Journal of Nanobiotechnology

RESEARCH Open Access

Dynamic QSAR modeling for predicting in vivo genotoxicity and inflammation induced by nanoparticles and advanced materials: a time-dose-property/response approach

Michalina Miszczak[1], Kabiruddin Khan[1], Pernille Høgh Danielsen[2], Keld Alstrup Jensen[2], Ulla Vogel[2], Roland Grafström[3,4] and Agnieszka Gajewicz-Skretna[1*]

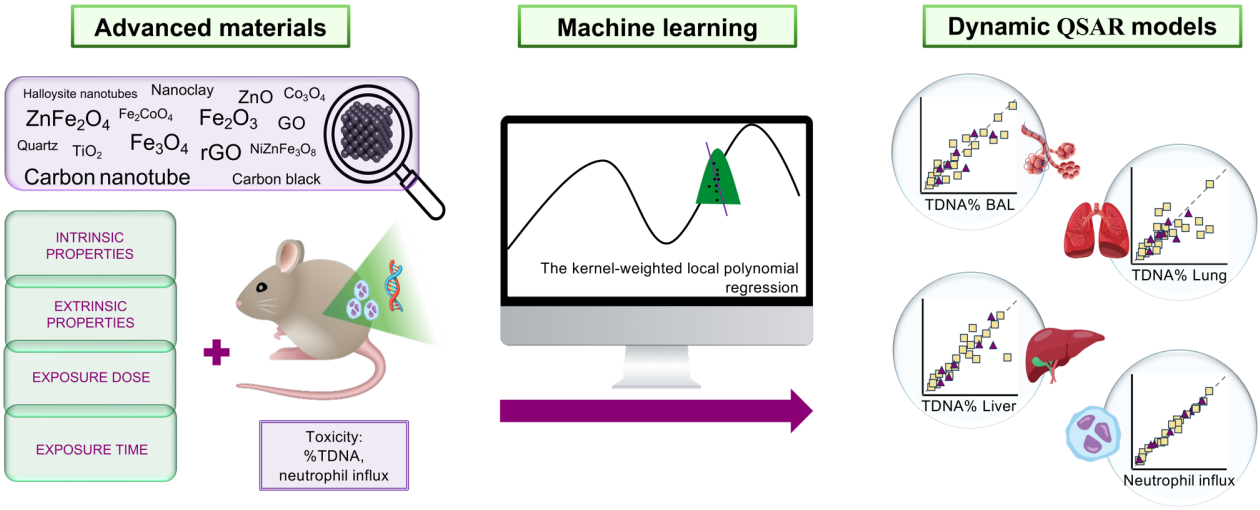
Abstract

Predicting the health risks of nanoparticles (NPs) and advanced materials (AdMa) is a critical challenge. Due to the complexity and time-consuming nature of experimental testing, there is a reliance on in silico methods such as quantitative structure-activity relationship (QSAR), which, while effective, often fail to capture the dynamic nature of material activity over time—essential for reliable risk assessment. This study develops dynamic QSAR models using machine learning to predict toxicological responses, such as inflammation and genotoxicity, following pulmonary exposure to 39 AdMa across various post-exposure time points and dose levels. By incorporating exposure time, administered dose, and material properties as independent variables, we successfully developed time-dose-property/response models capable of predicting AdMa-induced in vivo genotoxicity in bronchoalveolar lavage fluid cells, lung and liver tissue, and inflammation in terms of neutrophil influx into the lungs of mice. Key factors driving AdMa-induced toxicity were identified, including exposure dose, post-exposure duration time, aspect ratio, surface area, reactive oxygen species generation, and metal ion release. The time-dose-property/response modeling paradigm presented here provides a practical and robust approach for predicting in vivo genotoxicity and inflammation and supports the comprehensive risk assessment of morphologically diverse AdMa.

*Correspondence:
Agnieszka Gajewicz-Skretna
agnieszka.gajewicz@ug.edu.pl

Full list of author information is available at the end of the article

Graphical Abstract



Keywords Advanced materials, Genotoxicity, Inflammation, Computational toxicology, Data-driven modeling, Dynamic modeling, Time-dose-property/response modeling

Background

The ongoing advancement of modern technology has spurred the development and application of new nanoparticles (NPs) and advanced materials (AdMa), which exhibit exceptional nanoscale properties that surpass those of traditional materials [1–3]. This progress, however, has caused considerable concern regarding its potential effect on human health and the environment, particularly given their nanoscale dimensions and distinctive quantum mechanical properties [4, 5]. A substantial body of evidence from in vitro and in vivo animal studies has demonstrated that NPs and AdMa can elicit toxic effects—manifested as inflammatory responses and oxidative stress—which, in turn, are linked to adverse health effects, including cancer, fibrosis, and cardiovascular disease [6–12]. The toxic effects of NPs and AdMa are influenced by a number of factors, including size, morphology, exposure time, dose, and route of exposure [13–16]. However, it is important to note that the activation of molecular mechanisms and the likelihood of toxic effects are heavily influenced by dose level, exposure duration and post-exposure time [17–21].

Among all possible exposure routes, inhalation is considered to be the most significant for NPs and AdMa, as airborne exposure to particles of these materials can easily enter the respiratory system, deposit in the deep end of the airways and subsequently be distributed to other organs such as liver and spleen [10]. Due to the complexity of these materials and their potential to cause harm, it is essential to investigate the pathways by which they exert toxic effects. Adverse outcome pathways (AOPs) describe the causal relationships that link a sequence of biological events, starting with an initial molecular interaction between a material and a biological target, leading to adverse health outcomes. These pathways help identify early key events, such as inflammation and genotoxicity, that are directly linked to more serious outcomes such as cardiovascular disease, fibrosis, and cancer [22, 23]. Inflammation, as an early key event, plays a critical role in triggering disease processes, and its chronic activation can lead to long-term health effects. Genotoxicity, a characteristic commonly associated with carcinogens, is another key event that can promote cancer development [24]. Both inflammation and genotoxicity have been suggested as key characteristics of carcinogens and as key events in the proposed AOP for NPs and AdMa induced cancer [25, 26].

In the past, toxicity assessments predominantly relied on the outcomes of animal experimentation, which has been the gold standard in toxicological research for decades. Although this approach is largely preferred due to its effectiveness, it has significant drawbacks, such as lengthy procedures, high costs, and ethical concerns [27]. Moreover, in vivo studies need to be complemented with research targeting the molecular mechanisms underlying substance toxicity and then extrapolating this information to other species [28]. To overcome these limitations and adhere to the 3Rs principle (reduction, refinement, and replacement) [29], there has recently been a shift towards adopting new approach methodologies. These methodologies encompass various strategies, including in vitro, in chemico, and in silico methods, all

designed to enhance the predictive accuracy of nontesting approaches for risk assessment [30]. This shift aligns with the principles of the New Generation Risk Assessment (NGRA) framework, which has advocated for a phased and multidimensional approach to risk assessment [31, 32]. The implementation of contemporary strategies facilitates bridging gaps in existing data and curtailing the necessity for additional animal testing [11, 32, 33]. The overarching objective is to understand the mechanisms and pathways of toxicity while simultaneously reducing reliance on animal testing. Among the in silico methods endorsed by the Integrated Approaches to Testing and Assessment (IATA), quantitative structure-activity relationship (QSAR) models, as described in the Organisation for Economic Co-operation and Development (OECD) documentation [34], are particularly popular [35–37].

QSAR modeling is a data-driven, predictive modeling paradigm that uses mathematical equations to capture the causal relationship between structural differences in chemical compounds and their biological activity. These models are calibrated using experimental data and molecular descriptors that provide detailed information about physicochemical properties [38, 39]. Over the past two decades, computational nanotoxicology has gained considerable attention, particularly QSAR models used to evaluate the biological effects of NPs and AdMa. Early advances have laid a solid foundation for exploring the relationships between the structure of nanoparticles (e.g., metals, metal oxides, and carbon materials) and their biological activity or properties [40–47]. Among the pioneering systematic, computational efforts, Puzyn and co-authors [40] conducted one of the first studies to predict the cytotoxicity of metal oxide (MeOx) nanoparticles to E. coli bacteria. The authors established a causal relationship between the enthalpy of metal cation formation (ΔH_{Me+}) and the toxicity of MeOx nanoparticles, providing quantitative insights into how the structural parameters of NPs influence their toxicity [40]. Concurrently, Zhang and co-authors [43] demonstrated that the toxicity of MeOx nanoparticles can be predicted based on the overlap of their conduction band energy levels with the cellular redox potential. The authors established a robust in vitro and in vivo hazard ranking model that linked conduction band energy and particle dissolution to explain redox-related toxicological mechanisms in the lung, including oxidative stress and pulmonary inflammation [43]. Another study reported a nano-QSAR model for predicting the cell viability of superparamagnetic iron oxide nanoparticles (SPIONs). Using automated machine learning (Auto-ML) approaches, the toxicity response induced by SPIONs was modeled as a function of two key parameters: the magnetic core and particle size. The authors demonstrated that SPIONs with magnetite cores were non-toxic. In contrast, those with maghemite cores exhibited size-dependent toxicity, i.e., being toxic at smaller sizes and non-toxic when their size exceeded approximately 15 nm [48]. In another study, Kotzabasaki and colleagues [49] investigated the potentially causal relationship between the physicochemical properties of multi-walled carbon nanotubes (MWCNTs) and their genotoxic effects. Their high-performance classification models identified three key physicochemical factors: length, zeta average, and purity, which were strongly associated with MWCNT-induced genotoxicity. Recently, Merugu and colleagues [50] investigated how the structural properties of MWCNTs influence lung inflammation at the tissue level. To identify the structural properties of carbon nanotubes that perturb the acute phase response signaling pathway in the lungs of mice exposed to MWCNTs, the authors developed a QSAR model anchored in an AOP framework (AOP-anchored QSAR). The results indicated that carbon nanotubes with smaller aspect ratios and higher specific surface areas were less toxic, as evidenced by their reduced effects on the acute phase signaling pathway [50]. The majority of QSAR models developed in recent years are static, meaning they are tailored to specific test conditions of the endpoint, such as exposure duration, administered dose, and other factors. However, given the ongoing need to advance and refine in silico techniques for risk assessment, it is essential to account for the dynamic changes that chemicals and materials undergo over time. Prolonged exposure to AdMa may trigger molecular mechanisms that differ from those observed in short-term exposures, posing a challenge to the development of models that predict long-term effects due to the paucity of available data. Moreover, the development of separate models for specific time points and/or doses fails to capture the full temporal dynamics of these changes. To address this gap, the present study aims to demonstrate the effectiveness of machine learning (ML) in developing dynamic QSAR models that accurately reflect the evolving behavior of a diverse range of AdMa over time and across varying dose levels. By incorporating exposure time and administered doses as independent variables, we seek to comprehensively capture the temporal intricacies of AdMa toxicity, thereby enhancing our understanding of the origins, dynamics, and mechanisms driving AdMa-induced toxicity.

Materials and methods

Materials and toxicity data

Information about the nanomaterials and animal studies were previously published: nanoclays [51], cobalt ferrite NPs [52], DQ12 and titanium dioxide [53–55], carbon blacks [56, 57], zinc oxide [58], carbon nanotubes [8], halloysite nanotubes [59], iron oxides [60], and

graphene-based materials [52]. However, for reactive oxygen species (ROS), new harmonized data were generated using the acellular version of the 2´,7´dichloro-dihydrofluorescein diacetate (DCFH2-DA) assay following 3 h of incubation [61]. The ROS formation efficiency per mass unit and per specific surface area (SSA) were normalized to carbon black (Printex 90), which was included as benchmark material in each run. For dissolution and solubility, there was a large data gap, and available data had been generated using different methods and dissolution test media. Therefore, a ranking approach was developed with the intention to give a similar ranking of reported dissolution rates and 24-hour solubility of the core material in phagolysosomal fluids [using unpublished and published data from [62–64]. In addition, we made a separate scaling considering the release of transition elements and heavy metals after 24-hours pulmonary dissolution considering observed dissolution and/or elemental concentrations in the materials. The ranking approach is given in Table S1. The applied nanomaterial physicochemical properties are shown in Table 1.

These animal studies were conducted over a period of approximately 15 years, primarily using the same experimental setup. In brief, eight to ten-week-old female C57BL/6J BomTac mice from Taconic (Ejby, Denmark) ($N = 6–9$) were exposed to nanomaterials dispersed in 2% serum [65] by intratracheal instillation [66] at 2–3 dose levels and followed for 1, 3, 28, 90, and 180 days post-exposure. Not all materials were assessed at each time point, and limited data were available especially for the 180-day post-exposure time point due to variations in the design of specific studies. Dose selection was based on either data from the scientific literature or initial pilot studies designed to identify the highest dose that would avoid overt toxicity while ensuring compliance with the highest standards of animal welfare [20]. An overview of the administered doses and exposure time points for the animal studies is provided in Table S2 (Supplementary Materials). In most studies, 2% mouse serum (sibling serum) was used as a vehicle, but other validated vehicles were also used [65]. Vehicle-exposed negative controls and mice exposed to the benchmark particle carbon black Printex90 [53], were included in parallel. Mice were exposed via a single intratracheal instillation under isoflurane sedation, as previously described [66, 67]. Inflammation was assessed by quantifying neutrophil influx into bronchoalveolar lavage (BAL) fluid, as previously described [57]. Levels of DNA strand breaks were assessed in BAL fluid cells, as well as in lung and liver tissue, by measuring tail % DNA using the comet assay, as previously described [68].

Molecular descriptors

To capture the variations within a diverse set of AdMa, we used intrinsic and extrinsic properties obtained from the experimental characterization of materials as independent variables in QSAR modeling. We also considered theoretical descriptors derived from different molecular representations to ensure that no critical features potentially influencing AdMa toxicity were overlooked. Considering the significant structural diversity of AdMa—including differences in chemical composition, size, and complexity—we focused on descriptors that did not require the preoptimization of molecular geometry and were, therefore, independent of conformational issues. These descriptors, commonly referred to as 0D-2D descriptors, provided comprehensive constitutional and topological information, including details about the molecule's size, shape, symmetry, and branching. The 0D-2D descriptors were calculated using Dragon software (version 7.0) [69].

It is crucial to emphasize that the selection of molecular descriptors for machine learning modeling must be carefully tailored to the specific purpose, i.e., the endpoint(s) for which the models are developed. Consequently, the selected descriptors should be endpoint-oriented to ensure the relevance and predictive accuracy of the developed models. In the context of NPs and AdMa, genotoxicity and inflammatory effects are primarily driven by over production of ROS, which in turn leads to oxidative stress - a key mechanism of toxicity [56, 70–78]. In addition, the dissolution and subsequent release of toxic metals (including transition and heavy metals) play a critical role, influencing both the biological response(s) and the recovery process through clearance mechanisms [16, 79–81]. Beyond ROS generation and dissolution-related phenomena, genotoxic and inflammatory responses are also strongly influenced by the intrinsic physicochemical properties of NPs and AdMa, such as particle size, aspect ratio, shape, specific surface area (SSA), and surface-dependent properties [70, 73–78, 82]. Furthermore, as highlighted in the Introduction, the activation of cellular pathways and the overall likelihood of toxic effects are strongly influenced by both the administered dose and the post-exposure time point. Descriptors that capture all of these critical features were identified through a comprehensive review of the literature and expert knowledge. To further refine the selection of 0D-2D descriptors potentially relevant for modeling genotoxic and inflammatory effects, the backward elimination technique was used. This method involves an iterative process in which the least important features are progressively removed based on their influence on model performance metrics-specifically, R^2 and RMSE-assessed during model calibration. More specifically, the backward elimination technique starts with a full set of descriptors

Table 1 Physicochemical properties of nanoparticles and advanced materials

ID	Material name	Diameter (nm)	Length (nm)	Aspect ratio	BET_SSA	ROS/mass unit	Nanodimensions	Ranked release
Bentonite	Nanoclay	7.8	650.0	83.33	73.59	0.00003	1	0
Co_2O_3	Prom_Co_3O_4_MiLE19 A-9 1%	40.0	40.0	1.00	52.00	0.46245	3	3
DQ12	Quartz DQ12	225.0	225.0	1.00	10.10	0.00118	3	0
Fe_2CoO_4 1/1	Prom_Co1.5 Fe1.5_MiLE019 C-9 0.75%	40.0	40.0	1.00	132.00	0.00884	3	4
Fe_2CoO_4 1/3	Prom_Co2.25 Fe0.75_MiLE019 F-9 1%	40.0	40.0	1.00	49.00	0.00357	3	3
Fe_2CoO_4 3/1	Prom_Co0.75 Fe2.25_MiLE019 D-9 1%	40.0	40.0	1.00	103.00	0.07778	3	4
Fe_3O_4	Prom_Fe_3O_4_MiLE019 E-9 1.1%	40.0	40.0	1.00	50.00	0.00048	3	2
Flamruss101	Carbon black	95.0	95.0	1.00	23.00	0.10809	3	0
MKN-A015	TiO_2, anatase TiO_2 small particle	22.0	22.0	1.00	84.60	0.00049	3	0
MKN-A100	TiO_2, anatase TiO_2 large particle	31.0	31.0	1.00	73.50	0.00065	3	0
Nanofil® 3000 SE	Nanoclay (organically modified)	1.0	300.0	300.00	2.74	0.00593	1	0
Nanofil® 9	Nanoclay (organically modified)	1.0	300.0	300.00	24.07	0.00003	1	0
NM-110	ZnO	85.0	85.0	1.00	14.00	0.00151	3	5
NM-401	Carbon nanotube, MWCNT	67.0	4048.0	60.42	18.00	0.02944	2	2
NM-403	Carbon nanotube, MWCNT	12.0	443.0	36.92	135.00	1.83469	2	4
NN	Halloysite nanotube, Nanoclay, $Al_2Si_2O_5(OH)_4$	61.0	250.0	4.10	25.60	0.00080	2	0
NN-etched	Halloysite nanotube, Nanoclay, SiO_x	27.0	100.0	3.70	128.00	0.00053	2	0
NRCWE-001	TiO_2, rutile	10.0	10.0	1.00	99.00	0.00045	3	0
NRCWE-006	Carbon nanotube, MWCNT, Mitsui-7	74.0	5700.0	77.03	26.00	0.03491	2	3
NRCWE-018	Alpha Fe_2O_3 (Particle)	40.0	40.0	1.00	27.70	0.00017	3	2
NRCWE-019	Alpha Fe_2O_3 (Rod)	85.0	425.0	5.00	27.40	0.00044	2	2
NRCWE-020	$NiZnFe_3O_8$	20.0	20.0	1.00	104.00	0.00206	3	4
NRCWE-021	$ZnFe_2O_4$	22.5	22.5	1.00	87.70	0.00035	3	4
NRCWE-022	$NiFe_2O_4$	25.0	25.0	1.00	86.90	0.00564	3	4
NRCWE-025	TiO_2, rutile TiO_2	38.0	38.0	1.00	28.20	0.00090	3	0
NRCWE-040	Carbon nanotube, MWCNT, pristine	22.1	518.9	23.48	150.00	0.41297	2	3
NRCWE-041	Carbon nanotube, MWCNT, -OH	26.9	1005.0	37.36	152.00	1.00057	2	3
NRCWE-042	Carbon nanotube, MWCNT, -COOH	30.2	723.2	23.95	141.00	0.28229	2	3
NRCWE-043	Carbon nanotube, MWCNT, pristine	55.6	771.3	13.87	82.00	0.44217	2	4
NRCWE-044	Carbon nanotube, MWCNT, -OH	32.7	1330.0	40.67	74.00	0.33430	2	4
NRCWE-045	Carbon nanotube, MWCNT, -COOH	30.2	1553.0	51.42	119.00	0.52362	2	4
NRCWE-051	Carbon nanotube, SWCNT, pristine	1.0	17500.0	17500.00	442.60	0.22456	2	4
NRCWE-054	Carbon nanotube, SWCNT, -COOH	1.0	17500.0	17500.00	370.80	1.66761	2	4
NRCWE-055	Carbon nanotube, SWCNT, pristine, short	1.0	2000.0	2000.00	453.10	1.29452	2	4
NRCWE-056	Carbon nanotube, SWCNT, -OH, short	1.0	2000.0	2000.00	373.40	1.14844	2	4
NRCWE-058	Graphene Oxide	1.7	2500.0	1474.93	411.00	0.62487	1	0
NRCWE-059	Reduced Graphene Oxide	1.7	1500.0	884.96	411.00	0.33750	1	0
Printex90	Carbon black	14.0	14.0	1.00	338.00	1.00000	3	0
XE2B	Carbon black	30.0	30.0	1.00	1142.00	5.49524	3	0

and progressively eliminates features until only descriptors with a significant contribution to predictive accuracy remain. As previously noted, data across different post-exposure time points and administered doses were modeled together to develop dynamic models with broad applicability domains. This means that, in the calibration and validation process of each of the four QSAR models, two constant descriptors were used: post-exposure time (expressed in days) and doses (expressed in μg/mouse). List of considered descriptors is provided in Table S3.

Modeling technique

Nowadays, numerous data-driven ML algorithms have gained considerable acceptance and are now well-established, including linear regression, logistic regression, support vector machines, decision trees, random forests, naive Bayes, *k*-nearest neighbors, and artificial neural networks. Accordingly, the selection of the optimal ML-based algorithm should consider factors such as the data's size, distribution, complexity, dimensionality, and noise. Given the diverse AdMa set used in this study, which encompassed organic and inorganic materials, nonparametric similarity-based ML methods (SBMs) were considered particularly well suited for effectively mining the intricate relationships between the structure and properties of materials and their biological effects. In light of our previously positive experience with the locally weighted least squares kernel regression (KwLPR) algorithm [83, 84], we chose this method a priori. Conceptually, the KwLPR operates on the assumption that any function within a given dataset can be adequately approximated by a low-degree polynomial, with the kernel serving as a weighting function. This method combines the simplicity of linear regression with the flexibility of nonlinear regression, fitting a weighted linear model locally at each data point and considering only the nearest observations. The kernel function (K) and a smoothing parameter (bandwidth, h) determine the shape and size of the local neighborhood. A weighting function assigns relative importance to each neighboring point based on their distance from the query point, with the closest, most similar points receiving the highest weight, which diminishes with increasing distance [83, 84]. To provide the KwLPR algorithm with the requisite flexibility to model complex datasets, we used a variety of kernel types, including Gaussian, truncated Gaussian, Epanechnikov, and uniform, in conjunction with an array of parameters to adjust the local neighborhood size and shape. The primary parameter that influences the width of the kernel function, referred to as the smoothing parameter or bandwidth (h), is of paramount importance in determining the neighborhood size. Although several methods exist for specifying the h parameter, including fixed, adaptive nearest-neighbor, and generalized nearest-neighbor approaches, a fixed bandwidth that remains constant across the entire dataset is the most commonly used. However, a fixed bandwidth may not effectively capture spatial distribution heterogeneity, resulting in the formation of empty neighborhoods and underperformance at boundary points, a phenomenon known as the boundary effect. To address this issue, the smoothing parameter can be adjusted dynamically based on the local density of observations, which are referred to as adaptive or variable-bandwidth kernel density estimation methods. In the adaptive approach, the bandwidth changes as a function of the fitting data points; in regions of high density, a narrower kernel is used to reduce bias, while in low-density regions, a broader kernel is used to reduce variance. In this study, both fixed and adaptive nearest-neighbor bandwidth kernel density estimation methods were considered. It is also important to note that the effectiveness of other regression-based ML techniques, such as the classification and regression trees (CART), random forest (RF), support vector regression (SVR), *k*-nearest neighbors (kNN), and the distance weighted (weighted-kNN) *k*-nearest neighbor algorithm, were evaluated during the optimization of the dynamic QSAR models.

Statistical analyses were conducted using *R* software (version 4.3.2) [85], specifically using the *KwLPR.RMD* code [84], along with the '*np*' and '*ggplot2*' packages [86–88]. To ensure that the experimental data and modeling results are findable, accessible, interoperable, and reusable (FAIR), the input files, the R code supporting the reported results, and the corresponding output files are provided in the Supplementary Materials.

Model calibration and validation

Four data-driven models were developed to predict in vivo genotoxicity in BAL fluid cells, lung and liver tissue, and pulmonary inflammation in mice, respectively. In the first three cases, genotoxicity was quantified as the percentage of DNA in the tail (TDNA%) in the comet assay. In the fourth model, inflammation was represented by the total number of neutrophils in the BAL fluid, expressed as a decimal logarithm.

All of the data-driven models (derived from a richly diverse dataset of 39 AdMa) were simultaneously trained across various dose levels and exposure times. It is important to highlight that for each data point, raw toxicological data from animals exposed to vehicle alone were processed, and endpoint values were calculated as the average of groups consisting of six to nine animals. To conduct a comprehensive assessment of the predictive models' efficacy, the initial AdMa datasets were divided into training and validation sets (VS). Specifically, ~80% of the materials were allocated for model development (TS), while the remaining ~20% were set aside for model

evaluation (VS). A sorted, response-based division algorithm was used to ensure a balanced representation in both sets. Specifically, the VS was constituted by selecting every fourth or fifth data point from the endpoint vector, arranged in ascending order. The remaining data points formed the TS. The two data sets for each of the four models are provided in Tables S4-S7.

Validation was conducted in accordance with the standards set forth by the OECD guidelines [82], thereby ensuring a comprehensive assessment of the data-driven model's goodness-of-fit, robustness, and predictive abilities. The coefficient of determination (R^2) was used to assess the model fit, indicating the proportion of variability in the dependent variable explained by the independent variables in the training set (TS). Furthermore, the root mean square error of calibration ($RMSE_c$) was used to quantify deviations between the observed and predicted values. The model's flexibility was evaluated through internal validation, with the results quantified using the cross-validation coefficient (Q^2_{LOO}) and the root mean square error of cross-validation ($RMSE_{LOO}$). The true predictive performance of the model was examined with unseen data, i.e., data that were not involved in the model's calibration. External validation metrics included the external validation coefficient (Q^2_{ext}), the concordance correlation coefficient (CCC), the root mean square error of external validation ($RMSE_{ext}$), and the mean absolute error (MAE) [89–92].

To ensure that the QSAR models were not the result of random correlations and to confirm their statistical significance, we performed a Y-scrambling test [93–95]. In this procedure, 500 randomized models were calibrated using the same descriptor sets as in the original models, but with the response variable randomly shuffled in each iteration. In general, QSAR models derived from Y-randomization tests are expected to yield low R^2 and Q^2_{LOO} values. Occasionally, although rarely, higher values of R^2 and Q^2_{LOO} may be obtained due to random correlations and/or structural redundancy within the training data set. However, if the Y-randomization test consistently generates models with relatively high R^2 and Q^2_{LOO} values, this indicates that the chosen modeling approach is inadequate to generate a reliable QSAR model from the provided dataset [93–95].

Applicability domain

To define the scope and associated limitations of the model, its applicability domain was determined using the leverage approach [89, 96]. This calculates the Mahalanobis distance of each compound to the centroid of the TS within the features space. Thus, the leverage value (h_i) indicates the structural (dis)similarity of each compound to the center of the TS and assesses the degree of influence of the training data points on the predictive model. A Williams plot offers a visual representation of this relationship, plotting leverage values against cross-validated standardized residuals to identify response outliers and/or structurally distinct chemicals that fall outside the applicability domain. The critical leverage value (h^*), representing the cut-off distance from the centroid, is generally set at $h^* = 3(M+1)/N$, where M is the number of descriptors (i.e., independent variables) in the model, and N is the number of TS compounds [97]. The boundaries of the applicability domain with respect to the response space were defined by the limit of ± 3 standardized cross-validated residuals. Predictions were considered reliable and trustworthy, providing a high level of assurance only when the query compound fell within a square area bounded by ± 3 standard deviation units and the leverage threshold (h^*) [97]. However, as discussed by Jaworska and co-authors [97], high leverage values do not necessarily indicate outliers. Rather, they can vary significantly in their influence on the regression results depending on how closely they fit the established model. When high-leverage points fit the regression well, as reflected by low standardized residuals, they are typically referred to as "good high-leverage points" or "good influence points. Such points contribute positively to the model by increasing its stability, broadening the calibration range, and improving the accuracy of parameter estimates; therefore, they should be retained in the model. Conversely, high-leverage points associated with large residuals-indicating a poor fit to the regression-are classified as "bad high-leverage points" or "bad influence points. These points can adversely affect the model, potentially reducing predictive accuracy [97].

Results

Key features governing in vivo genotoxicity and inflammatory response of AdMa

The key descriptors for the BAL-cell genotoxicity model were the aspect ratio, defined as the ratio between the nanomaterial's length and diameter, and the specific surface area, measured using the Brunauer-Emmett-Teller method (BET_SSA) (Fig. 1).

In the case of the liver genotoxicity model, the most crucial descriptors were the aspect ratio, the production of ROS, and a nano-dimension descriptor that defined the AdMa shape (i.e., sheet or plate-like structures [1st nano-dimensional class], elongated or tube-like shapes [2nd nano-dimensional class], or spherical materials [3rd nano-dimensional class]). The key descriptors for the lung genotoxicity model were aspect ratio, BET_SSA, ROS, and ranked release, which defined the potential release of toxic metals. Among the key features associated with a neutrophil influx into BAL fluid, the aspect ratio and ROS were identified as the primary toxicity drivers. As previously mentioned, two additional

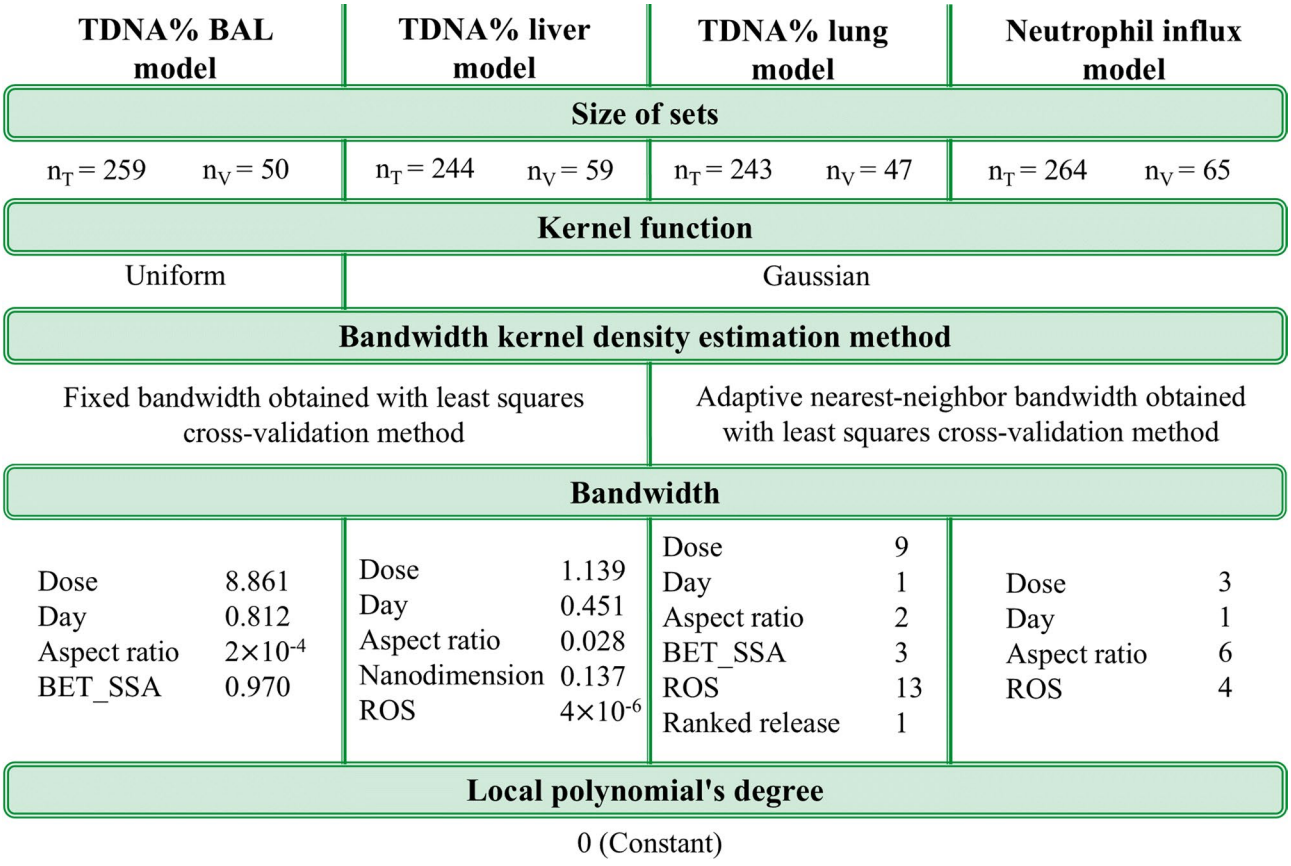


Fig. 1 KwLPR-based data-driven models for predicting in vivo genotoxicity for BAL fluid cells, lung and liver tissue, and neutrophil influx into BAL fluid in mice. BAL, bronchoalveolar lavage; BET_SSA, Brunauer-Emmett-Teller method; KwLPR, least squares kernel regression; ROS, reactive oxygen species; TDNA%, percentage of DNA strand breaks

descriptors common across all four models were post-exposure time (in days) and dose levels (in μg/mouse). We used ROS generation per mass unit in the $DCFH_2$-DA acellular ROS assay [98] as a measure of ROS activity. Since activity was measured in relative units, ROS activity was normalized to carbon black Printex90. As ROS is a surface-dependent response, its generation per mass unit will be larger for particles with a large specific surface area, as seen for the three carbon blacks in the present study. Therefore, it is possible that the effect of ROS seen here also reflects an effect of surface area. Surprisingly, when the ROS parameter normalized to mass (ROS/mass unit) was replaced with a combination of BET_SSA and ROS normalized to the surface area (ROS/SSA area), the model's predictive performance decreased slightly (by 4%). Moreover, removing BET_SSA entirely resulted in a significant decrease in the model's predictive value (by 12%), highlighting the importance of the specific surface area as a key predictor of inflammation [17, 53, 99, 100]. More detailed information can be found in the Supplementary Materials (Table S8).

Time-dose integrated ML-models for in vivo genotoxicity and inflammation by AdMa

Our comprehensive modeling of the TS data demonstrated that the fixed bandwidth approach yielded the most optimal fit for predicting in vivo genotoxicity for BAL fluid cells and liver tissue in mice (Fig. 1). This was indicated by the highest R^2 (close to one) and lowest $RMSE_C$ (close to zero) values. In the case of the remaining two models (i.e., for genotoxicity in lung tissue and neutrophil influx into BAL fluid), the adaptive nearest-neighbor bandwidth was identified as the optimal approach. In the modeling of three out of the four endpoints, a Gaussian kernel was used, whereas for the TDNA% BAL model, the neighborhood shape was defined based on the uniform kernel. This difference in the selection of the optimal kernel can be attributed to the considerably broader range of values observed for the modeled dependent variable for TDNA% BAL in comparison to the other endpoints (Figure S1). Furthermore, as illustrated in Fig. 1, all models were calibrated using the local constant polynomial, which was optimized during the leave-one-out validation process. Subsequently, the generalization ability of the data-driven models was

validated using unseen data beyond the TS. Figure 1; Table 2 provide detailed information on the bandwidth specifications and the quality metrics for the models. In addition, the details of the KwLPR-based, data-driven models developed using other considered hyperparameters, including methods for bandwidth selection, kernel regression estimators, types of kernel functions, and types of bandwidths, are summarized in Table S9.

The predictive models derived using the KwLPR algorithm exhibited high goodness-of-fit, with R^2 values ranging from 0.75 to 0.93 and strong predictive power, as indicated by Q^2_{ext} values ranging from 0.70 to 0.91 (Table 2). The scatterplots between the model's predictions and the experimentally measured TDNA% in cells from BAL fluid, the lungs and liver, and the total number of neutrophils in BAL fluid, respectively (Fig. 2), indicate that the majority of AdMa from both the TS and VS were closely aligned with the best-fit line, thereby confirming that the random errors associated with the data-driven models were within acceptable limits.

Furthermore, we verified the assumption of residual homoscedasticity in each of the four KwLPR-based models by assessing their adherence to a Gaussian probability distribution. The residual density plots (Figure S2) demonstrated that the residuals in the TS and VS formed two normally distributed peaks with averages close to zero. The internal validation results, which assessed the stability of the models, revealed some intriguing findings. For three of the four models—specifically, those predicting in vivo genotoxicity for BAL fluid cells, liver tissue, and neutrophil influx into BAL fluid in mice—the cross-validation coefficient exceeded the minimum threshold ($Q^2_{LOO} > 0.60$), with statistical metrics ranging from 0.62 to 0.69. However, as illustrated in Table 2, the model for predicting in vivo genotoxicity in lung cells did not exhibit sufficient robustness ($Q^2_{LOO} = 0.47 < 0.60$). The relatively low robustness of this model can be attributed to two factors: the methodological assumptions of the leave-one-out cross-validation method and the structure of the data set. The data set was notably diverse and demonstrated limited representativeness for materials analyzed at the longest exposure times of 90, 92, and 180 days. However, it is important to note that a high Q^2_{LOO} value does not unconditionally indicate that the model has satisfactory predictive ability. Literature reports have highlighted the necessity of external validation using an unseen TS to accurately assess a model's predictive ability. External validation results are considered the most critical metrics for evaluating a model's generalization performance [91, 92]. However, as previously discussed, the predictive capabilities of all four KwLPR-based models were confirmed (Table 2). Additional confirmation that the developed QSAR models were not the result of random correlations and exhibited statistical robustness was provided by the Y-scrambling validation. As shown in Figure S3, the R^2 and Q^2_{LOO} values for our original (i.e., "true") models were significantly higher than those obtained from the Y-randomization procedure, indicating that the predictive performance of the models was not due to chance.

Applicability domain: assessment of the reliability and relevance of in silico model predictions

The predictive uncertainty in the KwLPR-based models was further assessed through applicability domain evaluations. Using the leverage approach and a Williams plot (Fig. 3), the majority of the investigated AdMa were found to be within a squared area, inside ± 3 standard deviation units, and below the critical leverage value (h^*).

In the case of the KwLPR-based model for predicting in vivo genotoxicity in BAL fluid cells, a total of six outliers were identified among the 259 TS data points, and one outlier was identified among the 50 VS data points. These response outliers were identified based on their standardized residuals, which exceeded ± 3 standard deviations. These outliers were Printex90 (carbon black) at various dose levels and exposure times from the TS and NRCWE-043 (carbon nanotube) from the VS (Fig. 3a). Furthermore, 24 data points from the TS, including carbon black (XE2B), quartz (DQ12), carbon nanotubes (NRCWE-051, NRCWE-054), and anatase titanium dioxide (MKN-A100, MKN-A015), which varied in doses and exposure times, exhibited leverages greater than the threshold (h^*=0.058). Notwithstanding their high leverage values, these points exhibited small residuals, suggesting that they are 'good high leverage points'.

Table 2 Evaluation metrics for KwLPR-based, data-driven models

	R^2	$RMSE_c$	Q^2_{LOO}	$RMSE_{LOO}$	Q^2_{ext}	$RMSE_{ext}$	MAE	CCC
TDNA% BAL model	0.84	2.15	0.69	3.01	0.91	1.51	1.28	0.95
TDNA% liver model	0.93	0.48	0.62	1.14	0.82	0.64	0.30	0.91
TDNA% lung model	0.75	1.18	0.47	1.71	0.77	1.03	0.69	0.87
Neutrophil influx model	0.88	0.30	0.67	0.47	0.70	0.44	0.23	0.82

Evaluation metrics for KwLPR-based, data-driven models for predicting in vivo genotoxicity for BAL fluid cells, lung and liver tissue, and neutrophil influx into BAL fluid in mice

BAL, bronchoalveolar lavage; CCC, concordance correlation coefficient; KwLPR, least squares kernel regression; MAE, mean absolute error; Q^2_{ext}, external validation coefficient; Q^2_{LOO}, cross-validation coefficient; R^2, regression coefficient; $RMSE_C$, root mean square error of calibration; $RMSE_{ext}$, root mean square error of external validation; $RMSE_{LOO}$, root mean square error of cross-validation

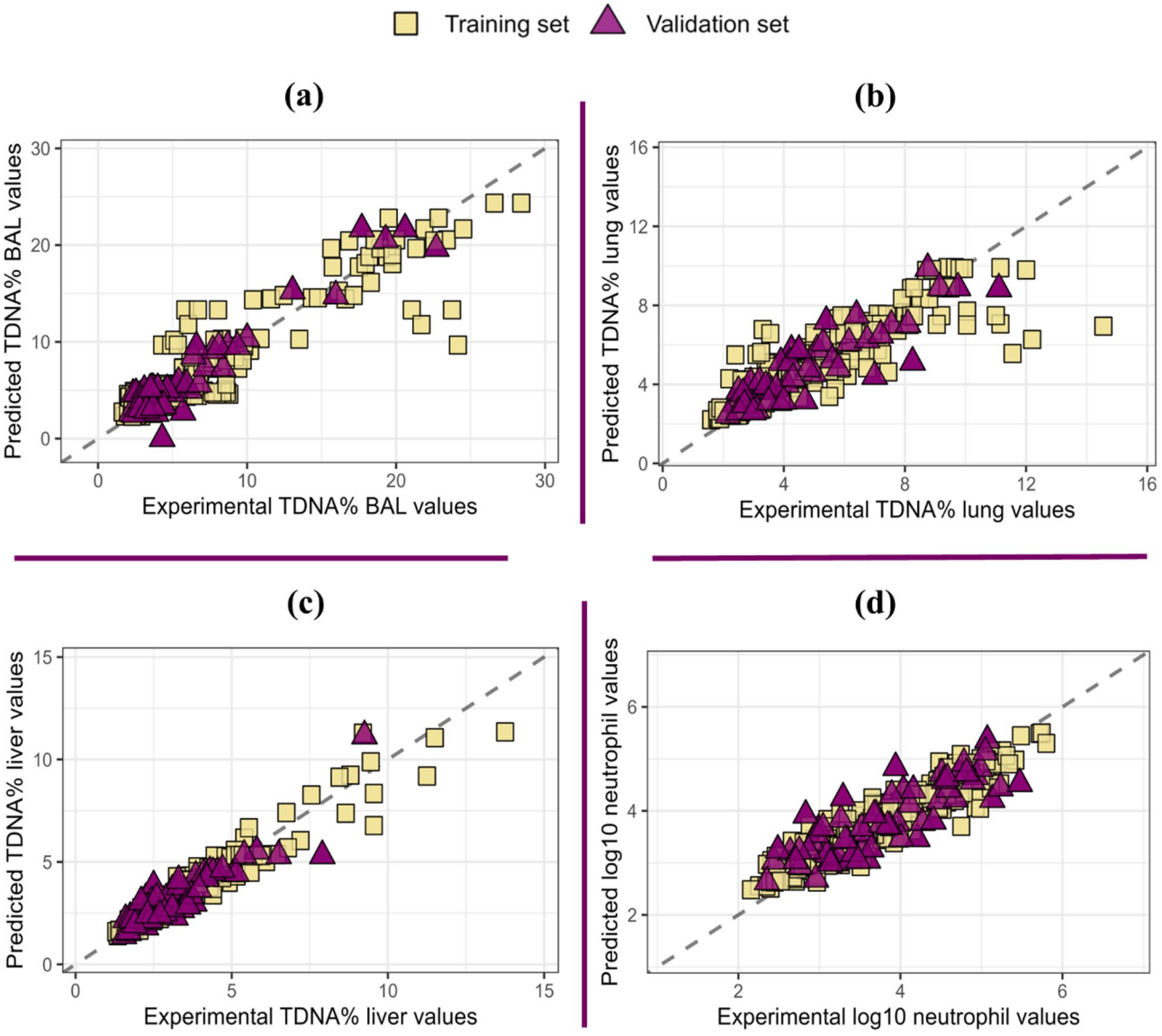


Fig. 2 Scatter plot of experimentally determined versus kernel-weighted local polynomial regression (KwLPR) predicted values for the percentage of DNA strand breaks (TDNA%) in cells from (**a**) bronchoalveolar lavage (BAL) fluid; (**b**) the lung; (**c**) the liver; and (**d**) the total number of neutrophils in BAL fluid, expressed as a decimal logarithm. The straight line indicates a perfect match between the experimental and predicted values. Squares represent predicted values for the training data points (TS), while triangles represent data for the validation data points (VS)

As discussed by Jaworska and colleagues [97], these compounds act as 'influence points' that stabilize and enhance a model's precision. However, two data points from the VS—carbon nanotube (NRCWE-055) and graphene oxide (NRCWE-058)—slightly exceeded the leverage threshold h^*. Therefore, it is recommended that predictions for these materials be treated with caution, as they may be less reliable. These materials potentially fall outside the applicability domain due to exhibiting extreme descriptor values. For example, NRCWE-051 and NRCWE-054 had the highest aspect ratio values among the materials under investigation. These were single-walled carbon nanotubes, where we were unable to measure the length and, therefore, used the mean length from the supplier [8]. XE2B exhibited the highest BET_SSA variable value among the materials under investigation, while MKN-A100 and MKN-A015 were assessed at the highest dose level (162 µg/mouse) and at the longest post-exposure time point (180 days).

We observed similar results across the other models. In the case of the KwLPR-based model predicting TDNA% in lung cells, four response outliers were identified from a TS of 243 data points, and one outlier was identified from a VS of 47 data points (Fig. 3b). These outliers in the Y-space were NRCWE-019 (Alpha Fe_2O_3 (rod)), NRCWE-022 ($NiFe_2O_4$), and NRCWE-025 (TiO_2

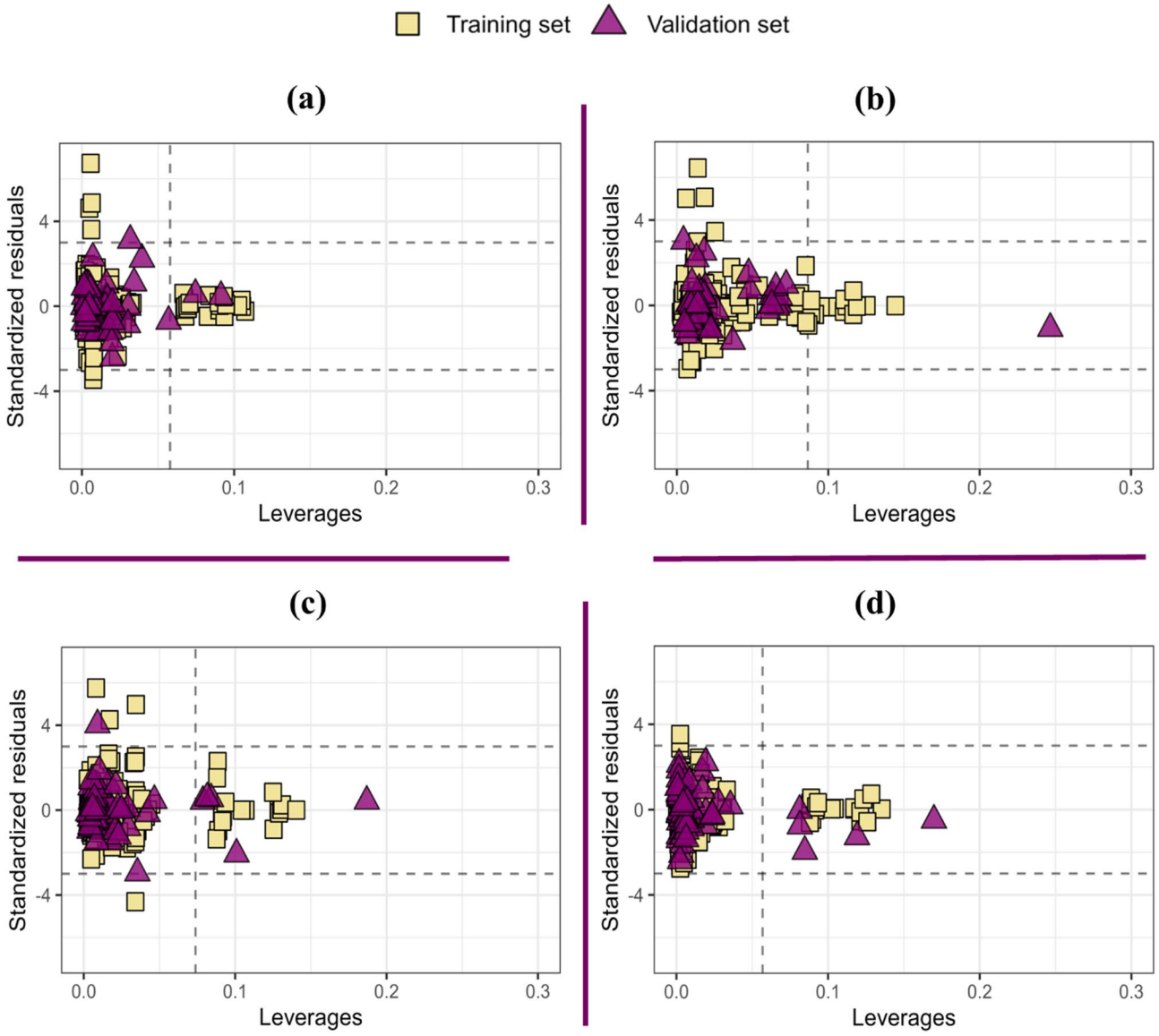


Fig. 3 Applicability domain of the kernel-weighted local polynomial regression (KwLPR) model estimated using the leverage approach for the percentage of DNA strand breaks (TDNA%) in cells from (**a**) bronchoalveolar lavage (BAL) fluid; (**b**) the lung; (**c**) the liver; and (**d**) the total number of neutrophils in BAL fluid, expressed as a decimal logarithm. Squares represent predicted values for the training data points (TS), triangles represent data for the validation data points (VS), while dashed lines represent ± 3 standard deviation units and the critical leverage value (h^*)

at two different dose levels) from the TS, and NRCWE-018 (Alpha Fe_2O_3 (particle)) from the VS. Furthermore, 14 TS nanomaterials, including XE2B (carbon black, quartz), NRCWE-051 and NRCWE-054 (carbon nanotubes), and MKN-A100 (TiO_2) at various dose levels and exposure times, were identified as influence points with $h > h^*$=0.086, yet with standardized residual values within the acceptable range of ± 3 standard deviations. One structural outlier, specifically NM-403 (a carbon nanotube), was identified among the AdMa in the VS. Further analysis of the molecular descriptor values used in the modeling process revealed that materials situated outside the boundaries of the applicability domain exhibited either the highest values (e.g., BET_SSA for XE2B; aspect ratio for NRCWE-054 and NRCWE-051; ROS for XE2B; dose level and longest exposure time for MKN-A100) or the lowest values (e.g., ROS for NRCWE-018) among all of the materials investigated.

For the KwLPR-based model predicting the TDNA% in liver cells, the leverage approach and the Williams plot (Fig. 3c) identified four response outliers from the TS of 244 data points and one outlier from the VS of 59 data points. The TS outliers included NRCWE-001 (TiO_2), NRCWE-006 (a carbon nanotube), and Nanofil®SE 3000 (a nanoclay at two different exposure days), while the VS outlier was Printex90 (carbon black). Furthermore, 19

data points in the TS had leverages greater than h^*=0.074 but exhibited small residuals. These included NRCWE-051 and NRCWE-054 (single-walled carbon nanotubes), DQ12 (quartz), XE2B (carbon black), and MKN-A100 (TiO_2) across various dose levels and exposure times. As previously stated, these high-leverage points significantly influenced the model's fitting performance. Moreover, five VS data points—NRCWE-051 and NRCWE-054 (carbon nanotubes), XE2B (carbon black), and MKN-A015 (TiO_2)—were identified as structural outliers, indicating that their predictions should be treated with caution. Similar to findings from two previously discussed models, the positioning of these materials outside the model's domain can be attributed to extreme descriptor values used in the modeling. To illustrate, NRCWE-051 and NRCWE-054 were associated with the highest aspect ratio values, XE2B was associated with the highest ROS value, while MKN-A015, MKN-A100, and DQ12 had the longest post-exposure time points.

For the KwLPR-based model predicting neutrophil influx into BAL fluid, only the predictions for NRCWE-001 and NRCWE-025 (TiO_2) were slightly beyond the upper control limit of + 3 standard deviation units, marking them as response outliers. Figure 3d indicates that there were no Y-space outliers within the VS materials. Additionally, 19 data points in the TS had leverages greater than h^*=0.057, yet their standardized residual values stayed within the acceptable range of ± 3 standard deviations. These data points included NRCWE-051 and NRCWE-054 (SWCNT), DQ12 (quartz), XE2B (carbon black), and MKN-A015 (TiO_2) across various doses and exposure durations. The Williams plot also indicated five structural outliers in the VS, namely NRCWE-051 and NRCWE-054 (SWCNT), MKN-A100 (TiO_2), and XE2B (carbon black). As with other models, the extreme descriptor values used in the modeling were the main reason for these materials falling outside the model domain, thereby reducing the reliability of the predictions. For comprehensive details on the KwLPR-based modeling results and the applicability domain evaluation of all four models, please see Tables S4-S7.

Exploration of ML methods for the re-optimization of in silico models for in vivo genotoxicity and inflammation by AdMa

As outlined in the Methods section, alternative modeling techniques were also explored during optimization of the predictive models. A comparative analysis of the evaluation metrics for data-driven models developed using CART, RF, kNN, distance-weighted kNN, and SVR with various kernel functions is presented in Table S10. All models were calibrated and externally validated using the same training/test sets and descriptor sets that were used to develop the KwLPR models. As indicated in Table S10, SVR with a spline kernel and distance-weighted kNN consistently showed the weakest performance among the evaluated techniques, regardless of the model response. QSAR models based on these techniques demonstrated the poorest goodness-of-fit and predictive capabilities, as evidenced by the lowest R^2 and Q^2 values, respectively. Conversely, models developed using the RF algorithm typically achieved the best overall performance, often comparable to that of the KwLPR models. The RF algorithm constructs a user-defined number of decision trees (500 in this case) by generating multiple bootstrapped subsets of the training data and averaging their predictions to estimate the dependent variable. Nevertheless, considering the known limitations of RF models - specifically, their tendency to overfit and the challenges associated with interpreting their results - and the fact that they did not surpass the performance of the KwLPR models, it was concluded that KwLPR remains the most optimal modeling approach for the analyzed datasets.

Discussion

A mechanistic interpretation of the models

As the KwLPR model does not provide traditional regression coefficients to quantify the contributions of each explanatory variable in estimating the model's response (here, in vivo genotoxicity and inflammation by AdMa), we conducted a feature importance analysis to address this limitation. This method assigns an importance score to input features based on their utility in predicting the target variable. Specifically, it involves randomly permuting the values of a given feature and then comparing the model's performance before and after this permutation. The feature importance score is calculated as the ratio of these two performance metrics. A higher score indicates greater importance, as the alteration of the feature's values has a significant effect on the model's performance. One efficient and practical method for visualizing feature importance is through the use of variable importance plots (VIPs). The analysis of VIPs (Figure S4) highlights the pivotal role of the aspect ratio in AdMa-induced toxicity. In three of the four models, specifically those assessing in vivo genotoxicity in BAL fluid cells and liver and lung tissue, the aspect ratio emerged as the most or second most significant factor driving AdMa genotoxicity. As for asbestos, the aspect ratio has been linked to a range of toxic effects, predominantly through inhalation exposure [101]. A wealth of studies has reported that high aspect ratio materials display heightened toxicity and a heightened potential to disrupt cellular signaling processes compared to low aspect ratio materials. This observation also holds for carbon nanotubes, nanoclays, halloysite nanotubes, and graphene-based materials, which were the focus of our study [53, 102, 103]. Fibers, especially those with diameters in the nano-range, can

readily penetrate cellular structures through inhalation, irrespective of their chemical composition. Furthermore, their clearance from the body is impeded, which may result in accumulation [103, 104]. Consequently, inhalation exposure to these materials is associated with the influx of neutrophils into BAL fluid, the induction of oxidative stress, and, ultimately, DNA damage in BAL fluid and lung cells, as observed in our study. DNA damage in liver cells is likely caused by the translocation of nanomaterials from the lung to the liver via systemic circulation [105]. In summary, this work and previous studies on carbon nanomaterials indicate that the aspect ratio of AdMa is a significant, if not the primary, factor contributing to their pulmonary pathogenicity.

Aside from aspect ratio, the shape of particles—referred to here as a nanodimension descriptor—plays a pivotal role in determining liver genotoxicity (Figure S4). This phenomenon is attributed to the shape-dependent ability of AdMa to penetrate and interact with cellular structures [106]. In particular, needle-like fibers such as thick and rigid carbon nanotubes were shown to disperse and exist as single fibers in lung tissue one year post-exposure, whereas carbon NPs and thin, entangled carbon nanotubes accumulated in granulomas [107]. Needle-like GaP nanowires have also been detected in lung tissue as single fibers [108]. The needle-like structure may be of less importance than the fiber shape in relation to carcinogenicity since both thin, entangled carbon nanotubes and thick, rigid, straight carbon nanotubes cause lung adenoma following pulmonary deposition in rats [109–111]. Notably, particle shape influences nanomaterial cellular uptake efficiency, with needle-/fiber-like or platelet-/sheet-like particles demonstrating greater penetration than sphere-like particles. This is due to their enhanced endocytic mechanisms, higher internalization rates, and more effective adhesion to target cell surfaces. For instance, Huang and colleagues' [112] research revealed that rod-shaped mesoporous silica NPs with a length of 450 nm were more effectively internalized by A375 human melanoma cells than 100 nm spherical particles. Similarly, Lee et al.'s [113] study on the effect of different shapes of iron oxide NPs on toxicity to murine macrophage cells found that rod-shaped Fe_2O_3 NPs induced more cell death than spherically shaped ones. This shape-dependent cellular uptake and resultant toxicity have also been experimentally demonstrated for numerous materials, including TiO_2, ZnO, carbon black, and carbon nanotubes [51, 53, 106, 114, 115]. In addition, nanomaterial shape may be a predictor for translocation from lung to liver and hence, genotoxicity in liver.

It is well-known that the toxicity of nanomaterials arises from interactions between their surface and cellular components [116]. Therefore, in addition to aspect ratio and shape, the specific surface area has been acknowledged as a pivotal factor affecting the toxicity of NPs. According to Schmid and Stoeger [99], surface area is responsible for approximately 80% of the observed variability in acute pulmonary toxicity tests with spherical NPs with low solubility. This can be attributed to the fact that poorly soluble materials with a high surface area generate higher levels of ROS, which can lead to cytotoxicity and genotoxicity. As illustrated in Figure S4, the specific surface area determined by the BET method emerged as the most significant factor driving AdMa in vivo genotoxicity in lung tissue and the second most significant factor in BAL fluid cells. This finding aligns with previous in vitro and in vivo studies indicating a correlation between BET_SSA values and pulmonary toxic responses to AdMa with low solubility, including TiO_2, halloysite nanotubes, nanoclays, and quartz [51, 53, 59]. For example, Ianni and colleagues demonstrated that reducing the surface area of bentonite through surface modifications resulted in diminished toxicological effects in mouse lungs [51]. Carbon-based nanomaterials are potent generators of ROS [14], and the surface generation of ROS has been proposed as a mechanism-of-action for carbon-induced genotoxicity [14, 56, 57, 103].

As shown in Figure S4, genotoxicity in lung tissue was also linked to the release of intracellular metal ions. An increasing body of research has demonstrated that metallic AdMa tend to release metal ions at various pH levels within cell organelles and subcellular structures, resulting in their accumulation in critical organs such as the liver, lungs, and kidneys. This accumulation may subsequently induce genotoxic and cytotoxic effects [117]. Initially regarded as passive, NPs such as ZnO, titanium dioxide, and aluminum oxide have garnered attention for their potential role in promoting oxidative damage and inflammation [118]. Among these, only TiO_2 is observed to be very poorly soluble with slow dissolution rates in phagolysosomal fluids [62, 63]. Jacobsen et al. [20] demonstrated that ZnO NPs internalized through the endocytic pathway have a greater capacity to release ions intracellularly, leading to enhanced toxicity compared to larger ZnO particles. Similarly, Mahto et al. [119] reported that quantum dot core/shell (CdSe/ZnSe) dispersed in aqueous media release cadmium ions that disrupt cellular homeostasis and lead to toxicity.

Further evaluation of the relative importance of the variables in the predictive models has revealed that ROS generation, either directly or indirectly, played a crucial role in the genotoxicity associated with metallic AdMa-induced toxicity. The excessive production of ROS can result in oxidative stress, which disrupts normal redox-regulated physiological functions. This, in turn, leads to DNA damage and the unregulated modulation of cellular signaling pathways involved in cell death, proliferation, and/or differentiation [5, 13]. As previously discussed

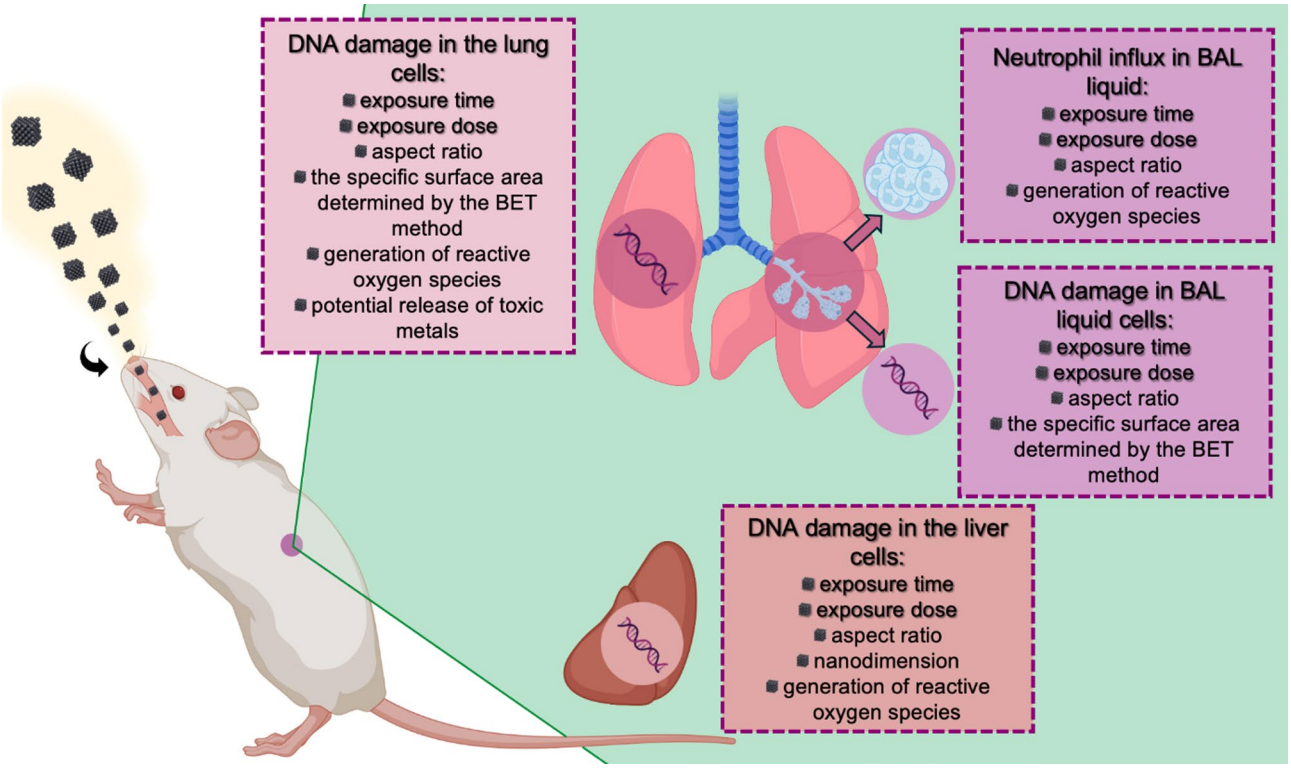


Fig. 4 Summary of key features governing the genotoxicity and inflammation of advanced materials in the present study

by Nel and colleagues [120], the generation of ROS and oxidative stress represent a well-established paradigm for explaining the toxic effects of inhaled NPs. Figure S4 illustrates that particle-dependent ROS generation was the most significant factor contributing to liver genotoxicity. Furthermore, ROS generation emerged as the second and fourth most significant factor contributing to neutrophil influx into BAL fluid and TDNA% in lung cells, respectively. These findings are consistent with those of previous research on graphene materials, carbon nanotubes, titanium dioxide, quartz, and carbon black, which demonstrated that increased ROS generation upon exposure to AdMa resulted in ROS-mediated toxicity [53, 105, 121, 122]. It is also important to acknowledge that nonmetallic particles have the potential to trigger oxidative stress. For instance, Shvedova et al. [123] demonstrated that, in addition to generating free radicals, the cellular uptake of carbon nanotubes resulted in the oxidation of polyunsaturated fatty acids, ultimately leading to cellular apoptosis. Guo et al. [124] reported that silica NPs can induce oxidative stress, disrupting the balance of nitric oxide/nitric oxide synthase, leading to inflammation and endothelial dysfunction. A summary of key factors influencing genotoxicity and inflammation associated with AdMa is shown in Fig. 4.

To achieve the primary objectives of the present study—overcoming the limitations of static QSAR models confined to single doses and post-exposure time points and deepening our understanding of dose-time-property/response dynamics in AdMa-induced toxicity—all KwLPR-based models were designed to incorporate varying doses and post-exposure time points from experimental testing as additional independent variables. Generally, an increase in dosage corresponds to an increase in the measured response. In addition, while the inherent toxicity of a particle remains unchanged over time, prolonged exposure can prolong the toxic effect on the organism, causing hazardous outcomes even at lower concentrations, as predicted by Harber's rule that the total dose is the product of concentration and exposure time. This was recently shown to be the case for the pulmonary retention of inhaled nanomaterials [104]. Chronic diseases such as cancer, cardiovascular disease, and chronic obstructive pulmonary disease take years to develop. Adverse outcome pathways delineate the causal relationships between key events leading to disease and may be used to identify early key events (such as inflammation) that are causally related to the adverse outcomes. Inflammation is an early key event causally related to cardiovascular disease, fibrosis, and cancer [22, 26]. Our findings indicate that, in addition to particle characteristics, both particle concentration and post-exposure duration are crucial in determining potential toxic effects. However, it is important to emphasize that the post-exposure duration in this study refers to a single exposure event during which particles were introduced

into the lungs of mice and then observed over several post-exposure time points. At any given time, the particles may still have been present, cleared, or undergone dissolution; however, they may still play a critical role in shaping potential toxic effects. The analysis of VIPs (Figure S4) demonstrated the substantial influence of post-exposure duration on AdMa-induced toxicity. In the model predicting neutrophil influx into BAL fluid, post-exposure time was identified as the most critical factor influencing AdMa-induced inflammation. For the TDNA% BAL and TDNA% lung models, post-exposure time was identified as the third most significant factor influencing AdMa genotoxicity, while for the TDNA% liver model, post-exposure duration was determined to be the fourth most significant factor. These results align with recent findings [59, 102, 125–127] such as those by Di Ianni et al. [51], who reported a time-dependent influx of neutrophils into BAL fluid in response to nanoclay materials. Figure S4 further confirms that AdMa-induced genotoxicity and inflammation were significantly influenced by the administered dose. In the neutrophil influx into the BAL fluid model, the feature importance of dose and ROS generation was approximately equal, ranking higher than the aspect ratio. Similarly, in the TDNA% liver model, three variables—nanodimension, exposure time, and dose—exhibited nearly identical feature importance scores, indicating a comparable effect on the model's response. These findings align with animal studies that have reported a dose-dependent effect in lung and liver tissues for graphene materials, zinc oxide, and carbon nanotubes [15, 101, 127]. A dose-dependent influx of neutrophils into BAL fluid was also observed for titanium dioxide and nanoclays [51, 53, 128].

Despite our efforts to incorporate theoretical descriptors encoding information on symmetry, topology, and molecular properties, none of the 0D-2D descriptors ultimately proved essential for calibrating the predictive models under discussion. There are several possible explanations for these results. First, dose and post-exposure time exert a direct and often dominant influence on biological responses. The inclusion of these parameters as independent variables introduces an additional layer of variability that may overshadow the more subtle effects related to structural features. As a result, the models may prioritize dose- and/or exposure-related variability over intrinsic structure-activity relationships. Second, the limited relevance of the 0D-2D descriptors in our models may have arisen from their inability to fully capture the complexity of the studied particles, thus failing to provide comprehensive insights into the molecular-level effects crucial for AdMa-induced genotoxicity and inflammation. Another potential explanation lies in the significant structural diversity and/or inhomogeneity of the analyzed materials, which encompassed various metal oxides, carbon materials, and sheet silicates. In local data-driven models that are restricted to a specific class of materials (e.g., metal oxides), one or more of these descriptors may prove to be important and useful. However, a limited amount of data precludes the feasibility of developing local models for all material classes. It is important to note that 3D descriptors, which encode information about particle reactivity (e.g., energy of the highest occupied molecular orbital, energy of the lowest unoccupied molecular orbital, ionization potential, electron affinity, energy gap, electronegativity, chemical potential, etc.), have the potential to provide valuable insights that lead to a more comprehensive understanding of AdMa-induced genotoxicity and inflammation. A notable limitation of using 3D descriptors in QSAR modeling is the necessity for preliminary geometry optimization of the molecular structure. This process is particularly time- and resource-intensive, especially for large systems such as carbon nanotubes. Consequently, 3D descriptors have predominantly been used in modeling metal-based particles and materials.

Comparison with existing data-driven predictive models

A direct comparison with previously reported models is challenging for a number of reasons. First, to the best of our knowledge, no publicly available models currently exist for predicting AdMa-induced genotoxicity and inflammation with such a broad and diverse model domain. Second, the majority of data-driven models reported in recent years are static, focusing on a specific dose and exposure time. Despite the well-established importance of dose level and exposure duration in toxicity studies and risk assessment, these factors are often overlooked in QSAR modeling. As a result, the dynamic modeling paradigm, which considers the effects of varying doses and exposure times, is not widely adopted. Therefore, to relate our modeling results to existing knowledge, we conducted a thorough comparative analysis at the mechanistic level, with particular emphasis on the relevance of the explanatory variables used in model development.

The descriptors used to derive KwLPR-based models can be classified into three categories, each corresponding to a specific aspect of the particles: (1) size, shape, and aspect ratio; (2) surface properties; and (3) reactivity and transformation referred to as the ease of metal ion release from particle surfaces (understood as dissolution). Among these features, the aspect ratio has emerged as the pivotal factor driving AdMa-induced genotoxicity and inflammation [129]. It is widely acknowledged that materials with a high aspect ratio tend to exhibit even higher toxicity compared to those with a low aspect ratio. This concept has been extensively validated through numerous studies, particularly those examining carbon

nanotubes and asbestos fibers. For example, Jagiello et al. [130] developed a single-descriptor QSAR model using a consensus linear regression technique to predict the biological event initiating the inflammation process described by the 'agranulocyte adhesion and diapedesis' pathway in the lungs of mice exposed to ten MWCNTs. In this model, the aspect ratio of MWCNTs was used as the sole independent variable. The authors concluded that an increase in the aspect ratio correlated with a decrease in the lower bound of the estimated benchmark dose for the agranulocyte adhesion and diapedesis pathway, indicating that this pathway is activated at the lowest doses of MWCNTs with the highest aspect ratios [130]. Le et al. [131] used quantitative computational models to investigate the relationships between the structure and properties of 45 ZnO NPs with diverse physicochemical modifications—specifically different dosing types and concentrations and surface coatings—and their biological effects on cell viability, membrane integrity, and oxidative stress. They assessed cellular damage or stress as endpoints by measuring the responses of human umbilical vein endothelial cells and human hepatocellular liver carcinoma cells (HepG2) to ZnO NPs. Their findings revealed that, in addition to the surface coating type and the nature and degree of the dosing, the aspect ratio of the particles significantly influenced the cell toxicity induced by ZnO NPs. These results further validate the hypothesis that high-aspect-ratio particles cause greater cellular damage than those with a lower aspect ratio [131]. It is well established that the small size of materials significantly increases their specific surface area, which is a crucial predictor of genotoxicity and lung inflammation. As previously discussed, the pulmonary responses triggered by inhaled nanosized particles are generally more severe than those caused by larger particles, primarily due to the higher surface area-to-particle mass ratio. Gernand and Casman [132] developed a random forest model and a random-forest-based, dose-response model to elucidate the relationship between carbon nanotubes properties and pulmonary toxicity in rodents exposed to 17 uncoated, nonfunctionalized carbon nanotubes. The authors' QSAR analysis highlighted the pivotal role of surface area in evaluating the pulmonary toxicity of carbon nanotubes. Other crucial variables identified by the authors included metallic impurities, carbon nanotubes length and diameter, and aggregate size [132]. Importantly, in the context of AdMa reactivity, an augmented particle surface area amplified the potential for ROS generation, exponentially increasing the number of electroactive sites on the AdMa and rendering them more susceptible to exposure [133]. Indeed, ROS generation has been identified as a key mechanism driving the genotoxicity of metal oxide NPs. For example, Mesárosová et al. [134] investigated the role of ROS formation on the genotoxic effects observed in A549 human lung adenocarcinoma epithelial cells and HEL 12469 human embryonic lung fibroblasts exposed to surface-modified magnetite NPs. As outlined in their review, ROS-mediated activity may be triggered by either iron ions released from the particle surface or the particles themselves. Specifically, iron ions released into the cytosol upon lysosomal enzymatic degradation can participate in Fenton and Fenton-like reactions, thereby mediating the generation of hydroxyl radicals. Alternatively, ROS can be generated through electron transfer and catalytic reactions occurring at the NPs' surface [134]. Furthermore, a QSAR analysis by Zhang and colleagues [39] aimed to predict the mixture toxicity of metal oxide NPs and confirmed that the released metal cations, ROS-induced oxidative damage, or a combination of both could be responsible for the observed toxicity. In a separate study, Haixia and colleagues [135] employed both in vivo and in silico approaches to assess the potential genotoxicity of various graphene-based materials, including graphene quantum dots, as well as their oxidized (graphene oxide) and unoxidized (pure graphene) forms. Through a functional density theory analysis, the authors demonstrated that graphene oxide induced ROS-mediated genotoxicity [135].

Unleashing the potential of dynamic (dose-time-property/response) models in computational toxicology

The dynamic modeling paradigm is becoming increasingly prominent due to its capacity to harmoniously integrate data from disparate sources, including the intrinsic and extrinsic attributes of AdMa, experimental conditions, exposure duration and dosage, and biological responses. This approach enables the prospective examination of multiple hypotheses, a capability that is increasingly regarded as an asset in scientific research. The greatest advantage of dynamic models (i.e., dose-time-property/response models) lies in their ability to interpolate effects across very short or different doses and time points [136]. This advantage arises from their capacity to provide invaluable insights into how material hazards evolve over time and under varying experimental conditions. This is particularly relevant in the case of extremely low doses and excessively long exposures, where experimental measurements may be difficult or impractical, often resulting in a lack of toxicological data. As shown above, the administered dose and in this case, post-exposure time may serve as key predictors in data-driven models for assessing AdMa-induced genotoxicity and inflammation. This conclusion is consistent with the findings of previous studies on dose-response, time-response, and dose-time-property/response models. For example, Ha et al. [137] developed a qualitative classification model using the random forest algorithm that

combined data from multiple individual studies on the cellular toxicity of metal oxide NPs. The model, trained on data from 216 publications, used 14 descriptors encompassing physicochemical, toxicological, and quantum-mechanical properties. Among the six most influential attributes of NP-induced cytotoxicity, the authors identified the administered dose, assay type, exposure time, and surface area of the NPs as the most significant factors [137]. In another study, Romeo and colleagues [138] conducted a meta-analysis using an ML approach to examine the toxicity of graphene-related materials (GRMs) on lung cells in vitro. The authors emphasized the significant influence of both material properties and experimental conditions on the intensity of the response, specifically the viability of lung cells exposed to GRMs. They identified four key features: GRM lateral size, number of layers, material functionalization, and dose [138]. Similarly, a QSAR analysis by Trinh and colleagues [139], which aimed to predict the cytotoxicity of 20 distinct types of MWCNTs on human lung cells, further confirmed the importance of exposure time and dose as essential predictive parameters. In that study, the authors used optimal descriptors, recorded as quasi-SMILES, to capture the physicochemical properties and experimental conditions of the MWCNTs based on 276 data records from previously published studies. The quasi-SMILES-based optimal descriptors included diameter, length, surface area, the in vitro toxicity assay, cell line, exposure time, and dose [139]. Furxhi and Murphy [140] developed a random forest model to predict the cellular viability induced by NPs, with a particular focus on brain tissue. To achieve this, they considered a number of factors, including the characteristics of the NPs, the experimental exposure conditions, and the in vitro parameters. Their research focused on the curation and analysis of data from experimental literature related to brain tissue toxicity to enhance the prediction of in vitro neurotoxicity. Through information gain analysis, the authors identified exposure dose, exposure duration, toxicological assay, cell type, and zeta potential as the five most significant factors for predicting neurotoxicity in vitro. They also emphasized that by incorporating in vitro experimental conditions—such as exposure dose, duration, and toxicological assay—as input parameters, the resulting model could be effectively applied to a wide range of in vitro brain tissue exposure scenarios [140]. The results of these studies have consistently identified exposure dose and/or time as relevant descriptor(s), thereby supporting our hypothesis regarding the significance of these features in AdMa-induced genotoxicity and inflammation. By simultaneously modeling concentration and exposure time, the KwLPR-based models presented here estimated response as a function of material properties, exposure dose, and time rather than fitting separate models for each concentration and time point. These findings further underscore the necessity for the comprehensive characterization and documentation of both materials and experimental conditions.

Conclusions

We successfully developed data-driven dynamic models capable of predicting in vivo genotoxicity and inflammation across various exposure doses, time points, and broad applicability domains. The resulting KwLPR-based models demonstrated strong predictive performance and fully complied with all OECD principles for QSAR model acceptance, as confirmed by rigorous internal cross-validation and external validation. The analysis of dose-time-property/response models, which also incorporated material characteristics, revealed the mechanisms of AdMa-induced in vivo genotoxicity in BAL fluid cells, lung and liver tissue, and neutrophil influx in mice that align with the current state of the art. In addition to experimental conditions such as exposure dose and duration, other factors, including aspect ratio, shape, surface area, ROS generation, and the propensity for metal ion release from particle surfaces, were identified as critical determinants of AdMa-induced genotoxicity and inflammation. The dose-time-property/response modeling paradigm was found to be an effective approach for predicting in vivo genotoxicity and inflammation, offering a robust tool for the risk assessment of morphologically diverse AdMa. With this research, we hope to take a step toward unleashing the potential of experimental conditions, such as exposure dose, duration, and/or others, as key predictive input parameters. We view this methodology as a promising tool for the hazard assessment of AdMa.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12951-025-03510-y.

Supplementary Material 1

Author contributions

Conceptualization: AGS; Data curation: MM, AGS, PHD, KAJ, UV; Methodology: AGS, MM, PHD, KAJ, UV; Formal analysis: MM, AGS; Investigation: MM, AGS; Visualization: MM; Project administration: UV, RG, AGS; Resources: KAJ, UV, RG, AGS; Funding acquisition: UV, RG, AGS; Supervision: PHD, KAJ, UV, AGS; Writing – original draft: MM, AGS; Writing – review & editing: MM, KK, PHD, KAJ, UV, RG, AGS.

Funding

The research has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 953183 (HARMLESS).

Data availability

The datasets supporting the conclusions of this article are available in the Supplementary Materials.

Declarations

Ethics approval and consent to participate
All animal experimental protocols were approved by the Danish Animal Experiment Inspectorate (permission no. 2006/561–1123, 2010/561–1779, 2012-15-2934-00223, 2015-15-0201-00465 and 2020-15-0201-00485) and received prior approval from the local Animal Ethics Council.

Consent for publication
All authors approved the final manuscript and the submission to this journal.

Competing interests
The authors declare no competing interests.

Author details
[1]Laboratory of Environmental Chemoinformatics, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, Gdansk, Poland
[2]National Research Centre for the Working Environment, Copenhagen, Denmark
[3]Misvik Biology, Division of Toxicology, Karjakatu 35 B, Turku 20520, Finland
[4]Institute of Environmental Medicine, Karolinska Insitutet, Stockholm 171 77, Sweden

Received: 27 February 2025 / Accepted: 28 May 2025
Published online: 06 June 2025

References

1. Gu C, Su X, Liu B, Zheng C, Wang S, Tian Y et al. Recent progress in advanced materials for electrochemical determination of phenolic contaminants. Microchem J. 2023;195.
2. Ling Z, Yang J, Zhang Y, Zeng D, Wang Y, Tian Y et al. Applications of advanced materials in the pretreatment and rapid detection of small molecules in foods: A review. Trends Food Sci Technol. 2023;141.
3. Liu X, Xiao M, Li Y, Chen Z, Yang H, Wang X. Advanced porous materials and emerging technologies for radionuclides removal from Fukushima radioactive water. Eco-Environment Health. 2023;2.
4. Arvidsson R, Peters G, Hansen SF, Baun A. Prospective environmental risk screening of seven advanced materials based on production volumes and aquatic ecotoxicity. NanoImpact. 2022;25.
5. El-Kady MM, Ansari I, Arora C, Rai N, Soni S, Verma DK et al. Nanomaterials: A comprehensive review of applications, toxicity, impact, and fate to environment. J Mol Liq. 2023;370.
6. Buchman JT, Hudson-Smith NV, Landy KM, Haynes CL. Understanding nanoparticle toxicity mechanisms to inform redesign strategies to reduce environmental impact. Acc Chem Res. 2019;52.
7. Di Ianni E, Møller P, Vogel UB, Jacobsen NR. Pro-inflammatory response and genotoxicity caused by clay and graphene nanomaterials in A549 and THP-1 cells. Mutat Res Genet Toxicol Environ Mutagen. 2021;872.
8. Solorio-Rodriguez SA, Williams A, Poulsen SS, Knudsen KB, Jensen KA, Clausen PA et al. Single-Walled vs. Multi-Walled carbon nanotubes: influence of Physico-Chemical properties on toxicogenomics responses in mouse lungs. Nanomaterials. 2023;13.
9. Zhang M, Li X, Lu Y, Fang X, Chen Q, Xing M et al. Studying the genotoxic effects induced by two kinds of bentonite particles on human B lymphoblast cells in vitro. Mutat Res Genet Toxicol Environ Mutagen. 2011;720.
10. Stone V, Miller MR, Clift MJD, Elder A, Mills NL, Møller P et al. Nanomaterials versus ambient ultrafine particles: an opportunity to exchange toxicology knowledge. Environ Health Perspect. 2017;125.
11. Hristozov D, Badetti E, Bigini P, Brunelli A, Dekkers S, Diomede L, et al. Next generation risk assessment approaches for advanced nanomaterials: current status and future perspectives. NanoImpact. 2024;35:100523.
12. Fadeel B, Bussy C, Merino S, Vázquez E, Flahaut E, Mouchet F, et al. Safety assessment of Graphene-Based materials. Focus on Human Health and the Environment. ACS Nano; 2018;12.
13. Egbuna C, Parmar VK, Jeevanandam J, Ezzat SM, Patrick-Iwuanyanwu KC, Adetunji CO et al. Toxicity of nanoparticles in biomedical application: nanotoxicology. J Toxicol. 2021;2021.
14. Di Ianni E, Jacobsen NR, Vogel UB, Møller P. Systematic review on primary and secondary genotoxicity of carbon black nanoparticles in mammalian cells and animals. Mutat Res Rev Mutat Res. 2022;790.
15. Jin M, Li N, Sheng W, Ji X, Liang X, Kong B et al. Toxicity of different zinc oxide nanomaterials and dose-dependent onset and development of parkinson's disease-like symptoms induced by zinc oxide nanorods. Environ Int. 2021;146.
16. Thu HE, Haider MA, Khan S, Sohail M, Hussain Z. Nanotoxicity induced by nanomaterials: A review of factors affecting nanotoxicity and possible adaptations. OpenNano. 2023;14.
17. Danielsen PH, Poulsen SS, Knudsen KB, Clausen PA, Jensen KA, Wallin H et al. Physicochemical properties of 26 carbon nanotubes as predictors for pulmonary inflammation and acute phase response in mice following intratracheal lung exposure. Environ Toxicol Pharmacol. 2024;107.
18. Ahmad J, Wahab R, Siddiqui A, Ahamed M. M. Cytotoxicity and apoptosis induction of zinc ferrite nanoparticle through the oxidative stress pathway in human breast cancer cells. J King Saud Univ Sci. 2024;36.
19. Husain M, Saber AT, Guo C, Jacobsen NR, Jensen KA, Yauk CL et al. Pulmonary instillation of low doses of titanium dioxide nanoparticles in mice leads to particle retention and gene expression changes in the absence of inflammation. Toxicol Appl Pharmacol. 2013;269.
20. Jacobsen NR, Stoeger T, van den Brule S, Saber AT, Beyerle A, Vietti G et al. Acute and subacute pulmonary toxicity and mortality in mice after intratracheal instillation of ZnO nanoparticles in three laboratories. Food Chem Toxicol. 2015;85.
21. Umezawa M, Onoda A, Korshunova I, Jensen ACØ, Koponen IK, Jensen KA et al. Maternal inhalation of carbon black nanoparticles induces neurodevelopmental changes in mouse offspring. Part Fibre Toxicol. 2018;15.
22. Gromelski M, Stoliński F, Jagiello K, Rybińska-Fryca A, Williams A, Halappanavar S et al. AOP173 key event associated pathway predictor–online application for the prediction of benchmark dose lower bound (BMDLs) of a transcriptomic pathway involved in MWCNTs-induced lung fibrosis. Nanotoxicology. 2022;16.
23. Halappanavar S, van den Brule S, Nymark P, Gaté L, Seidel C, Valentino S, Zhernovkov V, Høgh Danielsen P, De Vizcaya A, Wolff H, Stöger T, Boyadziev A, Poulsen SS, Sørli JB, Vogel U. Part Fibre Toxicol. 2020 May 25;17(1):16. doi: 10.1186/s12989-020-00344-4. PMID: 32450889.
24. DeMarini DM, Gwinn W, Watkins E, Reisfeld B, Chiu WA, Zeise L et al. IARC workshop on the key characteristics of carcinogens: assessment of end points for evaluating mechanistic evidence of carcinogenic hazards. Environ Health Perspect. 2025;133.
25. Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I et al. Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. Environ Health Perspect. 2016;124.
26. Nymark P, Karlsson HL, Halappanavar S, Vogel U. Adverse outcome pathway development for assessment of lung carcinogenicity by nanoparticles. Front Toxicol. 2021;3.
27. Huang HJ, Lee YH, Hsu YH, Liao C, Te, Lin YF, Chiu HW. Current strategies in assessment of nanotoxicity: alternatives to in vivo animal testing. Int J Mol Sci. 2021;22.
28. Pallocca G, Moné MJ, Kamp H, Luijten M, van de Water B, Leist M. Next-Generation risk assessment of Chemicals - Rolling out a Human-Centric testing strategy to drive 3R implementation: the RISK-HUNT3R project perspective. Altex. 2022;39.
29. Russell W, Burch R. The Principles of Humane Experimental Technique by, Russell WMS, Burch RL. John Hopkins Bloomberg School of Public Health. 1959.
30. Schmeisser S, Miccoli A, von Bergen M, Berggren E, Braeuning A, Busch W et al. New approach methodologies in human regulatory toxicology – Not if, but how and when! Environ Int. 2023;178.
31. Kavlock RJ, Bahadori T, Barton-Maclaren TS, Gwinn MR, Rasenberg M, Thomas RS. Accelerating the Pace of chemical risk assessment. Chem Res Toxicol. 2018;31.
32. ECHA (European Chemicals Agency). New Approach Methodologies in Regulatory Science. New Approach Methodologies in Regulatory Science. 2016.
33. Hatherell S, Baltazar MT, Reynolds J, Carmichael PL, Dent M, Li H et al. Identifying and characterizing stress pathways of concern for consumer safety in next-generation risk assessment. Toxicol Sci. 2020;176.
34. OECD. Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA). 2020 Oct.
35. Jeliazkova N, Bleeker E, Cross R, Haase A, Janer G, Peijnenburg W et al. How can we justify grouping of nanoforms for hazard assessment? Concepts and tools to quantify similarity. NanoImpact. 2022;25.

36. Murphy F, Jacobsen NR, Di Ianni E, Johnston H, Braakhuis H, Peijnenburg W et al. Grouping MWCNTs based on their similar potential to cause pulmonary hazard after inhalation: a case-study. Part Fibre Toxicol. 2022;19.
37. Braakhuis HM, Murphy F, Ma-Hock L, Dekkers S, Keller J, Oomen AG et al. An integrated approach to testing and assessment to support grouping and Read-Across of nanomaterials after inhalation exposure. Appl Vitro Toxicol. 2021;7.
38. Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V et al. QSAR without borders. Chem Soc Rev. 2020;49.
39. Zhang F, Wang Z, Peijnenburg WJGM, Vijver MG. Machine learning-driven QSAR models for predicting the mixture toxicity of nanoparticles. Environ Int. 2023;177:108025.
40. Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A et al. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat Nanotechnol. 2011;6.
41. Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ et al. Quantitative nanostructure-activity relationship modeling. ACS Nano [Internet]. 2010;4:5703–12. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2997621&tool=pmcentrez&rendertype=abstract
42. Liu R, Zhang HY, Ji ZX, Rallo R, Xia T, Chang CH et al. Development of structure-activity relationship for metal oxide nanoparticles. Nanoscale. 2013;5.
43. Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, et al. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano. 2012;6:4349–68.
44. Fjodorova N, Novic M, Gajewicz A, Rasulev B. The way to cover prediction for cytotoxicity for all existing nano-sized metal oxides by using neural network method. Nanotoxicology. 2017;11.
45. Gajewicz A, Schaeublin N, Rasulev B, Hussain S, Leszczynska D, Puzyn T et al. Towards Understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: hints from nano-QSAR studies. Nanotoxicology. 2015;9.
46. Gajewicz A. Development of valuable predictive read-across models based on real-life (sparse) nanotoxicity data. Environ Sci Nano. 2017;4.
47. Huang Y, Li X, Xu S, Zheng H, Zhang L, Chen J et al. Quantitative structure–activity relationship models for predicting inflammatory potential of metal oxide nanoparticles. Environ Health Perspect. 2020;128.
48. Kotzabasaki MI, Sotiropoulos I, Sarimveis H. QSAR modeling of the toxicity classification of superparamagnetic iron oxide nanoparticles (SPIONs) in stem-cell monitoring applications: an integrated study from data curation to model development. RSC Adv. 2020;10.
49. Kotzabasaki M, Sotiropoulos I, Charitidis C, Sarimveis H. Machine learning methods for multi-walled carbon nanotubes (MWCNT) genotoxicity prediction. Nanoscale Adv. 2021;3.
50. Merugu S, Jagiello K, Gajewicz-Skretna A, Halappanavar S, Willliams A, Vogel U et al. The impact of carbon nanotube properties on lung pathologies and atherosclerosis through acute inflammation: a new AOP-Anchored *in Silico* NAM. Small. 2025;21.
51. Di Ianni E, Møller P, Mortensen A, Szarek J, Clausen PA, Saber AT et al. Organo-modified nanoclays induce less inflammation, acute phase response, and genotoxicity than pristine nanoclays in mice lungs. Nanotoxicology. 2020;14.
52. Billing AM, Knudsen KB, Chetwynd AJ, Ellis LJA, Tang SVY, Berthing T et al. Fast and robust proteome screening platform identifies neutrophil extracellular trap formation in the lung in response to Cobalt ferrite nanoparticles. ACS Nano. 2020;14.
53. Danielsen PH, Knudsen KB, Štrancar J, Umek P, Koklič T, Garvas M et al. Effects of physicochemical properties of TiO2 nanomaterials for pulmonary inflammation, acute phase response and alveolar proteinosis in intratracheally exposed mice. Toxicol Appl Pharmacol. 2020;386.
54. Wallin H, Kyjovska ZO, Poulsen SS, Jacobsen NR, Saber AT, Bengtson S et al. Surface modification does not influence the genotoxic and inflammatory effects of TiO2 nanoparticles after pulmonary exposure by instillation in mice. Mutagenesis. 2017;32.
55. Saber AT, Mortensen A, Szarek J, Jacobsen NR, Levin M, Koponen IK et al. Toxicity of pristine and paint-embedded TiO 2 nanomaterials. Hum Exp Toxicol. 2019;38.
56. Di Ianni E, Møller P, Cholakova T, Wolff H, Jacobsen NR, Vogel U. Assessment of primary and inflammation-driven genotoxicity of carbon black nanoparticles in vitro and in vivo. Nanotoxicology. 2022;16.
57. Kyjovska ZO, Jacobsen NR, Saber AT, Bengtson S, Jackson P, Wallin H et al. DNA damage following pulmonary exposure by instillation to low doses of carbon black (Printex 90) nanoparticles in mice. Environ Mol Mutagen. 2015;56.
58. Hadrup N, Rahmani F, Jacobsen NR, Saber AT, Jackson P, Bengtson S et al. Acute phase response and inflammation following pulmonary exposure to low doses of zinc oxide nanoparticles in mice. Nanotoxicology. 2019;13.
59. Barfod KK, Bendtsen KM, Berthing T, Koivisto AJ, Poulsen SS, Segal E et al. Increased surface area of Halloysite nanotubes due to surface modification predicts lung inflammation and acute phase response after pulmonary exposure in mice. Environ Toxicol Pharmacol. 2020;73.
60. Hadrup N, Saber AT, Kyjovska ZO, Jacobsen NR, Vippola M, Sarlin E et al. Pulmonary toxicity of Fe2O3, ZnFe2O4, NiFe2O4 and NiZnFe4O8 nanomaterials: inflammation and DNA strand breaks. Environ Toxicol Pharmacol. 2020;74.
61. Jacobsen NR, Pojana G, White P, Møller P, Cohn CA, Korsholm KS et al. Genotoxicity, cytotoxicity, and reactive oxygen species induced by single-walled carbon nanotubes and C60 fullerenes in the FE1-Muta^{TM}mouse lung epithelial cells. Environ Mol Mutagen. 2008;49.
62. Keller JG, Graham UM, Koltermann-Jülly J, Gelein R, Ma-Hock L, Landsiedel R et al. Predicting dissolution and transformation of inhaled nanoparticles in the lung using abiotic flow cells: the case of barium sulfate. Sci Rep. 2020;10.
63. Holmfred E, Loeschner K, Sloth JJ, Jensen KA. Validation and demonstration of an Atmosphere-TemperaturepH-Controlled stirred batch reactor system for determination of (Nano)Material solubility and dissolution kinetics in physiological simulant lung fluids. Nanomaterials. 2022;12.
64. Zanoni I, Keller JG, Sauer UG, Müller P, Ma-Hock L, Jensen KA et al. Dissolution rate of nanomaterials determined by ions and particle size under lysosomal conditions: contributions to standardization of simulant fluids and analytical methods. Chem Res Toxicol. 2022;35.
65. Hadrup N, Bengtson S, Jacobsen NR, Jackson P, Nocun M, Saber AT, et al. Influence of dispersion medium on nanomaterial-induced pulmonary inflammation and DNA strand breaks: investigation of carbon black, carbon nanotubes and three titanium dioxide nanoparticles. Mutagenesis. 2017;32:581–97.
66. Jackson P, Lund SP, Kristiansen G, Andersen O, Vogel U, Wallin H et al. An experimental protocol for maternal pulmonary exposure in developmental toxicology. Basic Clin Pharmacol Toxicol. 2011;108.
67. Hadrup N, Guldbrandsen M, Terrida E, Bendtsen KMS, Hougaard KS, Jacobsen NR, et al. Intratracheal instillation for the testing of pulmonary toxicity in mice—Effects of instillation devices and feed type on inflammation. Anim Model Exp Med. 2025;8:378–86.
68. Jackson P, Pedersen LM, Kyjovska ZO, Jacobsen NR, Saber AT, Hougaard KS et al. Validation of freezing tissues and cells for analysis of DNA strand break levels by comet assay. Mutagenesis. 2013;28.
69. Talete srl. Milano, Italy http://talete.mi.it/. Dragon Software for Molecular Descriptor Calculation. 2014.
70. Di Ianni E, Erdem JS, Narui S, Wallin H, Lynch I, Vogel U, et al. Pro-inflammatory and genotoxic responses by metal oxide nanomaterials in alveolar epithelial cells and macrophages in submerged condition and air-liquid interface: an in vitro-in vivo correlation study. Toxicol in Vitro. 2024;100:105897.
71. Dai Y, Guo Y, Tang W, Chen D, Xue L, Chen Y, et al. Reactive oxygen species-scavenging nanomaterials for the prevention and treatment of age-related diseases. J Nanobiotechnol. 2024;22:252.
72. Shukla RK, Badiye A, Vajpayee K, Kapoor N. Genotoxic potential of nanoparticles: structural and functional modifications in DNA. Front Genet. 2021;12.
73. Di Giampaolo L, Zaccariello G, Benedetti A, Vecchiotti G, Caposano F, Sabbioni E et al. Genotoxicity and immunotoxicity of titanium dioxide-embedded mesoporous silica nanoparticles (Tio2@msn) in primary peripheral human blood mononuclear cells (pbmc). Nanomaterials. 2021;11.
74. Xuan L, Ju Z, Skonieczna M, Zhou PK, Huang R. Nanoparticles-induced potential toxicity on human health: applications, toxicity mechanisms, and evaluation models. MedComm. 2023;4.
75. Xuan Y, Zhang W, Zhu X, Zhang S. An updated overview of some factors that influence the biological effects of nanoparticles. Front Bioeng Biotechnol. 2023;11.
76. Aljabali AA, Obeid MA, Bashatwah RM, Serrano-Aroca Á, Mishra V, Mishra Y et al. Nanomaterials and their impact on the immune system. Int J Mol Sci. 2023;24.
77. Summer M, Ashraf R, Ali S, Bach H, Noor S, Noor Q, et al. Inflammatory response of nanoparticles: mechanisms, consequences, and strategies for mitigation. Chemosphere. 2024;363:142826.
78. Ursini CL, Cavallo D, Fresegna AM, Ciervo A, Maiello R, Tassone P et al. Evaluation of cytotoxic, genotoxic and inflammatory response in human alveolar and bronchial epithelial cells exposed to titanium dioxide nanoparticles. J Appl Toxicol. 2014;34.

79. Havelikar U, Ghorpade KB, Kumar A, Patel A, Singh M, Banjare N, et al. Comprehensive insights into mechanism of nanotoxicity, assessment methods and regulatory challenges of nanomedicines. Discover Nano. 2024;19:165.
80. Bi J, Mo C, Li S, Huang M, Lin Y, Yuan P et al. Immunotoxicity of metal and metal oxide nanoparticles: from toxic mechanisms to metabolism and outcomes. Biomater Sci. 2023;11.
81. Balfourier A, Marty AP, Gazeau F. Importance of metal biotransformation in cell response to metallic nanoparticles: A transcriptomic Meta-analysis study. ACS Nanosci Au. 2023;3.
82. Sukhanova A, Bozrova S, Sokolov P, Berestovoy M, Karaulov A, Nabiev I. Dependence of nanoparticle toxicity on their physical and chemical properties. Nanoscale Res Lett. 2018;13.
83. Gajewicz-Skretna A, Furuhama A, Yamamoto H, Suzuki N. Generating accurate in Silico predictions of acute aquatic toxicity for a range of organic chemicals: towards similarity-based machine learning methods. Chemosphere. 2021;280.
84. Gajewicz-Skretna A, Kar S, Piotrowska M, Leszczynski J. The kernel-weighted local polynomial regression (KwLPR) approach: an efficient, novel tool for development of QSAR/QSAAR toxicity extrapolation models. J Cheminform. 2021;13.
85. core Team R. R: A Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
86. Hayfield T, Racine JS. The np Package. Relation. 2009.
87. Hayfield T, Racine JS. Nonparametric econometrics: the Np package. J Stat Softw. 2008;27.
88. Wickham H. ggplot2: elegant graphics for data analysis by, Wickham H. Springer New York. 2009.
89. OECD. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)Sar] Models. Transport. 2007.
90. De P, Kar S, Ambure P, Roy K. Prediction reliability of QSAR models: an overview of various validation tools. Arch Toxicol. 2022;96.
91. Gramatica P. Principles of QSAR models validation: internal and external. Environ Chem. 2007;26.
92. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model. 2002;20.
93. Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci. 2003;22.
94. Tropsha A. Best practices for QSAR model development, validation, and exploitation. Mol Inf. 2010;29.
95. Mitra I, Saha A, Roy K. Exploring quantitative structure-activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol Simul. 2010;36.
96. Rakhimbekova A, Madzhidov TI, Nugmanov RI, Gimadiev TR, Baskin II, Varnek A. Comprehensive analysis of applicability domains of QSPR models for chemical reactions. Int J Mol Sci. 2020;21.
97. Jaworska J, Nikolova-jeliazkova N, Aldenberg T. QSAR applicability domain Estimation by projection of the training set in descriptor space: A review. Methods. 2005; 33.
98. Boyles M, Murphy F, Mueller W, Wohlleben W, Jacobsen NR, Braakhuis H et al. Development of a standard operating procedure for the DCFH2-DA acellular assessment of reactive oxygen species produced by nanomaterials. Toxicol Mech Methods. 2022;32.
99. Schmid O, Stoeger T. Surface area is the biologically most effective dose metric for acute nanoparticle toxicity in the lung. J Aerosol Sci. 2016;99.
100. Cosnier F, Seidel C, Valentino S, Schmid O, Bau S, Vogel U et al. Retained particle surface area dose drives inflammation in rat lungs following acute, subacute, and subchronic inhalation of nanomaterials. Part Fibre Toxicol. 2021;18.
101. Di Ianni E, Erdem JS, Møller P, Sahlgren NM, Poulsen SS, Knudsen KB et al. In vitro-in vivo correlations of pulmonary inflammogenicity and genotoxicity of MWCNT. Part Fibre Toxicol. 2021;18.
102. Hadrup N, Saber AT, Kyjovska ZO, Jacobsen NR, Vippola M, Sarlin E, et al. Pulmonary toxicity of Fe2O3, ZnFe2O4, NiFe2O4 and NiZnFe4O8 nanomaterials: inflammation and DNA strand breaks. Environ Toxicol Pharmacol. 2020;74:103303.
103. Tran CL, Tantra R, Donaldson K, Stone V, Hankin SM, Ross B et al. A hypothetical model for predicting the toxicity of high aspect ratio nanoparticles (HARN). J Nanopart Res. 2011;13.
104. Hadrup N, Vogel U, Jacobsen NR. Biokinetics of carbon black, multi-walled carbon nanotubes, cerium oxide, silica, and titanium dioxide nanoparticles after inhalation: a review. Nanotoxicology. 2024;18:678–706.
105. Modrzynska J, Berthing T, Ravn-Haren G, Jacobsen NR, Weydahl IK, Loeschner K et al. Primary genotoxicity in the liver following pulmonary exposure to carbon black nanoparticles in mice. Part Fibre Toxicol. 2018;15.
106. Yang H, Liu C, Yang D, Zhang H, Xi Z. Comparative study of cytotoxicity, oxidative stress and genotoxicity induced by four typical nanomaterials: the role of particle size, shape and composition. J Appl Toxicol. 2009;29.
107. Knudsen KB, Berthing T, Jackson P, Poulsen SS, Mortensen A, Jacobsen NR et al. Physicochemical predictors of Multi-Walled carbon Nanotube–induced pulmonary histopathology and toxicity one year after pulmonary deposition of 11 different Multi-Walled carbon nanotubes in mice. Basic Clin Pharmacol Toxicol. 2019;124.
108. Berthing T, Lard M, Danielsen PH, Abariute L, Barfod KK, Adolfsson K et al. Pulmonary toxicity and translocation of gallium phosphide nanowires to secondary organs following pulmonary exposure in mice. J Nanobiotechnol. 2023;21.
109. Rittinghausen S, Hackbarth A, Creutzenberg O, Ernst H, Heinrich U, Leonhardt A et al. The carcinogenic effect of various multi-walled carbon nanotubes (MWCNTs) after intraperitoneal injection in rats. Part Fibre Toxicol. 2014;11.
110. Saleh DM, Alexander WT, Numano T, Ahmed OHM, Gunasekaran S, Alexander DB et al. Comparative carcinogenicity study of a thick, straight-type and a thin, tangled-type multi-walled carbon nanotube administered by intratracheal instillation in the rat. Part Fibre Toxicol. 2020;17.
111. Saleh DM, Luo S, Ahmed OHM, Alexander DB, Alexander WT, Gunasekaran S et al. Assessment of the toxicity and carcinogenicity of double-walled carbon nanotubes in the rat lung after intratracheal instillation: a two-year study. Part Fibre Toxicol. 2022;19.
112. Huang X, Teng X, Chen D, Tang F, He J. The effect of the shape of mesoporous silica nanoparticles on cellular uptake and cell function. Biomaterials. 2010;31.
113. Lee JH, Ju JE, Kim B, Il, Pak PJ, Choi EK, Lee HS et al. Rod-shaped iron oxide nanoparticles are more toxic than sphere-shaped nanoparticles to murine macrophage cells. Environ Toxicol Chem. 2014;33.
114. Cai X, Dong J, Liu J, Zheng H, Kaweeteerawat C, Wang F et al. Multi-hierarchical profiling the structure-activity relationships of engineered nanomaterials at nano-bio interfaces. Nat Commun. 2018;9.
115. Yamashita K, Yoshioka Y, Higashisaka K, Morishita Y, Yoshida T, Fujimura M et al. Carbon nanotubes elicit DNA damage and inflammatory response relative to their size and shape. Inflammation. 2010;33.
116. Halappanavar S, Sharma M, Solorio-Rodriguez S, Wallin H, Vogel U, Sullivan K et al. Substance interaction with the pulmonary resident cell membrane components leading to pulmonary fibrosis. OECD Series on Adverse Outcome Pathways, No. 33, OECD Publishing, Paris, https://doi.org/10.1787/10372cb8-en. 2023.
117. Cho W, Duffin R, Poland CA, Howie SEM, Macnee W, Bradley M. Metal oxide nanoparticles induce unique inflammatory footprints in the lung: important implications for nanoparticle testing. Environ Health. 2010;118.
118. Manuja A, Kumar B, Kumar R, Chhabra D, Ghosh M, Manuja M et al. Metal/metal oxide nanoparticles: toxicity concerns associated with their physical state and remediation for biomedical applications. Toxicol Rep. 2021;8.
119. Mahto SK, Yoon TH, Rhee SW. A new perspective on in vitro assessment method for evaluating quantum Dot toxicity by using microfluidics technology. Biomicrofluidics. 2010;4.
120. Nel A, Xia T, Mädler L, Li N. Toxic potential of materials at the nanolevel. Science. 2006;311.
121. Bengtson S, Kling K, Madsen AM, Noergaard AW, Jacobsen NR, Clausen PA et al. No cytotoxicity or genotoxicity of graphene and graphene oxide in murine lung epithelial FE1 cells in vitro. Environ Mol Mutagen. 2016;57.
122. Jackson P, Kling K, Jensen KA, Clausen PA, Madsen AM, Wallin H et al. Characterization of genotoxic response to 15 multiwalled carbon nanotubes with variable physicochemical properties including surface functionalizations in the FE1-Muta(TM) mouse lung epithelial cell line. Environ Mol Mutagen. 2015;56.
123. Shvedova AA, Pietroiusti A, Fadeel B, Kagan VE. Mechanisms of carbon nanotube-induced toxicity: focus on oxidative stress. Toxicol Appl Pharmacol. 2012;261.
124. Guo C, Xia Y, Niu P, Jiang L, Duan J, Yu Y et al. Silica nanoparticles induce oxidative stress, inflammation, and endothelial dysfunction in vitro via activation of the MAPK/Nrf2 pathway and nuclear factor-κB signaling. Int J Nanomed. 2015;10.
125. Poulsen SS, Saber AT, Mortensen A, Szarek J, Wu D, Williams A et al. Changes in cholesterol homeostasis and acute phase response link pulmonary exposure to multi-walled carbon nanotubes to risk of cardiovascular disease. Toxicol Appl Pharmacol. 2015;283.

126. Købler C, Poulsen SS, Saber AT, Jacobsen NR, Wallin H, Yauk CL et al. Time-Dependent subcellular distribution and effects of carbon nanotubes in lungs of mice. PLoS ONE. 2015;10.
127. Poulsen SS, Bengtson S, Williams A, Jacobsen NR, Troelsen JT, Halappanavar S et al. A transcriptomic overview of lung and liver changes one day after pulmonary exposure to graphene and graphene oxide. Toxicol Appl Pharmacol. 2021;410.
128. Bendtsen KM, Brostrøm A, Koivisto AJ, Koponen I, Berthing T, Bertram N et al. Airport emission particles: exposure characterization and toxicity following intratracheal instillation in mice. Part Fibre Toxicol. 2019;16.
129. Visani de Luna LA, Loret T, He Y, Legnani M, Lin H, Galibert AM et al. Pulmonary toxicity of Boron nitride nanomaterials is aspect ratio dependent. ACS Nano. 2023;17.
130. Jagiello K, Halappanavar S, Rybińska-Fryca A, Willliams A, Vogel U, Puzyn T. Transcriptomics-Based and AOP-Informed Structure–Activity relationships to predict pulmonary pathology induced by multiwalled carbon nanotubes. Small. 2021;17.
131. Le TC, Yin H, Chen R, Chen Y, Zhao L, Casey PS et al. An experimental and computational approach to the development of ZnO nanoparticles that are safe by design. Small. 2016;12.
132. Gernand JM, Casman EA. A meta-analysis of carbon nanotube pulmonary toxicity studies-how physical dimensions and impurities affect the toxicity of carbon nanotubes. Risk Anal. 2014;34.
133. Verdon R, Stone V, Murphy F, Christopher E, Johnston H, Doak S et al. The application of existing genotoxicity methodologies for grouping of nanomaterials: towards an integrated approach to testing and assessment. Part Fibre Toxicol. 2022;19.
134. Mesárošová M, Kozics K, Bábelová A, Regendová E, Pastorek M, Vnuková D et al. The role of reactive oxygen species in the genotoxicity of surface-modified magnetite nanoparticles. Toxicol Lett. 2014;226.
135. Haixia X, de Barros AO, da Mello SE, VC F, Sozzi-Guo F, Müller C, Gemini-Piperni S et al. Graphene: Insights on Biological, Radiochemical and Ecotoxicological Aspects. J Biomed Nanotechnol. 2021;17.
136. Raies AB, Bajic VB. In Silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci. 2016;6.
137. Ha MK, Trinh TX, Choi JS, Maulina D, Byun HG, Yoon TH. Toxicity classification of oxide nanomaterials: effects of data gap filling and PChem Score-based screening approaches. Sci Rep. 2018;8.
138. Romeo D, Louka C, Gudino B, Wigström J, Wick P. Structure-activity relationship of graphene-related materials: A meta-analysis based on mammalian in vitro toxicity data. NanoImpact. 2022;28.
139. Trinh TX, Choi JS, Jeon H, Byun HG, Yoon TH, Kim J. Quasi-SMILES-Based Nano-Quantitative Structure-Activity relationship model to predict the cytotoxicity of multiwalled carbon nanotubes to human lung cells. Chem Res Toxicol. 2018;31.
140. Furxhi I, Murphy F. Predicting in vitro neurotoxicity induced by nanoparticles using machine learning. Int J Mol Sci. 2020;21.

Publisher's note