3686-3701.

19. Hernández, L.G., et al., *Can carcinogenic potency be predicted from in vivo genotoxicity data? a meta-analysis of historical data.* Environmental and Molecular Mutagenesis, 2011. **52**(7): p. 518-528.

20. Notter, D.A., D.M. Mitrano, and B. Nowack, *Are Nanosized or Dissolved Metals More Toxic in the Environment? A Meta-Analysis.* Environmental Toxicology and Chemistry, 2014. **33**(12): p. 2733-2739.

21. Ioannidis, J. and J. Lau, *Pooling research results: benefits and limitations of meta-analysis.* The Joint Commission journal on quality improvement, 1999. **25**(9): p. 462-469.

22. Spector, T.D. and S.G. Thompson, *The potential and limitations of meta-analysis.* Journal of epidemiology and community health, 1991. **45**(2): p. 89.

CHAPTER 2

A Dose-Response-Recovery Clustering Algorithm for Categorizing Carbon Nanotube Variants into Toxicologically Distinct Groups

The contents of this chapter have been published in "Computational Toxicology" under the title *"A dose-response-recovery clustering algorithm for categorizing carbon nanotube variants into toxicologically distinct groups" (https://doi.org/10.1016/j.comtox.2019.02.003).*

ABSTRACT

This article presents a dose-response clustering algorithm for the purpose of revealing toxicologically distinct clusters of carbon nanotubes (CNTs). Current exposure guidelines such as those published by the National Institute for Occupational Safety and Health (NIOSH) consider all types of CNTs and carbon nanofibers (CNFs) as a single substance, even though experimental data have demonstrated significant differences in CNT toxicity. The unique combinations of physical and chemical characteristics cause variations in the observed dose-response-recovery that is substantial enough to consider them as a collection of different substances rather than a single substance. This paper presents an algorithm capable of grouping CNTs into toxicologically distinct clusters assisting in the identification of physicochemical differences between the clusters, and different proposed exposure limits for each of the CNT classes. Based on a dataset composed of peer-reviewed *in vivo* experimental studies in rodents, the CNT variants are divided into sub-groups based on their dose-response-recovery similarity and the Akaike Information Criterion (AIC) of the family of models. Results indicate the presence of 4 toxicologically unique

CNT classes based on 5 toxicity endpoints selected. Certain physicochemical attributes vary significantly between clusters and are more likely to define the categories than others. The potency of the clusters is derived from their associated dose-response-recovery relationship parameters. The clusters with largest potency were found to be between 80 to 400 times more potent than the cluster with the lowest potency indicative of a large spread between the values across all the responses. The absence of key characterization data for some of the CNT variants analyzed in this study prevents the designation of physical characteristic-based labels which could have assisted in identifying the key factors affecting the toxic potential of CNTs. The standardized potency of the 4 CNT clusters was compared, 2 clusters showed increased levels of immune response activity, 1 cluster only showed increased cell damage indicators and, 1 cluster displayed both elevated immune response and cell damage indicators based on the short-term endpoints measured in the test subjects.

Abbreviations

CNT, carbon nanotubes; SWCNT, single-walled carbon nanotubes; MWCNT, multi-walled carbon nanotubes; BAL, bronchoalveolar lavage; MAC, macrophage count in BAL fluid; TP, total protein concentration in BAL fluid; LDH, lactose dehydrogenase; PMN, polymorphonuclear neutrophil; TC, total cell count.

Keywords

Dose-Response, Recovery, Pulmonary Toxicity, CNT, Hierarchical Clustering, Potency

1. INTRODUCTION

The development of new nanomaterials (NMs) is progressing at a pace that exceeds the capabilities of traditional toxicological testing, which can be slow and expensive.[1] Some of these NMs might eventually pose an occupational hazard or become an environmental pollutant presenting a risk to the people working with them and to those in the surroundings. NIOSH has proposed a recommended exposure limit (REL) of 1 $\mu g/m^3$ for Carbon nanotubes (CNTs) and nanofibers (CNFs) based on human equivalent no-observed-adverse-effect-levels (NOAEL) and lowest-observed-adverse-effect-levels (LOAEL) estimates of 3.5 - 18 $\mu g/m^3$ and uncertainty factors of 20 or 60.[2] Pauluhn (2010) proposed an occupational exposure limit (OEL) for one specific variety of multi-walled CNTs (0.05 mg/m^3) based on NOAEL of 0.1 mg/m^3 and adjustment factor of 2, the resultant OEL is 50 times higher than the NIOSH REL.[3] Similarly, several other researchers have documented their NOAEL and corresponding adjustment factors. These NOAELs and adjustment factors all vary between 0.1 to 0.13 mg/m^3 and 2 to 50 respectively.[2] Given the diversity in responses to different varieties of CNTs further specification of the exposure limits by CNT categories may be beneficial by incentivizing the production and use of less toxic varieties of CNTs in the future.

Experimental toxicology studies are often time consuming and require considerable investment. A viable solution to the problem is testing the NMs on a priority basis—targeting the more toxic variants first. The development of computational approaches targeted at predicting the toxicity of the NMs based on their attributes can be used as a means to guide nanomaterial synthesis towards less hazardous variants. Experimental evaluation is subject to a certain degree of error or bias and in some cases the results may be difficult to replicate. To overcome some of these inefficiencies, forecasting methods have been proposed to predict the potential toxicity and environmental impact of new NMs.[4] Computational modeling of nanotoxicity is valuable for its efficiency of time and expense and the flexibility offered. Quantitative structure activity

relationships (QSARs) are one of the modeling schemes utilized; these models hypothesize that the activity of any nanostructure is a function of its physicochemical properties.[5] QSARs have been employed to predict the solubility, Young's modulus, and partition coefficient of nanomaterials, showcasing their applicability to NMs.[6, 7] Puzyn et al. demonstrated the use of QSARs to predict the mechanism of cytotoxicity of 17 different metal oxide nanoparticles.[4] A QSAR toxicity study performed on Carbon nanotubes (CNTs) concluded that increase in contact (via chemical treatments) between CNTs and the bacterial cells leads to enhanced cytotoxicity of CNTs.[8] Fourches et al. have also published an application of QSAR's to predict the biological activity of synthesized nanoparticles and prioritizing their design with minimal associated risk.[9] Investigation into the mechanism of nanotoxicity of functionalized CNTs using QSAR models and the study of 6 different endpoints has been conducted previously[10], further studies explore the role played by the molecular descriptors and the physicochemical properties in the selection and evaluation of QSAR models.[11]

Meta-analysis is a complementary method useful for harnessing the collection of published nanotoxicology data available, and testing hypotheses that were not possible within the individual experiments. Meta-analyses employ statistical and/or machine learning approaches such as regression, neural networks, decision trees, and support vector machines to identify important associations.[12] Since meta-analysis curates data from multiple sources, the individual variation or appearance of unusual results across the complete dataset are evaluated in light of the whole. Traditional meta-analysis, which involves literature-based data collection, can in some cases be ineffective due to the presence of missing data caused by disparity in NM characterization methods.[13] Meta-analysis has been previously used to study CNT toxicity to assess whether cellular penetration is a possibility in *in vitro* experiments; the study concluded that the possibility exists and recommended minimizing the exposure.[14] Gernand and Casman performed a meta-analysis of CNT pulmonary toxicity studies to assess how the physical and

chemical attributes of a CNT variant affect its toxicity.[13] The analysis employed a random forest regression modeling approach to determine the significance of the attributes and their effect on the response; their study concluded that metallic impurities, length, diameter and aggregate size significantly impact the toxicity of the NMs. However, that approach does not take into consideration dose-response pairing and instead treats the total dose as equivalent to a particle property for the CNTs. Modeling the dose-response curves for each nanoparticle would enable clustering of the CNTs and also identify the attributes contributing to the separation of different classes and their impact on the toxicity.

The benefits of meta-analysis techniques include its ability to identify the sources of diversity across the various participating studies[15], it can be used to improve the precision of the estimates due to the increase in the size of the population.[16] The method however is not without its limitations: it has no control over the homogeneity of the participating studies[15]; it cannot improve the quality of the data used and consequently the final result; and it could suffer from "publication bias", where researchers fail to report or exclude certain studies to present a more coherent result which could impact the analysis.[16]

Attempts at using computational modeling to predict the effects on toxicity of changes in particle properties (QSARS, meta-analysis etc.) have been limited in their scope to date. These methods have either been exclusive to only metal oxide particles, particles of a particular size, cytotoxicity assays, or structural analysis of single particles.[4, 9, 17-20] There is a need to quickly identify those NMs which pose greater risk than others to maximize their benefit while simultaneously minimizing the risk associated with them. This could be accomplished by quantifying the current knowledge base of in vivo and in vitro studies to reveal how different NMs compare to one another and how changes in their properties affect their toxicity. The rapid increase in the number of NMs synthesized reveals a need to cluster together NMs displaying similar effects as a means of grouping the NMs on the basis of their toxicity. To address the

question of how different nanomaterials can be considered the same substance toxicologically, this study proposes a new algorithmic approach applied to meta-analysis of extant CNT toxicity data.

The primary objectives of this study can be summarized as: (1.) to develop a new algorithmic approach to cluster together NMs based on the similarity of their dose-response-recovery relationship; and (2.) to evaluate the relative toxic potency of CNT clusters and associated variations in the physicochemical properties of the clusters.

Progress in this direction could be significant in the analysis of new and emerging nanoscale materials which can be categorized based on their attributes and potency, this form of categorization could assist in the prioritization of toxicological testing. Prioritizing animal testing would facilitate the development of new nanomaterials while keeping a check on the toxic variants. The evaluation of exposure limits for the new clusters is also necessary to assess their agreement with the existing single limit recommendation. Evaluation of new exposure limits would also assist in refining the current exposure limits established at workplaces.

2. METHODS

2.1 Data Selection

The dataset for the preliminary clustering analysis utilizes a collection of carbon nanotube (CNT) toxicity studies curated from 22 different publications. The publications reported the results on *in vivo* CNT pulmonary exposures in rodents (rats and mice). These publications were identified through online search engines including Web of Science, Google Scholar, and PubMed using search terms such as "*in vivo*", "CNT", "BAL" and "Pulmonary". To be included in this dataset, studies had to perform independent characterization of the CNTs, *in vivo* pulmonary exposure of those CNTs to rodents (*in vitro* studies were not included), and

quantitative measures of response in Broncho alveolar lavage (BAL) fluid. Each study is a collection of several exposure groups (experiments) identified by their total dose, mode of exposure, length of the post-exposure recovery period, and the physical and chemical characteristics of the CNT particles. The studies included in the final analysis satisfied a key criterion: the BAL fluid measures for the rodents must be measured for at least 1 dose level excluding the control and 1 or more post exposure periods. All publications used as the source for this analysis have been listed in Table 2-1, additional information is provided in the supplementary data.

Five quantitative toxicity endpoints were reported in a sufficient number of publications to be included in this analysis including: the total cell count (TCC), macrophage count (MAC), total protein concentration (TP), polymorphonuclear neutrophil count (PMN) and lactate dehydrogenase concentration (LDH). The exposure characteristics that are of interest are the total dose, post-exposure recovery period, and the CNT particle characteristics (such as the median length, diameter, specific surface area, and % metallic impurities). The response variables are expressed as folds (multiples) of the control group's measure i.e., if the response observed in the experiments is 4 times what the control group displayed the fold of control value is 4. Previous work has indicated that this response normalization is sufficient to consider different exposure modes (instillation, aspiration, and inhalation) as similar.

Table 2-1. Compilation of the data (with sources) used for the CNT clustering process. (NR designation indicates "Not Reported").

StudyID	Primary Author	CNT Type	No. of endpoints (of interest) reported	No. of exposure groups	CNT Length (nm)	CNT Diameter (nm)	Average Specific Surface Area (m^2/g)	Average Impurities (%)	Average MMAD (nm)
1001	Pauluhn[21]	MWCNT	5	13	200 – 300	10	253	1.4	1925
1002	Ma-Hock[22]	MWCNT	5	7	100 – 10000	5 – 15	275	10	1200
1003	Muller[23]	MWCNT	3	9	700 – 5900	9.7 – 11.3	342.5	2.1	NR
1004	Shvedova[24]	SWCNT	4	4	100 – 1000	0.8 – 1.2	508	18	4200
1005	Shvedova[24]	SWCNT	2	7	100 – 1000	0.8 – 1.2	508	18	4200
1006	Nygaard[25]	SWCNT, MWCNT	2	7	500 – 200000	4.05 – 15.04	341.3	5	NR
1007	Warheit[26]	SWCNT	3	9	1000	1.4	NR	10	NR
1009	Park[27]	SWCNT	3	5	2000 – 10000	1.2	NR	10	NR
1010	Teeguarden[28]	SWCNT	3	2	500 – 2000	0.4 – 1.2	1040	0.3	NR
1011	Elgrabli[29]	MWCNT	1	16	500 – 2000	20 – 50	NR	NR	NR
1012	Mercer[30]	SWCNT	2	4	NR	NR	NR	2	690
1013	Porter[31]	MWCNT	2	13	3860	8.8 – 89.2	NR	0.5	NR
1016	Ellinger-Ziegelbauer[32]	MWCNT	3	10	200 - 300	10 – 16	253	NR	2533
1017	Shvedova[33]	SWCNT	4	16	NR	1 – 4	1040	0.3	NR
1021	Ge[34]	MWCNT	4	8	NR	NR	NR	5.95	NR

2.2 Model Definition

The dose-response model is a key parameter for the clustering analysis; it describes the rate of change of the response variable as the dosage levels are varied. A single dose-response model cannot be utilized to explain the behavior of all substances and organisms due to the complexity of the situation, leading to the development of benchmarked dose-response models that are able to cater to a wide variety of situations in the field of toxicology. Slob (2001) suggested a 3-parameter model available for modeling continuous data. But this model did not reflect the decrease in the response levels over time when no longer exposed, which is a key characteristic to exposures in complex organisms like rodents. We propose a revised model to analyze the complete dose-response-recovery relationship by the addition of a fourth parameter reflecting the decay in the response over time (i.e., the animal recovers). The relation between the response (y), dose (x) and post-exposure period (t) is,

$$y = a[c - (c - 1)e^{-bx}] - dt$$

The 4 parameters, a (signifies the response at dose = 0), b (the toxic potency of the nanoparticles), c (the maximum relative shift in response), and d (slope of the response decay) are the key parameters which can be used to quantify and reflect the potential of the particle to be a hazard.

Recovery is likely to have an exponential relationship as the animal returns to homeostasis following the end of exposure. However, since such a relationship would require the addition of two additional model parameters, rather than one for the linear term, and given the available data fit the linear representation of recovery better, we make the parsimonious choice here with the single caveat that the model is valid for all recovery times, t, until the value of the response returns to 1 (i.e., no change from the control group).

2.3 Clustering Algorithm

The proposed algorithm is aimed at segregating the dataset into groups or "clusters" based on the similarity of their dose-response-recovery relationship. The process is similar to decision trees or hierarchical clusters and mimics those structures. The algorithm is capable of fitting the studies from dataset to the model proposed using non-linear regression. The process involves splitting the dataset into clusters at each stage and progressing till we obtain the best overall combination of clusters. The performance metric used to evaluate the goodness of a cluster is the Akaike Information Criterion (AIC), the AIC is calculated based on the number of parameters (k) in a model and its log-likelihood (L), $AIC = 2k - 2\ln(L)$.

The likelihood function of a model having a given set of parameters is defined as the probability of that model achieving an outcome given those parameters, maximizing the logarithm of the likelihood function reduces the error between the expected and observed data. The AIC can be regarded as a measure of the information lost; hence lower values of AIC are highly preferred. The overall AIC of a cluster is obtained by summing together the AIC values of its component studies. AIC needs to be modified to include corrections that can be applied when dealing with small datasets, this corrected form is denoted as AICc.

$AICc = AIC + \frac{2k(k+1)}{n-k-1}$, where n is the number of observations and k is the number of model parameters.

For the purpose of this study, we assume the likelihood function of the model follows a normal distribution and adopt the AICc as our measure of performance. The algorithm uses an exhaustive search approach and evaluates all possibilities (i.e., all possible combinations) before choosing the best available option. For example, if the dataset were to contain 5 particle studies, step 2 in the process would determine that there are 4 "configurations" possible to separate the 5 particles into 2 clusters, but each of those has multiple scenarios, for example a cluster of 1

particle and a second containing 4 particles expressed as (1 4) has 5 different possible scenarios, while (2 3) has 4 possible scenarios, while (1 4) and (4 1) are basically equivalent mathematically, but different in how the tree looks visually. The next 2 steps in the process evaluate which grouping of studies is ideal for each of the 4 scenarios by generating a combination matrix of every possible grouping of the studies and selecting the situation with the lowest AICc value. The process is stopped if the dataset has been split completely into its individual studies and the best choice for the number of clusters is chosen the stage where the total AICc value is the least.

3. RESULTS AND DISCUSSION

A complete clustering tree developed using the algorithm is shown in Figure 2-1. This tree was generated using the multiple change in LDH in BAL fluid from the control group as the response variable; similar trees were generated for the remaining 4 responses. The tree provides insight into the size of the clusters formed and based on its constituent particles we are able to draw further inferences on their observed potency. The definition of each cluster is a specific dose-response-recovery model that describes the variation of the response for changes in dose and recovery time. Figure 2-2 displays a visual representation of a single cluster model illustrating the variation of the response for both dose and recovery.

The algorithm categorized the pulmonary effects of different CNT variants based on the similarity of their dose-response-recovery relationships. This process improves the information retention of the complete model as evidenced by the decrease in the AICc of the clusters at each stage observed in Figure 2-3. Figure 2-3 is a comparison of the variation in AICc values for all the 5 response indicators used in the study, it also serves to determine the ideal number of clusters for each indicator.

The choice of the numbers clusters is based on the tradeoff between increasing the number clusters that are toxicologically unique and the extra information that is conveyed by these additional clusters. The AICc is the best metric for comparing the amount of information that is provided by models of varying complexity or as in our case comparing configurations of the tree with varying degrees of splitting (e.g., more or fewer branches). Based on the evidence shown in Figure 2-3 it is clear that 4 clusters are the best choice on the basis of simplicity as each of the responses achieves their minimum at that stage. This implies that the choice of 4 unique clusters is able to best represent the difference between the clusters from a toxicological standpoint and further splitting to include more clusters categories does not add sufficient information to the existing group of models to overcome the increased complexity.

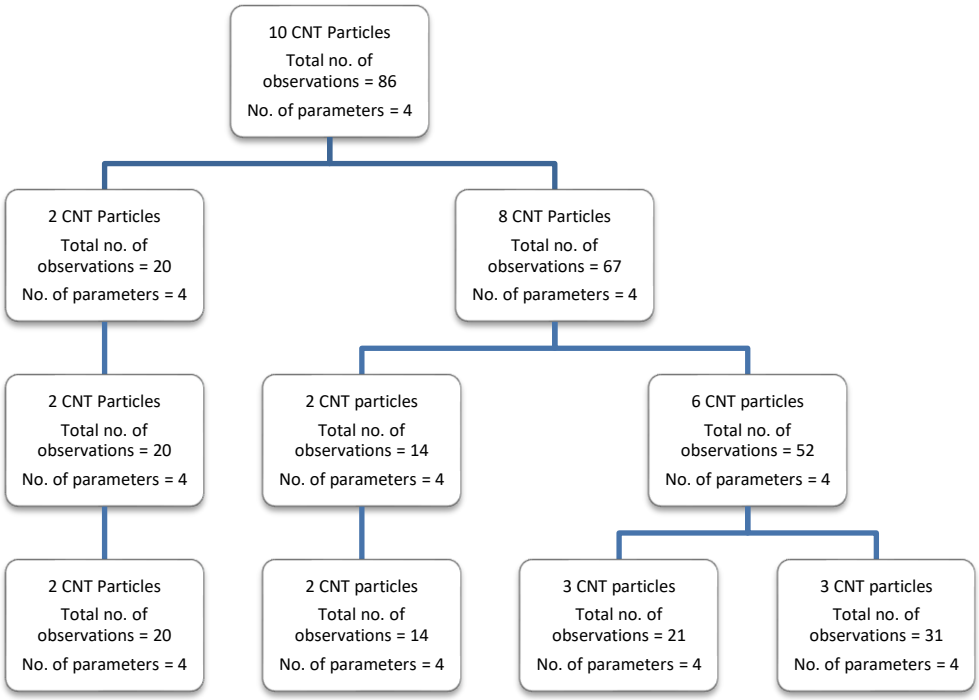

Figure 2-1. Clustering tree for CNT pulmonary toxicity using LDH as response. The tree illustrates the progress of the algorithm from a single cluster containing all the CNT variants into 4 new toxicologically distinct clusters along with the distribution of particles across them.

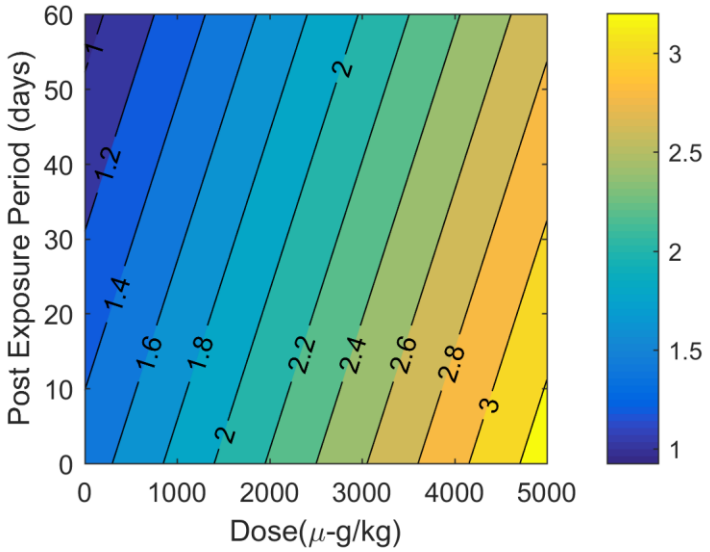

Figure 2-2. Contour plot of cluster model for LDH. The yellow region indicates the high dose and low post exposure response (elevated end-points). The bluer regions correspond to lower dosage and increasing post exposure periods (decreasing response/potential return to homeostasis).

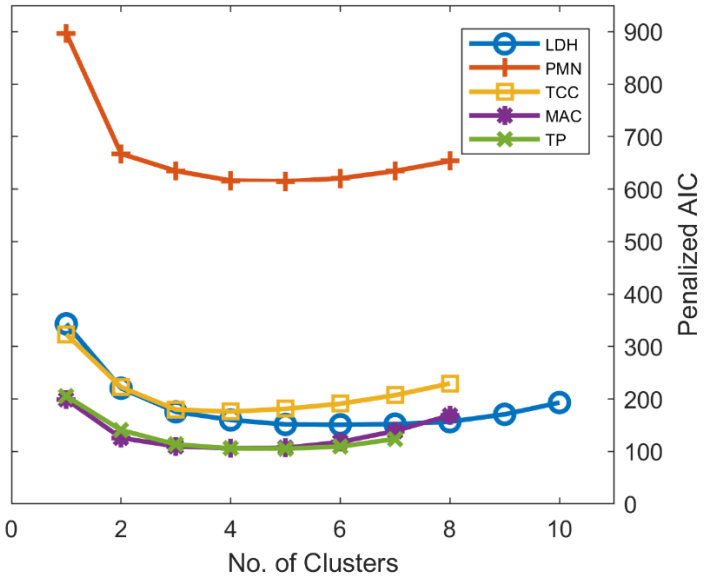

Figure 2-3. Corrected Akaike Information Criterion (AICc) as a function of the number of clusters in the mode. The figure illustrates the optimum number of clusters is 4 based on the minimum achieved for all endpoints indicating there are 4 distinct groups of CNTs.

Understanding that these clusters exist based on the similarity of their *in vivo* toxic effects is important but does not explain why these clusters exist. The next step required to answer this question is an examination of the CNT characteristics within each of the clusters to define their physical and chemical characteristics and generate a hypothesis as to why one cluster or class of CNTs may be more or less toxic than another. The reasoning behind not grouping the particles based on their physical characteristics as the initial step is that physically and/or chemically similar nanomaterials need not necessarily be similar in their dose-response-recovery relationships as well, depending on which characteristics are used to define similarity. Identifying the CNT particle attributes that do and do not vary significantly between the clusters aids in determining which properties are key contributors in the separation at each stage and also towards the relative toxicity between the clusters. Examination of the properties of all the particles in the clusters could assist in potentially identifying the physical characteristics unique to each cluster.

Table 2-1 showcases the relevant available characteristics associated with each particle included as part of this study. We can note here that there is some missing data on the physical characteristics and their distributions amongst the tested CNT variants to make determinations of statistically significant category labels. The endpoints reported by the source publications varied in some capacity for each CNT variant tested leading to a different set of CNT variants available for clustering at each endpoint measure. For example, from Table 2-1 we can observe that only sources 1001 and 1002 report all the 5 endpoints of interest to this paper, while source 1011 reports only 1 endpoint of interest. This variability influences the clusters, based on the toxic endpoints measured the members of the clusters change. Additionally, other, unmeasured or hidden characteristics unavailable to us may actually be mechanistically responsible for their variations in toxicity. Further experimental study is needed to understand why these clusters exist.

Table 2-2 reflects on the consistency of the clustering process; we can see that certain variants tend to exist with particular variants within clusters across the endpoints analyzed. The

clusters have been highlighted for ease of identification. We can observe that CNT variants (1001, 1002, 1009) occur across 2 endpoints, 1009 did not report LDH as an endpoint in their study. Similarly, CNT variants (1003, 1007) are clustered together across 2 responses. The particle characteristics detailed in Table 2-1 would allow for provisional designation of physical property-based labels to the clusters. The contrast in the CNT variants across the clusters would cause an overlap of the physical characteristics. The overlap coupled with the limited nature of the currently available data does not permit statistical tests to determine if the distributions of the length and diameter of the CNT clusters are significantly different. For example, CNT variant 1021 is present across 3 responses as an individual cluster but the source publication failed to report characterization data which could have assisted in identifying why it is isolated. The cluster configuration (1004, 1005, 1013) has also been highlighted because (1004 and 1005) are derived from the same source but were separated because their method of exposure differed (inhalation vs instillation), this could potentially be an indicator that the mode of exposure does not affect the response induced.

The last step of the clustering process is the examination of the “potency” of each cluster to provide insight into the severity and relative toxicity of CNT groups. The “potency” of a cluster is directly related to its toxic potential and derived from the dose-response-recovery model fitted to it. When comparing 2 particles with identical or similar characteristics, the particle with the higher potency can be regarded as possessing a greater toxic potential. The significance of using 4 clusters to categorize the CNTs is that it allows for ease of identification and future assignment of newly synthesized CNTs into more- or less-toxic categories based on their physical properties. Characterization data obtained for newly synthesized particles could be used in a similar fashion as discussed previously to establish sub-groups within clusters. Table 2-3 showcases the potency measures for the clusters derived across the five endpoints. The disparity and variation in the potency measures for the individual CNTs also reiterates the need for sub-

groups amongst the CNTs to ease future categorization. For example, the variation in LDH response models show a maximum potency approximately 400 times that of the minimum for the 10 particles analyzed.

Table 2-2. Clusters generated across the 5 endpoints analyzed. The highlighted clusters are to illustrate the consistency of the clustering process. Certain clusters (1021) are found to exist independently or in conjunction with one another (1001, 1002, 1009 & 1003, 1007).

LDH Clusters	TCC Clusters	MAC Clusters	PMN Clusters	TP Clusters
1001, 1002	1001, 1002, 1009	1001, 1002, 1009	1001, 1013	1001
1003, 1007	1007	1012	1003, 1007, 1013	1002, 1003, 1017
1016, 1017, 1021	1006, 1016	1021	1021	1021
1004, 1005, 1013	1004, 1011	1004, 1006, 1017	1002, 1005	1004, 1016

Table 2-3. Potency values for response across cluster categories (b x 10^{-10}). The disparity between the highest and lowest values for potency can be observed. The highest value is more than 400 times larger than the lowest across all the responses which further stresses the need for sub-groups

Cluster ID	LDH	PMN	Protein	Macrophages	Total cell count
A	147 ± 61.1	52.6 ± 715.2	248 ± 244.1	26.9 ± 28	50.6 ± 143.6
B	9.6 ± 61.1	797 ± 715.2	541 ± 244.1	74.6 ± 28	3.5E-04 ± 143.6
C	24.3 ± 61.1	51.7 ± 715.2	15.2 ± 244.1	15.3 ± 28	321 ± 143.6
D	40.5 ± 61.1	1580 ± 715.2	16.9 ± 244.1	13.6 ± 28	43.1 ± 143.6

Figures 2-4 and 2-5 are graphical presentations of the variation of the average potency of the clusters against the response indicators and the natural log of the response potencies in each cluster respectively. Higher values for potency result in greater increase in the response for a unit

increase in dose. For example, the high value of potency for the PMN neutrophil response levels imply that for even small increases in dose the PMN levels in the subject would show significantly larger increase compared to the other 4 responses. PMN neutrophils are essentially the white blood cells present in the body and are a core component of the immune system. Sharp changes in the white blood cell levels are usually associated with the onset of an infection. Identifying the clusters and particles which contribute to increased potency levels for the PMN could be useful in mitigating the onset of infections. Similar analogies can also be drawn for the remaining responses. Figure 2-6 compares the standardized potency of the clusters across the 5 responses, standardizing the potency makes it easier to compare the various response groups directly and determine their relative extent of threat. It can be seen in Fig. 6 that the particles belonging to clusters “A” display signs associated with cell membrane damage while clusters labeled “C and “D” show elevated levels of response indicators associated with immune activity. The cluster designated as “B” is noteworthy as it is the only cluster that shows signs of both cell membrane damage and immune response activity.

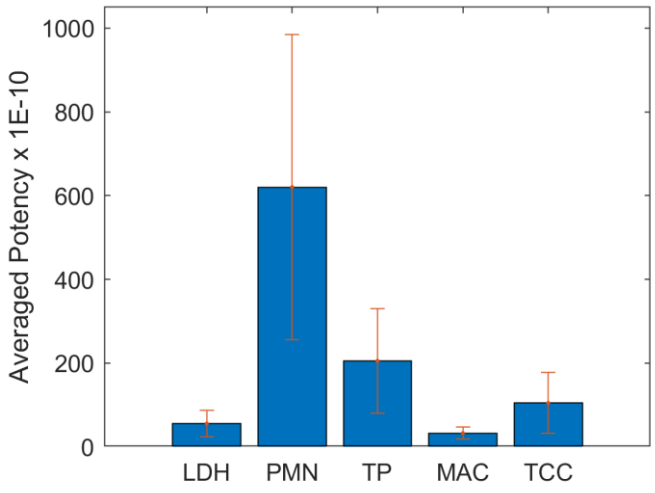

Figure 2-4. The cluster averaged potency for each response. The values are indicative of the relative increase in measured response for the same given dosage.

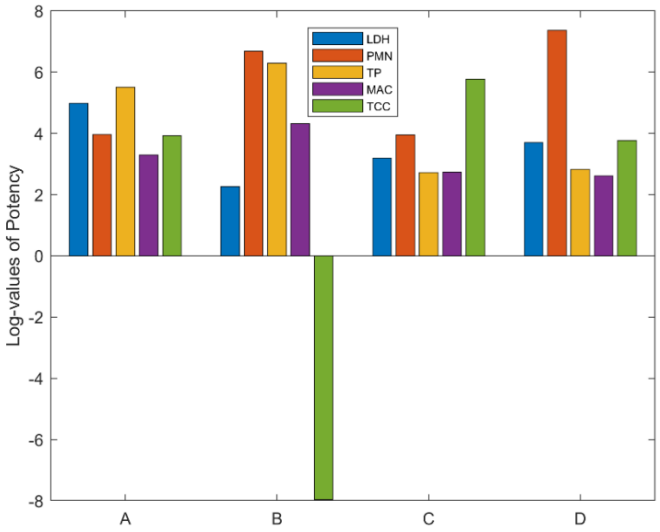

Figure 2-5. The variation between natural logs of the potency of the responses in each cluster provides insight towards the observed effects each cluster of particles has on the response. A single CNT particle constituent of Cluster B possessed low potency as derived from the dose-response-recovery model.

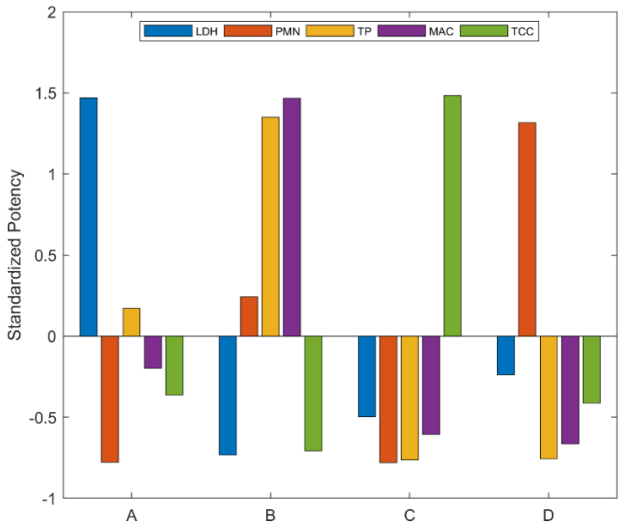

Figure 2-6. The variation of the standardized potency of the responses across the clusters allows for easier interpretation of the observed effects of exposure to the particle clusters. Cluster A shows signs of increased cell damage. Cluster B shows signs of both increased immune response and cell damage while clusters C & D only displayed signs of increased immune response activity.

4. CONCLUSIONS

Previous literature indicated that there are observed effects of particle size and subsequent aggregation of particles on their toxicity. The literature also features differing views from authors on the cytotoxicity of CNTs of varying dimensions. Kang et al (2008) found the smaller single walled (SW) CNTs to be more cytotoxic than multi-walled (MW) CNTs, the SWCNTs studied (2000 nm in length) are similar to the group of particles from cluster "B" suggested by the algorithm.[20] In contrast, Nygaard et al (2009) has suggested that the observed allergic immune response of MWCNTs is stronger than SWCNTs.[25] There is also evidence to suggest that larger particle sizes and aggregates cause increased levels of toxicity, Muller et al (2005) additionally found that larger aggregates lead to increased PMN response which corresponds to the observed increase in standardized potency of PMN for cluster "D".[23] The size distribution of the particles also plays a role as ground CNTs were found to be more active than unground CNTs.[6]

This study demonstrates the development and application of an algorithm to create a clustering tree for meta-analysis of nanomaterial *in vivo* pulmonary toxicity in rodents. The algorithm divides the nanomaterials based on the similarity of their dose response into clusters with each cluster having its own physical and chemical characteristics. The algorithm used is proficient in creating clusters using the available data and can accommodate small datasets and missing characterization data without total loss of function. The argument for using this algorithm over other available alternatives is that the objective of this process on grouping similar particles takes dose-response and recovery patterns into account rather than treating those parameters as equivalent to particle properties, or the assumption that all necessary physicochemical characteristics are already known. In so doing, it generates hypotheses regarding what mechanisms or characteristics might explain why certain particles cause similar or dissimilar

pulmonary effects. All CNTs included in this study performed an independent characterization of the CNTs utilized, and the details of such characterization procedures are present in the referenced publications in Table 2-1. While, this is not an ideal situation from the perspective of conformity of methods and instruments, this is the largest set of *in vivo* pulmonary CNT toxicity data available to date, and it is large enough to permit a meta-analysis such as that presented in this paper. Given these limitations, this research should be considered hypothesis-generating rather than hypothesis-confirming.

This work and the identified clusters propose that certain characteristics are significant or insignificant to the identification of CNT toxic potency. The study was able to identify and group the CNT particles into 4 toxicologically distinct clusters, the clusters were further categorized into potential classes defined by their physical properties. The clustering methodology employed was validated using a leave-one-out meta-analysis approach. 90% of the validation process iterations resulted in 4 clusters as the best solution and the remaining 10% of iterations had 4 clusters within 1 standard error of the minimum. The validation of the clustering process emphasizes the consistency of the algorithm and the clusters generated. The lack of sufficient characterization data does not prevent the use of this algorithm, although it may limit the definitive description of the distinct clusters after the fact. Nanomaterials behaving in a similar/dissimilar pattern can be identified and grouped together and their differing attributes can be identified and related to their potency.

The analysis of the potency of the CNT clusters using the response variables was able to reveal that certain group (cluster “B”) of CNT particles have the potential to cause increased cell damage and immune response intensity due to their elevated levels for LDH, Total protein and Macrophage content. Increased levels of Total protein and PMN were also observed in subjects exposed to CNT particles from cluster “A”. When correlating the potency of the clusters and particles with the physical characteristics (length and diameter) of the corresponding clusters

using linear regression, the results showed that there is no evidence of a significant relationship between the potency and the predictors. This could imply that the potency of the particle might not depend on the physical properties alone and the chemical characteristics need to be taken into consideration as well, expansion to include more properties to explain the relationship could be meaningful. These results could also be interpreted as sign that a simple relationship between potency and the physical and/or chemical variables is non-existent and it is more complex than anticipated earlier and requires further investigation to be ascertained.

AUTHOR INFORMATION

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript and contributed equally.

Funding Sources

Funding for this work was provided by NIOSH, through grant 1 R03 OH010956-01. This work has also been supported 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).

ACKNOWLEDGMENT

Funding for this work was provided by NIOSH, through grant 1 R03 OH010956-01. This work has also been supported 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).

REFERENCES

1. USEPA, *Economic Analysis of the Proposed Change in Data Requirements Rule for Conventional Pesticides.* US Environmental Protection Agency, 2004.

2. NIOSH, *Occupational Exposure to Carbon Nanotubes and Nanofibers.* National Institute for Occupational Safety and Health, 2013.

3. Pauluhn, J., *Multi-walled carbon nanotubes (Baytubes (R)): Approach for derivation of occupational exposure limit.* Regulatory Toxicology and Pharmacology, 2010. **57**(1): p. 78-89, 10.1016/j.yrtph.2009.12.012.

4. Puzyn, T., et al., *Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles.* Nature Nanotechnology, 2011. **6**(3): p. 175-178, Doi 10.1038/Nnano.2011.10.

5. Rasulev, B., et al., *Nano-QSAR: Advances and Challenges.* Towards Efficient Designing of Safe Nanomaterials: Innovative Merge of Computational Approaches and Experimental Techniques, 2013. **25**: p. 220-256, Doi 10.1039/9781849735476-00220.

6. Donaldson, K., et al., *Carbon nanotubes: A review of their properties in relation to pulmonary toxicology and workplace safety.* Toxicological Sciences, 2006. **92**(1): p. 5-22, 10.1093/toxsci/kfj130.

7. Toropov, A.A. and J. Leszczynski, *A new approach to the characterization of nanomaterials: Predicting Young's modulus by correlation weighting of nanomaterials codes.* Chemical Physics Letters, 2006. **433**(1-3): p. 125-129, 10.1016/j.cplett.2006.11.010.

8. Sayes, C.M., et al., *Comparative pulmonary toxicity assessments of C-60 water suspensions in rats: Few differences in fullerene toxicity in vivo in contrast to in vitro profiles.* Nano Letters, 2007. **7**(8): p. 2399-2406, 10.1021/nl0710710.

9. Fourches, D., D.Q.Y. Pu, and A. Tropsha, *Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles.* Combinatorial Chemistry & High Throughput Screening, 2011. **14**(3): p. 217-225, Doi 10.2174/138620711794728743.

10. Esposito, E.X., et al., *Exploring possible mechanisms of action for the nanotoxicity and protein binding of decorated nanotubes: interpretation of physicochemical properties from optimal QSAR models.* Toxicology and Applied Pharmacology, 2015. **288**(1): p. 52-62, 10.1016/j.taap.2015.07.008.

11. Shao, C.Y., et al., *Dependence of QSAR Models on the Selection of Trial Descriptor Sets: A Demonstration Using Nanotoxicity Endpoints of Decorated Nanotubes.* Journal of Chemical Information and Modeling, 2013. **53**(1): p. 142-158, 10.1021/ci3005308.

12. Gernand, J.M. and E.A. Casman, *Machine Learning for Nanomaterial Toxicity Risk Assessment.* Ieee Intelligent Systems, 2014. **29**(3): p. 84-88.

13. Gernand, J.M. and E.A. Casman, *A Meta-Analysis of Carbon Nanotube Pulmonary Toxicity Studies-How Physical Dimensions and Impurities Affect the Toxicity of Carbon Nanotubes.* Risk Analysis, 2014. **34**(3): p. 583-597, Doi 10.1111/Risa.12109.

14. Genaidy, A., et al., *Health effects of exposure to carbon nanofibers: Systematic review, critical appraisal, meta analysis and research to practice perspectives.* Science of The Total Environment, 2009. **407**(12): p. 3686-3701, http://dx.doi.org/10.1016/j.scitotenv.2008.12.025.

15. Ioannidis, J. and J. Lau, *Pooling research results: benefits and limitations of meta-analysis.* The Joint Commission journal on quality improvement, 1999. **25**(9): p. 462-469.

16. Spector, T.D. and S.G. Thompson, *The potential and limitations of meta-analysis.* Journal of epidemiology and community health, 1991. **45**(2): p. 89.

17. Sayes, C. and I. Ivanov, *Comparative Study of Predictive Computational Models for Nanoparticle-Induced Cytotoxicity.* Risk Analysis, 2010. **30**(11): p. 1723-1734, DOI 10.1111/j.1539-6924.2010.01438.x.

18. Liu, R., et al., *Classification NanoSAR Development for Cytotoxicity of Metal Oxide Nanoparticles.* Small, 2011. **7**(8): p. 1118-1126, DOI 10.1002/smll.201002366.

19. Fourches, D., et al., *Quantitative Nanostructure-Activity Relationship Modeling.* Acs Nano, 2010. **4**(10): p. 5703-5712, Doi 10.1021/Nn1013484.

20. Kang, S., M.S. Mauter, and M. Elimelech, *Physicochemical determinants of multiwalled carbon nanotube bacterial cytotoxicity.* Environmental Science & Technology, 2008. **42**(19): p. 7528-7534, Doi 10.1021/Es8010173.

21. 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.* Toxicological Sciences, 2010. **113**(1): p. 226-242, 10.1093/toxsci/kfp247.

22. Ma-Hock, L., et al., *Inhalation Toxicity of Multiwall Carbon Nanotubes in Rats Exposed for 3 Months.* Toxicological Sciences, 2009. **112**(2): p. 468-481, 10.1093/toxsci/kfp146.

23. Muller, J., et al., *Respiratory toxicity of multi-wall carbon nanotubes.* Toxicology and Applied Pharmacology, 2005. **207**(3): p. 221-231, 10.1016/j.taap.2005.01.008.

24. Shvedova, A.A., et al., *Inhalation vs. aspiration of single-walled carbon nanotubes in C57BL/6 mice: inflammation, fibrosis, oxidative stress, and mutagenesis.* American

Journal of Physiology-Lung Cellular and Molecular Physiology, 2008. **295**(4): p. L552-L565, 10.1152/ajplung.90287.2008.

25. Nygaard, U.C., et al., *Single-Walled and Multi-Walled Carbon Nanotubes Promote Allergic Immune Responses in Mice.* Toxicological Sciences, 2009. **109**(1): p. 113-123, 10.1093/toxsci/kfp057.

26. Warheit, D.B., et al., *Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats.* Toxicological Sciences, 2004. **77**(1): p. 117-125, 10.1093/toxsci/kfg228.

27. Park, E.J., et al., *A single intratracheal instillation of single-walled carbon nanotubes induced early lung fibrosis and subchronic tissue damage in mice.* Archives of Toxicology, 2011. **85**(9): p. 1121-1131, 10.1007/s00204-011-0655-8.

28. Teeguarden, J.G., et al., *Comparative Proteomics and Pulmonary Toxicity of Instilled Single-Walled Carbon Nanotubes, Crocidolite Asbestos, and Ultrafine Carbon Black in Mice.* Toxicological Sciences, 2011. **120**(1): p. 123-135, 10.1093/toxsci/kfq363.

29. Elgrabli, D., et al., *Induction of apoptosis and absence of inflammation in rat lung after intratracheal instillation of multiwalled carbon nanotubes.* Toxicology, 2008. **253**(1-3): p. 131-136, 10.1016/j.tox.2008.09.004.

30. Mercer, R.R., et al., *Alteration of deposition pattern and pulmonary response as a result of improved dispersion of aspirated single-walled carbon nanotubes in a mouse model.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2008. **294**(1): p. L87-L97, 10.1152/ajplung.00186.2007.

31. Porter, D.W., et al., *Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes.* Toxicology, 2010. **269**(2-3): p. 136-147, 10.1016/j.tox.2009.10.017.

32. Ellinger-Ziegelbauer, H. and J. Pauluhn, *Pulmonary toxicity of multi-walled carbon nanotubes (Baytubes (R)) relative to alpha-quartz following a single 6 h inhalation exposure of rats and a 3 months post-exposure period.* Toxicology, 2009. **266**(1-3): p. 16-29, 10.1016/j.tox.2009.10.007.

33. Shvedova, A.A., et al., *Unusual inflammatory and fibrogenic pulmonary responses to single-walled carbon nanotubes in mice.* Am J Physiol Lung Cell Mol Physiol, 2005. **289**(5): p. L698-708, 10.1152/ajplung.00084.2005.

34. Ge, C.C., et al., *Acute pulmonary and moderate cardiovascular responses of spontaneously hypertensive rats after exposure to single-wall carbon nanotubes.* Nanotoxicology, 2012. **6**(5): p. 526-542, 10.3109/17435390.2011.587905.

CHAPTER 3

Examining the In Vivo Pulmonary Toxicity of Engineered Metal Oxide Nanomaterials using a Genetic Algorithm-based Dose-Response-Recovery Clustering Model

The contents of this chapter have been published in "Computational Toxicology" under the title *"*Examining the *In Vivo* Pulmonary Toxicity of Engineered Metal Oxide Nanomaterials Using a Genetic Algorithm-Based Dose-Response-Recovery Clustering Model*"* *(https://doi.org/10.1016/j.comtox.2019.100113).*

ABSTRACT

This article presents the examination of the pulmonary toxicity of engineered metal oxide nanoparticles using a dose-response-recovery clustering model for the purpose of revealing toxicologically distinct clusters. Current recommended exposure limits published by the National Institute for Occupational Safety and Health (NIOSH) consider "ultrafine" metal oxide particles to pose significantly increased health risks as compared to larger particles. The unique combination of physical and chemical characteristics afforded by the metal oxide nanoparticles has enabled them to find use in various large-scale industrial setting leading to a risk of exposure for current and future workers. This paper presents an algorithmic examination of the metal oxide nanoparticles by their categorization into toxicologically distinct clusters. Based on a dataset composed of peer-reviewed *in vivo* experimental studies in rodents, the metal oxide nanoparticle variants are divided into sub-classes based on their dose-response-recovery similarity and the

Akaike Information Criterion (AIC) of the family of models. Results indicate the presence of 4 toxicologically unique classes based on 5 toxicity endpoints selected. The cluster with greatest potency was found to be 400000 times more potent than the cluster with the lowest potency indicative of substantial variation across all the responses. The absence of coherent characterization data for the metal oxide nanoparticle variants analyzed in this study prevents the designation of significant physical characteristic-based labels which could have assisted in identifying the key factors affecting the toxic potential of the metal oxide nanoparticles. The standardized potency of the 4 metal oxide nanoparticle clusters was compared: 3 clusters (I, II and IV) showed signs of elevated immune system activity and cell membrane damage. These clusters were primarily composed of iron oxide nanoparticles (cluster I) silica (cluster II), and titanium dioxide (cluster IV). One cluster (III) showed comparatively reduced toxicity across all 5 responses; this cluster primarily featured zinc oxide and cerium oxide nanoparticles.

Abbreviations

MONP, Metal oxide nanoparticles; BAL, bronchoalveolar lavage; MAC, macrophage count in BAL fluid; TP, total protein concentration in BAL fluid; LDH, lactose dehydrogenase; PMN, polymorphonuclear neutrophil; TC, total cell count.

Keywords

Dose-Response, Recovery, Pulmonary Toxicity, Metal oxides, Hierarchical Clustering, Potency

1. INTRODUCTION

Toxicological testing necessary for the establishment of safe exposure levels to emerging pollutants can be time consuming and resource intensive [1]. When the emerging pollutant is an

engineered nanoparticle, whose chemical and physical nature can be changed at the will of the designer, the pace of development for new nanoparticle variants is too rapid for such safety testing to keep up. Some of these NMs might eventually prove to be a serious occupational hazard or become an environmental pollutant presenting risk to the people working with them and to those who use the technology or live near places they are used. Anticipating the toxicological impact and exposure response to newly synthesized nanoparticles is difficult. Metal oxide nanoparticles (MONP's) are particularly popular in industrial environments due to the diverse array of benefits offered by virtue of their small size. For example, ultrafine (<100nm) metal oxides such as aluminum oxide and cerium oxide are employed as fuel additives to lower emissions and promote more efficient combustion.[2] Similarly, nano-titanium dioxide is a popular component of multiple food, personal care, and consumer products.[3] Exposure to ultrafine particles induces both acute and chronic inflammatory response, nano-ceria has been known to disperse with diesel exhaust leading to potential pulmonary fibrosis.[4, 5] Further evidence of pulmonary response exposure to ultrafine particles has been documented for various metal oxide particles.[6-9] NIOSH has set recommended exposure limits (REL) of 2400 $\mu g/m^3$ and 300 $\mu g/m^3$ for fine-sized (respirable) and ultrafine-sized titanium dioxide. A REL for respirable crystalline silica has also recently been revised 50 $\mu g/m^3$, but there is no current recommendation for ultrafine silica. It is yet unclear whether new RELs (and eventually associated regulatory limits based on the recommendations) will be set at similarly reduced levels for most other metal oxide nanoparticles or not. These limits are derived from the extrapolation of the effects of pulmonary exposure in Rats to Humans.[10, 11]

Experimental toxicology testing studies are often time consuming and require considerable investment with collective costs potentially running up to billions of dollars.[12] A study published in 2011 about the long term development of NMs projected that up to 6 million individuals could potentially be exposed by 2020.[13] One potential solution to this problem is

prioritizing the testing the NMs targeting the more toxic varieties first. The development of computational approaches targeted at predicting the toxicity of the NMs based on their attributes can be used as a means to guide nanomaterial synthesis towards less hazardous variants. Experimental evaluation is subject to a certain degree of error or bias and in some cases the results may be difficult to replicate. To overcome some of these inefficiencies, forecasting methods have been proposed to predict the potential toxicity and environmental impact of new NMs.[14] Computational modeling of nanotoxicity is valuable for its efficiency of time and expense and the flexibility offered. Quantitative structure activity relationships (QSARs) are one of the modeling schemes currently utilized; these models hypothesize that the activity of any nanostructure is a function of its physicochemical properties.[15, 16] Puzyn et al. demonstrated the use of QSARs to predict the mechanism of cytotoxicity of 17 different metal oxide nanoparticles.[14] QSAR modeling and novel descriptors have previously used in conjunction to predict the cytotoxicity of metal oxide nanoparticles using E. coli exposure studies.[17] The cell viability of human lung and skin cells upon exposure to metal oxide nanomaterials has also been predicted using QSAR based modeling.[18]

Meta-analysis is a useful analytical method for harnessing the available published nanotoxicology data, and testing hypotheses that were not possible within the original experiments. Meta-analyses employ statistical and/or machine learning approaches such as regression, neural networks, decision trees, and support vector machines to identify important associations.[19] Since meta-analysis curates data from multiple sources, the individual variation or appearance of unusual results across the complete dataset are evaluated in light of the whole. Traditional meta-analysis, which involves literature-based data collection, can in some cases be ineffective in this case due to the presence of missing data caused by disparity in NM characterization methods.[20] A meta-analysis published in 2013 on the health effects of exposure to nano-TiO_2 concluded that the particles were retained in several key organs in the

body.[21] Meta-analysis has also been used to assess the toxicity of NMs when they are dissolved—a possibility with high-solubility metal oxide nanoparticles.[22] The prediction of carcinogenic potency using historical *in vivo* genotoxicity data has also been accomplished using meta-analysis.[23] The benefits of meta-analysis techniques include its ability to identify the sources of diversity across the various participating studies[24], it can be used to improve the precision of the estimates due to the increase in the size of the population.[25] The method however is not without its limitations: it has no control over the homogeneity of the participating studies[24]; it cannot improve the quality of the data used and consequently the final result; and it could suffer from "publication bias", where researchers fail to report or exclude certain studies to present a more coherent result which could impact the analysis.[25]

Multiple researchers have documented the adverse effects associated with exposure to metal oxide nanoparticles. Peng et al (2014) compared the pulmonary toxicity of 2 types of nano-ceria and concluded that smaller agglomerates were more potent due to higher lung deposition.[5] Kadoya et al (2011) compared micro-sized and sub-micron sized nickel oxide and deduced that sub-micron sized particles induce stronger effects in the BAL fluid.[26] Titanium dioxide and silica have been extensively documented for their toxicological response with multiple researchers concluding that large doses of ultrafine Titania and silica induces immune response in subjects.[6, 7, 9, 27-36] Similarly, articles documenting the effects of exposure to zinc, iron and other metal oxides have described the negatives effects associated with their exposure.[8, 37-42] Attempts at using computational modeling to predict the effects on toxicity of changes in particle properties (QSARS, meta-analysis etc.) have been limited in their scope to date. These methods have either been exclusive to only metal oxide particles, particles of a particular size, cytotoxicity assays, or structural analysis of single particles.[14, 43-47] There is a need to quickly identify and prioritize the testing of NMs which pose greater risk than others to maximize their benefit while simultaneously minimizing the risk associated with them. This could be accomplished by

quantifying the current knowledge base of *in vivo* and *in vitro* studies to reveal how different NMs compare to one another and how changes in their properties affect their toxicity. The rapid increase in the number of synthesized NMs variants reveals a need to cluster the variants displaying similar effects as a means of generating hypotheses about the mechanisms of their pulmonary toxicity. To address the question of how different metal oxide nanomaterials can be considered the same substance toxicologically, this study proposes a new algorithmic approach applied to meta-analysis of extant metal oxide nanoparticle *in vivo* toxicity data.

2. METHODS

2.1 Data Selection

The data utilized for the meta-analysis of the *in vivo* pulmonary toxicity of metal oxide nanoparticles is curated from 30 different peer-reviewed journal articles. These publications were identified through online search engines including the Web of Science, Google Scholar, and PubMed using the search terms: "*in vivo*", "Metal oxide", "BAL" and "Pulmonary". The publications are all independent "*in vivo*" studies conducted on metal oxide nanoparticles, they are a collection of exposure groups (experiments) identified by their total dosage, method of exposure (inhalation, instillation, aspiration etc.), post-exposure recovery period, and an independent characterization profile (physical and chemical properties) of the particles. The data extracted from the publications identified and included as part of the dataset additionally satisfied a key criterion before inclusion: the measures of response in Bronchoalveolar lavage (BAL) fluid for the rodents must be recorded for a minimum of 3 different levels (including the control level) of dose or exposure. All publications used as the source for this analysis have been provided in Table 3-1.

The five quantitative response measures from the BAL fluid which are the focus of this meta-analysis include: Total cell count (TCC), Macrophage count (MAC), Total protein concentration (TP), Polymorphonuclear neutrophil count (PMN) and Lactate dehydrogenase concentration (LDH). These 5 measures were reported in sufficient capacity across the publications to be included in this analysis. The measured response reported in each publication is expressed as a fold (multiple) of its respective control group's measure.

2.2 Model Definition

The dose-response model used for the clustering analysis is an important feature of the process; it reflects the variation of the response to changes in the applied dose. Benchmarked dose-response models have been previously developed to analyze the response to multiple engineered substances. Slob (2001) has previously published a family of 3-parameter models available for analyzing continuous data, but these models do not reflect the variation in the measured response when exposure has ceased, and the test subject has begun to recover. This paper proposes a revision to one of Slob's model, the addition of a fourth parameter that describes the decrease of the response measure over time (i.e., the animal recovers) to provide a more complete description of the dose-response-recovery relationship. The relation between the response measured (y), dosage applied (x) and post-exposure recovery period (t) is,

$$y = a[c - (c-1)e^{-bx}] - dt \qquad \text{Eq. 1}$$

The 4 parameters, *a* (signifies the response at dose = 0), *b* (the toxic potency of the nanoparticles), *c* (the maximum relative shift in response), and *d* (slope of the response decay) are the key parameters which have been used to describe the potential of the particle to be a hazard.

Recovery is likely to have an exponential relationship as the animal returns to homeostasis following the end of exposure. However, since such a relationship would require the addition of two additional model parameters, rather than one for the linear term, and given the

available data fit the linear representation of recovery better, we make the parsimonious choice here with the single caveat that the model is valid for all recovery times, t, until the value of the response returns to 1 (i.e., no change from the control group).

2.3 Clustering Algorithm

The proposed algorithm is aimed at assigning the metal oxide nanoparticles into "clusters" based on the similarity of their dose-response-recovery relationships. The methodology is a derivation of the traditional top-down hierarchical clustering process and mimics its structure. The algorithm utilizes non-linear regression to fit the proposed dose-response-recovery model (Eq. 1) to the available data. The process involves using an exhaustive search approach driven by combinatorics to identify the ideal number of cluster constituents at each stage. A detailed explanation of the entire clustering process has been covered in a previous article by the authors and can be referenced for further clarity on the process.[48]

The clustering approach yields multiple possibilities in terms of the total number of clusters and the membership within those clusters which are assessed using a performance metric to avoid overfitted solutions. The performance metric used here to evaluate the goodness of a cluster is the Akaike Information Criterion (AIC), the AIC is calculated based on the number of parameters (k) in a model and its log-likelihood (L),

$$AIC = 2k - 2\log L \qquad \text{Eq. 2}$$

The likelihood function of a model is defined as the probability of that model achieving an outcome given a set of parameters, maximizing the likelihood function reduces the error between the expected and observed data. We assume the likelihood function of the model follows a normal distribution and its logarithmic form represented mathematically as,

$$logL(\mu, \sigma) = -n\log(\sigma) - \frac{n}{2}log(2\pi) - \frac{1}{2\sigma^2}\sum_1^n (Y_i - \mu)^2 \qquad \text{Eq. 3}$$

Y_i, represents the model derived response data, μ and σ represent the mean of the experimentally measured response and its associated standard deviation, n represents the number of exposure groups present in the model.

The overall AIC of a cluster is obtained by summing together the AIC values of its component models. Lower values of AIC are preferred when comparing the clusters as they indicate the model is closer to the data in those cases. AIC needs to be modified to include corrections that can be applied when dealing with small datasets, this corrected form is denoted as the AICc.

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \qquad \text{Eq. 4}$$

Where n is the number of observations and k is the number of model parameters. The AICc is the best metric for comparing the amount of information conveyed by distinct models of varying complexity or as in our case comparing multiple cluster configurations.[49] We adopt the AICc as our measure of performance since it issues a strong penalty to those clusters where there is evidence of overfitting.

2.4 Genetic Algorithm

An exhaustive clustering algorithm would be inefficient when deployed on large datasets. The current size and presumably the future increasing availability of data for the metal oxide nanoparticle response measures would result in high computational time requirements. Figure 3-1 shows the steep increase in the total number of cluster configurations analyzed as the number of particles expands.

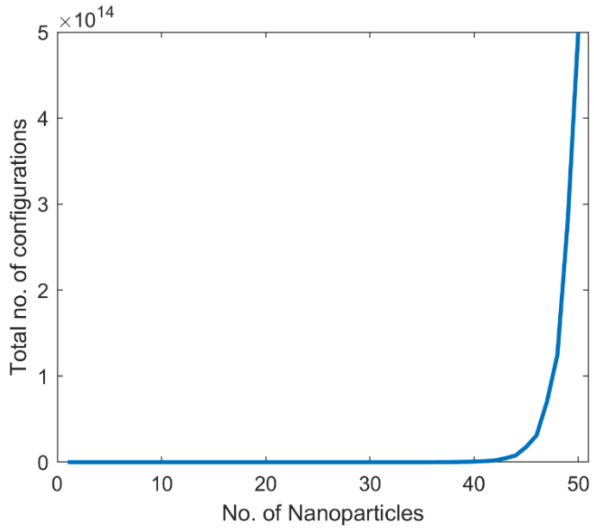

Figure 3-1. Variation between the "Total no. of configurations" analyzed vs "No. of Nanoparticles". The plot demonstrates the increasing computational load as more particles are added to the analysis.

The search space for the metal oxide dataset has a large number of possible configurations (order of magnitude: 10^{10}), given the variety of unique nanoparticles included, employing an exhaustive search algorithm on this dataset yielded unrealistic completion times. An alternative approach to search the available space to identify the solution is to use Genetic algorithms (GA). Genetic algorithms use an adaptive and heuristic approach towards determining the solution to a search problem, the underlying principle to their methodology is the process of natural selection which is the driving factor behind biological evolution.

The original clustering algorithm has been incorporated within a genetic algorithm framework which facilities identification of the ideal cluster constituents from the large search space. Each stage of the clustering process is subjected through a genetic algorithm to identify the ideal dissociation of the data into clusters. The AICc score of the clusters are used to compare and select the best cluster combination for the complete data. The objective function minimized by the algorithm can be mathematically expressed as,

$$Y = \sum_{i=1}^{n} AIC_i \quad \text{Eq. 5}$$

Where, Y represents the overall AIC score of the clusters used to represent the data, the individual cluster AIC is computed using Eq. 2 referenced earlier, n is the number of clusters at each stage of the evaluation.

For example, if the dataset were to contain 35 particle variants, the next step in the process would determine using combinatorics the multiple possible ways to separate the 35 particle variants into 2 clusters, but each of those has multiple scenarios. Assume for illustration a cluster of 1 particle and a second containing 34 particles expressed as (1 34) can be accomplished in 35 ($^{35}C_1$) unique ways, while a (2 33) split has 595 ($^{35}C_2$) unique groupings and so on. These are just a few of the "configurations" available when deciding which is the ideal grouping possible, an exhaustive search through all the possibilities is time consuming.

The GA was implemented using MATLAB R2017b[50]. The algorithm uses a "Bit String" population input i.e., the "genes" in the algorithm correspond to various particle variants encoded as 0 or 1, and the "chromosomes" that compose the population are all the possible solutions to the optimization problem. The mutation and crossover functions for the GA are set to be "Uniform" with a 1% rate of mutation (uniform and random gene replacement throughout the progression) and "Scattered" (randomized crossover points to generate new chromosomes) respectively. The new "generations" (population of every iteration) are synthesized from the parent chromosomes using an 80% rate of crossover. The stopping criteria for the algorithm is set at 100 "Stall Generations" (i.e., when the best solution does not change for 100 generations as computed by the algorithm).

One limitation of the GA is the possibility of identifying a local solution to the AICc minimization problem rather than the global solution. There are testing methods to significantly reduce the possibility of this result. Multiple iterations of the overall GA process could be conducted, and a statistical distribution could be generated of the derived results, the alternative

would be to compare the results of the GA with the results of a brute force algorithmic search. The results from the brute force algorithm would align with the global solution due to the rigorous nature of the search mechanism.

An alternative to the GA is another stochastic search method known as Simulated Annealing. Simulated annealing is an approach to solving optimization problems such as this one: the identification of similar toxicological groups with the minimum AIC. However, given the specifics of this problem including the very large sample space and the need to employ parallel processing, simulated annealing was not anticipated to provide any meaningful benefits in terms of processing time or the fitness of the identified solutions. Several other studies have come to similar conclusions under similar circumstances[18, 49, 51-53].

Table 3-1. Compilation of the data (with sources) used for the MONP clustering process. (NR designation indicates "Not Reported").

SourceID	Primary Author	MONP Types Tested	No. of endpoints (of interest) reported	Total No. of exposure groups	MONP Typical Size (nm)	Average Specific Surface Area (m^2/g)
T1	Nemmar[49]	Titanium	2	3	4 - 6	49.4
T2	Oberdorster[29]	Titanium	4	7	20 - 250	NR
T3, S1	Warheit[50]	Titanium, Silica	4	36	5.8 – 6.1 (T3), 1500 (S1)	51.5
T4	Renwick[51]	Titanium	4	6	29 - 250	28.2
T5, S2	Rehn[52]	Titanium, Silica	4	30	20 (T5), 900 (S2)	NR
T6	Grassian[32]	Titanium	3	16	3.5 – 17.8	130
T7, S3	Warheit[53]	Titanium, Silica	4	88	25 (T7), 200 – 2000 (S3)	23.6
T8, S4	Warheit[35]	Titanium, Silica	3	20	100 (T8), 534 (S4)	26.4
T9, S5	Kobayashi[54]	Titanium, Silica	3	41	4.9 – 154.2	233.6
T10, T13	Gustafsson[6]	Titanium	3	16	21	NR
T11	Oyabu[28]	Titanium	3	20	14	102
T12	Roberts[55]	Titanium	3	12	21.75	NR
T14	Silva[9]	Titanium	1	14	24 - 28	113
S6	Gosens[56]	Silica	4	4	NR	NR
S7	Cho[57]	Silica	4	16	14	NR
S8	Creutzenberg[31]	Silica	5	6	76	21.2
S9	Roursgaard[34]	Silica	3	21	100 - 1600	20.7
N7	Morimoto[58]	Nickel	2	5	20	NR
F1	Ban[37]	Iron	2	11	35 - 147	22.5
F2	Pirela[59]	Iron	4	4	19.6	41.5
F3	Zhu[40]	Iron	4	13	22 - 280	28.8

F4	Katsnelson[39]	Iron	3	4	10 - 1000	NR
S10, Z1	Sayes[30]	Silica, Zinc	3	36	90 - 452	26.1
Ce1	Toya[60]	Cerium	4	17	200 - 3900	NR
N1	Morimoto[61]	Nickel	1	15	20	104.6
Z2, F5	Xia[42]	Zinc, Iron	2	15	8 – 20.2	97.5
S11, Z3	Warheit[36]	Silica, Zinc	3	28	90 – 111 (Z3)	10.9
T15, S12, Ce2, Z4, N2, Cu1	Cho[7]	Titanium, Silica, Cerium, Zinc, Nickel, Copper	3	24	10 - 35	93.7
Ce3	Ma[62]	Cerium	3	18	20	NR
Ce4	Peng[5]	Cerium	4	12	4	NR
T16, S13	Roursgaard[27]	Titanium, Silica	3	28	8 - 1600	112
T17, S14, A1	Lindenschmidt[33]	Titanium, Silica, Aluminum	5	45	2200 - 5300	NR
Ce5	Park[63]	Cerium	3	7	NR	NR
Ce6	Wingard[64]	Cerium	3	4	8	44
Ce7	Xue[65]	Cerium	3	15	6.6	86.8
Ce8	Ma[4]	Cerium	2	6	10.14	90
Ce9	Minarchick[66]	Cerium	2	4	4	81.4
T18, Co1	Dick[67]	Titanium, Cobalt	3	6	20	43.4
S15, M1	Gelli[41]	Silica, Magnesium	1	15	50 - 63000	22.5 (M1)
T19, S16, Co2	Zhang[68]	Titanium, Silica, Cobalt	5	25	20 - 1850	45.6
T20, N3	Horie[69]	Titanium, Nickel	2	16	7 - 1000	NR
T21, S17, N4	Kadoya[26]	Titanium, Silica, Nickel	1	25	27 - 2700	16.5
T22, S18, N5	Ogami[70]	Titanium, Silica, Nickel	2	25	27 - 2700	16.5
N6, Z5, Cu2	Cho[38]	Nickel, Zinc, Copper	2	8	10 - 50	56.3
Z6	Cho[8]	Zinc	4	9	10.7	48.2
N8	Nishi[71]	Nickel	2	15	20	NR

3. RESULTS AND DISCUSSION

3.1 Composition-Blind Clustering Results

The clustering analysis was conducted on the Metal Oxide Nanoparticles (MONP) dataset. The dataset for the MONP's contains several studies where more than 1 unique set of nanoparticles is analyzed. These studies were identified and separated to ensure each study reflects a single particle i.e., if a study documents the results from testing multiple distinct nanoparticles, the study is subdivided to reflect the number of particles tested. For example, the third row of Table 3-1 refers to a single publication that included data for two nanoparticle variants, one silica nanoparticle, and one titanium dioxide nanoparticle. The dataset was analyzed for the 5 toxicological endpoints discussed earlier and their ideal cluster combinations were derived, each cluster is defined by a distinct dose-response-recovery model. The algorithm aligned the pulmonary effects of different MONP variants based on the similarity of their dose-response-recovery relationships. Figure 3-2 is a comparison of the variation in AICc values for all 5 response indicators used in the study, it also serves to determine the ideal number of clusters for each indicator. The decrease in the AICc of the clusters at each stage observed in Figure 3-2 is an indicator that the clustering process works (information retained increases as the number of clusters increases).

The decision on the ideal number of clusters that can be used to represent the MONP dataset is based on the trade-off between a penalty for increasing the number of toxicologically unique clusters (i.e., increasing model complexity) and the benefits of new information gleaned from them (i.e., reduced model error). Analyzing the variation in AICc documented in Figure 3-2, a single consensus cluster size could not be agreed upon across the 5 responses for the MONPs due to the varying size of the number of particles analyzed for each response.

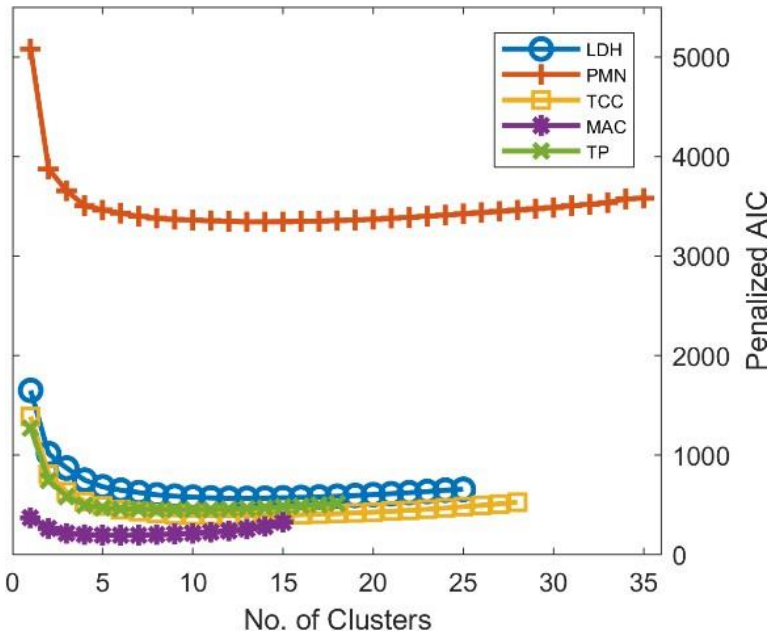

Figure 3-2. Penalized Akaike Information Criterion (AICc) as a function of the number of clusters. The figure illustrates the variation in the number of particles analyzed for each response and hence the lack of a consensus optimum number across the 5 responses

The clusters generated by the algorithm exist on the basis of similarity between their toxicological profiles but that does not fully explain the existence of the clusters. To further investigate the clusters, we examine and study the particle variants within each of the clusters. Contrasting the physical and chemical characteristics of the different particle configurations could lead to the generation of hypotheses as to why one cluster or class of nanoparticles may be more or less toxic than another. The reasoning the MONP dataset constituents are not segregated initially based on their physical/chemical properties is that physically and/or chemically similar nanomaterials need not necessarily be similar in their dose-response-recovery relationships as well, depending on which characteristics are used to define similarity. Identifying the particle properties that do and do not vary significantly between the clusters assists in determining which

properties are key contributors in the separation at each stage and also towards the relative toxicity between the clusters. Table 3-1 is an account of all the publications that have been used as data sources for the MONP. Table 3-1 also provides a comprehensive look at the information available to the authors during the process, namely the types of particles tested by each publication, the number of endpoints measured, and the physical characteristics reported. We can observe from Table 3-1, that not all the source publications report all 5 endpoints and some sources are also missing or lacking in sufficient physical characterization data. The disparity in endpoint measures reported by the source publications is one cause of cluster membership that is not consistent across the 5 endpoints.

The MONP groupings generated by the clustering process were examined to identify and highlight any trends between the clusters and their particle properties. The optimal number of clusters for the PMN, LDH and TCC responses was found to be 8 based on the AICc values, the MAC and TP endpoints had an optimal cluster quantity of 4. The lack of a consensus amongst the 5 endpoints analyzed could be associated with the number of particle variants analyzed for each endpoint. It can be observed from Figure 3-2 that the 3 endpoints with a higher minimum cluster size also had larger number of particle variants with reported measurements from the BAL fluid. Table 3-2 displays the particle properties and membership composition of the MONP clusters for the PMN response. The “Particle Typical Size” measurement is the mean diameter for the spherical MONPs that were analyzed. There is no obvious trend between the primary particle size and the cluster composition for this particular response independent of particle composition. Analysis of the 4 other responses for the MONPs revealed the same. This could imply that creating category labels for the MONP clusters purely using either the physical properties and/or chemical composition might not be possible with the available data and a different approach might be required.

Table 3-2. Mean "typical particle size" for MONP clusters along with constituent cluster composition for PMN response. The composition of the clusters indicates that multiple metal oxide particles can be categorized together by virtue of their dose- response-recovery relationships

Cluster	Mean Particle size (nm)	Cluster Composition
1	1228.4	Crystalline Silica, Amorphous Silica
2	3133.3	Titanium, Aluminum, Copper
3	14.6	Titanium
4	232.8	TiO_2, NiO, Crystalline Silica, ZnO, CeO_2
5	107.7	TiO_2, Fe_2O_3
6	164.4	TiO_2, Amorphous Silica, ZnO
7	13.9	CeO_2, Crystalline Silica
8	154	TiO_2, Crystalline Silica, ZnO, CeO_2

3.2 Results of Clustering by Particle Composition

The initial clustering process treated each of the individual MONP variants analyzed from each study as a unique substance. This leads to instances of similar substances appearing in multiple clusters as evidenced in Table 3-2. For example, crystalline silica nanoparticle variants appear in 4 of the 8 clusters. The information provided in Table 3-2 for the MONP clusters does not seem to provide evidence of any discernable pattern between the particle properties (such as size, or aggregation, or solubility, or bond strength) and membership for the respective clusters. The treatment of the particle variants from each study as unique entities results in a skewed distribution across the clusters, for example Titanium Dioxide (TiO_2) was the most commonly tested particle and appears in 6 out of the 8 clusters as seen in Table 3-2. To eliminate the

multiple instances of the same substance occurring across the clusters in a secondary analysis, the particle variants were grouped in the dataset based on their chemical composition and the clustering process was repeated. The repeat of the clustering process post composition-grouping allows for different distinctions to be made between clusters.

Figure 3-3 is a comparison of the AICc variation between chemically isolated particles across the 5 response indicators. The composition grouping process reduced the number of particle groups to between 7 and 9 per response. We can observe from the figure that the ideal number of clusters across the 5 responses is 4, implying that there are 4 toxicologically unique clusters amongst the different metal oxide particles. The choice of 4 clusters as the ideal set is based on weighing the value of additional information conveyed against the increasing complexity (> 5 clusters), the 4-cluster model provide the best balance between information conveyed and overall complexity. The clustering methodology employed was validated using a leave-one-out meta-analysis approach. 95% of the validation process iterations resulted in 4 clusters as the best solution and the remaining 5% of iterations had 4 clusters within 1 standard error of the minimum. The validation of the clustering process emphasizes the consistency of the algorithm and the clusters generated. Tables 3-4 and 3-5 display the newly clustered metal oxide particle variants (that have been grouped chemically) and the associated potency of the respective clusters. It can be observed that there is a level of consistency to clustering process in Table 3-4 that was not visible before in Table 3-2. Table 3-3 illustrates the degree of consistence of the specific observations (dose-recovery groups with the same particle variant) amongst the 5 endpoints. We can see that most particle types were tested for more than one endpoint with the exception of those evaluated for changes in macrophage count.

As mentioned in section 2.2 and Eq. 1, the recovery of an animal following the end of toxic exposure is expected to follow an exponential relationship as the subject returns to homeostasis. However, limitations in the available data set did not support the differentiation between this

expected relationship and a linear model. In Figure 3-4, one can observe the comparison between the linear and exponential models in the AIC results for each toxic endpoint modeled. These results are very similar with AIC values for the linear recovery model being either nearly equivalent or up to 5% lower than those of the exponential model. Given these results there is no positive statistical justification to use the more complex exponential model at this time, though future data collection may change this assessment.

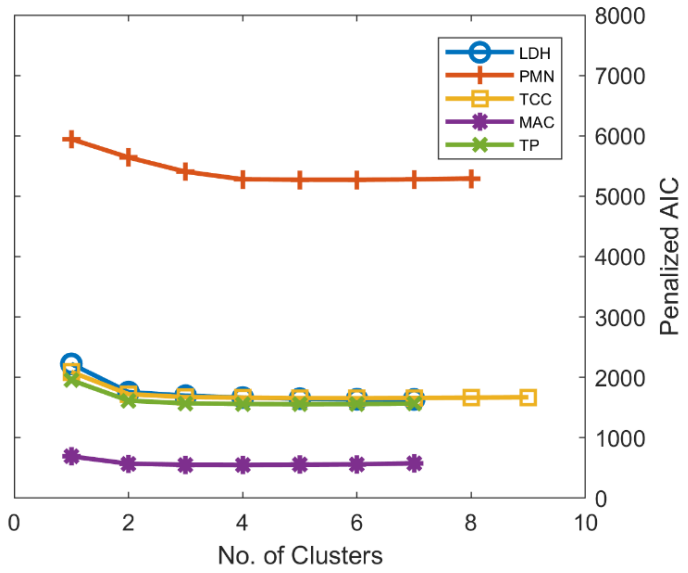

Figure 3-3. Corrected Akaike Information Criterion (AICc) as a function of the number of clusters in the mode. The figure illustrates the optimum number of clusters is 4 based on the minimum achieved for all endpoints indicating there are 4 distinct groups of MONPs.

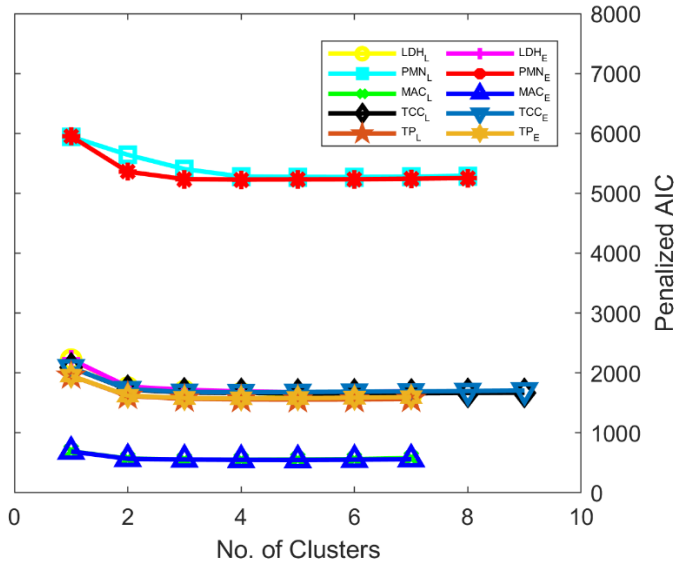

Figure 3-4. Corrected Akaike Information Criterion (AICc) comparison between models with a linear recovery (L) versus models with an exponential recovery (E). The lack of any discernable difference between the AICc values is evidence to back our assumption of a linear recovery.

Table 3-3. Fractional overlap of exposure group data between the measured endpoints from the BAL fluid of the animals.

Measured Endpoint	TCC	PMN	LDH	TP	MAC
TCC	1	0.70	0.61	0.58	0.37
PMN		1	0.73	0.55	0.29
LDH			1	0.59	0.25
TP				1	0.42
MAC					1

Table 3-4. Clusters generated across the 5 endpoints for the various metal oxides (data sources are indicated in parentheses). The bolded elements illustrate particle composition "anchors" across the clustering results.

Cluster ID	LDH Clusters	TCC Clusters	MAC Clusters	PMN Clusters	TP Clusters
I	**Iron** *(F2, F3, F5)*	**Iron** *(F1, F2)*, Zinc *(Z1, Z4-Z6)*, Cerium *(Ce1, Ce2, Ce5, Ce6)*	Cerium *(Ce1, Ce3, Ce5, Ce6)*	**Iron** *(F2, F3, F5)*	**Iron** *(F3)*
II	**Silica** *(S1, S3-S6, S8, S10, S11, S14-S16)*	**Silica** *(S1- S3, S5- S8, S10, S12, S14, S16, S18)*, Copper *(Cu1, Cu2)*, Nickel *(N2, N5-N8)*	**Silica** *(S2, S6-S9, S13, S14, S16)*, Nickel *(N1)*	**Silica** *(S1, S3-S5, S7, S10-S12, S14, S16)*, Copper *(Cu1, Cu2)*, Aluminum *(A1)*	**Silica** *(S1, S3, S4, S7-S9, S11, S13, S14, S16, S17)*, Cerium *(Ce1, Ce2, Ce4)*
III	Zinc *(Z1-Z3, Z6)*, Cerium *(Ce1, Ce3, Ce4, Ce7-Ce9)*, Aluminum *(A1)*	Aluminum *(A1)*, Cobalt *(Co1, Co2)*	Cobalt *(Co1, Co2)*	Zinc *(Z1–Z6)*, Cerium *(Ce2, Ce3, Ce6, Ce8)*, Nickel *(N2, N6)*	Zinc *(Z3, Z4, Z6)*, Nickel *(N2-N4)*
IV	**Titanium** *(T3, T6–T9, T11, T12, T17, T19, T20)*, Magnesium *(M1)*	**Titanium** *(T2–T7, T9-T11, T13, T15, T17–T19, T22)*	**Titanium** *(T1, T2, T5, T10, T12, T13, T16–T19)*, Iron *(F2, F4)*, Aluminum *(A1)*	**Titanium** *(T1-T5, T7-T12, T14, T15, T17)*	**Titanium** *(T3–T8, T16, T17, T19–T21)*, Aluminum *(A1)*

Table 3-5. Potency values for response across cluster categories (b x 10^{-14}). The disparity between the highest and lowest values for potency can be observed. The highest value is more than 400,000 times larger than the lowest across all the responses which further stresses the need for sub-groups.

Cluster ID	LDH	PMN	Total Protein	Macrophages	Total Cell Count
I	105 ± 9.7E+05	40.5 ± 47.2	1060 ± 383	313 ± 120	1960 ± 666
II	2.36E+06 ± 9.7E+05	102 ± 47.2	553 ± 383	30.4 ± 120	464 ± 666
III	6.7E+05 ± 9.7E+05	5.18 ± 47.2	166 ± 383	85 ± 120	611 ± 666
IV	1.1E+06 ± 9.7E+05	103 ± 47.2	867 ± 383	146 ± 120	825 ± 666

3.2.1 Observed Differences in Toxicity

The potency values for the 5 responses reported in Table 3-5 displays the quantification of the relative differences in toxicity between the clusters. Higher values of potency reflect greater toxic potential for the set of nanoparticle variants in the cluster. Based on the values in Table 3-5 we can observe that there is a substantial difference between the potency of the particle variants for LDH. The variation between the potency of the particle variants and associated clusters supports the need for categories amongst metal oxide particles, and a significant limitation of treating them as the same substance in a risk assessment or regulatory framework (e.g., such as “ultrafine particulates”). The physical and chemical properties described in Table 3-1 for the metal oxide nanoparticle variants could potentially be utilized to derive characteristic-based definitions for the clusters. However, the lack of consistency in the endpoints reported by the source publications coupled with the skew in the particle variants tested (22 articles reported Titanium dioxide while only 5 reported Iron and 6 reported Zinc) leads to a low degree of consistency in the toxicological and particle characterization data across the clusters which precludes the possibility of deriving statistically significant property-based labels. A more consistent testing and reporting procedure would assist in identifying the mechanistic variations

in toxicity due to physical characteristics and deriving labels for the clusters. Such recommendations have been published for nanoparticle characterization, but there has so far been limited programming of toxicological research regarding specific particle variants and dose levels.[77-79]

Figure 3-5 is a graphical presentation of the natural log of the potency across the clusters. Higher values for potency result in greater increase in the response for a unit increase in dose. For example, the high value of potency for the LDH response levels imply that for even small increases in nanoparticle dose the LDH levels in the subject would show significantly larger increases compared to the other 4 responses. LDH or Lactate Dehydrogenase is an enzyme and a core component of the energy production system in the body, it is involved in the conversion of "Nicotinamide adenine dinucleotide (NAD)" to its reduced form "NAD + Hydrogen (NADH)". LDH is not normally located outside cells. LDH appears outside cells in fluids like blood or BAL fluid when cell membranes are leaking or have been ruptured. Sharp changes in the LDH levels are usually associated with tissue damage. Identifying the clusters and particles which contribute to increased potency levels for the LDH could be useful in early detection and mitigation of tissue damage.

Figure 3-6 compares the standardized potency of the clusters across the 5 responses, standardizing the potency makes it easier to compare the various response groups directly and determine their relative extent of threat. It can be seen in Fig. 2-6 that the particles belonging to clusters 'I', 'II' and 'IV' display signs associated with cell membrane damage and tissue damage/immune response. The cluster designated as 'III' is noteworthy as it is the only cluster that displays reduced response activity below the typical (i.e., mean) levels, this is indicative that the particles categorized as belonging to cluster 'III' possess less toxic potential than the particles in the other 3 clusters. Cluster 'III' is comprised primarily of Zinc oxide nanoparticles (3 out of 5

responses), previous literature has indicated that nano-Zinc oxide is non-toxic based on exposure studies conducted on mice.[80] Sayes et al. (2007) provided insight into the *in vivo* toxicity of nano-Zinc oxide and noted that acute toxicity response does exist, they were able to conclude that the effects were transient in nature and resolved themselves.[30] Nano-Zinc oxide is also highly commercialized and extensively used in sunscreen[81], a previous study was able to conclude that Zinc oxide nanoparticles greater than 30 nm in size exhibit the same properties as their non-nano counterparts.[82] While these results cannot conclusively assist in classifying Zinc oxide nanoparticles as non-toxic, they do offer some insight into the relatively low toxic potency of cluster 'III'.

Also, the normalized response indicator levels between clusters 'I' and 'II' are inverse to one another, cluster 'II' displays elevated LDH and PMN levels whereas cluster 'I' nanoparticles showed reduced LDH and PMN levels and vice versa with respect to the other 3 endpoints measured. Cluster 'I' is primarily rich in Iron nanoparticles, while cluster 'II' is mainly composed of Silica nanoparticles. Previous studies have compared the acute toxicity of Iron oxide and Silica nanoparticles, one study noted that no dose-related changes were observed in rats that were orally administered the nanoparticles.[83] Comparison of the individual cluster components across the 5 responses shows they differ in their chemical composition but more information is required before we hypothesize about the specific toxicological mechanism behind the differences in potency.

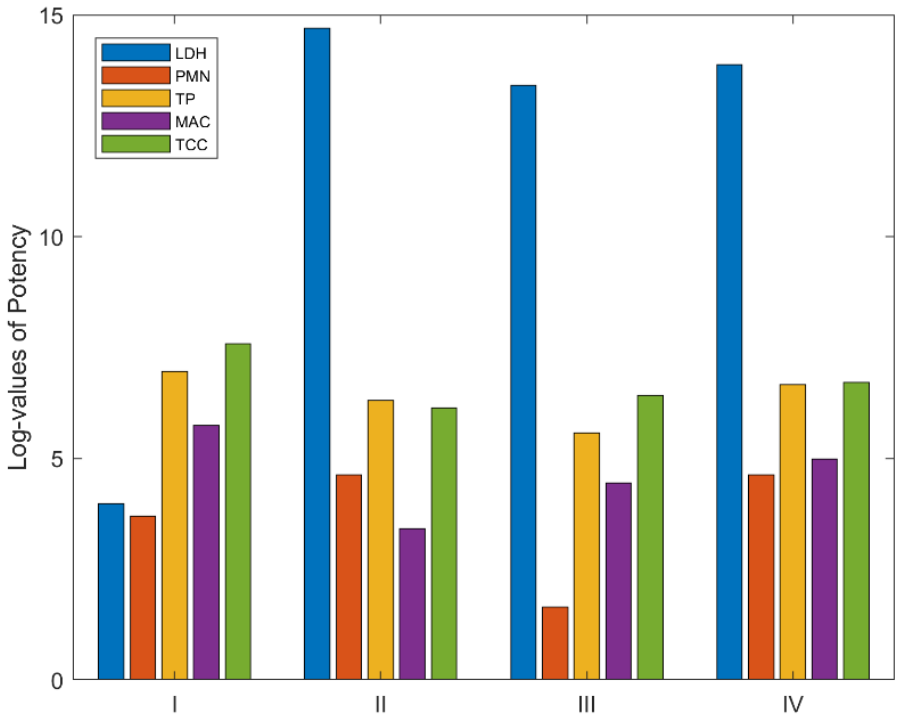

Figure 3-5. The variation between natural logs of the potency of the responses in each cluster provides insight towards the observed effects each cluster of particles has on the response. LDH response levels were highly elevated for clusters ‘II’, ‘III’ and ‘IV’.

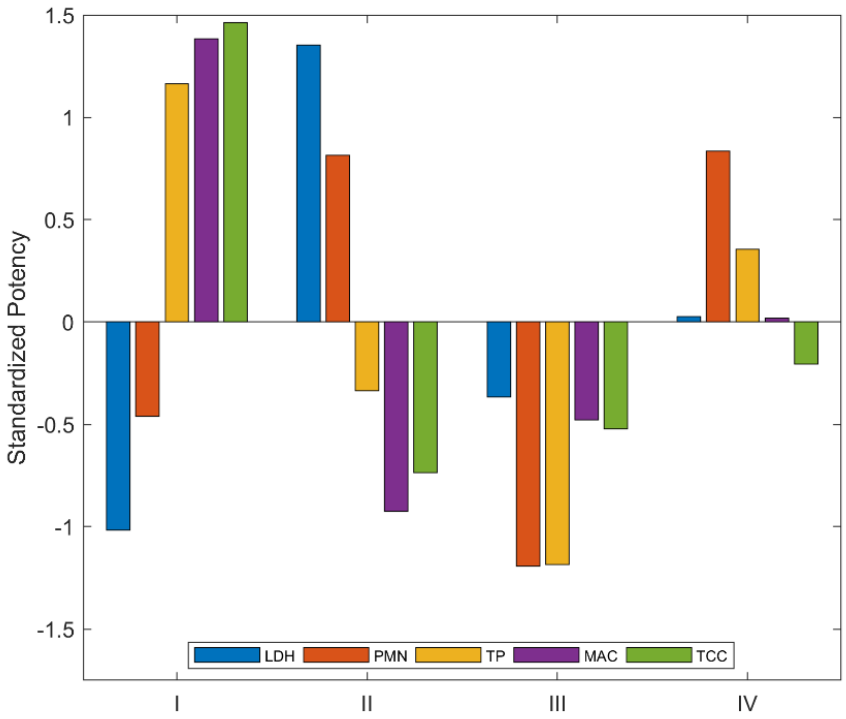

Figure 3-6. The variation of the standardized potency of the responses across the clusters allows for easier interpretation of the observed effects of exposure to the particle clusters. Cluster 'I', 'II" and 'IV' shows signs of increased cell damage and elevated immune response. Cluster 'III' shows below average toxic potency across all 5 measured endpoints

3.2.2 Possible Policy Implications

These results suggest that it may be possible and desirable to establish toxicological categories of engineered metal oxide nanoparticles and establish regulatory policy around those categories rather than seeking to establish a nano-specific exposure limit for each of the currently regulated solid particles. There are currently 10 of metal oxide particle exposure limits existing in OSHA regulations as of the date of this paper.[84]

If regulatory agencies would follow the pattern of the NIOSH REL for titanium dioxide and establish exposure limits that vary with metal oxide particle size for each of these particles,

the burden and cost of demonstrating compliance or non-compliance would be substantially increased. While this paper has not included ease of detection as a parameter for defining clusters, such an analysis could be conducted to develop categories that did account for this factor.

Beyond possible implications for defining exposure limits on a categorical rather than a specific-particle-by-specific-particle basis, these categories could also provide useful for the purposes of categorizing particular engineered nanoparticle variants for increased scrutiny as to their safety. Even in the absence of specific national regulations, new chemical products including engineered metal oxide nanoparticle products need to be approved for use. Those particles composed of constituents in the more toxic clusters should be prioritized for additional data gathering for safety purposes above those in the least toxic cluster. For example, based on these results, iron oxide, cerium oxide and silicon dioxide nanoparticles would be prioritized over zone, nickel, or aluminum oxide.

4. CONCLUSIONS

This analysis suggests that it is possible to define clusters or categories of metal oxide nanoparticles that behave similarly in terms of their dose-response-recovery effect on animals, which may be helpful in the future to maintain safety while also encouraging development of these materials without excessive burdens on future regulators. Further, this analysis indicates that 3 out of the 4 identified composition-based clusters display elevated response levels consistent with cellular damage (increased TP), onset of cytotoxicity (increased LDH) and immune activity (increased MAC, PMN and TCC). These are consistent with previously published literature on the toxicological effects of metal oxide nanoparticles.

This study demonstrates deployment of a hierarchical clustering algorithm to perform a meta-analysis of the *in vivo* pulmonary toxicity of engineered nanomaterials (specifically metal-oxide nanoparticles) in rodents. The argument for utilizing this algorithm over other available alternatives is the emphasis placed on the dose- response and recovery patterns rather than treating the available parameters as equivalent to particle properties, or assuming that all necessary physicochemical characteristics are already known. The process assists in generating hypotheses in regard to what mechanisms or characteristics might explain why certain particles cause similar or dissimilar pulmonary effects. The two iterations of the clustering on the MONP dataset produced 2 different set of results. The initial analysis treated the multiple MONPs analyzed as unique variations and performed the clustering leading to the identification of 8 toxicologically distinct clusters. There was a lack of relation between the characteristics of the MONP clusters and their composition, which limited the definitive description of the distinct clusters. The second iteration was performed after chemically grouping the various MONPs, the results were much more promising with 4 toxicologically distinct clusters being identified.

This work and the identified clusters propose that certain characteristics/nanoparticles are significant or insignificant to the identification of toxic potency. The clusters were further categorized into potential classes defined by their potency. The lack of sufficient characterization data does not prevent the use of this algorithm, although it may limit the definitive description (property-based labels) of the distinct clusters after the fact. Nanomaterials behaving in a similar/dissimilar pattern can be identified and grouped together and their differing attributes can be identified and related to their potency.

The analysis of the potency of the MONP clusters using the response variables was able to reveal that all but one group (cluster “III”) of metal oxide particles have the potential to cause increased cell damage and immune response intensity due to their elevated levels for the response

indicators analyzed. The toxicological behavior of the particles belonging cluster 'III' is worthy of further investigation to explore the underlying cause, it could be representative of certain particles in tandem displaying the capability to mitigate their respective biological effects. Overall, this study has shown that the potency of the nanoparticle clusters might not depend just on the physical and/or chemical characteristics, expansion to include more properties to explain the relationship could be meaningful. These results could also be interpreted as a sign that a simple relationship between potency and the physical and/or chemical characteristics is either non-existent or that it is more complex than anticipated and requires further investigation.

AUTHOR INFORMATION

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript and contributed equally.

Funding Sources

Funding for this work was provided by NIOSH, through grant 1 R03 OH010956-01. This work has also been supported 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).

ACKNOWLEDGMENT

Funding for this work was provided by NIOSH, through grant 1 R03 OH010956-01. This work has also been supported 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).

REFERENCES

1. USEPA, *Economic Analysis of the Proposed Change in Data Requirements Rule for Conventional Pesticides.* US Environmental Protection Agency, 2004.

2. Prabu, A. and R.B. Anand, *Emission control strategy by adding alumina and cerium oxide nano particle in biodiesel.* Journal of the Energy Institute, 2016. **89**(3): p. 366-372.

3. Weir, A., et al., *Titanium Dioxide Nanoparticles in Food and Personal Care Products.* Environmental Science & Technology, 2012. **46**(4): p. 2242-2250.

4. Ma, J.Y.C., et al., *Interactive effects of cerium oxide and diesel exhaust nanoparticles on inducing pulmonary fibrosis.* Toxicology and Applied Pharmacology, 2014. **278**(2): p. 135-147.

5. Peng, L., et al., *Comparative Pulmonary Toxicity of Two Ceria Nanoparticles with the Same Primary Size.* International Journal of Molecular Sciences, 2014. **15**(4): p. 6072-6085.

6. Gustafsson, A., et al., *Lung exposure of titanium dioxide nanoparticles induces innate immune activation and long-lasting lymphocyte response in the Dark Agouti rat.* J Immunotoxicol, 2011. **8**(2): p. 111-21.

7. Cho, W.S., et al., *Metal Oxide Nanoparticles Induce Unique Inflammatory Footprints in the Lung: Important Implications for Nanoparticle Testing.* Environmental Health Perspectives, 2010. **118**(12): p. 1699-1706.

8. Cho, W.S., et al., *Progressive severe lung injury by zinc oxide nanoparticles; the role of Zn2+ dissolution inside lysosomes.* Particle and Fibre Toxicology, 2011. **8**.

9. Silva, R.M., et al., *Biological Response to Nano-Scale Titanium Dioxide (Tio2): Role of Particle Dose, Shape, and Retention.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 2013. **76**(16): p. 953-972.

10. NIOSH, *Occupational Exposure to Titanium Dioxide.* National Institute for Occupational Safety and Health, 2011.

11. Yassin, A., F. Yebesi, and R. Tingle, *Occupational exposure to crystalline silica dust in the United States, 1988-2003.* Environmental Health Perspectives, 2005. **113**(3): p. 255-260.

12. Choi, J.Y., G. Ramachandran, and M. Kandlikar, *The Impact of Toxicity Testing Costs on Nanomaterial Regulation.* Environmental Science & Technology, 2009. **43**(9): p. 3030-3034.

13. Roco, M.C., *The long view of nanotechnology development: the National Nanotechnology Initiative at 10 years.* Journal of Nanoparticle Research, 2011. **13**(2): p. 427-445.

14. Puzyn, T., et al., *Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles.* Nature Nanotechnology, 2011. **6**(3): p. 175-178.

15. Rasulev, B., et al., *Nano-QSAR: Advances and Challenges.* Towards Efficient Designing of Safe Nanomaterials: Innovative Merge of Computational Approaches and Experimental Techniques, 2013. **25**: p. 220-256.

16. Bakhtyari, N.G., et al., *Prediction of Genotoxicity of Nano Metal Oxides by Computational Methods: A New Decision Tree QSAR Model.* Environmental and Molecular Mutagenesis, 2014. **55**: p. S43-S43.

17. Pan, Y., et al., *Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors.* Rsc Advances, 2016. **6**(31): p. 25766-25775.

18. Sexton, R.S., R.E. Dorsey, and J.D. Johnson, *Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing.* European Journal of Operational Research, 1999. **114**(3): p. 589-601.

19. Gernand, J.M. and E.A. Casman, *Machine Learning for Nanomaterial Toxicity Risk Assessment.* Ieee Intelligent Systems, 2014. **29**(3): p. 84-88.

20. Gernand, J.M. and E.A. Casman, *A Meta-Analysis of Carbon Nanotube Pulmonary Toxicity Studies-How Physical Dimensions and Impurities Affect the Toxicity of Carbon Nanotubes.* Risk Analysis, 2014. **34**(3): p. 583-597.

21. Chang, X.H., et al., *Health effects of exposure to nano-TiO2: a meta-analysis of experimental studies.* Nanoscale Research Letters, 2013. **8**: p. 1-10.

22. Notter, D.A., D.M. Mitrano, and B. Nowack, *Are Nanosized or Dissolved Metals More Toxic in the Environment? A Meta-Analysis.* Environmental Toxicology and Chemistry, 2014. **33**(12): p. 2733-2739.

23. Hernández, L.G., et al., *Can carcinogenic potency be predicted from in vivo genotoxicity data? a meta-analysis of historical data.* Environmental and Molecular Mutagenesis, 2011. **52**(7): p. 518-528.

24. Ioannidis, J. and J. Lau, *Pooling research results: benefits and limitations of meta-analysis.* The Joint Commission journal on quality improvement, 1999. **25**(9): p. 462-469.

25. Spector, T.D. and S.G. Thompson, *The potential and limitations of meta-analysis.* Journal of epidemiology and community health, 1991. **45**(2): p. 89.

26. Kadoya, C., et al., *Analysis of Bronchoalveolar Lavage Fluid Adhering to Lung Surfactant-Experiment on Intratracheal Instillation of Nickel Oxide with Different Diameters.* Industrial Health, 2012. **50**(1): p. 31-36.

27. Roursgaard, M., et al., *Acute and Subchronic Airway Inflammation after Intratracheal Instillation of Quartz and Titanium Dioxide Agglomerates in Mice.* Thescientificworldjournal, 2011. **11**: p. 801-825.

28. Oyabu, T., et al., *Dose-dependent pulmonary response of well-dispersed titanium dioxide nanoparticles following intratracheal instillation.* Journal of Nanoparticle Research, 2013. **15**(4).

29. Oberdorster, G., et al., *Role of the alveolar macrophage in lung injury: studies with ultrafine particles.* Environ Health Perspect, 1992. **97**: p. 193-9.

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

31. Creutzenberg, O., et al., *Toxicity of a Quartz with Occluded Surfaces in a 90-Day Intratracheal Instillation Study in Rats.* Inhalation Toxicology, 2008. **20**(11): p. 995-1008.

32. Grassian, V.H., et al., *Inflammatory response of mice to manufactured titanium dioxide nanoparticles: Comparison of size effects through different exposure routes.* Nanotoxicology, 2007. **1**(3): p. 211-226.

33. Lindenschmidt, R.C., et al., *The Comparison of a Fibrogenic and 2 Nonfibrogenic Dusts by Bronchoalveolar Lavage.* Toxicology and Applied Pharmacology, 1990. **102**(2): p. 268-281.

34. Roursgaard, M., et al., *Time-response relationship of nano and micro particle induced lung inflammation. Quartz as reference compound.* Human & Experimental Toxicology, 2010. **29**(11): p. 915-933.

35. Warheit, D.B., et al., *Pulmonary exposures to Sepiolite nanoclay particulates in rats: resolution following multinucleate giant cell formation.* Toxicol Lett, 2010. **192**(3): p. 286-93.

36. Warheit, D.B., C.M. Sayes, and K.L. Reed, *Nanoscale and Fine Zinc Oxide Particles: Can in Vitro Assays Accurately Forecast Lung Hazards following Inhalation Exposures?* Environmental Science & Technology, 2009. **43**(20): p. 7939-7945.

37. Ban, M., et al., *Effect of submicron and nano-iron oxide particles on pulmonary immunity in mice.* Toxicology Letters, 2012. **210**(3): p. 267-275.

38. Cho, W.S., et al., *Differential pro-inflammatory effects of metal oxide nanoparticles and their soluble ions in vitro and in vivo; zinc and copper nanoparticles, but not their ions, recruit eosinophils to the lungs.* Nanotoxicology, 2012. **6**(1): p. 22-35.

39. Katsnelson, B., et al., *Some Peculiarities of Pulmonary Clearance Mechanisms in Rats after Intratracheal Instillation of Magnetite (Fe3O4) Suspensions with Different Particle Sizes in the Nanometer and Micrometer Ranges: Are We Defenseless against Nanoparticles?* International Journal of Occupational and Environmental Health, 2010. **16**(4): p. 508-524.

40. Zhu, M.T., et al., *Comparative study of pulmonary responses to nano- and submicron-sized ferric oxide in rats.* Toxicology, 2008. **247**(2-3): p. 102-111.

41. Gelli, K., M. Porika, and R.N.R. Anreddy, *Assessment of Pulmonary Toxicity of MgO Nanoparticles in Rats.* Environmental Toxicology, 2015. **30**(3): p. 308-314.

42. Xia, T.A., et al., *Decreased Dissolution of ZnO by Iron Doping Yields Nanoparticles with Reduced Toxicity in the Rodent Lung and Zebrafish Embryos.* Acs Nano, 2011. **5**(2): p. 1223-1235.

43. Fourches, D., D.Q.Y. Pu, and A. Tropsha, *Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles.* Combinatorial Chemistry & High Throughput Screening, 2011. **14**(3): p. 217-225.

44. Sayes, C. and I. Ivanov, *Comparative Study of Predictive Computational Models for Nanoparticle-Induced Cytotoxicity.* Risk Analysis, 2010. **30**(11): p. 1723-1734.

45. Liu, R., et al., *Classification NanoSAR Development for Cytotoxicity of Metal Oxide Nanoparticles.* Small, 2011. **7**(8): p. 1118-1126.

46. Fourches, D., et al., *Quantitative Nanostructure-Activity Relationship Modeling.* Acs Nano, 2010. **4**(10): p. 5703-5712.

47. Kang, S., M.S. Mauter, and M. Elimelech, *Physicochemical determinants of multiwalled carbon nanotube bacterial cytotoxicity.* Environmental Science & Technology, 2008. **42**(19): p. 7528-7534.

48. Ramchandran, V. and J.M. Gernand, *A dose-response-recovery clustering algorithm for categorizing carbon nanotube variants into toxicologically distinct groups.* Computational Toxicology, 2019. **11**: p. 25-32.

49. Halim, A.H. and I. Ismail, *Combinatorial Optimization: Comparison of Heuristic Algorithms in Travelling Salesman Problem.* Archives of Computational Methods in Engineering, 2019. **26**(2): p. 367-380.

50. Inc., T.M., *MATLAB R2017b.* Natick, MA, USA.

51. Adewole, A., et al., *A comparative study of simulated annealing and genetic algorithm for solving the travelling salesman problem.* International Journal of Applied Information Systems (IJAIS), 2012. **4**(4): p. 6-12.

52. Andresen, M., et al., *Simulated annealing and genetic algorithms for minimizing mean flow time in an open shop.* Mathematical and Computer Modelling, 2008. **48**(7-8): p. 1279-1293.

53. Kerr, A. and K. Mullen, *A comparison of genetic algorithms and simulated annealing in maximizing the thermal conductance of harmonic lattices.* Computational Materials Science, 2019. **157**: p. 31-36.

54. Nemmar, A., K. Melghit, and B.H. Ali, *The acute proinflammatory and prothrombotic effects of pulmonary exposure to rutile TiO2 nanorods in rats.* Exp Biol Med (Maywood), 2008. **233**(5): p. 610-9.

55. Warheit, D.B., et al., *Pulmonary instillation studies with nanoscale TiO2 rods and dots in rats: toxicity is not dependent upon particle size and surface area.* Toxicol Sci, 2006. **91**(1): p. 227-36.

56. Renwick, L.C., et al., *Increased inflammation and altered macrophage chemotactic responses caused by two ultrafine particle types.* Occup Environ Med, 2004. **61**(5): p. 442-7.

57. Rehn, B., et al., *Investigations on the inflammatory and genotoxic lung effects of two types of titanium dioxide: untreated and surface treated.* Toxicology and Applied Pharmacology, 2003. **189**(2): p. 84-95.

58. Warheit, D.B., et al., *Pulmonary toxicity study in rats with three forms of ultrafine-TiO2 particles: differential responses related to surface properties.* Toxicology, 2007. **230**(1): p. 90-104.

59. Kobayashi, N., et al., *Comparative pulmonary toxicity study of nano-TiO2 particles of different sizes and agglomerations in rats: Different short- and long-term post-instillation results.* Toxicology, 2009. **264**(1-2): p. 110-118.

60. Roberts, J.R., et al., *Toxicological Evaluation of Lung Responses After Intratracheal Exposure to Non-Dispersed Titanium Dioxide Nanorods.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 2011. **74**(12): p. 790-810.

61. Gosens, I., et al., *Impact of agglomeration state of nano- and submicron sized gold particles on pulmonary inflammation.* Particle and Fibre Toxicology, 2010. **7**.

62. Cho, W.S., et al., *Inflammatory mediators induced by intratracheal instillation of ultrafine amorphous silica particles.* Toxicology Letters, 2007. **175**(1-3): p. 24-33.

63. Morimoto, Y., et al., *Inflammogenic effect of well-characterized fullerenes in inhalation and intratracheal instillation studies.* Particle and Fibre Toxicology, 2010. **7**.

64. Pirela, S., et al., *Effects of copy center particles on the lungs: a toxicological characterization using a Balb/c mouse model.* Inhalation Toxicology, 2013. **25**(9): p. 498-508.

65. Toya, T., et al., *Pulmonary Toxicity Induced by Intratracheal Instillation of Coarse and Fine Particles of Cerium Dioxide in Male Rats.* Industrial Health, 2010. **48**(1): p. 3-11.

66. Morimoto, Y., et al., *Expression of inflammation-related cytokines following intratracheal instillation of nickel oxide nanoparticles.* Nanotoxicology, 2010. **4**(2): p. 161-176.

67. Ma, J.Y., et al., *Cerium oxide nanoparticle-induced pulmonary inflammation and alveolar macrophage functional change in rats.* Nanotoxicology, 2011. **5**(3): p. 312-325.

68. Park, E.J., et al., *Induction of Inflammatory Responses in Mice Treated with Cerium Oxide Nanoparticles by Intratracheal Instillation.* Journal of Health Science, 2010. **56**(4): p. 387-396.

69. Wingard, C.J., et al., *Mast cells contribute to altered vascular reactivity and ischemia-reperfusion injury following cerium oxide nanoparticle instillation.* Nanotoxicology, 2011. **5**(4): p. 531-545.

70. Xue, L.X., et al., *Pulmonary Toxicity of Ceria Nanoparticles in Mice After Intratracheal Instillation.* Journal of Nanoscience and Nanotechnology, 2013. **13**(10): p. 6575-6580.

71. Minarchick, V.C., et al., *Pulmonary Cerium Dioxide Nanoparticle Exposure Differentially Impairs Coronary and Mesenteric Arteriolar Reactivity.* Cardiovascular Toxicology, 2013. **13**(4): p. 323-337.

72. Dick, C.A.J., et al., *The role of free radicals in the toxic and inflammatory effects of four different ultrafine particle types.* Inhalation Toxicology, 2003. **15**(1): p. 39-52.

73. Zhang, Q.W., et al., *Differences in the extent of inflammation caused by intratracheal exposure to three ultrafine metals: Role of free radicals.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 1998. **53**(6): p. 423-438.

74. Horie, M., et al., *Comparison of acute oxidative stress on rat lung induced by nano and fine-scale, soluble and insoluble metal oxide particles: NiO and TiO2.* Inhalation Toxicology, 2012. **24**(7): p. 391-400.

75. Ogami, A., et al., *Pathological features of different sizes of nickel oxide following intratracheal instillation in rats.* Inhalation Toxicology, 2009. **21**(8-11): p. 812-818.

76. Nishi, K., et al., *Expression of cytokine-induced neutrophil chemoattractant in rat lungs by intratracheal instillation of nickel oxide nanoparticles.* Inhalation Toxicology, 2009. **21**(12): p. 1030-1039.

77. Cho, E.J., et al., *Nanoparticle Characterization: State of the Art, Challenges, and Emerging Technologies.* Molecular Pharmaceutics, 2013. **10**(6): p. 2093-2110.

78. Jiang, J.K., G. Oberdorster, and P. Biswas, *Characterization of size, surface charge, and agglomeration state of nanoparticle dispersions for toxicological studies.* Journal of Nanoparticle Research, 2009. **11**(1): p. 77-89.

79. Hassellov, M., et al., *Nanoparticle analysis and characterization methodologies in environmental risk assessment of engineered nanoparticles.* Ecotoxicology, 2008. **17**(5): p. 344-361.

80. Wang, B., et al., *Acute toxicological impact of nano- and submicro-scaled zinc oxide powder on healthy adult mice.* Journal of Nanoparticle Research, 2008. **10**(2): p. 263-276.

81. Smijs, T.G. and S. Pavel, *Titanium dioxide and zinc oxide nanoparticles in sunscreens: focus on their safety and effectiveness.* Nanotechnol Sci Appl, 2011. **4**: p. 95-112.

82. Auffan, M., et al., *Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective.* Nature Nanotechnology, 2009. **4**(10): p. 634-641.

83. Yun, J.W., et al., *Comparative toxicity of silicon dioxide, silver and iron oxide nanoparticles after repeated oral administration to rats.* J Appl Toxicol, 2015. **35**(6): p. 681-93.

84. (OSHA), O.S.H.A., *TABLE Z-1 Limits for Air Contaminants.* Toxic and Hazardous Substances. **Occupational Safety and Health Standards**.

CHAPTER 4

Evaluation of Risk and Uncertainty for Model-Predicted NOAELs of Engineered Nanomaterials Based on Dose-Response-Recovery Clusters

The contents of this chapter have been published in the ASME "Journal of Risk and Uncertainty, Part B: Mechanical Engineering" under the title *"Evaluation of Risk and Uncertainty for Model-Predicted NOAELs of Engineered Nanomaterials Based on Dose-Response-Recovery Clusters"* (*https://doi.org/10.1115/1.4055157*).

ABSTRACT

Experimental toxicology studies for the purposes of setting occupational exposure limits for aerosols have drawbacks including excessive time and cost which could be overcome or limited by the development of computational approaches. A quantitative, analytical relationship between the characteristics of emerging nanomaterials and related *in vivo* toxicity can be utilized to better assist in the subsequent mitigation of exposure toxicity by design. Predictive toxicity models can be used to categorize and define exposure limitations for emerging nanomaterials. Model-based no-observed-adverse-effect-level (NOAEL) predictions were derived for toxicologically distinct nanomaterial clusters, referred to as MP-NOAELs. The lowest range of MP-NOAELs for the polymorphonuclear neutrophil (PMN) response observed by CNTs was found to be 21 – 35 μg/kg (cluster "A"), indicating that the CNT belonging to cluster "A" showed the earliest signs of adverse effects. Only 25% of the MP-NOAEL values for the CNTs can be quantitatively defined at present. The lowest observed MP-NOAEL range for the metal oxide nanoparticles was Cobalt oxide nanoparticles (Cluster III) for the Macrophage (MAC)

response at 54 – 189 μg/kg. Nearly 50% of the derived MP-NOAEL values for the metal oxide nanoparticles can be quantitatively defined based on current data. A sensitivity analysis of the MP-NOAEL derivation highlighted the dependency of the process on the shape and type of the fitted dose response model, its parameters, dose selection and spacing, and the sample size analyzed.

1. INTRODUCTION

Inhalation exposure to engineered nanomaterials presents the potential for an emerging occupational hazard. The unique combinations of properties of these small particulates may produce unique effects in individuals who are exposed to them in significant quantities. However, engineers, chemists, and materials scientists have an opportunity to guide in the future directions of this technology towards less harmful paths, if we can discern those safer choices before we are faced with a growing set of occupational illnesses that could arise from this industry.

Experimental toxicology studies, the best way we have to anticipate the effects of human exposure, are often time consuming and require considerable investment. A viable solution to the problem is testing the NMs on a priority basis—targeting the more toxic variants first. The development of computational approaches targeted at predicting the toxicity of the NMs based on their attributes can be used as a means to isolate potent nanomaterial variants and guide the future synthesis of less hazardous variants. Precaution against hazardous nanomaterials can be taken by the imposition of recommended exposure limits (RELs) and permissible exposure limits (PELs), NIOSH and OSHA are responsible for their development respectively. NIOSH has currently proposed recommended exposure limits (REL) of 1 $\mu g/m^3$ for Carbon nanotubes (CNTs) and RELs of 50 $\mu g/m^3$ for respirable Crystalline Silica, these limits are derived from the extrapolation

of pulmonary exposure in Rats to Humans.[1-3]. NIOSH treats metal oxide nanoparticles in a similar fashion, the current proposed REL for Titanium dioxide (TiO_2) currently stands at 2.4 mg/m^3 and 0.3 mg/m^3 for fine and ultrafine (includes nanoscale) species of TiO_2[1]. It remains uncertain as to whether these recommendations will prove feasible as occupational exposure limits due to the difficulty of detecting nanoparticle concentrations at low levels in the context of realistic occupational environments.

The establishment of a recommended exposure limit is typically accomplished through identification of the No Observed Adverse Effect Level (NOAEL) in a toxicological study. The NOAEL is the highest dose for which no discernable adverse or negative change of health in the test subject could be detected. Since test subjects are typically animals or sometimes just cells, a series of adjustment factors or uncertainty factors are used to reduce the experimentally derived NOAEL to an established recommended exposure level or reference dose. Depending on the quantity and quality of the data, these adjustment factors can range from 3 to 1000 or more.

Though not yet used for the establishment of government recommendations in the US, several researchers have documented NOAELs [4] and corresponding adjustment factors for other nanomaterial variants that could presumably be used in such a way. The NOAELs and adjustment factors in these studies have varied between 0.1 to 0.37 mg/m^3 and 2 to 50 respectively[2, 5]. The NOAEL has traditionally played a key role in the determination of RELs. Given the diversity in responses to different varieties of nanoparticles, further subdivision of the exposure limits may be useful by incentivizing the production and use of less toxic varieties of nanoparticles in the future. Experimental evaluation of toxicity and exposure limits are subject to a certain degree of error or bias and in some cases the results may be difficult to replicate. To overcome some of these inefficiencies, forecasting methods have been proposed to predict the potential toxicity and environmental impact of new NMs[6-11].

Attempts at using computational modeling to predict the effects on toxicity of changes in particle properties (QSARS, meta-analysis etc.) have been limited in their scope to date. The rapid increase in the variety of engineered nanoparticles synthesized reveals a need to enhance the currently available state of knowledge and investigate exposure patterns[12]. A 10-year study in 2011 estimated that roughly 6 million people could be at risk of exposure to NMs by 2020 [13]. Categorizing NMs displaying similar effects as a means of grouping the NMs on the basis of their toxicity is a potential solution to the existing problem. To address the question of how different nanomaterials can be and still be considered the same substance an algorithmic approach applied to meta-analysis of extant *in vivo* nanomaterial toxicity data was accomplished [14]. Progress in this direction could be significant in the analysis of new and emerging NMs which can be pre-screened and prioritized for more careful evaluation if predicted to be more hazardous. *In-vivo* screening of nanoparticles has previously been attempted on isopods [15]. The screening concluded that duration of exposure had greater effect than the quantity consumed. The financial factor impacting the potential regulation of NMs cannot be ignored with cost estimations for hazard testing of nanoparticles projected to range between $250 million to $1.2 billion [16]. The evaluation of updated exposure limits for the new NMs is also necessary to assess their agreement with the existing recommendations for workplace exposure. The challenges associated with the assessment of the risk posed by nanoparticles has been documented extensively [17, 18]. An alternative to the traditionally determined NOAEL is the “Benchmark Dose” (BMD) which in similar fashion to the NOAEL is a single number that is statistically derived and can be used to quantify the exposure risk to substances [19-22]. A detailed and critical comparison between the BMD and NOAEL has been conducted previously which suggests that NOAEL derivation methodology is antiquated and dependent on several experimental factors such as dose selection, dose spacing and sample size [23]. The advantages associated with NOAELs lie in its ease of derivation and

ability to work all types of exposure data. The BMD process takes into account multiple factors of which the shape of the dose-response curve is an important factor, there are multiple available BMD models to choose from depending on the variation between the dose and response [23]. This paper illustrates the derivation of a model predicted NOAEL (MP-NOAEL) which in similar fashion to the BMD utilizes a fitted dose-response model as a critical component of the methodology.

2. METHODS

2.1 Data Selection

The data utilized for the meta-analysis of the *in vivo* pulmonary toxicity of the engineered nanoparticles is curated from peer-reviewed journal articles. The publications are all independent studies, a study can be defined as a complete experiment that consists of several levels of exposure and recovery times to one or more different nanoparticle variants (a batch of nanoparticles defined by a set of chemical and physical properties). Studies are comprised of multiple exposure groups, each group is a set of animals characterized by their exposure to a specific dosage, method of exposure (inhalation, instillation, aspiration etc.), post-exposure recovery length, and their resultant toxicological measurements. There are five quantitative response measures from the BAL fluid which are the focus of our meta-analysis: Total cell count (TC), Macrophage count (MAC), Total protein concentration (TP), Polymorphonuclear neutrophil count (PMN) and Lactate dehydrogenase concentration (LDH). These 5 measures were reported in a sufficient number of publications to be included in this analysis. The measured response reported in each publication is expressed as a fold (multiple) of its respective control group's measure, this has been done in accordance with previous work which suggests that to

compare different exposure modes (instillation, aspiration, and inhalation) a normalized response is needed. For the purposes of this study, *in vivo* pulmonary toxicity data was chosen from 2 major divisions of nanomaterials namely Carbon nanotubes (CNTs) and Metal Oxide Nanoparticles (MONPs). Table 4-1 below is a comprehensive summary of the data sources utilized for the analysis presented in this publication.

Table 4-1. Compilation of the data (with sources) used for the analysis presented in this publication.

StudyID	Primary Author	Nanomaterial Type	No. of endpoints (of interest) reported	No. of exposure groups
1001	Pauluhn[24]	MWCNT	5	13
1002	Ma-Hock[25]	MWCNT	5	7
1003	Muller[26]	MWCNT	3	9
1004	Shvedova[27]	SWCNT	4	4
1005	Shvedova[27]	SWCNT	2	7
1006	Nygaard[28]	SWCNT, MWCNT	2	7
1007	Warheit[29]	SWCNT	3	9
1009	Park[30]	SWCNT	3	5
1010	Teeguarden[31]	SWCNT	3	2
1011	Elgrabli[32]	MWCNT	1	16
1012	Mercer[33]	SWCNT	2	4
1013	Porter[34]	MWCNT	2	13
1016	Ellinger-Ziegelbauer[35]	MWCNT	3	10
1017	Shvedova[36]	SWCNT	4	16
1021	Ge[37]	MWCNT	4	8
T1	Nemmar[38]	Titanium	2	3
T2	Oberdorster[39]	Titanium	4	7
T3, S1	Warheit[40]	Titanium, Silica	4	36

T4	Renwick[41]	Titanium	4	6
T5, S2	Rehn[42]	Titanium, Silica	4	30
T6	Grassian[43]	Titanium	3	16
T7, S3	Warheit[44]	Titanium, Silica	4	88
T8, S4	Warheit[45]	Titanium, Silica	3	20
T9, S5	Kobayashi[46]	Titanium, Silica	3	41
T10, T13	Gustafsson[47]	Titanium	3	16
T11	Oyabu[48]	Titanium	3	20
T12	Roberts[49]	Titanium	3	12
T14	Silva[50]	Titanium	1	14
S6	Gosens[51]	Silica	4	4
S7	Cho[52]	Silica	4	16
S8	Creutzenberg[53]	Silica	5	6
S9	Roursgaard[54]	Silica	3	21
N7	Morimoto[55]	Nickel	2	5
F1	Ban[56]	Iron	2	11
F2	Pirela[57]	Iron	4	4
F3	Zhu[58]	Iron	4	13
F4	Katsnelson[59]	Iron	3	4
S10, Z1	Sayes[60]	Silica, Zinc	3	36
Ce1	Toya[61]	Cerium	4	17
N1	Morimoto[62]	Nickel	1	15
Z2, F5	Xia[63]	Zinc, Iron	2	15

S11, Z3	Warheit[64]	Silica, Zinc	3	28
T15, S12, Ce2, Z4, N2, Cu1	Cho[65]	Titanium, Silica, Cerium, Zinc, Nickel, Copper	3	24
Ce3	Ma[66]	Cerium	3	18
Ce4	Peng[67]	Cerium	4	12
T16, S13	Roursgaard[68]	Titanium, Silica	3	28
T17, S14, A1	Lindenschmidt[69]	Titanium, Silica, Aluminum	5	45
Ce5	Park[70]	Cerium	3	7
Ce6	Wingard[71]	Cerium	3	4
Ce7	Xue[72]	Cerium	3	15
Ce8	Ma[73]	Cerium	2	6
Ce9	Minarchick[74]	Cerium	2	4
T18, Co1	Dick[75]	Titanium, Cobalt	3	6
S15, M1	Gelli[76]	Silica, Magnesium	1	15
T19, S16, Co2	Zhang[77]	Titanium, Silica, Cobalt	5	25
T20, N3	Horie[78]	Titanium, Nickel	2	16
T21, S17, N4	Kadoya[79]	Titanium, Silica, Nickel	1	25
T22, S18, N5	Ogami[80]	Titanium, Silica, Nickel	2	25
N6, Z5, Cu2	Cho[81]	Nickel, Zinc, Copper	2	8
Z6	Cho[82]	Zinc	4	9
N8	Nishi[83]	Nickel	2	15

2.2 Model and Algorithm Definition

The dose-response model used for the clustering analysis is an important feature of the process; it reflects the variation of the response to changes in the applied dose. A single dose-response model precludes the possibility of explaining the behavior of all engineered substances, to overcome the limitations of a single model the development of benchmarked dose-response models that are able to cater to a wide variety of situations. Slob (2001) has previously published a 3-parameter model available for modeling continuous data.[84] A revision has been proposed to Slob's model, the addition of a fourth parameter that assimilates the decay of the response over time (i.e. the animal recovers) to showcase a more complete dose- response-recovery relationship.[14] The revised relationship between the response measured (y), dosage applied (x) and post-exposure recovery period (t) is,

$$y = a[c - (c-1)e^{-bx}] - dt \qquad \text{[Eq. 1]}$$

The 4 parameters, *a* (signifies the response at dose = 0), *b* (the toxic potency of the nanoparticles), *c* (the maximum relative shift in response), and *d* (slope of the response decay) are the key parameters which can be used to quantify and reflect the potential of the particle to be a hazard.

An algorithm was developed to facilitate segregation of the dataset into groups or "clusters" based on the similarity of their dose-response-recovery profiles. The methodology is derived from the hierarchical clustering process and utilizes non-linear regression to fit the model proposed to the data collected. The approach yields multiple possibilities which are assessed using AIC as the performance metric to exclude overfitted solutions, the AIC is calculated based on the number of parameters (k) in a model and its log-likelihood (LogL).

$$AIC = 2k - 2LogL \qquad \text{[Eq. 2]}$$

Lower values of AIC are preferred when comparing the clusters as they indicate the model is closer to the data in those cases. AIC needs to be adjusted when dealing with small datasets, this corrected form is denoted as AICc.

$AICc = AIC + \frac{2k(k+1)}{n-k-1}$, where n is the number of observations and k is the number of model parameters.

The AICc weighs both the complexity and performance factors to provide the most balance, it helps in filtering about overfitted solutions that provide marginal improvement while increasing the complexity of the model. A detailed discussion behind the process and approach can be found in a previous publication by the same author's.[14]

2.3 Model-Predicted NOAELs

This publication presents a methodology to derive Model-Predicted NOAEL (MP-NOAEL) values for clusters of both the CNTs and MONPs. The process utilizes the measured control group response for each particle tested and the standard deviation associated with the control group response to derive the MP-NOAEL. For the purpose of our analysis, we assume our MP-NOAEL response to be the dose that corresponds to the response measured at 2 standard deviations above the control group response. The dose-response curve (Eq. 1) for each cluster can be used to estimate the dose that corresponds to any given value of the measured response, the recovery term from Eq. 1 is set to 0 as there is no observed adverse effect onset at "t" = 0. Since, the dose-response-recovery curve for each cluster is different we are able to obtain model predicted NOAELs which are unique to each cluster. The NOAEL values predicted by this methodology might not necessarily agree with what eventual experiments predict. Some of the factors responsible for the differences can be the inherent variance arising from the multiple studies included in our analysis, the NOAELs being a function of the dose levels tested (variation in dose levels tested between publications leads to diverse NOAELs), diversity in the animal

species used in the exposure studies, and multiple nanoparticle sources. These factors are just a few factors which could be contributing to difference between model-predicted MP-NOAELs and experimentally determined NOAELs.

The type of model used to model the dose-response variation directly impacts the MP-NOAEL calculation since the MP-NOAEL dose value is back-calculated using the MP-NOAEL response, altering the type of dose-response model used results in directly modification of the MP-NOAEL values. The default option for all clusters is the model represented by Eq. 1, in some special cases it would be favorable to substitute the traditionally exponential dose-response model for a simple linear dose-response relationship ($y = mx + c + \varepsilon$) or a simple exponential model ($y = a * e^{-bx}$), where 'y' is the measured response and 'x' is the corresponding delivered dose. The substitution is necessary due to compatibility issues between the original dose-response model and the data encapsulated by the particular cluster. Common examples where model substitutions become necessary include situations where there is a lack of data, such cases might be better modeled using a linear model than an exponential model. Similarly, when the empirical evidence suggests non-conformity and poor model fit across all model choices, we can choose the simplest model. The 95% confidence interval of the parameter estimates for the dose-response model from Eq. 1 are used to derive the corresponding confidence limits of the MP-NOAELs; the parameter estimates themselves are based on how well the model conforms to the available data.

The novelty behind the approach described by this publication lies in the data collection process and model-driven approach adopted to derive NOAEL limits. The addition of a temporal component to the dose-response relationship to model response decay, choosing to cluster the nano-substances together based on their dose-response-recovery profiles rather than physico-chemical characteristics, and working on an assembled in-vivo dataset are just some prominent examples where the adopted methodology for deriving NOAELs differs from the traditional. This methodology has the potential to provide a diverse range of NOAELs for similar substances

categorized by their relative toxicity rather than a single all-encompassing limit. Stratification of the NOAEL limits would be highly effective to isolate toxic variants and prioritize their testing. While there are differences between the exposure limits determined using the method highlighted in this paper and the BMD methodology, the NOAELs suggested by the authors are meant to supplement the existing information in the domain rather than supplant the existing benchmarks.

3. RESULTS AND DISCUSSION

3.1 Clustering Results

The clustering analysis was conducted on 2 datasets of engineered nanoparticles, CNTs and MONPs. The dataset for the MONPs is comprised of multiple studies where more than 1 unique set of nanoparticles is analyzed, such studies were identified and separated to ensure each study reflects a single particle. Both sets of data were analyzed for the 5 endpoints discussed earlier and their ideal cluster combinations were derived. An ideal cluster size of 4 was consistently achieved across the 5 responses for the CNTs, there is no consensus cluster size that can be agreed upon across the 5 responses for the MONPs due to the varying size of the number of particles analyzed for each response. A detailed reasoning behind the ideal choice of clusters being 4 has been addressed previously by the authors.[14]

The existence of these clusters is not based solely on the similarity of their *in vivo* toxic effects. Detailed examination of the particles within the clusters is necessary to study the characteristics of the constituent particles and generate hypotheses to ascertain why one cluster or class of nanoparticles may be more or less toxic than another. The particles were not separated into groups based on their physical/chemical characteristics before the clustering process. The reasoning being that physically and/or chemically similar nanomaterials need not necessarily be

similar in their dose-response-recovery relationships as well, depending on which characteristics are used to define similarity. Differences in the relative toxicity of the clusters could be correlated to significant variation in the particle attributes between them, these attributes are also hypothesized as key contributors in the separation of the clusters at each stage.

3.2 Characterization Evaluation

Examination of the properties of all the particles in the CNT and MONP clusters revealed that the CNT particles can be identified by their physical characteristics. Table 4-2 is an illustration of the distribution of physical properties (median length and median diameter) and impurity content of 4 CNT clusters using PMN as a response. We can observe upon inspection of the physical properties the potential to define provisional thresholds which can be used to assign the clusters to different categories. By virtue of these thresholds, we can designate rudimentary labels to the clusters for ease of reference. Setting thresholds of 15 nm for the diameter leads to clusters 2 and 3 being designated as “Thin” clusters whereas clusters 1 and 4 can be regarded as “Thick”. Using 2200 nm as a threshold for length we can label the 4 clusters as “Long” and “Short” based on their median length. A similar approach adopted yielded consistent results across the other 4 responses analyzed although with varying threshold limits for the length and diameter. A single uniform range of values for length and diameter for all the response variables could not be established due to the shift in the constituent members of a cluster across the responses. It should be noted that these labels are provisional and cannot be verified as statistically meaningful across the 5 responses analyzed due to insufficient characterization data amongst the tested CNTs. The labels have been highlighted here to point to the presence of a pattern which could possibly be examined in thoroughly in the future.

Similarly, we analyzed the MONP clusters for the PMN response in the same fashion as the CNTs to identify and highlight any trends between the clusters and their particle properties.

The ideal number of clusters for the PMN response were found to be 8 based on the AICc values. Table 4-3 showcases the particle properties and cluster composition of the MONP clusters for the PMN response. The “Particle Typical size” measurement is the mean diameter for the spherical MONPs that were analyzed. We can observe that there is no perceivable trend between the particle size and the cluster composition for this particular response. Analysis of the 4 other responses for the MONPs revealed the same. This could imply that unlike the CNT particles, categorizing the metal oxide nanoparticles purely using their physical properties might not be possible and hence a more detailed approach might be required which considers both the physical and chemical characteristics of the particles.

Table 4-2. Median physical attributes of the CNT clusters for PMN response. The attributes can be used to group the clusters based on their physical properties. Provisional thresholds can be set at a median length of 2200 nm and median diameter of 10 nm to generate labels for the clusters.

Cluster ID	Median Length (nm)	Median Diameter (nm)	% Impurities	Model Predicted NOAEL (µg/kg)
A	2800±678.4	5.5±1.4	14	21 – 35
B	2150±559.9	4.3±1.2	3.26	-1712 – 11825
C	2090±369.1	29.3±4.1	0.95	255 – 401
D	2441±366.3	16.7±3.0	5.95	-51 – -42

Table 4-3. Mean "typical particle size" for MONP clusters along with constituent cluster composition for PMN response. The composition of the clusters indicates that multiple metal oxide particles can be categorized together by virtue of their dose- response-recovery relationships.

Cluster ID	Mean Particle size (nm)	Cluster Composition	Model Predicted NOAEL (µg/kg)
I	1228.4	Crystalline Silica, Amorphous Silica	22.3
IV	232.8	TiO_2, NiO, Crystalline Silica, ZnO, CeO_2	49.7
VII	13.9	CeO_2, Crystalline Silica	117
II	3133.3	TiO_2, Al_2O_3, CuO	214.4
III	14.6	TiO_2	440.6
VI	164.4	TiO_2, Amorphous Silica, ZnO	1460
VIII	154	TiO_2, Crystalline Silica, ZnO, CeO_2	1993.1
V	107.7	TiO_2, Fe_2O_3	16255

3.3 Model Predicted NOAELs

No-observed-adverse-effect-levels (NOAELs) can be predicted for each of the clusters generated using the algorithm. The NOAELs can be defined as the highest experimental dose that does not produce an adverse effect.[85] Table 4-2 and 4-3 are the measures of the predicted NOAELs for the CNTs and MONPs respectively using PMN as a response. Increase in the PMN response is an indicator of inflammation, they are also preferred due to their relative sensitivity during inhalation- based experiments.[86, 87] Large values for the NOAEL of a substance imply that higher dosage is required before adverse effects are observed in an organism and hence lower the potency of the substance.

The initial clustering process treated the MONPs analyzed from each study as a unique particle. This leads to different instances of the same substance appearing in multiple clusters as evidenced in Table 4-3. The information provided in Table 4-3 for the MONP clusters do not seem to provide evidence of any discernable trend between the particle properties, cluster composition and the predicted NOAELs for the respective clusters. A repeat of the clustering process was performed by isolating the various particles on a chemical basis. Table 4-4 contains the measures of particle size and predicted NOAELs for the new set of clusters derived along with their respective compositions.

Table 4-4. Mean particle size and predicted NOAELs for the newly formed clusters based on chemical isolation of the MONPs. U indicates that a range of values for the cluster NOAEL is undefined.

Cluster ID	Mean Particle size (nm)	Cluster Composition	Model Predicted NOAEL (µg/kg)
III	48.3	ZnO, CeO_2, NiO	U
I	105.3	Iron Oxides (Fe^{2+}, Fe^{3+})	U
II	1340.8	Silica, CuO, Al_2O_3, Co_3O_4	-2700 – -1567
IV	193.4	TiO_2	514 – 837

We can observe from Table 4-4 that not only are the optimum number of clusters reduced from 8 to 4 for the MONP dataset compared to the earlier iteration but also that two of the clusters contain isolated chemicals (TiO_2 and Iron Oxides). A key observation between the 2 clustering variants is the presence of irrational MP-NOAEL values for the second clustering iteration, the cause of these values and their possible treatment has been addressed in the following section.

3.4 Sensitivity Analysis

The model-predicted NOAEL estimates discussed earlier for the CNTs and MONPs are point estimates that are determined using the dose-response model. The dose-response model plays a crucial role in both the clustering process and the determination of MP-NOAELs. The confidence interval for the MP-NOAEL values can be visualized using a “confidence spread”. Evaluating the confidence spread of the dose-response curve would aid in visual assessment of the fit of the model to the experimental data, a tight spread can be viewed as representative of low error and a good fit to the data. Conversely, larger spreads imply larger error in the parameter

estimates. Figures 4-1 and 4-2 are helpful in visually symbolizing the confidence spread for the dose-response models associated with each cluster across all 5 responses. Based, on the evidence presented in Figure 4-1 we can see that only a handful of CNT clusters (5 out of 20) display the signs associated with good fit and low error.

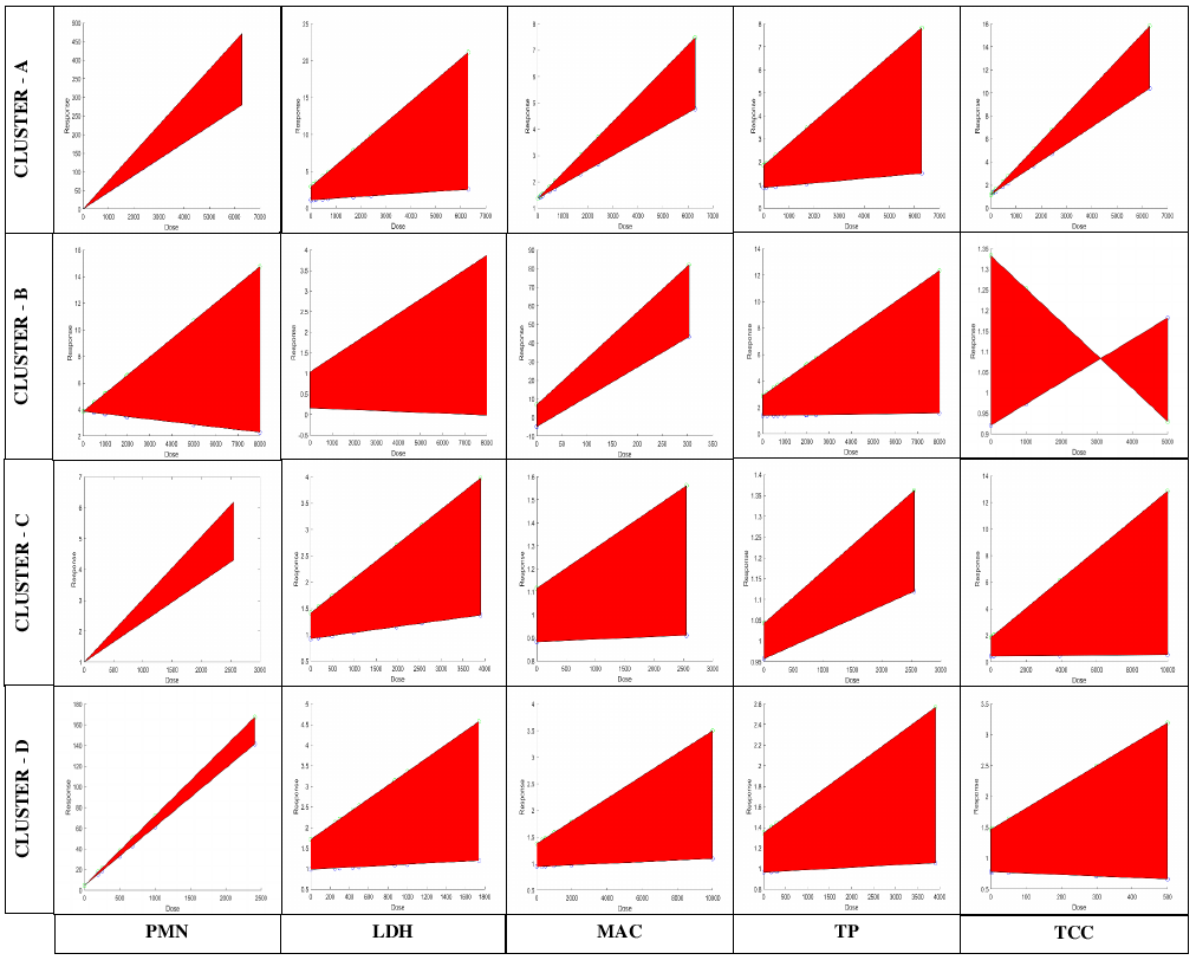

Figure 4-1. Confidence spread for the 4 clusters across the 5 responses for the CNT particles. Tighter spreads correlate to lower mean squared error computations. Larger spreads are associated with increased error in parameter estimation.

The cause for greater uncertainty associated with the estimated dose-response model parameters could be attributed to several reasons. We can see from Figure 4-1, there are 3 clusters ('A', 'C' and 'D') for the PMN response which show and conform to the typical dose-response model relationship. Amongst the other clusters only Cluster 'A' for the "MAC" and "TCC" responses are consistent with good fit, low error representations. The majority of the remaining clusters show larger spreads which could imply that the fitted model is not the right choice for the data evaluated. Examining the models for the larger spread cases we observed the regression model parameters were all deemed significant at a 0.05 level of significance. The possibility of the model despite passing the test of significant not being the right choice for the data was considered and alternatives were tested, the alternatives did not show significant improvements. The lack of insignificant parameters for the dose-response model could also indicate that the data itself could be a source of error. Detailed exploration of the available data revealed that there is indeed a lack of data in those cases showing large spread, the quality of data could also be a factor leading to larger error. The data collection process is also a factor especially in the case of Cluster 'B" under the TCC response. The source publication only reported the results of evaluation at only 2 dose levels which were well spaced from one another leading to large error.[29]

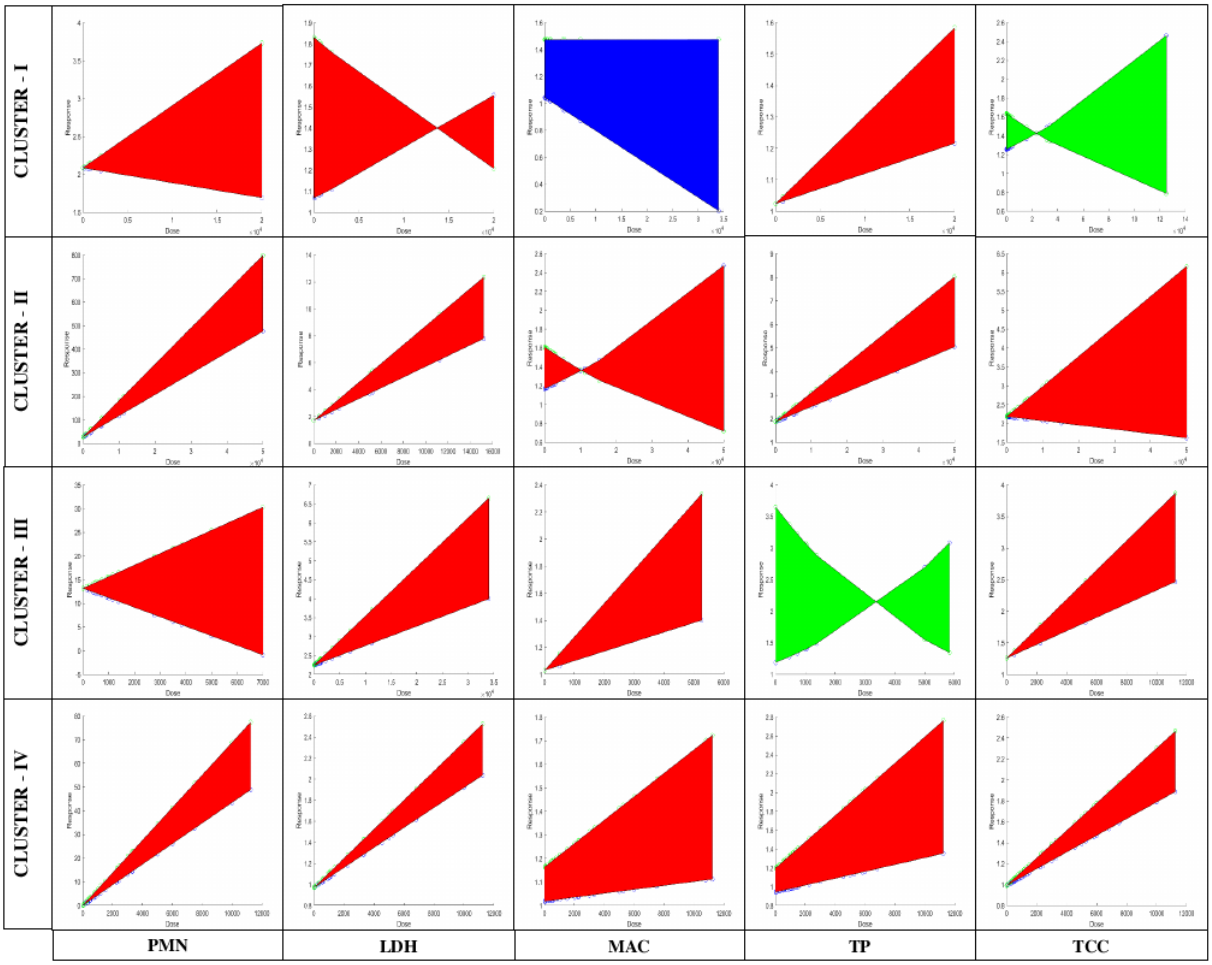

Figure 4-2. Dose-response including the 95% confidence interval of the response for the 4 clusters across the 5 responses for the MONP particles. Tighter spreads correlate to lower mean squared error computations. Larger spreads are associated with increased error in parameter estimation. The color of shade is used to present the type of model used in evaluate the dose-response variation. Red is used to represent the traditional exponential dose-response model discussed in this chapter. Blue is used for a linear representation of the dose-response variation. Green is used to represent a "simple" exponential model was substituted for the traditional dose-response model.

The evidence presented in Figure 4-2 shows much of the same issues plaguing the clustering process for the MONPs that did in the case of CNTs. The increase in the number of sources for the MONPs reflects a greater number of particles analyzed and therefore an overall increase in the number of data points available for analysis for each cluster. Figure 4-2 shows that half of the clusters (10 out of 20) overall are consistent with expectations. This could directly be a

result of an increase in available data for the overall process. Detailed inspection of the data available to the clusters with large uncertainty spreads revealed that the underlying issue could be due to the data collection process, there was a large difference in the range of dose-levels tested. For example, cluster 'I' under the "PMN" response for the MONP particles showed that the lowest dose tested for was at 800 μg/kg while the highest was 20000 μg/kg with no recorded results at dose levels in between. The lack of available data between the 2 dose levels tested denies a detailed look at the complete progression of the response across the dose range. This disparity in dose-levels is even more apparent in cluster 'I" under the "MAC" response, the large spacing between the doses forced a change in the type of dose-response model used, a linear dose-response relationship was substituted for that particular cluster. Cluster "I" under the "TCC" response and cluster "III" under the "TP" response, in both cases an exponential model was used. Additional reasons for the large spreads at higher doses could be attributed to variability in the biological response of the animals sourced and tested, and the lack of data covering a wide range of dose levels. The uncertainty could also be a property of the underlying data used for the analysis, this hypothesis could be verified in the future using newer data sources and checking if the same level of error is observed.

Tables 4-5 and 4-6 below are the model predicted NOAEL value ranges for the 20 clusters generated by the clustering process for the CNTs and MONPs respectively. The NOAEL calculation is not impacted by the dose-response modeling process directly as it corresponds to the highest dose delivered that does not induce any response in the subject. Errors present in the dose-response model similarly do not affect the experimentally derived NOAEL. The process described earlier for the determination of the model predicted NOAELs is highly sensitive to the parameters for the dose-response model whereas the parameters themselves are estimated based on the data used.

Table 4-5. Model based Predicted NOAEL (MP-NOAEL) ranges (95% level of confidence) for Carbon Nanotube particles. The negative range of values indicate the source publications analyzed lack test results at low dose levels. U indicates that the evaluated NOAELs are undefined. The bolded range of values are examples where the confidence spread appears rational, yet the derived NOAEL range includes irrational values (e.g., negative values). All values presented below are in (µg/kg).

Cluster ID	LDH	PMN	Total Protein	Macrophages	Total cell count
A	U	21 – 35	U	**-484 – -270**	37 – 59
B	U	**-1712 – 11825**	U	**-362 – 39**	U
C	U	255 – 401	1378 – 4049	1163 – 38883	U
D	U	**-51 – -42**	U	U	U

Table 4-6. Model based Predicted NOAEL (MP-NOAEL) ranges (95% level of confidence) for Metal Oxide Nanoparticles. The negative range of values indicate the source publications analyzed lack test results at low dose levels. U indicates that the evaluated NOAELs cannot be defined. The bolded range of values are examples where the confidence spread appears statistically valid, yet the derived NOAEL range contains irrational limits (e.g., negative values). All values presented below are in (µg/kg).

Cluster ID	LDH	PMN	Total Protein	Macrophages	Total cell count
I	U	U	13576 – 39616	U	U
II	570 – 998	**-2700 – -1567**	1464 – 2384	U	U
III	**-10340 – -4102**	U	U	54 – 189	2860 – 6179
IV	9369 – 13647	514 – 837	7576 – 26209	9419 – 73834	7334 – 12062

We can observe from Tables 4-5 and 4-6 that the clusters with viable estimations of the MP-NOAEL are most of the same clusters that showed a tight confidence spread. Most of the clusters are missing their designated MP-NOAEL because the calculated range contained negative values and was deemed undefined, since the minimum possible NOAEL value is 0

which corresponds to a case of no exposure. NOAELs of zero (0) do exist for substances that have no threshold of safe exposure. However, a range of entirely negative predicted MP-NOAELs would indicate that some adverse response is present at a dose of zero. Such an occurrence must be either an artifact of limited data to reliably establish model parameters or an artifact of the experimental design. The common link between the clusters with "U" designations is that they displayed large spreads of uncertainty in the model parameters.

There are certain MP-NOAEL measures that have been highlighted in both the tables that are derived from clusters with a tight spread. The main reasoning behind the occurrence of negative values in the range for the NOAELs is the position of the MP-NOAEL response relative to the spacing of the confidence spread. For example, cluster 'D' under the PMN response for the CNTs displays a tight confidence spread consistent with a good fit and low error yet its estimated MP-NOAEL is negative, this is because the MP-NOAEL response falls below the confidence spread spacing and hence their intersection can only possibly occur at negative dose values. The relative shift is also observed in cluster 'B' for the PMN response and clusters 'A' and 'B' for the "MAC" response in the CNTs, cluster 'II' for the "PMN" response and cluster 'III' for the "LDH" response in the MONPs. This shift in the confidence spread occurs mainly in response to the fitted model parameters which are highly dependent on the data analyzed and specifically the dose spacing. Clusters with viable MP-NOAEL values were comprised of source publications that reported the BAL response at multiple low dose (<200 μg/kg) levels and where the spacing of the dose values tested was more even. A potential solution to the "relative shift" issue is to set bounds of the estimated parameters to ensure that the MP-NOAEL response is always within range, this methodology was successful and was used to derive the estimates for the MP-NOAEL for the "PMN" response for the CNTs and MONPs (Data in Tables 4-3 – 4-5), extension of the same technique to the other responses resulted in model parameters which failed the test of significance at the 0.05 level.

The utility of these predicted NOAELs and associated confidence ranges for these nanomaterials is two-fold. First, these values provide a simple means for evaluating our current understanding of nanomaterial safety from the perspective of inhalation exposures, which is of interest for nanomaterial manufacturers as well as health and safety research organizations. For materials or clusters of materials where the data is current insufficient to predict a NOAEL, there is a need to prioritize the continued collection of *in vivo* toxicological data especially for those clusters anticipated to be more toxicologically potent based on the point estimates. Second, these values permit anticipation of future exposure limits. While recommendations have only been published for CNTs and nano-TiO2, these values allow nanomaterial development scientists and toxicologists and policy analysts to anticipate at what levels such future recommendations and potentially future regulatory limits will be set. For instance, since the MP-NOAEL for the silica-containing cluster is approximately half of the Titania-containing cluster, one should currently expect the future recommended exposure limit for nano-silica to be about half of the NIOSH REL for nano-TiO2. Organizations working with such materials could take this provisional guidance into account to help ensure worker safety while the available dataset continues to evolve.

4. CONCLUSIONS

The development and application of a hierarchical clustering algorithm to perform a meta-analysis of engineered nanomaterials *in vivo* pulmonary toxicity in rodents has been documented previously[14]. NOAELs were predicted for the clusters using their respective dose-response curves and correlated to which cluster of particles can be categorized as relatively more or less potent. The model predicted NOAELs were derived for the clusters belonging to a single response as point estimates. These estimates reflect the available data and the dose-response

model used. A detailed sensitivity analysis of the clusters and their predicted NOAELs suggests that the model predicted NOAELs are dependent on the dose range evaluated for a response, the error in the model parameter estimates and the form of the model chosen to represent the dose-response relationship. Reducing the overall observed error between the fitted dose-response model and the data would be beneficial to the derivation of model predicted NOAELs. The MP-NOAEL predictions could be further enhanced and developed by incorporating multiple functional model forms to create a database of models capable of explaining wide ranging dose-response behavior. Multiple clusters showed substantial uncertainty which results from limitations in data availability across the 5 responses analyzed. The uncertainty can be mitigated by constantly expanding the available data to be analyzed and ensuring that BAL responses to particles are recorded at multiple dosage levels. Similarly, researchers could be encouraged to test particles at multiple dose levels further expanded from their initial designs, the additional time and financial investment required might preclude that possibility.

ACKNOWLEDGMENTS

Funding for this work was provided by NIOSH, through grant 1 R03 OH010956-01. This work has also been supported by the National Science Foundation (NSF) and the Environmental Protection Agency (EPA) under NSF Cooperative Agreement EF0830093, Center for the Environmental Implications of NanoTechnology (CEINT).

REFERENCES

1. NIOSH, *Occupational Exposure to Titanium Dioxide.* National Institute for Occupational Safety and Health, 2011.

2. NIOSH, *Occupational Exposure to Carbon Nanotubes and Nanofibers.* National Institute for Occupational Safety and Health, 2013.

3. Yassin, A., F. Yebesi, and R. Tingle, *Occupational exposure to crystalline silica dust in the United States, 1988-2003.* Environmental Health Perspectives, 2005. **113**(3): p. 255-260.

4. Dorato, M.A. and J.A. Engelhardt, *The no-observed-adverse-effect-level in drug safety evaluations: use, issues, and definition(s).* Regul Toxicol Pharmacol, 2005. **42**(3): p. 265-74.

5. Nakanishi, J., et al., *Risk Assessment of the Carbon Nanotube Group.* Risk Anal, 2015. **35**(10): p. 1940-56.

6. Puzyn, T., et al., *Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles.* Nature Nanotechnology, 2011. **6**(3): p. 175-178.

7. Burello, E. and A.P. Worth, *QSAR modeling of nanomaterials.* Wiley Interdisciplinary Reviews-Nanomedicine and Nanobiotechnology, 2011. **3**(3): p. 298-306.

8. Fourches, D., et al., *Quantitative Nanostructure-Activity Relationship Modeling.* Acs Nano, 2010. **4**(10): p. 5703-5712.

9. Fourches, D., D.Q.Y. Pu, and A. Tropsha, *Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles.* Combinatorial Chemistry & High Throughput Screening, 2011. **14**(3): p. 217-225.

10. Gernand, J.M. and E.A. Casman, *Machine Learning for Nanomaterial Toxicity Risk Assessment.* Ieee Intelligent Systems, 2014. **29**(3): p. 84-88.

11. Gernand, J.M. and E.A. Casman, *A Meta-Analysis of Carbon Nanotube Pulmonary Toxicity Studies-How Physical Dimensions and Impurities Affect the Toxicity of Carbon Nanotubes.* Risk Analysis, 2014. **34**(3): p. 583-597.

12. Schulte, P.A., et al., *Current state of knowledge on the health effects of engineered nanomaterials in workers: a systematic review of human studies and epidemiological investigations.* Scand J Work Environ Health, 2019. **45**(3): p. 217-238.

13. Roco, M.C., *The long view of nanotechnology development: the National Nanotechnology Initiative at 10 years.* Journal of Nanoparticle Research, 2011. **13**(2): p. 427-445.

14. Ramchandran, V. and J.M. Gernand, *A dose-response-recovery clustering algorithm for categorizing carbon nanotube variants into toxicologically distinct groups.* Computational Toxicology, 2019. **11**: p. 25-32.

15. Drobne, D., A. Jemec, and Z.P. Tkalec, *In vivo screening to determine hazards of nanoparticles: Nanosized TiO2.* Environmental Pollution, 2009. **157**(4): p. 1157-1164.

16. Choi, J.Y., G. Ramachandran, and M. Kandlikar, *The Impact of Toxicity Testing Costs on Nanomaterial Regulation.* Environmental Science & Technology, 2009. **43**(9): p. 3030-3034.

17. MacPhail, R.C., E.A. Grulke, and R.A. Yokel, *Assessing nanoparticle risk poses prodigious challenges.* Wiley Interdisciplinary Reviews-Nanomedicine and Nanobiotechnology, 2013. **5**(4): p. 374-387.

18. *Feasibility and Challenges of Human Health Risk Assessment for Engineered Nanomaterials*, in *Engineered Nanoparticles and the Environment: Biophysicochemical Processes and Toxicity*.

19. Haber, L.T., et al., *Benchmark dose (BMD) modeling: current practice, issues, and challenges.* Critical Reviews in Toxicology, 2018. **48**(5): p. 387-415.

20. Filipsson, A.F., et al., *The benchmark dose method - Review of available models, and recommendations for application in health risk assessment.* Critical Reviews in Toxicology, 2003. **33**(5): p. 505-542.

21. Starr, T.B., J.I. Goodman, and D.G. Hoel, *Uses of benchmark dose methodology in quantitative risk assessment.* Regulatory Toxicology and Pharmacology, 2005. **42**(1): p. 1-2.

22. Setzer, R.W. and C.A. Kimmel, *Use of NOAEL, benchmark dose, and other models for human risk assessment of hormonally active substances.* Pure and Applied Chemistry, 2003. **75**(11-12): p. 2151-2158.

23. Davis, J.A., J.S. Gift, and Q.J. Zhao, *Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1.* Toxicol Appl Pharmacol, 2011. **254**(2): p. 181-91.

24. 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.* Toxicological Sciences, 2010. **113**(1): p. 226-242.

25. Ma-Hock, L., et al., *Inhalation Toxicity of Multiwall Carbon Nanotubes in Rats Exposed for 3 Months.* Toxicological Sciences, 2009. **112**(2): p. 468-481.

26. Muller, J., et al., *Respiratory toxicity of multi-wall carbon nanotubes.* Toxicology and Applied Pharmacology, 2005. **207**(3): p. 221-231.

27. Shvedova, A.A., et al., *Inhalation vs. aspiration of single-walled carbon nanotubes in C57BL/6 mice: inflammation, fibrosis, oxidative stress, and mutagenesis.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2008. **295**(4): p. L552-L565.

28. Nygaard, U.C., et al., *Single-Walled and Multi-Walled Carbon Nanotubes Promote Allergic Immune Responses in Mice.* Toxicological Sciences, 2009. **109**(1): p. 113-123.

29. Warheit, D.B., et al., *Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats.* Toxicological Sciences, 2004. **77**(1): p. 117-125.

30. Park, E.J., et al., *A single intratracheal instillation of single-walled carbon nanotubes induced early lung fibrosis and subchronic tissue damage in mice.* Archives of Toxicology, 2011. **85**(9): p. 1121-1131.

31. Teeguarden, J.G., et al., *Comparative Proteomics and Pulmonary Toxicity of Instilled Single-Walled Carbon Nanotubes, Crocidolite Asbestos, and Ultrafine Carbon Black in Mice.* Toxicological Sciences, 2011. **120**(1): p. 123-135.

32. Elgrabli, D., et al., *Induction of apoptosis and absence of inflammation in rat lung after intratracheal instillation of multiwalled carbon nanotubes.* Toxicology, 2008. **253**(1-3): p. 131-136.

33. Mercer, R.R., et al., *Alteration of deposition pattern and pulmonary response as a result of improved dispersion of aspirated single-walled carbon nanotubes in a mouse model.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2008. **294**(1): p. L87-L97.

34. Porter, D.W., et al., *Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes.* Toxicology, 2010. **269**(2-3): p. 136-147.

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

36. Shvedova, A.A., et al., *Unusual inflammatory and fibrogenic pulmonary responses to single-walled carbon nanotubes in mice.* Am J Physiol Lung Cell Mol Physiol, 2005. **289**(5): p. L698-708.

37. Ge, C.C., et al., *Acute pulmonary and moderate cardiovascular responses of spontaneously hypertensive rats after exposure to single-wall carbon nanotubes.* Nanotoxicology, 2012. **6**(5): p. 526-542.

38. Nemmar, A., K. Melghit, and B.H. Ali, *The acute proinflammatory and prothrombotic effects of pulmonary exposure to rutile TiO2 nanorods in rats.* Exp Biol Med (Maywood), 2008. **233**(5): p. 610-9.

39. Oberdorster, G., et al., *Role of the alveolar macrophage in lung injury: studies with ultrafine particles.* Environ Health Perspect, 1992. **97**: p. 193-9.

40. Warheit, D.B., et al., *Pulmonary instillation studies with nanoscale TiO2 rods and dots in rats: toxicity is not dependent upon particle size and surface area.* Toxicol Sci, 2006. **91**(1): p. 227-36.

41. Renwick, L.C., et al., *Increased inflammation and altered macrophage chemotactic responses caused by two ultrafine particle types.* Occup Environ Med, 2004. **61**(5): p. 442-7.

42. Rehn, B., et al., *Investigations on the inflammatory and genotoxic lung effects of two types of titanium dioxide: untreated and surface treated.* Toxicology and Applied Pharmacology, 2003. **189**(2): p. 84-95.

43. Grassian, V.H., et al., *Inflammatory response of mice to manufactured titanium dioxide nanoparticles: Comparison of size effects through different exposure routes.* Nanotoxicology, 2007. **1**(3): p. 211-226.

44. Warheit, D.B., et al., *Pulmonary toxicity study in rats with three forms of ultrafine-TiO2 particles: differential responses related to surface properties.* Toxicology, 2007. **230**(1): p. 90-104.

45. Warheit, D.B., et al., *Pulmonary exposures to Sepiolite nanoclay particulates in rats: resolution following multinucleate giant cell formation.* Toxicol Lett, 2010. **192**(3): p. 286-93.

46. Kobayashi, N., et al., *Comparative pulmonary toxicity study of nano-TiO2 particles of different sizes and agglomerations in rats: Different short- and long-term post-instillation results.* Toxicology, 2009. **264**(1-2): p. 110-118.

47. Gustafsson, A., et al., *Lung exposure of titanium dioxide nanoparticles induces innate immune activation and long-lasting lymphocyte response in the Dark Agouti rat.* J Immunotoxicol, 2011. **8**(2): p. 111-21.

48. Oyabu, T., et al., *Dose-dependent pulmonary response of well-dispersed titanium dioxide nanoparticles following intratracheal instillation.* Journal of Nanoparticle Research, 2013. **15**(4).

49. Roberts, J.R., et al., *Toxicological Evaluation of Lung Responses After Intratracheal Exposure to Non-Dispersed Titanium Dioxide Nanorods.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 2011. **74**(12): p. 790-810.

50. Silva, R.M., et al., *Biological Response to Nano-Scale Titanium Dioxide (Tio2): Role of Particle Dose, Shape, and Retention.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 2013. **76**(16): p. 953-972.

51. Gosens, I., et al., *Impact of agglomeration state of nano- and submicron sized gold particles on pulmonary inflammation.* Particle and Fibre Toxicology, 2010. **7**.

52. Cho, W.S., et al., *Inflammatory mediators induced by intratracheal instillation of ultrafine amorphous silica particles.* Toxicology Letters, 2007. **175**(1-3): p. 24-33.

53. Creutzenberg, O., et al., *Toxicity of a Quartz with Occluded Surfaces in a 90-Day Intratracheal Instillation Study in Rats.* Inhalation Toxicology, 2008. **20**(11): p. 995-1008.

54. Roursgaard, M., et al., *Time-response relationship of nano and micro particle induced lung inflammation. Quartz as reference compound.* Human & Experimental Toxicology, 2010. **29**(11): p. 915-933.

55. Morimoto, Y., et al., *Inflammogenic effect of well-characterized fullerenes in inhalation and intratracheal instillation studies.* Particle and Fibre Toxicology, 2010. **7**.

56. Ban, M., et al., *Effect of submicron and nano-iron oxide particles on pulmonary immunity in mice.* Toxicology Letters, 2012. **210**(3): p. 267-275.

57. Pirela, S., et al., *Effects of copy center particles on the lungs: a toxicological characterization using a Balb/c mouse model.* Inhalation Toxicology, 2013. **25**(9): p. 498-508.

58. Zhu, M.T., et al., *Comparative study of pulmonary responses to nano- and submicron-sized ferric oxide in rats.* Toxicology, 2008. **247**(2-3): p. 102-111.

59. Katsnelson, B., et al., *Some Peculiarities of Pulmonary Clearance Mechanisms in Rats after Intratracheal Instillation of Magnetite (Fe3O4) Suspensions with Different Particle Sizes in the Nanometer and Micrometer Ranges: Are We Defenseless against Nanoparticles?* International Journal of Occupational and Environmental Health, 2010. **16**(4): p. 508-524.

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

61. Toya, T., et al., *Pulmonary Toxicity Induced by Intratracheal Instillation of Coarse and Fine Particles of Cerium Dioxide in Male Rats.* Industrial Health, 2010. **48**(1): p. 3-11.

62. Morimoto, Y., et al., *Expression of inflammation-related cytokines following intratracheal instillation of nickel oxide nanoparticles.* Nanotoxicology, 2010. **4**(2): p. 161-176.

63. Xia, T.A., et al., *Decreased Dissolution of ZnO by Iron Doping Yields Nanoparticles with Reduced Toxicity in the Rodent Lung and Zebrafish Embryos.* Acs Nano, 2011. **5**(2): p. 1223-1235.

64. Warheit, D.B., C.M. Sayes, and K.L. Reed, *Nanoscale and Fine Zinc Oxide Particles: Can in Vitro Assays Accurately Forecast Lung Hazards following Inhalation Exposures?* Environmental Science & Technology, 2009. **43**(20): p. 7939-7945.

65. Cho, W.S., et al., *Metal Oxide Nanoparticles Induce Unique Inflammatory Footprints in the Lung: Important Implications for Nanoparticle Testing.* Environmental Health Perspectives, 2010. **118**(12): p. 1699-1706.

66. Ma, J.Y., et al., *Cerium oxide nanoparticle-induced pulmonary inflammation and alveolar macrophage functional change in rats.* Nanotoxicology, 2011. **5**(3): p. 312-325.

67. Peng, L., et al., *Comparative Pulmonary Toxicity of Two Ceria Nanoparticles with the Same Primary Size.* International Journal of Molecular Sciences, 2014. **15**(4): p. 6072-6085.

68. Roursgaard, M., et al., *Acute and Subchronic Airway Inflammation after Intratracheal Instillation of Quartz and Titanium Dioxide Agglomerates in Mice.* Thescientificworldjournal, 2011. **11**: p. 801-825.

69. Lindenschmidt, R.C., et al., *The Comparison of a Fibrogenic and 2 Nonfibrogenic Dusts by Bronchoalveolar Lavage.* Toxicology and Applied Pharmacology, 1990. **102**(2): p. 268-281.

70. Park, E.J., et al., *Induction of Inflammatory Responses in Mice Treated with Cerium Oxide Nanoparticles by Intratracheal Instillation.* Journal of Health Science, 2010. **56**(4): p. 387-396.

71. Wingard, C.J., et al., *Mast cells contribute to altered vascular reactivity and ischemia-reperfusion injury following cerium oxide nanoparticle instillation.* Nanotoxicology, 2011. **5**(4): p. 531-545.

72. Xue, L.X., et al., *Pulmonary Toxicity of Ceria Nanoparticles in Mice After Intratracheal Instillation.* Journal of Nanoscience and Nanotechnology, 2013. **13**(10): p. 6575-6580.

73. Ma, J.Y.C., et al., *Interactive effects of cerium oxide and diesel exhaust nanoparticles on inducing pulmonary fibrosis.* Toxicology and Applied Pharmacology, 2014. **278**(2): p. 135-147.

74. Minarchick, V.C., et al., *Pulmonary Cerium Dioxide Nanoparticle Exposure Differentially Impairs Coronary and Mesenteric Arteriolar Reactivity.* Cardiovascular Toxicology, 2013. **13**(4): p. 323-337.

75. Dick, C.A.J., et al., *The role of free radicals in the toxic and inflammatory effects of four different ultrafine particle types.* Inhalation Toxicology, 2003. **15**(1): p. 39-52.

76. Gelli, K., M. Porika, and R.N.R. Anreddy, *Assessment of Pulmonary Toxicity of MgO Nanoparticles in Rats.* Environmental Toxicology, 2015. **30**(3): p. 308-314.

77. Zhang, Q.W., et al., *Differences in the extent of inflammation caused by intratracheal exposure to three ultrafine metals: Role of free radicals.* Journal of Toxicology and Environmental Health-Part a-Current Issues, 1998. **53**(6): p. 423-438.

78. Horie, M., et al., *Comparison of acute oxidative stress on rat lung induced by nano and fine-scale, soluble and insoluble metal oxide particles: NiO and TiO2.* Inhalation Toxicology, 2012. **24**(7): p. 391-400.

79. Kadoya, C., et al., *Analysis of Bronchoalveolar Lavage Fluid Adhering to Lung Surfactant-Experiment on Intratracheal Instillation of Nickel Oxide with Different Diameters.* Industrial Health, 2012. **50**(1): p. 31-36.

80. Ogami, A., et al., *Pathological features of different sizes of nickel oxide following intratracheal instillation in rats.* Inhalation Toxicology, 2009. **21**(8-11): p. 812-818.

81. Cho, W.S., et al., *Differential pro-inflammatory effects of metal oxide nanoparticles and their soluble ions in vitro and in vivo; zinc and copper nanoparticles, but not their ions, recruit eosinophils to the lungs.* Nanotoxicology, 2012. **6**(1): p. 22-35.

82. Cho, W.S., et al., *Progressive severe lung injury by zinc oxide nanoparticles; the role of Zn2+ dissolution inside lysosomes.* Particle and Fibre Toxicology, 2011. **8**.

83. Nishi, K., et al., *Expression of cytokine-induced neutrophil chemoattractant in rat lungs by intratracheal instillation of nickel oxide nanoparticles.* Inhalation Toxicology, 2009. **21**(12): p. 1030-1039.

84. Slob, W., *Dose-response modeling of continuous endpoints.* Toxicological Sciences, 2002. **66**(2): p. 298-312.

85. Dorato, M.A. and J.A. Engelhardt, *The no-observed-adverse-effect-level in drug safety evaluations: Use, issues, and definition(s).* Regulatory Toxicology and Pharmacology, 2005. **42**(3): p. 265-274.

86. Pauluhn, J., *Multi-walled carbon nanotubes (Baytubes (R)): Approach for derivation of occupational exposure limit.* Regulatory Toxicology and Pharmacology, 2010. **57**(1): p. 78-89.

87. Oberdorster, G., *Lung Particle Overload - Implications for Occupational Exposures to Particles.* Regulatory Toxicology and Pharmacology, 1995. **21**(1): p. 123-135.

Chapter 5

CONCLUSIONS

The rapid development of nanomaterials necessitates advancement in nanoparticle toxicology testing methodologies to keep pace and prevent collateral damage from exposure to nanomaterials. The toxic potential of newly synthesized nanomaterials needs to be assessed in a quick manner to enable prioritized testing of nanomaterials deemed potent. Computational approaches offer the resources and capability to assist in making time efficient decisions with regards to mitigating the synthesis of hazardous species of nanomaterials.

A hierarchical clustering methodology was developed that is capable of categorizing nanomaterials based on the similarity of their dose-response and recovery relationships. The algorithm was designed with the objective of modeling the dose-response and recovery of *in-vivo* pulmonary toxicity exposure to nanomaterials across 5 toxicological endpoints chosen. *In-vivo* toxicity tests are better suited to capture the biological response to nanomaterial exposure than *in-vitro* tests. The high variance in the *in-vivo* testing results was overcome using a meta-analysis-based approach, the datasets were assembled using the reported results from multiple peer-reviewed publications. 2 distinct datasets of nanomaterials were analyzed using the clustering algorithm, CNTs and MONPs. The clustering methodology was successfully able to determine that there are 4 toxicologically distinct clusters amongst both sets of nanomaterials analyzed for the 5 endpoints. The constituent particle attributes and associated potency of the derived clusters was examined to identify what features significantly impact the toxicity of the cluster.

Patterns were observed in the physical properties between the CNT clusters leading to the assignment of provisional labels, the repetition of data across the endpoints and the lack of sufficient characterization data prevents the designation of statistically meaningful labels for the CNT clusters. Physical characterization-based distinction could not be made for the MONP

clusters. Detailed analysis of the potency of the derived clusters for CNTs revealed that all 4 CNT clusters demonstrated signs associated with cell damage or increased immune activity, certain groups of CNT particles (cluster "B") are capable of causing both effects upon exposure. Analysis of the potency of the MONP clusters revealed that 3 out of the 4 clusters (I, II, IV) showed signs of inflammation and cell damage, these clusters were primarily composed of Iron oxide and Silica nanoparticles. The 4th cluster consisting primarily of Zinc oxide nanoparticles was found to show the least toxic potential based on our analysis, survey of the published literature revealed differing outlooks on the toxicity of nano-Zinc oxide. Comparison of the individual cluster components across the 5 responses for the MONPs showed they differ in their chemical composition; more information would be required before the differences in potency can be correlated to the mechanistic toxicity.

A dose-response model-based NOAEL was derived for the clusters generated for both sets of nanomaterials. The MP-NOAEL derivation process takes the form and type of the dose-response model into account. By taking into account the shape of the dose-response model some of the shortcomings of the traditional experimentally derived NOAEL process could potentially be overcome. The PMN based MP-NOAEL estimate derived for the MONP cluster consisting of nano-Zinc oxide was found to be the lowest, this result is in agreement with the potency assessment previously. A sensitivity analysis conducted on the MP-NOAEL process revealed the dependency of the predicted NOAELs on the dose range evaluated, dose spacing and sample size available for testing. These shortcomings are consistent with those of the traditional NOAEL derivation process. Improvements to the data testing and collection process coupled with the establishment of a database of different dose-response models would make it feasible to derive more accurate MP-NOAEL estimates.

This study presents one of the few purely *in-vivo* pulmonary toxicity based meta-analysis of nanomaterials. The results presented can be used to generate hypotheses about the relationship

between the physicochemical features of the nanomaterials and their toxicity. The data used by the analysis conveyed in this document is fully based on early and short-term biological response to nanomaterial exposure, the availability of data to analyze acute and short-term exposures exceeds the pathological data. This study also poses potential policy implications, the establishment of toxicological categories within nanomaterial types could potentially lessen the burden and cost of toxicological testing for manufacturers and regulators. Categorical exposure limits such as the MP-NOAEL described by this study could assist in prioritizing the testing process for engineered nanoparticles based on their affiliation with the more potent clusters, these limits also allow for the anticipation of future exposure limits.

VITA

Vignesh Ramchandran was born and raised in Maharashtra, India. He holds a Bachelor of Technology degree in Chemical Engineering from the National Institute of Technology, Tiruchirappalli, India (2011). He earned his PhD degree in Energy and Mineral Engineering (with a Minor in Statistics) at The Pennsylvania State University under the supervision of Dr. Jeremy M. Gernand. His research interests include the application of Statistical Learning, Data Science, and Machine Learning to develop data-driven solutions to real-world problems. He is currently working as a Senior Data Scientist at a leading Financial Services provider in India, where he is focused on building underwriting models and scorecards to mitigate lending risk.