NanoImpact 25 (2022) 100376

Contents lists available at ScienceDirect

NanoImpact

journal homepage: www.elsevier.com/locate/nanoimpact

Combined *in vitro-in vivo* dosimetry enables the extrapolation of *in vitro* doses to human exposure levels: A proof of concept based on a meta-analysis of *in vitro* and *in vivo* titanium dioxide toxicity data

Daina Romeo [a], Bernd Nowack [b], Peter Wick [a,*]

[a] *Empa, Swiss Federal Laboratories for Materials Science and Technology, Particles-Biology Interactions Laboratory, Lerchenfeldstrasse 5, 9014 St. Gallen, Switzerland*
[b] *Empa, Swiss Federal Laboratories for Materials Science and Technology, Technology and Society Laboratory, Lerchenfeldstrasse 5, 9014 St. Gallen, Switzerland*

ARTICLE INFO

Editor: Phil Demokritou

Keywords:
Occupational exposure limits
Risk assessment
Hazard
Benchmark dose

ABSTRACT

Evaluating the potential risks of nanomaterials on human health is fundamental to assure their safety. To do so, Human Health Risk Assessment (HHRA) relies mostly on animal studies to provide information about nanomaterials toxicity. The scarcity of such data, due to the shift of the nanotoxicology field away from a phenomenological, animal-based approach and towards a mechanistic understanding based on *in vitro* studies, represents a challenge for HHRA. Implementing *in vitro* data in the HHRA methodology requires an extrapolation strategy; combining *in vitro* dosimetry and lung dosimetry can be an option to estimate the toxic effects on lung cells caused by inhaled nanomaterials. Since the two dosimetry models have rarely been used together, we developed a combined dosimetry model (CoDo) that estimates the air concentrations corresponding to the *in vitro* doses, extrapolating in this way *in vitro* doses to human doses. Applying the model to a data set of *in vitro* and *in vivo* toxicity data about titanium dioxide, we demonstrated CoDo's multiple applications. First, we confirmed that most *in vitro* doses are much higher than realistic human exposures, considering the Swiss Occupational Exposure Limit as benchmark. The comparison of the Benchmark Doses (BMD) extrapolated from *in vitro* and *in vivo* data, using the surface area dose metric, showed that despite both types of data had a quite wide range, animal data were overall more precise. The high variability of the results may be due both to the dis-homogeneity of the original data (different cell lines, particle properties, etc.) and to the high level of uncertainty in the extrapolation procedure caused by both model assumptions and experimental conditions. Moreover, while the surface area BMDs from studies on rodents and rodent cells were comparable, human co-cultures showed less susceptibility and had higher BMDs regardless of the titanium dioxide type. Last, a Support Vector Machine classification model built on the *in vitro* data set was able to predict the BMD-derived human exposure level range for viability effects based on the particle properties and experimental conditions with an accuracy of 85%, while for cytokine release *in vitro* and neutrophil influx *in vivo* the model had a lower performance.

1. Introduction

The evaluation of engineered nanomaterials (ENM) potential toxicity to human health is a fundamental step to assure a safe integration of this technology in society. In this direction, Human Health Risk Assessment (HHRA) aims at estimating the risk posed by a substance, e.g. an ENM, to the human population, accounting for the potential of exposure and the hazard of the substance. The identification of the hazard requires quantitative toxicological information either from epidemiological studies, or, in their absence, from animal studies. Such dependency on *in vivo* studies is though a limiting factor for a timely assessment of new ENM, since such studies are resource-consuming and ethically concerning, and their accuracy and reproducibility have shown limitations (Gottmann et al., 2001; Basketter et al., 2004). Instead, the nanotoxicology field is evolving towards a combined approach involving mechanistic studies conducted *in vitro*, often generating a great amount of information (e.g. omics technology), and bioinformatics and *in silico* modelling to manage, mine, and integrate the experimental knowledge across disciplines (van Vliet, 2011; Hartung, 2009).

Whereas most toxicity and screening studies are now conducted *in vitro* using human cells, such data cannot directly substitute animal studies in HHRA; instead, an *in vitro* to *in vivo* extrapolation (IVIVE)

* Corresponding author.
E-mail addresses: Daina.Romeo@empa.ch (D. Romeo), Bernd.Nowack@empa.ch (B. Nowack), Peter.Wick@empa.ch (P. Wick).

https://doi.org/10.1016/j.impact.2021.100376
Received 9 June 2021; Received in revised form 10 December 2021; Accepted 10 December 2021
Available online 15 December 2021
2452-0748/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

strategy is needed to link cell responses to whole organism responses. In a previous study (Romeo et al., 2020), we identified a combination of *in vitro* dosimetry and lung dosimetry as a mature way for IVIVE of toxicity data about the effects of inhaled particles on the lung. The focus on this exposure route and target organ is of particular relevance for HHRA as inhalation is considered one of the most important entry routes of nanomaterials, especially in the workplace (Praphawatvet et al., 2020).

In vitro dosimetry simulates the deposition of particles in submerged *in vitro* systems, providing a more accurate dose than the simple concentration of particles in the media (Cohen et al., 2015). The behavior of the particles depends on the particle properties themselves and the experimental conditions, which have to be accurately measured (DeLoid et al., 2014; Deloid et al., 2017); diffusion, sedimentation, and (if applicable) dissolution processes are then modeled to estimate the amount of particles deposited on the cells (Thomas et al., 2018; DeLoid et al., 2015). Using the deposited dose has been shown to improve the correlation between *in vitro* and *in vivo* toxicity data (Pal et al., 2015a; Thrall et al., 2019). Lung dosimetry simulates the deposition of particles in the human or animal lung, thus identifying the amount of particles accumulating in different sections of the respiratory system, net of clearance removal processes (Asgharian et al., 2016). Lung dosimetry has been used both to extrapolate animal deposited doses to humans (Jung et al., 2018) but also to estimate relevant *in vitro* doses based on human exposure levels (Khatri et al., 2013; Gangwal et al., 2011; Smith et al., 2021). In a few cases, both models were used together: in the works from Demokritou et al. (2013) and Teeguarden et al. (2014), lung and *in vitro* dosimetry are used to compare *in vivo* and *in vitro* results by extrapolating both data to physiologically-equivalent effective doses (i.e. the deposited amount of particle per surface area or cell), respectively for cerium oxide and iron oxide nanoparticles. Pal et al. (2015b) instead proposed a procedure to monitor, sample, and characterize nanoparticles released on the workplace and then apply the dosimetry models to estimate the deposited doses in the human lung and the corresponding *in vitro* doses to use for toxicity testing; printer-emitted nanoparticles and incinerated polyurethane-carbon nanotubes composites are presented as case studies. In all of these cases, the nanoparticle aerosol was well-characterized, and the same particles were used *in vitro* as either tested *in vivo* or measured on the workplace. However, only in few cases a particle is tested at the same time *in vitro* and *in vivo*, and most often cells are exposed to primary particles rather than sampled particulate, making it challenging to link the *in vitro* dose to a human-relevant exposure level.

We believe that using realistic doses should be a priority in *in vitro* toxicity testing, even when no clear exposure scenario is available as benchmark (e.g. the animal test or the emission sampling data). However, despite dosimetry considerations not being new, the use of both dosimetry models to select relevant *in vitro* doses and compare *in vitro* and *in vivo* toxicity data is still not common practice. On a practical level, applying the two models "by hand" can be time consuming, making it difficult to apply them consistently beyond the single case study. To facilitate the application of these models we developed a combined dosimetry model (CoDo) that estimates the air concentrations for humans corresponding to *in vitro* doses. We show the potential of our model via a case study about titanium dioxide, verifying how many of the doses used *in vitro* are in a realistic range, estimating and comparing human Benchmark Doses (BMD) and BMD-derived human exposure levels from *in vitro* and *in vivo* data, and testing the possibility of estimating BMD-derived human exposure level ranges from the particle characteristics and the experimental conditions. The BMD represents the dose level at which a certain response level is observed, for example a 1% increase in disease incidence compared to control in epidemiological studies, and is derived by fitting a dose-response curve over experimental data (Davis et al., 2011). We chose the BMD as basis of comparison of ENM toxicity as such approach is recognised by the scientific and regulatory communities as an advanced method to estimate safe exposure levels in HHRA (Haber et al., 2018; Committee et al., 2017). While the BMD is expressed as dose per lung surface area (mg/cm^2 lung), the BMD-derived human exposure levels indicate the exposure concentration (mass of particles per volume of air, mg/m^3) over a defined exposure time corresponding to the BMD; by integrating the fate of the particle in the lung, such indicator allows a comparison with the occupational exposure levels, which are expressed in the same unit.

2. Methods

2.1. Combined dosimetry model

The combined dosimetry model (CoDo) was developed using Python programming language (Van Rossum et al., 2009) to simulate the exposure concentrations corresponding to the doses used in *in vitro* studies in submerged systems. It works by integrating *in vitro* dosimetry and lung dosimetry, and assuming that the deposited dose per area *in vitro* corresponds to the deposited dose per area in the lung (Fig. 1).

The input data include experimental parameters about the *in vitro* system and lung parameters that define the hypothetical human exposure scenario; the required parameters and the parameters that, if not specified by the user, are calculated by the model are shown in Table S1.

For the simulation of the deposition of particles *in vitro* we integrated the one-dimensional Distorted Grid (DG) model into CoDo; the behavior of the particles is simulated via subsequent rounds of sedimentation and diffusion repeated over small discrete time intervals, as described in DeLoid et al. (2015). By default, a reflective well bottom ("non-sticky") is selected, meaning that the particles reaching the bottom of the well (by default a 10 μm interaction layer) are subjected to weak non-specific interactions with the cells, and can be re-suspended due to diffusion forces; this choice is supported by the observations of DeLoid et al. (2015), who suggested that a reflective boundary condition is most likely for metal and metal oxide particles. A "sticky" bottom, i.e. a condition where particles have strong affinity with the cells and can therefore be removed from the system, can be selected; in this case an adsorption dissociation constant for agglomerate binding to cells of 10^{-9} is used as default; a different value for the adsorption dissociation constant can be entered by the user to represent intermediate levels of stickiness. The *in vitro* dosimetry simulation reports the deposited mass, surface area, and number of particles per cm^2 of *in vitro* plate.

For the calculation of the air concentrations corresponding to the *in vitro* doses the user can choose between two different deposition scenarios: a conservative estimate which assumes 100% deposition of the particles in the lung, and the use of a lung dosimetry model to estimate the retained dose based on the particle characteristics and exposure parameters. Regardless of the choice, four different exposure scenarios are considered: the same exposure time as *in vitro*, five days of exposure on the workplace (eight hours a day, five days a week), one year of exposure on the workplace, and 35 years of exposure on the workplace (the average working life in the European Union as of 2019, rounded-down (Eurostat, 2021)). The output of the model is, for each exposure scenario, the air concentration corresponding to the *in vitro* doses.

When 100% particle deposition is assumed, the calculation of the air concentration depends on the total amount of particles inhaled, which in turn depends on the breathing parameters and the exposure scenario:

$$\text{Air concentration} = \frac{(\text{Deposited dose per area} \cdot \text{AlveoliSA})}{(\text{Exposure time} \cdot \text{BF} \cdot \text{TV})}$$

where the "Deposited dose per area" has been calculated via *in vitro* dosimetry, the "Alveoli SA" (alveoli surface area) is 792,000 cm^2 for the average man and 559,000 cm^2 for the average woman (Brown et al., 2013), the "Exposure time" depends on the exposure scenario (e.g. 2400 min for five days of exposure on the workplace), the "BF" (breathing frequency) is 12 breaths/min for the average man and 14 breaths/min for the average woman, and the "TV" (tidal volume) is 625 mL for the average man and 464 mL for the average woman (Brown et al., 2013).

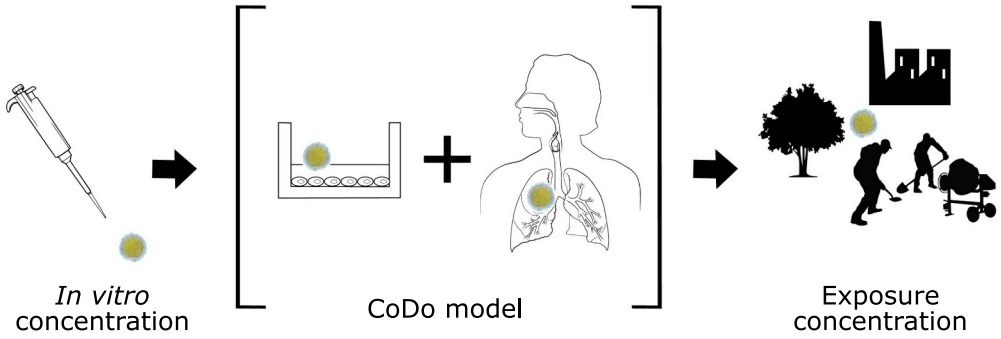


Fig. 1. CoDo model integrates *in vitro* dosimetry and lung dosimetry to estimate the exposure concentrations corresponding to *in vitro* doses.

Instead, if the lung dosimetry is used, the deposited mass of particles per area is divided by the fraction of particles retained in the lung per cm^2 at the end of the exposure time. The fraction of particles retained in the lung per cm^2 is calculated as the alveolar retention fraction divided by the lung surface area (792,000 cm^2 for the average man and 559,000 cm^2 for the average woman (Brown et al., 2013)). The alveolar retained fraction is calculated by automatically interacting with the MPPD model (Asgharian et al., 2012), including clearance processes. It should be noted that for short exposure times the deposited and retained doses correspond, while over longer exposure times the clearance process has a significant impact on the retained dose (Oyabu et al., 2013; Van Rijt et al., 2016).

For the human exposure simulation, the user can choose, through the "type of particle in air" parameter, to either indicate the aerodynamic diameter of the agglomerate in air, or to consider the primary particle in air, or to consider an agglomerate which has the same size and fractal dimension as the agglomerate measured *in vitro*, but where the pores are empty instead of filled with media. The effective density is automatically recalculated from the *in vitro* effective density (agg_density), the primary particle density (pp_density) and the media density via the formula (see SI for a demonstration of the equation):

$$\text{Air agglomerate density} = \frac{\text{agg\_density} - \text{media density}}{\text{pp\_density} - \text{media density}} \cdot \text{pp\_density}$$

While this scenario is not realistic, as agglomeration in air and in cell-culture media is driven by different processes (Anaraki et al., 2020; Schneider et al., 2009), we include it to represent exactly the same agglomerate to which the cells are exposed.

The effect of the "sticky bottom" parameter, the deposition scenarios, and the lung dosimetry type of particle in air parameter was tested with the titanium dioxide data set (Section 2.2), by varying the parameters one by one and comparing the model outputs.

2.2. Titanium dioxide data collection

A literature search of *in vitro* titanium dioxide toxicity data was performed on Scopus using the keywords "titanium dioxide *in vitro* toxicity" and "titanium dioxide *in vitro* inflammation", considering the time frame 2015–2020. The 249 results were further screened to select those that: (a) evaluated the effects on lung cells, (b) used human cells and eventually also murine macrophages (RAW264.7 cell line), (c) used spherical nanoparticles, (d) included endpoints on viability, reactive oxygen species production, and/or cytokine release (IL-6, IL-1β, TNFα, IL-8), (e) included the parameters needed to apply CoDo (reported in Table S1). Five additional papers published between 2012 and 2014 were included as well due to their completeness, resulting in 217 dose-response data sets extracted from 23 publications (see Supplementary files).

In vivo titanium dioxide toxicity data was collected via a literature search on Scopus using different combinations of the keywords "titanium dioxide", "*in vivo*", "rat", "mouse", "lung inflammation", "lung toxicity", and by screening review papers for references to *in vivo* studies. The criteria for inclusion were: (a) particles delivered via pulmonary administration route (e.g. via inhalation or intratracheal instillation), (b) reported particle primary size and/or aerodynamic diameter (for inhalation) or agglomerate diameter in media (for instillation), (c) at least two doses tested in addition to the negative control, (d) at least one endpoint among Bronchoalveolar lavage fluid (BALF) cytology, Lactate dehydrogenase (LDH) in BALF, reduced glutathione (GSH) in BALF, cytokine levels (IL-6, IL-1 β, TNFα, IFNγ) in BALF. 368 dose-response data sets were extracted from 28 publications (see Supplementary files).

2.3. Comparison with occupational exposure limits

The *in vitro* data set consisted of 484 dose values; multiple assays performed in the same experimental conditions in the same study were not double-counted. CoDo was applied choosing as lung dosimetry parameters the average man and the primary particle in air; as comparison, the conservative scenario assuming 100% deposition in the lung was also evaluated. For the *in vitro* dosimetry, a non-sticky bottom was chosen as the most realistic condition (DeLoid et al., 2015). The calculated air concentrations for the different exposure scenarios were then compared with the Swiss Occupational Exposure Limit (OEL) for titanium dioxide, equal to 3 mg/m^3, which is among the most conservative limits in the European area, lacking a unique value at EU level (GESTIS - International Limit Values For Chemical Agents, 2021).

2.4. Comparison of in vitro and in vivo benchmark doses and BMD-derived human exposure levels

Fig. 2 shows the procedure followed to calculate the BMD values (mass of particle deposited per cm^2 lung) and BMD-derived human exposure levels (corresponding air concentration) using respectively the deposited doses and the human-extrapolated doses from *in vitro* and *in vivo* data. For the *in vitro* data, we selected the air concentrations obtained via CoDo considering a non sticky bottom, the primary particle in air, and five days of exposure on the workplace. For the *in vivo* data set, the deposited dose in the lung of the animals was assumed to be 100% in the case of instillation, while for inhalation the retained dose was calculated via MPPD by including clearance processes and post-exposure time. Each deposited/retained dose was then extrapolated to the corresponding air concentration needed to obtain the same deposited dose per lung surface in the average man over five days of exposure. The Benchmark Dose (BMD) was calculated for each dose-response data set (*in vitro* and *in vivo*) with at least two doses in addition to control, using the PROAST software (Slob, 2018; Varewyck and Verbeke, 2017). A Benchmark response (BMR) of 20% was chosen for viability endpoints, ROS production and cytokine release (*in vitro*), and neutrophil (PMN) influx in BALF in absolute numbers and in percentage of the total cell amount, LDH in BALF and cytokines in BALF (*in vivo*). Such change is

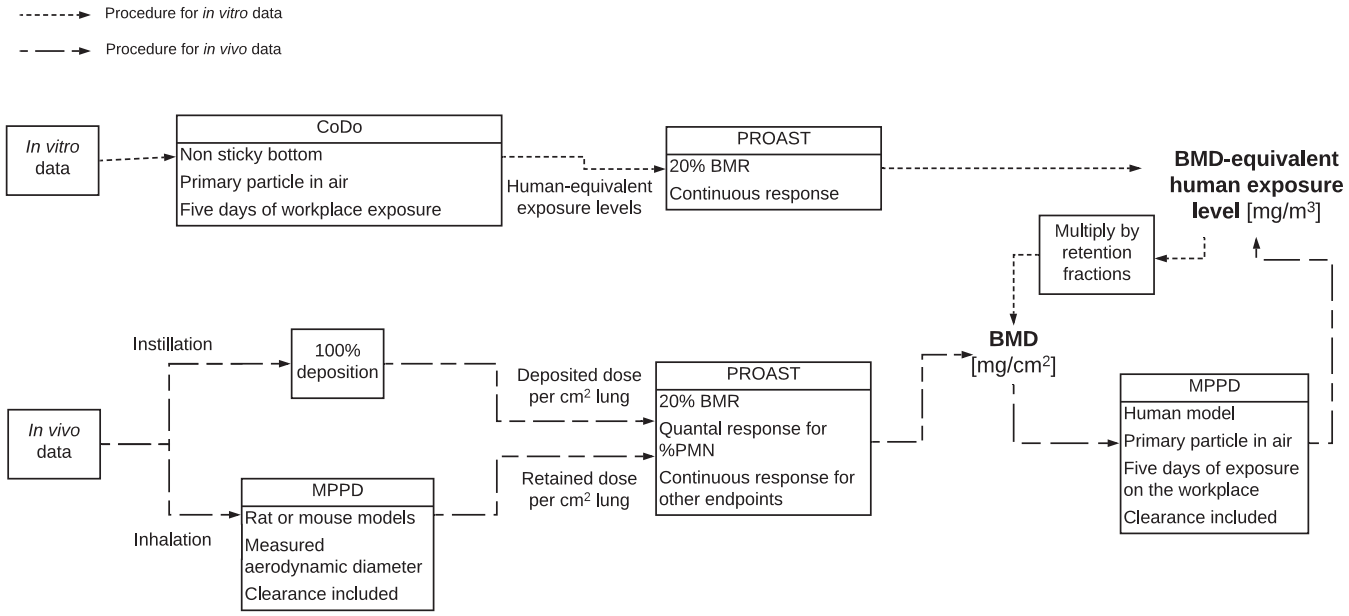


Fig. 2. The procedure followed to estimate the BMD and BMD-derived human exposure level values from *in vitro* and *in vivo* data. For *in vitro* data, CoDo is used to extrapolate the doses to human, which are then used together with the corresponding effects as dose-response data in input to the BMD calculation via PROAST, obtaining the BMD-derived human exposure levels. The values are then multiplied by the lung retention fractions to calculate the BMD in retained dose per lung surface area. For *in vivo* data, two different procedures are followed depending on the used exposure method. For instillation, data is included in the analysis only if the post-exposure is at max three days, and the deposited dose per alveoli surface is calculated by assuming 100% deposition. For inhalation, no boundary is set for the post-exposure time, which ranged from zero to sixteen days, since in this case the dose retained per cm^2 of alveoli can be calculated using the MPPD model, integrating clearance processes and post-exposure times in the simulation. At this point, PROAST is used to calculate the BMD from the dose-response data. The human dose is then estimated via MPPD to obtain the same deposited/retained dose per area as in the animal, obtaining the BMD-derived human exposure level over five days of exposure.

considered the threshold for cytotoxicity (ISO, 2009), and a sign of low inflammation (Noël et al., 2013; Vranic et al., 2017), and, in general, corresponds to the dose in which the slope of the dose-response curve changes the most in the low-dose region (Sand et al., 2006). The BMD in mg/cm^2 lung can be converted to BMD-derived human exposure levels in mg/m^3 by dividing the former value by the deposition fraction over five days of exposure, and vice versa.

For the comparison of BMD values, the *in vivo* data set was restricted to at maximum one week of exposure, and, in the case of particle administration via instillation to at maximum 72 hours of post-exposure time. In this way, similar exposure times are compared *in vitro* and *in vivo*. The exclusion of data with long post-exposure times when only the deposited dose but not the retained dose of particles was available is in line with the work from Cosnier et al. (2021), where it was shown that PMN influx had a strong dependence on the post-exposure time when considering the deposited dose, but not when using the retained dose. Using the deposited dose in place of the retained dose in the case of instillation is an acceptable approximation since the impact of clearance in the first post-exposure days has been shown to be small (Oyabu et al., 2013; Van Rijt et al., 2016).

Surface area was used as dose metric due to its higher predictivity of dose-response relationships for inhaled particles (Schmid and Stoeger, 2016).

2.5. SVM classification of in vitro and in vivo BMD-derived human exposure levels

Three Linear Support Vector Machine (SVM) classifiers (Yue et al., 2003) were built on the *in vitro* and *in vivo* data sets, considering respectively the BMD-derived human exposure levels for viability, for cytokine release *in vitro*, and for PMN influx effects, as such endpoints represented the majority of the data (respectively 55, 59 and 72 values). Considered inputs for the classification were, for the *in vitro* data set, the diameter of the primary particle, the diameter of the agglomerate, the exposure time *in vitro*, the assay used, the specific surface area of the particle, the type of particle (anatase, rutile, or any mixture), the presence or absence of serum in media, and the cell type. For the *in vivo* data set, we considered the diameter of the primary particle, the exposure length, the specific surface area of the particles, the post-exposure time before the effects had been measured, the type of particle, the animal species, the administration route (simplified as either inhalation or instillation), and the sex of the animal. Numerical inputs were normalized using min-max transformation while nominal inputs were transformed into numerical dummy variables (i.e. with 0 or 1 value) via One-Hot encoding.

The classifier used a one-vs-rest multi-class strategy and minimized the hinge loss (Hsu and Lin, 2002); the best number and combination of parameters was identified via a sequential feature selection algorithm by maximizing the leave-one-out cross-validation (LOOCV) accuracy (Chandrashekar and Sahin, 2014). LOOCV consists in iteratively training the model on all the data set except for one sample, which is used for validation; the accuracy of the model is calculated as the number of times the model correctly classified the validation sample, expressed as percent of the total number of classifications (Wong, 2015). This validation method provides an accurate estimation of the model performance, and is particularly appropriated to be used with small data sets, as it is computationally expensive (Wong, 2015). A grid search algorithm was used to select the best cost parameter "c" based again on the LOOCV accuracy. A different number of classes were considered, with the main goal of distinguishing BMD-derived human exposure levels in a realistic concentration range and higher concentrations. The optimal number of classes was chosen based not only on the maximization of the LOOCV overall accuracy, but also to maximize the f1-score of the first class (i.e. the lowest BMD-derived human exposure level range), with f-1 being the weighted average of the model precision (True positives/Total positives) and sensitivity (True positives/(True positives

+ False negatives).

2.6. Statistical analysis

The difference between Benchmark Doses (and BMD-derived human exposure levels) calculated from *in vitro* and *in vivo* data was evaluated via Welch's t-test with Bonferroni correction, using Python library SciPy (Virtanen et al., 2020).

3. Results and discussion

3.1. Effect of parameters on CoDo results

Table 1 reports the effect of the sticky bottom parameter, the deposition type parameter, and the type of particle in air parameter on the air concentrations estimated from the database of titanium dioxide *in vitro* doses. For the bottom stickiness parameter, the selection of sticky conditions resulted in the majority of cases in an increased deposition of particles in the well bottom, which therefore meant a higher concentration of particles in air was needed to obtain higher deposited amounts in the lung. The median of the sticky over non-sticky ratio was 6.6, i.e. the sticky conditions showed six times the deposition than in non-sticky conditions. The range is however very wide, ranging from no difference in deposition (ratio = 1) to 70 times more deposition (ratio = 70). Such differences are driven by the contribution of diffusion versus sedimentation processes on the deposition of particles; in fact, the biggest differences were observed for small particles forming relatively small agglomerates (generally below 250 nm), while with bigger agglomerates and bigger primary particles the differences in deposition were more modest, or even negligible in the case of micro-sized agglomerates.

Using the lung dosimetry model considering the primary particle in air results in higher corresponding air concentration than if 100% of the particles is assumed to deposit in the lung, as expected. The range of the ratio between air concentrations calculated using dosimetry or assuming full deposition is very wide, ranging from 3.7 times to 159.1 times, indicating that the particle size has a strong effect on the deposition fraction and that at max around a third of the particles effectively deposit in the lung.

The impact of the type of particle parameter is much more contained, with a median of the agglomerate over primary particle air concentration of 2.3, and an interquartile range of 1.6 to 2.9. In most cases, considering agglomerates in air resulted in lower deposition of the particles in the lung, and therefore higher air concentrations were estimated to obtain the same deposited dose as calculated from the *in vitro* doses. This can be explained by the difference in pulmonary deposition according to the particles size (Fig. S1), which sees a declining trend for particles bigger than 30 nm. Only with very small primary particles (e.g. 5 nm), the deposition of the agglomerates was higher (agglomerates over primary particle ratio < 1).

Another source of uncertainty in the simulation of *in vitro* dosimetry is the calculation method for the agglomerate effective density; as demonstrated by DeLoid et al. (2014), the experimental volumetric centrifugation method is a more accurate method than the estimation of the parameter via Sterling equation, with the latter being either in agreement with or overestimating the measured agglomerate effective density. However, this parameter is not often measured, except for those studies that apply a dosimetry model; this is why CoDo uses the Sterling equation when an agglomerate effective density is not provided. To avoid this source of uncertainty, we recommend to follow the protocol by Deloid et al. (2017) for the characterization of the particle properties in the *in vitro* system.

Table 1
The effect of the stickiness, the deposition type, and particle type in air parameters on the estimated air concentration of particles. The showed medians and quartile ranges refer to the ratio between the air concentrations calculated considering a sticky bottom versus a non-sticky bottom for the stickiness parameter, lung dosimetry considering the primary particle in air versus 100% deposition for the deposition type, and air agglomerates versus primary particles in air for the particle type in air parameter.

Parameter	Compared result	Median value	First–third quartile range	Min–max range
Stickiness of bottom	Sticky/non sticky	6.6	1.4–27.6	1.0–70.1
Deposition type	Dosimetry considering primary particle/100% deposition	4.1	4.0–6.0	3.7–159.1
Particle type in air	Agglomerates/ primary particles	2.3	1.6–2.9	0.1–4.0

3.2. Comparison of in vitro doses and occupational exposure limit

The comparison of *in vitro* doses with the OEL value indicates that most *in vitro* doses are representative of long human exposures. Fig. 3 shows the distribution of the *in vitro* doses based on the ratio between the extrapolated air concentration and the Occupational Exposure Limit when considering the same exposure time for humans as the *in vitro* experiment. In Fig. 3a, which shows the results when applying the lung dosimetry model considering the primary particle in air, 11% of the doses are below the OEL (i.e. have an air concentration over OEL ratio between zero and one), and another 20% are between one and ten times the OEL. Even with the conservative assumption of 100% deposition in the lung only 24% of the dose are below the OEL, while 50% are more than ten times the exposure limits (Fig. 3b). Instead, when extrapolating the *in vitro* doses to a year-long human exposure on the workplace (Fig. 4), three quarters of the doses are less than ten times the OEL, with 51% below the OEL itself. The results indicate that only low lung deposited doses (corresponding to the lower *in vitro* doses) are reached in law-abiding workplaces after short human exposure times, while most *in vitro* doses depict deposited levels reached after a year of workplace exposure. It should be noted, however, that cells exposed to a single dose do not have the same bio-response as when exposed to repeated doses over a longer exposure time, even if the cumulative dose is the same (Annangi et al., 2016; Torres et al., 2020; Thurnherr et al., 2011; Mukherjee et al., 2020).

Ten years ago, Gangwal et al. (2011) used a similar, but reversed, approach to suggest realistic *in vitro* dose ranges based on occupational exposure levels. Their study assumed total deposition of particles *in vitro*, which we found true only for the biggest agglomerates and/or for longer exposure times, indicating that their *in vitro* concentrations would in most cases be underestimated. A critique, at the time, was that estimating the highest *in vitro* concentrations from the deposited amount of particles over 45 years of exposure would result in very high doses to be administered all at once, compared to a long-term accumulation (Oberdörster et al., 2012). While CoDo includes as well such long-term exposure scenarios (35 years), which can be useful to clearly identify extreme doses (as in Fig. S2a), we recommend to choose the dose range for acute studies based on the short-term exposure levels, considering that legal thresholds may be exceeded either because of concentration spikes (as the OEL is calculated as an average concentration over the exposure time), but also because of non-compliance or lack of regulation. For longer repeated exposures, the one year exposure level may be used as upper benchmark concentration. In both cases, it is important to verify that the chosen doses do not conflict with experimental constraints, such as assay interference (Ong et al., 2014). Moreover, the impact of different stickiness conditions should be considered, as the stickier the bottom the higher the risk of exceeding realistic conditions, as shown by comparing Fig. S3b with Fig. S2b, which shows the distribution of *in vitro* doses when considering one year of exposure and a sticky bottom.

A last consideration should be made about the OEL used as basis for

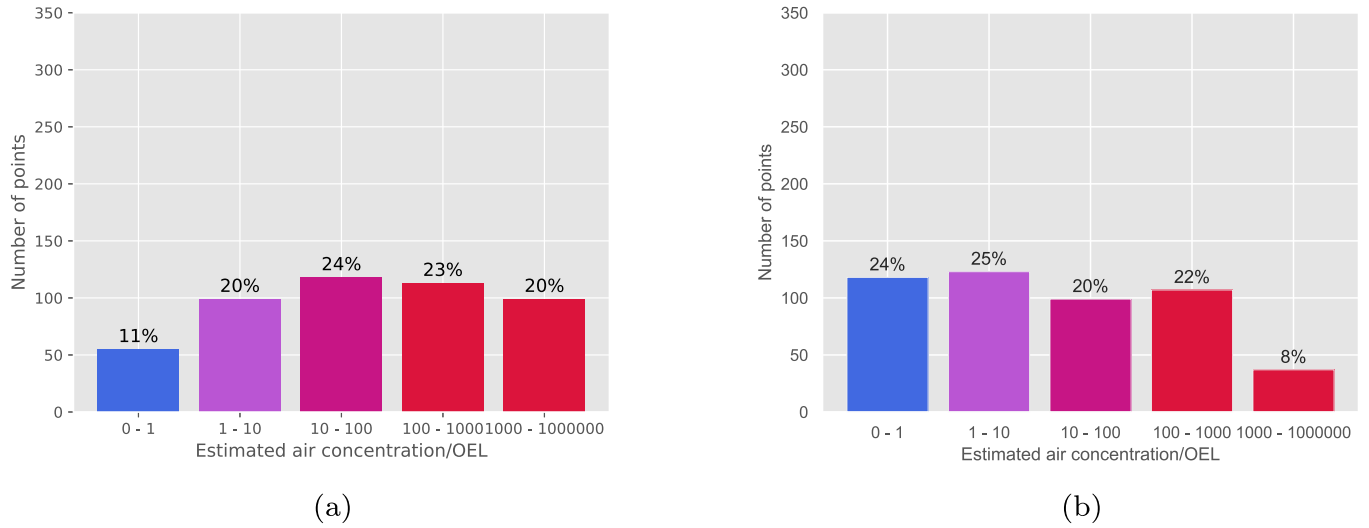


Fig. 3. The distribution of the *in vitro* doses based on their extrapolated air concentrations over OEL ratio considering the same exposure time as *in vitro*, when using lung dosimetry and assuming the primary particle in air (a), and with the conservative assumption of 100% deposition in the lung (b). The *y* axis ("Number of points") reports the absolute number of data points, i.e. doses, while the percentage of doses belonging to each range is indicated over each bar. The bottom of the well is considered non-sticky. $N = 484$.

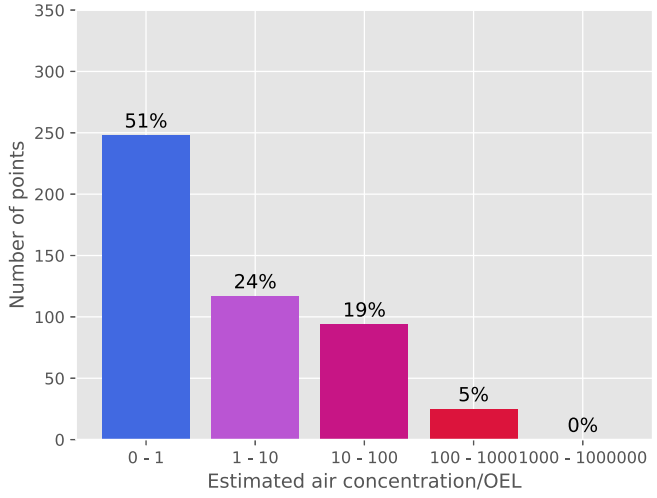


Fig. 4. The distribution of the *in vitro* doses based on their extrapolated air concentrations over OEL ratio, considering one year of exposure on the workplace. The *y* axis ("Number of points") reports the absolute number of data points, i.e. doses, while the percentage of doses belonging to each range is indicated over each bar. The bottom of the well is considered non-sticky. $N = 484$.

comparison; depending on the country, a slightly different exposure limit may apply, and this would affect the classification of the *in vitro* data; this depends in part on the fraction of particles the limit applies to, since in certain cases a specific limit for the nano-sized fraction exists, while in others the limit refers to the inhalable fraction. Using for example the NIOSH limit for ultrafine TiO_2 particles, which is 0.3 mg/m^3, would result in even more *in vitro* doses to be above the legal limit. This because the OEL is set to protect workers' health and is derived from animal data, to which safety factors are applied to account for uncertainties in the extrapolation; the more stringent limit for the ultrafine fraction of titanium dioxide reflects the higher risk posed by the nanomaterial compared to its bulk counterpart (Schulte et al., 2010).

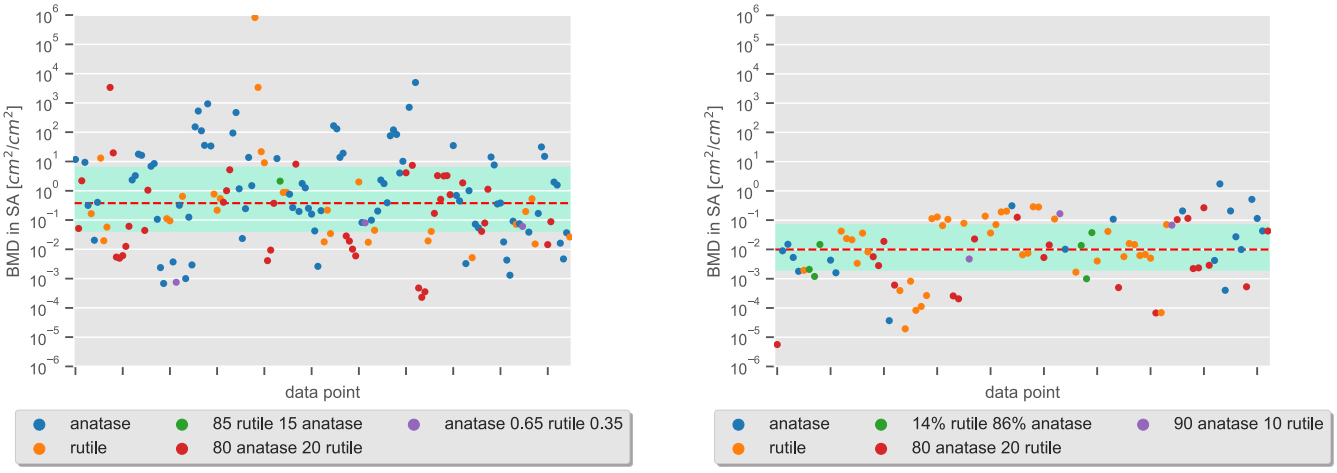
3.3. *In vitro and in vivo benchmark dose and BMD-derived human exposure level comparison*

The comparison of the BMDs in surface area dose extrapolated from *in vitro* and *in vivo* data showed for both data sets a wide range of values, extending over more than four orders of magnitude (Fig. 5). No clear trend for specific titanium dioxide types emerged, nor any difference based on the endpoint considered (Figs. S3 and S4). The *in vitro* BMD median value was around 0.38 cm^2/cm^2, with an interquartile range of roughly two orders of magnitude ($4.1 \cdot 10^{-2}$–6.42 cm^2/cm^2), while *in vivo* BMDs were generally lower, with the median at 0.01 cm^2/cm^2 and the interquartile range between $1.99 \cdot 10^{-3}$ and $7.0 \cdot 10^{-2}$ cm^2/cm^2. Overall, *in vivo* data were one order of magnitude more precise than *in vitro* ones, both when considering 50% and 90% of the data centered on the median (assuming that the 5% lowest and highest values might be outliers).

The differences between *in vitro* and *in vivo* BMDs may be attributed to the differences between human (cells) and animals, between *in vitro* and *in vivo* endpoints, but also to the different level of uncertainty of *in vitro* dosimetry and animal lung dosimetry.

A more detailed comparison of animal and cell-based BMD values in surface area dose for inflammatory endpoints showed similar values for rat, mouse, and murine macrophages, while human cells turned out to be less susceptible to inflammatory effects, as shown in Fig. 6. Both *in vitro* and *in vivo* endpoints, respectively the release of pro-inflammatory cytokines and the neutrophil influx in BALF, are indicators of inflammation, and have been suggested as one of the most promising endpoints for IVIVE (Donaldson et al., 2008). The rat BMD values were generally lower than the mouse ones (despite the difference not being statistically significant), in line with the observations of comparative studies (Bermudez et al., 2004; Warheit, 2011). For anatase TiO_2, the murine cell line showed values similar to the animal data, while human co-cultures of A549 epithelial cells and THP1 macrophages had a BMD range significantly higher than the animal one (370 times higher than the rat, 96 times higher than the mouse). The dendritic cells showed BMD values similar to the mouse data, but higher than the rat and significantly lower than the co-culture. THP-1 monocytes had a median value similar to the co-culture one,and were statistically different from the rat data.

Comparing the animal and human co-culture BMD for NM105/P25 TiO_2, a mixture of 80% anatase and 20% rutile, primary particle size 21 nm, confirmed the significantly higher BMD for the co-culture



(a) N=158 (b) N=93

Fig. 5. The distribution of the human BMD values extrapolated from the *in vitro* data set (a), and the *in vivo* data set (b), for different titanium dioxide types, expressed in surface area dose. The red dashed line represents the median value, and the aquamarine area the interquartile range of the distribution (25–75%).(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

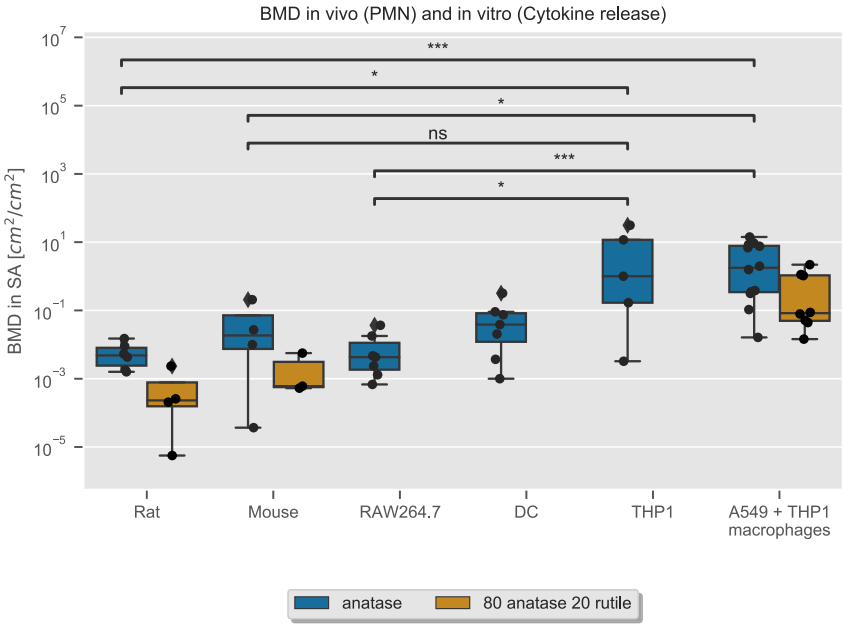


Fig. 6. Comparison of the rat and mouse BMD in surface area for PMN influx in BALF and the BMD in surface area for cytokine release (IL-6, IL-1β, TNFα, IL-8) of multiple cell lines: the murine macrophage RAW264.7, human dendritic cell (DC) monoculture, human monocytes THP-1, and a coculture of human epithelial cells A549 and human macrophages differentiated from THP-1 cells. The colored boxes represent the interquartile range of BMD values, the single points are outliers, calculated as points exceeding 1.5 times the interquartile range past the high or low quartile (represented by the whiskers). * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, *** $0.0001 < p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

compared to the rat (362 times higher) and the mouse (138 times higher), even though we couldn't ascertain the similarity of murine macrophages and animal BMD due to lack of data.

When comparing the data it should be kept in mind that the calculated BMD values have a high uncertainty, as multiple experimental factors may affect the results; this was clear for example in the P25 data, which, despite using the same particle, showed a very high variability both *in vitro* and *in vivo*. A limitation of the analysis is the scarcity of comparable *in vitro* and *in vivo* studies; for example, as observable in Fig. 5, most animal studies tested rutile TiO_2, while *in vitro* studies focused more on the anatase form. This prevented a more robust and comprehensive comparison of BMDs across species and cell lines.

3.4. Testing the surface area dose metric hypothesis

In our comparison of BMD values we did not observe any relevant difference when using the mass dose or the surface area dose. As surface area has been often reported as the most relevant metric for acute pulmonary toxicity (Schmid and Stoeger, 2016; Monteiller et al., 2007), we wanted to verify whether our results depended on the use of the human exposure concentration estimated via CoDo, or whether they were due to the high variability of the data in terms of particles properties and experimental conditions. Therefore, we applied our model to the data set from Rushton et al. (2010), for which the authors observed a linear relationship between the *in vitro* and *in vivo* steepest slope of the dose-response curve of eight different particles, when considering the surface area dose (Fig. 7a). The linear relationship was maintained when using the human air concentration in surface area as dose (Fig. 7b), indicating that extrapolating to humans via CoDo does not "hide" the relationship between *in vitro* and *in vivo* effects when they are observed using surface area doses. We could though not test the existence of this relationship on our TiO_2 data set due to the scarcity of corresponding *in*

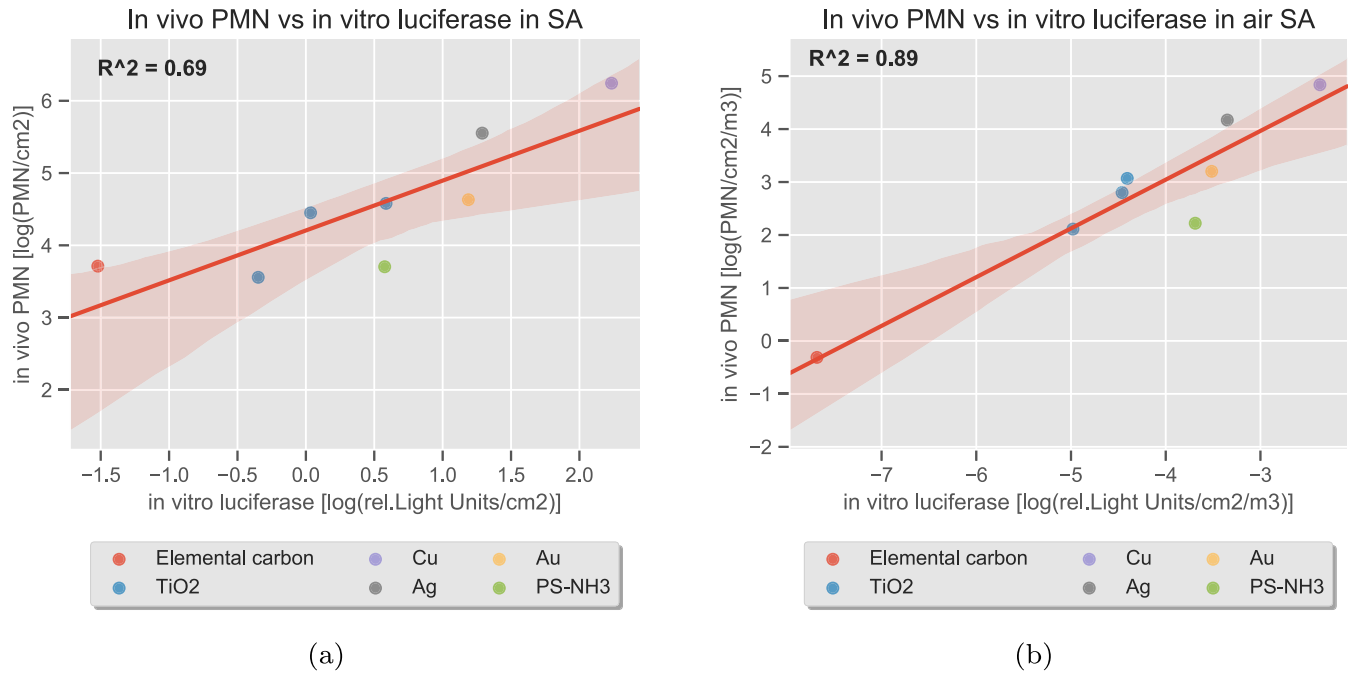


Fig. 7. Correlations between *in vitro* luciferase-transfected human type II lung epithelial cell line (A549 Luc-1) luciferase response and *in vivo* inflammatory response (number of neutrophils). (a) Responses normalized to the instilled surface area dose (*in vivo*) and particle concentration in media in surface area dose (*in vitro*). (b) Responses normalized to the human-extrapolated air concentrations in surface area dose.

vitro and *in vivo* data.

3.5. SVM classification models

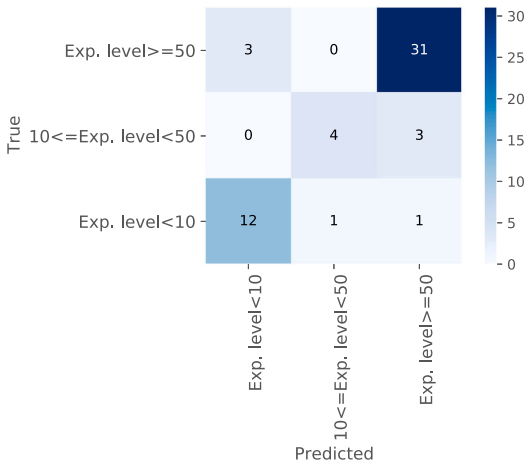
The optimal SVM model built on the *in vitro* viability data set used six features to classify the data in three BMD-derived human exposure level ranges, considering five days of workplace exposure. The selected features are: the primary particle diameter, the agglomerate diameter, the presence or absence of serum in media, the anatase type, the dendritic cell (DC) cell type, and the LDH assay. Since the classifier is built on the provided data set, the relevance of different features is affected by how these features are represented in the training data. For example, while the exposure time is a known parameter affecting toxicity, it was not a good predictor in our model because the training data it was built on included mostly exposure times of 24 hours, i.e. the effect of the exposure time on toxicity was not well-represented in the data set. The data was classified in the classes: BMD-derived human exposure level < 10 mg/m^3, 10 mg/m^3 $\leq$ BMD-derived human exposure level < 50 mg/m^3, BMD-derived human exposure level ≥ 50 mg/m^3. The model has an overall accuracy of 85% (Table 2 and Fig. 8a), correctly classifying most data points belonging to the lowest and highest classes, i.e. BMD-derived human exposure level ≤ 10 mg/m^3 and BMD-derived human exposure level > 50 mg/m^3. The analysis of the coefficients (Figs. S5, S6, and S7) shows the importance in particular of the LDH assay, the primary particle diameter, the anatase type and the presence/absence of serum in media for the classification into the three different classes.

For cytokine release, the optimal model classified the data in three classes: BMD-derived human exposure level < 10 mg/m^3, 10 mg/m^3 $\leq$ BMD-derived human exposure level < 50 mg/m^3, BMD-derived human exposure level ≥ 50 mg/m^3, and used six features: the presence or absence of serum in the media, the agglomerate diameter, the exposure time, the 65% anatase 35% rutile titanium dioxide type, and multiple cell types, i.e. A549+THP1 macrophages, human bronchial epithelial cell (16HBE)+ THP1+Human Lung Microvascular Endothelial Cells (Hlmvec), murine macrophages (RAW264.7), human monocyte-derived macrophages (hMDM), Human lung epithelial cells (BEAS-2B), and dendritic cells (DC). The SVM model had a worse performance than for viability endpoints, with a total accuracy of 66% (Fig. 8b and Table 2. The type of cells used and the presence or absence of serum were important features for the classification (Figs. S8, S9, and S10).

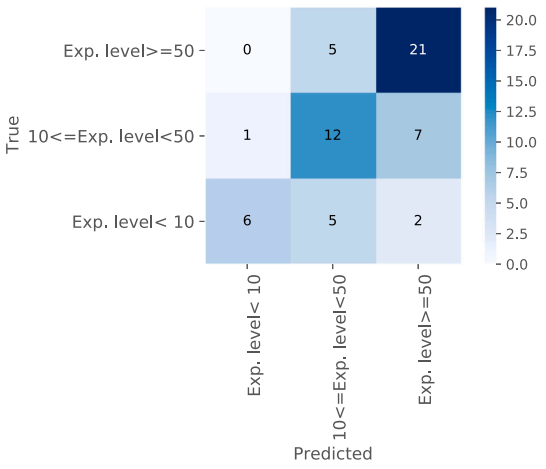
Table 2
The performance of the SVM classifiers for viability and cytokine release built on the *in vitro* data. Exp. level indicates the BMD-derived human exposure level considering five days of workplace exposure. Precision measures the number of true positives over the sum of true and false positives. Sensitivity measures the ratio of true positives over true positives and false negatives. F1-score is a weighted average of precision and sensitivity. Accuracy is the ratio of correctly classified data points.

Class	Precision	Sensitivity	F1-score	Number of data points
Viability				
Exp. level < 10 mg/m^3	0.80	0.86	0.83	14
$10 \leq$ Exp. level $<$ 50 mg/m^3	0.80	0.57	0.67	7
Exp. level ≥ 50 mg/m^3	0.89	0.91	0.90	34
Accuracy		0.85		55
Cytokine release (IL-6, IL-1β, TNFα, IL-8)				
Exp. level < 10 mg/m^3	0.86	0.46	0.60	13
$10 \leq$ Exp. level $<$ 50 mg/m^3	0.55	0.60	0.57	20
Exp. level ≥ 50 mg/m^3	0.70	0.81	0.75	26
Accuracy		0.66		59

For the *in vivo* data, the optimal classifier utilized only three features, i.e. the primary particle diameter, the length of exposure, and the rutile type, to classify the data in three classes for the neutrophil influx in BALF endpoint: BMD-derived human exposure level < 5 mg/m^3, 5 mg/m^3 $\leq$ BMD-derived human exposure level < 8 mg/m^3, BMD-derived human exposure level ≥ 8 mg/m^3. Fig. 9 and Table 3 report the accuracy, precision, sensitivity and F1-score for the different classes. The model correctly classifies the lowest class in 91% of the cases, while the



(a) Viability



(b) Cytokine release (IL-6, IL-1β, TNFα, IL-8)

Fig. 8. Confusion matrix of the SVM classification model built on the *in vitro* data set for (a) viability endpoints (55 data points), and (b) cytokine release endpoints (59 data points). The columns indicate in which BMD-derived human exposure level range (Exp. level in the figure) each data point was classified, i.e. the predicted class, while the rows indicate in which class the data point really belongs. The anti-diagonal (i.e. the diagonal from top right to bottom left) indicates the number of data points correctly classified in each class.

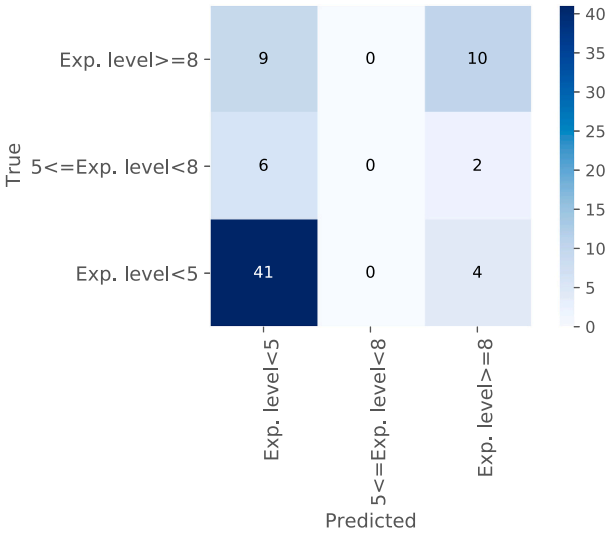


Fig. 9. Confusion matrix of the SVM classification model built on the *in vivo* data set (72 data points), for the neutrophil influx in BALF. The columns indicate in which BMD-derived human exposure level range (Exp. level in the figure) each data point was classified, i.e. the predicted class, while the rows indicate in which class the data point really belongs. The anti-diagonal (i.e. the diagonal from top right to bottom left) indicates the number of data points correctly classified in each class.

precision, i.e. the number of true positives over the sum of true and false positives, is 73%, versus 63% of the baseline (probability to randomly guess correctly the belonging to the first class). The misclassification of the data belonging to the middle class suggests that, in addition to the limited amount of data points, the data has a lot of noise and/or some data points are outliers. To strengthen the model, and in particular the middle class predictions, future experiments would have to focus on replicating the conditions of the middle class data set. The analysis of the coefficients in Figs. S11, S12, and S13 show that the primary particle diameter and the rutile type are the most important features, while the total length of exposure contributes marginally.

Table 3
The performance of the SVM classifier for PMN influx in BALF built on the *in vivo* data. Exp. level indicates the BMD-derived human exposure level considering five days of workplace exposure. Precision measures the number of true positives over the sum of true and false positives. Sensitivity measures the ratio of true positives over true positives and false negatives. F1-score is a weighted average of precision and sensitivity. Accuracy is the ratio of correctly classified data points.

Class	Precision	Sensitivity	F1-score	Number of data points
Exp. level $<$ 5 mg/m^3	0.73	0.91	0.81	45
5 $\leq$ Exp. level $<$ 8 mg/m^3	0.00	0.00	0.00	8
Exp. level $\geq$ 8 mg/m^3	0.62	0.53	0.57	19
Accuracy		0.71		72

3.6. Limitations

Great uncertainties exist in the calculation of human exposure levels corresponding to *in vitro* doses. As presented in Section 3.1, both *in vitro* and lung dosimetry parameters affect the results. While the user may test the different conditions and calculate a range rather than a single value, there are other sources of uncertainties on which less control can be exerted at modeling level. For example, the model assumes that the agglomerate in the *in vitro* system are stable over time, and while a dispersion procedure has been published by Deloid et al. (2017), it is not always followed (especially in older studies). Moreover, considering only the median diameter is a simplification of the real size distribution of the particles, and depending on the method used to measure this property a different value may be obtained, for example DLS has been shown to be greatly affected by the presence of large particles compared to NTA (Filipe et al., 2010). Concerning lung dosimetry, the lack of information about the particle agglomeration state in the air and the correspondence between an airborne particle and the particle tested *in vitro* represents a challenge. The approach proposed by Pal et al. (2015b) solves this issue by considering the emission of particles on the workplace as the starting point: the aerosol is collected and used in *in vitro* tests, after applying lung and *in vitro* dosimetry to assure correct dosing. In this way, the cells are exposed to the airborne particles, which though could still be altered by e.g. the interaction with the media. However,

CoDo takes the opposite approach, i.e. starting from the *in vitro* dose and moving up towards an hypothetical exposure scenario, thus not requiring any environmental sampling or knowledge of specific exposure conditions. This makes the model more accessible to the nanotoxicology community, but also increases the uncertainty of the results.

All these sources of uncertainty should be kept in mind also when calculating BMD and BMD-derived human exposure levels; additionally, further variability is introduced for example by differences in *in vitro* and *in vivo* experimental conditions that may affect the toxicity of the particles, such as the presence of serum in the media (Vranic et al., 2017). Last, choosing representative endpoints *in vitro* remains an important issue for the extrapolation of *in vitro* effects to human-relevant endpoints (Romeo et al., 2020).

4. Conclusions

Applying the theoretical framework developed in our previous publication (Romeo et al., 2020), we developed a combined dosimetry model (CoDo) that estimates the human exposure concentrations corresponding to the doses used *in vitro*. Our analysis of titanium dioxide data confirms that most *in vitro* doses are still quite high, being representative of long exposure times. CoDo can be used retrospectively to assess the doses used in *in vitro* studies, but also prospectively to select realistic doses during the design of an experiment.

The wide range covered by the human surface area BMDs extrapolated from *in vitro* and *in vivo* data suggests that both data sources have a large inter-study variability, but *in vivo* data produce more consistent results. By comparison, *in vitro* values were on average thirty-eight times higher than *in vivo* ones; when looking specifically at different TiO_2 types and different cell lines, we observed comparable BMDs from rodents and murine cell experiments, while human macrophages and co-cultures showed a lower susceptibility to inflammatory effects. However, the scarcity of comparable data across species, cell lines, and particle types hinders the evaluation of the effects of these factors on the BMD range.

The SVM classification model built on the *in vitro* data set for viability endpoints was able to predict with good accuracy the range of the BMD-derived human exposure level, based on a limited number of particle properties and experimental parameters.

Combined dosimetry demonstrated to be a successful strategy for IVIVE, and CoDo a useful tool when working with big data sets, allowing a meta analysis of titanium dioxide toxicity data.

Model and data availability

CoDo model, its supporting files, and the User Guide are openly available in Zenodo at https://doi.org/10.5281/zenodo.4889169. The data supporting the findings of this study are available within the article's supplementary materials.

Funding

This research is part of the project NANORIGO, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 814530. This publication reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. Owen Price for his support in the use of the MPPD model, Dr. Cordula Hirsch for her insights about *in vitro* dosimetry, Dr. Roland Hischier for the fruitful discussions, and Daniele Lain for his help with coding.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.impact.2021.100376.

References

Anaraki, N.I., Sadeghpour, A., Iranshahi, K., Toncelli, C., Cendrowska, U., Stellacci, F., Dommann, A., Wick, P., Neels, A., 2020. New approach for time-resolved and dynamic investigations on nanoparticles agglomeration. Nano Res. 13, 2847–2856.
Annangi, B., Rubio, L., Alaraby, M., Bach, J., Marcos, R., Hernández, A., 2016. Acute and long-term in vitro effects of zinc oxide nanoparticles. Arch. Toxicol. 90, 2201–2213.
Asgharian, B., Price, O., McClellan, G., Corley, R., Einstein, D.R., Jacob, R.E., Harkema, J., Carey, S.A., Schelegle, E., Hyde, D., Kimbell, J.S., Miller, F.J., 2012. Development of a rhesus monkey lung geometry model and application to particle deposition in comparison to humans. Inhal. Toxicol. 24, 869–899.
Asgharian, B., Miller, F.J., Price, O., Schroeter, J.D., Einstein, D.R., Corley, R.A., Bentley, T., 2016. Modeling particle deposition in the pig respiratory tract. J. Aerosol Sci. 99, 107–124.
Basketter, D.A., York, M., McFadden, J.P., Robinson, M.K., 2004. Determination of skin irritation potential in the human 4-h patch test. Contact Dermatitis 51, 1–4.
Bermudez, E., Mangum, J.B., Wong, B.A., Asgharian, B., Hext, P.M., Warheit, D.B., Everitt, J.I., 2004. Pulmonary Responses of Mice, Rats, and Hamsters to subchronic inhalation of ultrafine titanium dioxide particles. Toxicol. Sci. 77, 347–357.
Brown, J.S., Gordon, T., Price, O., Asgharian, B., 2013. Thoracic and respirable particle definitions for human health risk assessment. Partic. Fib. Toxicol. 10, 12.
Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. Comput. Elect. Eng. 40, 16–28.
Cohen, J.M., DeLoid, G.M., Demokritou, P., 2015. A critical review of in vitro dosimetry for engineered nanomaterials. Nanomedicine 10, 3015–3032.
Committee, E.S., Hardy, A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, K.H., More, S., Mortensen, A., Naegeli, H., Noteborn, H., et al., 2017. Update: use of the benchmark dose approach in risk assessment. EFSA J. 15, e04658.
Cosnier, F., Seidel, C., Valentino, S., Schmid, O., Bau, S., Vogel, U., Devoy, J., Gaté, L., 2021. Retained particle surface area dose drives inflammation in rat lungs following acute, subacute, and subchronic inhalation of nanomaterials. Partic. Fib. Toxicol. 18, 1–21.
Davis, J.A., Gift, J.S., Zhao, Q.J., 2011. Introduction to benchmark dose methods and US EPA's benchmark dose software (BMDs) version 2.1. 1. Toxicol. Appl. Pharmacol. 254, 181–191.
DeLoid, G., Cohen, J.M., Darrah, T., Derk, R., Rojanasakul, L., Pyrgiotakis, G., Wohlleben, W., Demokritou, P., 2014. Estimating the effective density of engineered nanomaterials for in vitro dosimetry. Nat. Commun. 5, 3514.
DeLoid, G.M., Cohen, J.M., Pyrgiotakis, G., Pirela, S.V., Pal, A., Liu, J., Srebric, J., Demokritou, P., 2015. Advanced computational modeling for in vitro nanomaterial dosimetry. Partic. Fib. Toxicol. 12, 32.
Deloid, G.M., Cohen, J.M., Pyrgiotakis, G., Demokritou, P., 2017. Preparation, characterization, and in vitro dosimetry of dispersed, engineered nanomaterials. Nat. Protoc. 12, 355–371.
Demokritou, P., Gass, S., Pyrgiotakis, G., Cohen, J.M., Goldsmith, W., McKinney, W., Frazer, D., Ma, J., Schwegler-Berry, D., Brain, J., Castranova, V., 2013. An in vivo and in vitro toxicological characterisation of realistic nanoscale CeO_2 inhalation exposures. Nanotoxicology 7, 1338–1350.
Donaldson, K., Borm, P., Oberdorster, G., Pinkerton, K.E., Stone, V., Tran, C., 2008. Concordance between in vitro and in vivo dosimetry in the proinflammatory effects of low-toxicity, low-solubility particles: the key role of the proximal alveolar region. Inhal. Toxicol. 20, 53–62.
Eurostat, 2021. Duration of working life - statistics - Statistics Explained. Eurostat, the statistical office of the European Union, 2920 Luxembourg, Luxembourg.
Filipe, V., Hawe, A., Jiskoot, W., 2010. Critical evaluation of nanoparticle tracking analysis (nta) by nanosight for the measurement of nanoparticles and protein aggregates. Pharm. Res. 27, 796–810.
Gangwal, S., Brown, J.S., Wang, A., Houck, K.A., Dix, D.J., Kavlock, R.J., Hubal, E.A.C., Gangwa, S., Brown, J.S., Wang, A., Houck, K.A., Dix, D.J., Kavlock, R.J., Cohen Hubal, E.A., 2011. Informing selection of nanomaterial concentrations for toxcast in vitro testing based on occupational exposure potential. Environ. Health Perspect. 119, 1539–1546.
https://www.dguv.de/ifa/gestis/gestis-internationale-grenzwerte-fuer-chemische-substanzen-limit-values-for-chemical-agents/index-2.jsp, 2021–. (Accessed 20 January 2021).
Gottmann, E., Kramer, S., Pfahringer, B., Helma, C., 2001. Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. Environ. Health Perspect. 109, 509–514.

Haber, L.T., Dourson, M.L., Allen, B.C., Hertzberg, R.C., Parker, A., Vincent, M.J., Maier, A., Boobis, A.R., 2018. Benchmark dose (bmd) modeling: current practice, issues, and challenges. Crit. Rev. Toxicol. 48, 387–415.
Hartung, T., 2009. A Toxicology for the 21st century-mapping the road ahead. Toxicol. Sci. 109, 18–23.
Hsu, C.-W., Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. IEEE Transac. Neural Networks 13, 415–425.
ISO 10993-5:2009, 2009. Biological Evaluation of Medical Devices-Part 5: Tests For in Vitro Cytotoxicity, Standard. International Organization for Standardization, Geneva, CH.
Jung, F., Nothnagel, L., Gao, F., Thurn, M., Vogel, V., Wacker, M.G., 2018. A comparison of two biorelevant in vitro drug release methods for nanotherapeutics based on advanced physiologically-based pharmacokinetic modelling. Eur. J. Pharm. Biopharm. 127, 462–470.
Khatri, M., Bello, D., Pal, A.K., Cohen, J.M., Woskie, S., Gassert, T., Lan, J., Gu, A.Z., Demokritou, P., Gaines, P., 2013. Evaluation of cytotoxic, genotoxic and inflammatory responses of nanoparticles from photocopiers in three human cell lines. Partic. Fib. Toxicol. 10, 1–22.
Monteiller, C., Tran, L., MacNee, W., Faux, S., Jones, A., Miller, B., Donaldson, K., 2007. The pro-inflammatory effects of low-toxicity low-solubility particles, nanoparticles and fine particles, on epithelial cells in vitro: the role of surface area. Occup. Environ. Med. 64, 609–615.
Mukherjee, S.P., Gupta, G., Klöditz, K., Wang, J., Rodrigues, A.F., Kostarelos, K., Fadeel, B., 2020. Next-generation sequencing reveals differential responses to acute versus long-term exposures to graphene oxide in human lung cells. Small 16, 1907686.
Noël, A., Charbonneau, M., Cloutier, Y., Tardif, R., Truchon, G., 2013. Rat pulmonary responses to inhaled nano-TiO2: effect of primary particle size and agglomeration state. Partic. Fib. Toxicol. 10, 1–18.
Oberdörster, G., 2012. Nanotoxicology: in vitro-in vivo dosimetry. Environ. Health Perspect. 120, A13 author reply A13.
Ong, K.J., MacCormack, T.J., Clark, R.J., Ede, J.D., Ortega, V.A., Felix, L.C., Dang, M.K., Ma, G., Fenniri, H., Veinot, J.G., et al., 2014. Widespread nanoparticle-assay interference: implications for nanotoxicity testing. PLoS One 9, e90650.
Oyabu, T., Morimoto, Y., Hirohashi, M., Horie, M., Kambara, T., Lee, B.W., Hashiba, M., Mizuguchi, Y., Myojo, T., Kuroda, E., 2013. Dose-dependent pulmonary response of well-dispersed titanium dioxide nanoparticles following intratracheal instillation. J. Nanopart. Res. 15, 1–11.
Pal, A.K., Bello, D., Cohen, J., Demokritou, P., 2015a. Implications of in vitro dosimetry on toxicological ranking of low aspect ratio engineered nanomaterials. Nanotoxicology 9, 871–885.
Pal, A.K., Watson, C.Y., Pirela, S.V., Singh, D., Chalbot, M.-C.G., Kavouras, I., Demokritou, P., 2015b. Linking exposures of particles released from nano-enabled products to toxicology: an integrated methodology for particle sampling, extraction, dispersion, and dosing. Toxicol. Sci. 146, 321–333.
Praphawatvet, T., Peters, J.I., Williams III, R.O., 2020. Inhaled nanoparticles-an updated review. Int. J. Pharm. 587, 119671.
Romeo, D., Salieri, B., Hischier, R., Nowack, B., Wick, P., 2020. An integrated pathway based on in vitro data for the human hazard assessment of nanomaterials. Environ. Int. 137, 105505.
Rushton, E.K., Jiang, J., Leonard, S.S., Eberly, S., Castranova, V., Biswas, P., Elder, A., Han, X., Gelein, R., Finkelstein, J., Oberdörster, G., 2010. Concept of assessing nanoparticle hazards considering nanoparticle dosemetric and chemical/biological response metrics. J. Toxicol. Environ. Health Part A 73, 445–461.
Sand, S., von Rosen, D., Victorin, K., Falk Filipsson, A., 2006. Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. Toxicol. Sci. 90, 241–251.
Schmid, O., Stoeger, T., 2016. Surface area is the biologically most effective dose metric for acute nanoparticle toxicity in the lung. J. Aerosol Sci. 99, 133–143.
Schneider, T., Jensen, K.A., 2009. Relevance of aerosol dynamics and dustiness for personal exposure to manufactured nanoparticles. J. Nanopart. Res. 11, 1637–1650.
Schulte, P., Murashov, V., Zumwalde, R., Kuempel, E., Geraci, C., 2010. Occupational exposure limits for nanomaterials: state of the art. J. Nanopart. Res. 12, 1971–1987.
Slob, W., 2018. Joint project on benchmark dose modelling with rivm. EFSA Support. Publ. 15, 1497E.
Smith, J.N., Skinner, A.W., 2021. Translating nanoparticle dosimetry from conventional in vitro systems to occupational inhalation exposures. J. Aerosol Sci. 155, 105771.
Teeguarden, J.G., Mikheev, V.B., Minard, K.R., Forsythe, W.C., Wang, W., Sharma, G., Karin, N., Tilton, S.C., Waters, K.M., Asgharian, B., Price, O.R., Pounds, J.G., Thrall, B.D., 2014. Comparative iron oxide nanoparticle cellular dosimetry and response in mice by the inhalation and liquid cell culture exposure routes. Partic. Fib. Toxicol. 11, 46.
Thomas, D.G., Smith, J.N., Thrall, B.D., Baer, D.R., Jolley, H., Munusamy, P., Kodali, V., Demokritou, P., Cohen, J., Teeguarden, J.G., 2018. Isd3: a particokinetic model for predicting the combined effects of particle sedimentation, diffusion and dissolution on cellular dosimetry for in vitro systems. Partic. Fib. Toxicol. 15, 1–22.
Thrall, B.D., Kodali, V., Skerrett, S., Thomas, D.G., Frevert, C.W., Pounds, J.G., Teeguarden, J.G., 2019. Modulation of susceptibility to lung bacterial infection by engineered nanomaterials: in vitro and in vivo correspondence based on macrophage phagocytic function. NanoImpact 14, 100155.
Thurnherr, T., Brandenberger, C., Fischer, K., Diener, L., Manser, P., Maeder-Althaus, X., Kaiser, J.-P., Krug, H.F., Rothen-Rutishauser, B., Wick, P., 2011. A comparison of acute and long-term effects of industrial multiwalled carbon nanotubes on human lung and immune cells in vitro. Toxicol. Lett. 200, 176–186.
Torres, A., Dalzon, B., Collin-Faure, V., Rabilloud, T., 2020. Repeated vs. acute exposure of raw264. 7 mouse macrophages to silica nanoparticles: a bioaccumulation and functional change study. Nanomaterials 10, 215.
Van Rijt, S.H., Bölükbas, D.A., Argyo, C., Wipplinger, K., Naureen, M., Datz, S., Eickelberg, O., Meiners, S., Bein, T., Schmid, O., et al., 2016. Applicability of avidin protein coated mesoporous silica nanoparticles as drug carriers in the lung. Nanoscale 8, 8058–8069.
Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA.
van Vliet, E., 2011. Current standing and future prospects for the technologies proposed to transform toxicity testing in the 21st century. ALTEX-Alter. Animal Exp. 28, 17–44.
Varewyck, M., Verbeke, T., 2017. Software for benchmark dose modelling. EFSA Support. Publ. 14, 1170E.
Virtanen, P., Gommers, R., Oliphant, T.E., et al., 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
Vranic, S., Gosens, I., Jacobsen, N.R., Jensen, K.A., Bokkers, B., Kermanizadeh, A., Stone, V., Baeza-Squiban, A., Cassee, F.R., Tran, L., Boland, S., 2017. Impact of serum as a dispersion agent for in vitro and in vivo toxicological assessments of TiO2 nanoparticles. Arch. Toxicol. 91, 353–363.
Warheit, D.B., 2011. Pulmonary bioassay methods for evaluating hazards following exposures to nanoscale or fine particulate materials. Assess. Nanopartic. Risks Human Health 99–108.
Wong, T.-T., 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit. 48, 2839–2846.
Yue, S., Li, P., Hao, P., 2003. Svm classification: its contents and challenges. Appl. Math. - J. Chinese Univ. 18, 332–342.