Predicting nanomaterials pulmonary toxicity in animals by cell culture models: Achievements and perspectives

Emilio Di Ianni¹ | Nicklas Raun Jacobsen¹ | Ulla Vogel¹,² | Peter Møller³

¹National Research Centre for the Working Environment, Copenhagen, Denmark
²National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark
³Department of Public Health, Section of Environmental Health, University of Copenhagen, Copenhagen, Denmark

Correspondence
Peter Møller, Department of Public Health, Section of Environmental Health, University of Copenhagen, Copenhagen, Denmark.
Email: pemo@sund.ku.dk

Funding information
Danish Government FFIKA, Focused Research Effort on Chemicals in the Working Environment; European Union’s Horizon 2020 Research and Innovation Program (GRACIOUS), Grant/Award Number: 760840

Edited by: Nancy Monteiro-Riviere, Associate Editor and Gregory M. Lanza, Co-Editor-in-Chief

Abstract
Animal experiments are highly relevant models for the assessment of toxicological effects of engineered nanomaterials (ENMs), due to lack of biomonitoring and epidemiological studies. However, the expanding number of ENMs with different physico-chemical properties strains this approach, as there are ethical concerns and economical challenges with the use of animals in toxicology. There is an urgent need for cell culture models that predict the level of toxicological responses in vivo, consequently reducing or replacing the use of animals in nanotoxicology. However, there is still a limited number of studies on in vitro—in vivo correlation of toxicological responses following ENMs exposure. In this review, we collected studies that have compared in vitro and in vivo toxic effects caused by ENMs. We discuss the influence of cell culture models and exposure systems on the predictability of in vitro models to equivalent toxic effects in animal lungs after pulmonary exposure to ENMs. In addition, we discuss approaches to qualitatively or quantitatively compare the effects in vitro and in vivo. The magnitude of toxicological responses in cells that are exposed in submerged condition is not systematically different from the response in cells exposed in air—liquid interface systems, and there appears to be similar ENMs hazard ranking between the two exposure systems. Overall, we show that simple in vitro models with cells exposed to ENMs in submerged condition can be used to predict toxic effects in vivo, and identify future strategies to improve the associations between in vitro and in vivo ENMs-induced pulmonary toxicity.

This article is categorized under:
Toxicology and Regulatory Issues in Nanomedicine > Toxicology of Nanomaterials

KEYWORDS
advanced cell culture models, DNA damage, inflammation, nanomaterials, predictivity

Received: 16 November 2021 Revised: 10 February 2022 Accepted: 10 March 2022
DOI: 10.1002/wnan.1794

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. WIREs Nanomedicine and Nanobiotechnology published by Wiley Periodicals LLC.

WIREs Nanomed Nanobiotechnol. 2022;14:e1794.
https://doi.org/10.1002/wnan.1794
1 | INTRODUCTION

Engineered nanomaterials (ENMs) are increasingly used for industrial, consumer, and medical applications, resulting in increased human exposure. Greatest exposure occurs in the working environment, with inhalation being the main route of exposure. Because of their small size, inhaled ENMs deposit in the distal part of the respiratory tree regions of the lung and generate inflammation, oxidative stress, and damage to biomolecules such as lipids and DNA (Møller et al., 2014). In turn, these effects activate a cascade of events leading to diseases such as fibrosis, lung cancer, and atherosclerosis (Hadrup et al., 2020; Halappanavar et al., 2020; Stone et al., 2017). Diseases such as cancer and atherosclerosis may develop after long-term exposure by either repeated exposure to materials with “high” clearance rate (e.g., carbon black) or even single exposure to materials with “low” clearance rate (e.g., carbon nanotubes). The dose determines the debut time and severity of the disease. In general, in vitro models cannot yet recreate the formation of a tumor or atheroma because these clinical findings develop through multiple steps. In addition, there is as yet no single intermediate step that has proven to be necessary or a sufficient causal factor for ENM-induced diseases. In vitro models are based on the current knowledge on the mechanism of action for ENM-induced health effects.

Safety assessment of ENMs is an essential element in safe-by-design approaches, which introduces toxicological testing in early steps of research and development to prevent adverse health effects in exposed workers. Animal experiments is the traditional way of studying toxicological effects of ENMs since evidence from epidemiological studies is not (yet) available for these new materials. Furthermore, available epidemiological studies of occupational exposure to high volume production nanomaterials such as carbon black and titanium dioxide suffer from limitations such as lack of information on particle size and thus, the magnitude of nanoparticle exposure (Saber et al., 2019). A typical experiment performed according to Organization for Economic Co-operation and Development (OECD) guidelines for testing of chemicals entails multiple doses and time-points, each group having up to 10 animals, resulting ultimately in large numbers of animals used to assess a substance (e.g., 120 animals per ENM). Study designs can be developed to reduce the number of animals, but it still sums up to a rather large number of animals, thus colliding with the principles of Replacement, Reduction, and Refinement (3Rs) of animal testing, which have been developed to perform more humane animal research. Alternatives to animal testing is an ongoing theme in nanotoxicology. There are huge benefits of developing robust in vitro models for screening toxicological effects of ENMs in cell cultures before embarking on more advanced and expensive animal models. These benefits can be summarized as follows:

1. Animal testing may only be needed if ENMs induce toxic effects in cell cultures (i.e., screening of hazardous ENMs).
2. Cell culture models can be used to estimate the required group size in animal studies.
3. Cell culture studies can be used to select the dose range in animal studies.
4. For non-toxic ENMs in cell cultures, it may not be necessary to perform a full dose–response relationship (i.e., a no observed adverse effect level can be established using information about realistic human exposures and fewer doses and time-points in an animal study).

Despite the large number of studies assessing the toxicity of ENMs in cell cultures, there are few studies that have investigated whether equivalent responses are obtained in cell cultures and animals after exposure to the same ENMs. Regulatory authorities still rely on in vivo studies of ENM toxicity to establish exposure levels with tolerable risk of health effects for workers. There is a need for reliable in vitro testing systems that predict in vivo responses following ENM exposure.

Studies on the relationship between toxicological effects in cell cultures and animals may follow different paths, depending on the assessed research question. The simplest question is whether the responses in vitro and in vivo are the same (i.e., concordance based on binary categorization of responses). However, this question is typically complicated by the fact that it is the dose of ENMs that mainly determines the level of effect. Thus, very high doses may produce a statistically significant effect, but may have little biological relevance. It introduces a somewhat arbitrary factor of selecting an upper boundary of doses, which may or may not be regarded as realistic doses. A more elegant approach is to compare the dose–response relationship in cell culture and animal studies, although this entails some assumptions regarding the shape of the dose–response curves. This analysis uses either categorical or continuous data from cell culture models to predict the response in animals and ultimately in humans. Another important issue is the type of cell culture model. The traditional in vitro experiment entails a monoculture of immortalized or primary cells. Cell lines are robust, easy to culture, and are relatively cheap. Thus, they are broadly used, which increases possibilities of interstudy comparison of toxicological effects of ENMs. However, immortalized cells generally have aberrant karyotype and
increased genomic instability as compared with the tissue cells from which they originated. On the contrary, early passages of primary cells have the same genotypic and phenotypic characteristics as tissue cells and better represent the healthy microenvironment. Yet primary cells are typically costly to produce, have low proliferation rate, less robust, and usually more laborious to work with.

Advanced cell culture models are explored in particle toxicology as more physiologically relevant exposure systems as compared with monocultures of cells exposed to ENMs in submerged condition. These include co-cultures (i.e., more than one cell type), 3D cultures (systems where cells are allowed to form a three-dimensional structure rather than a monolayer) and special exposure systems such as air–liquid interface (ALI). The latter includes particle exposure in vertical (i.e., sedimentation of particles from a cloud) and horizontal direction (i.e., sedimentation of particles from radial flow of air over the cellular layer). There are two commercially developed ALI systems—called INVITROCELL and CULTEX—as well as a number of non-commercial exposure systems (Upadhyay & Palmberg, 2018). Further description of difference between cell culture models has been elaborated previously (Evans et al., 2017). It is still unclear whether these advanced cell culture models are superior to the traditional monoculture systems. In this regard, cost–benefit considerations should be kept in mind as simple monoculture systems are relatively cheap and the added value of more complex systems needs to be carefully assessed. Figure 1 shows the differences between traditional monocultures, co-cultures, 3D cultures and ALI. The increasing complexity of the cell culture systems are categorized into four groups corresponding to the physiological relevance relative to the exposure condition in human airway epithelial cells.

The scope of this review is to discuss advances in the field of in vitro nanotoxicology, with emphasis on exposure systems and cell models to mimic the lung cell architecture, yet some aspects concerning the in vitro–in vivo extrapolation are applicable to other physiological compartments such as gastrointestinal tract, brain, and skin. There is not a sufficient number of studies on toxicity caused by ENMs in humans, therefore hazard identification relies currently on in vitro and in vivo studies. The review encompasses three sections that address key questions in research on in vitro–in vivo comparisons of health hazard identification by exposure to ENMs:

1. Can cell culture models predict pulmonary toxicity in animals?
2. Does the magnitude of response in cell cultures correspond to the magnitude of pulmonary toxicity in animals?
3. Are advanced cell culture models better predictors of pulmonary toxicity in animals as compared with monocultures?

2 | METHODS

We have searched the PubMed database for studies on ENMs that have compared in vitro with in vivo endpoints upon exposure to ENMs. Inclusion criteria of studies have been: (1) ENMs have been characterized with regard to particle size, surface area, and so on; 2) at least three ENMs have been assessed in cells and animals; (3) comparison of the in vitro–in vivo relationship was a primary objective of the study. Eighteen studies have been included in the review (listed
chronologically in Table 1). The comparisons are based on metal oxides and SiO₂ (13 studies), and carbonaceous nanomaterials including elemental carbon, carbon black, multiwalled carbon nanotubes, and graphene oxide (9 studies). The endpoints encompass biomarkers of cytotoxicity, oxidative stress, inflammation, and DNA damage. In the review, we distinguish between “endpoint” (i.e., specific biomarker) and “response” (i.e., inflammation). The same response can be assessed using different endpoints in cells and animals.

The review does not systematically address factors for differences in toxicological effects of different ENMs. Rather, we assume the ENMs in each study have been selected to introduce a gradient in hazardous properties for the purpose of comparing toxicological responses in cells and animals. Thus, one study may have selected material A and B, and another study material C and D; the ENMs cannot be compared, but the difference between them can. In this respect, it does not affect the analysis that material A–D have different mediators of toxicity such as solubility, particle surface area, reactive oxygen species (ROS) production potential, and frustrated phagocytosis. In fact, it is far more concerning if researchers have selected ENMs that are too similar, because it might make it difficult to obtain a wide difference in toxicological responses in both cell cultures and animals. The studies described here used different cell types, mostly epithelial cells and macrophages, including primary cells and, in the majority of the studies, cancer cell lines. Of these studies, only three compared responses in the ALI system and lung tissue of animals. We have segregated the studies according to the level of complexity of the in vitro cell models (Di Ianni, Erdem, et al., 2021; Loret et al., 2018). Monoculture of cell lines in submerged conditions is regarded as a “simple” test system (lowest level). Test systems using co-cultures, primary cells, or ALI are more complex than the simple monoculture system. The most complex test system encompass co-cultures of primary cells in ALI. Table 2 outlines the different classes of complexity and the belonging of the studies in each of these categories.

The in vivo data were generated from exposure of mice or rats via inhalation, oropharyngeal aspiration, and intratracheal instillation. The endpoints correlated were related to cytotoxicity, inflammation, and oxidative stress, and only two studies included DNA damage. The original papers and the analyses in the review use either correlations or concordance as tools to compare the relationship between responses in cell cultures and animals. The meaning and use of these and other terms are described below.

2.1 Correlation

Test using continuous data on the predictor (endpoint in vitro) and outcome scale (endpoint in vivo). The data may come from effects at fixed doses in cells or animals, or from the slopes of individual dose–response curves in cell cultures and animal models. The results are usually assessed by correlation analysis using assumption of a qualitative (e.g., Spearman correlation test) or quantitative (e.g., linear regression or Pearson correlation test) relationship between responses in cell cultures and animals.

2.2 Concordance

Test that determines the proportion of correct outcomes of a test method (i.e., correctly identified positives and negatives). The analysis also generates information about positive predictivity, negative predictivity, sensitivity, and specificity. This is based on binary segregation of results into level of responses according to predefined criteria such as a certain fold-increase or statistical significance. Concordance is often used interchangeably with “accuracy,” but the latter term is also used to describe the closeness of agreement between test results and accepted reference values. There is not a threshold of acceptable concordance of a test in genetic toxicology. However, the Ames test—the most used regulatory test for genotoxicity—has 70% concordance (Zeiger, 2019). The concordance of carcinogenicity in mice and rats has been shown to be 75% (Gold et al., 1991).

2.3 Dose

In traditional toxicology dose refers to the mass of a poisonous agent relative to the mass of the biological system (e.g., mg/kg bodyweight), whereas concentration refers to the mass relative to the volume. Most studies on ENMs are designed to expose cells or animals to fixed masses of particles (i.e., the nominator is a mass unit). The mass is often
| Authors            | Nanoparticles                                                                 | Cell models and exposure system endpoints | Animal models and exposure endpoints | Outcome                                                                 |
|--------------------|-------------------------------------------------------------------------------|-------------------------------------------|-------------------------------------|-------------------------------------------------------------------------|
| Sayes et al. (2007)| Carbonyl iron (CI, 800–3000 nm), crystalline silica (CS, 1600 nm), precipitated amorphous silica (AS, 1000–3000 nm), nano-ZnO (50–70 nm), fine-ZnO (<1000 nm) | Rat lung epithelial cells (L2) primary alveolar macrophages (AMs) L2-AMs co-culture Submerged (4, 24, and 48 h) Inflammation (MIP-2, TNF-α, and IL-6), cytotoxicity (LDH, MTT) | Fisher 344 rats Intratracheal Instillation (24 h, 1 week, and 3 months) Neutrophil influx and LDH in BALF lung histology | No correlation between in vitro and in vivo endpoints across the time points The method developed later by Rushton et al. (2010) was applied to data generated in this study. This showed strong correlations between LDH or MIP-2 in AMs or L2-AMs co-culture, respectively, and neutrophil influx after doses were normalized by surface area |
| Warheit et al. (2009)| ZnO (nanosize; 50–70 nm; fine size: <1000 nm)                                | Lung epithelial cell line (L2), primary alveolar macrophages (AMs), or L2-AMs co-culture Submerged (4 and 24 h) LDH, MIP-2, and TNF-α | Male Crl:CD(SD)IGS BR rats Inhalation and intratracheal instillation (24 h, 3 days, and 8 days) Neutrophil influx and LDH in BALF | No correlation between in vitro and in vivo proinflammatory responses |
| Rushton et al. (2010)| Elemental carbon (41 nm); TiO₂ (20-250 nm); Cu (40 nm); Ag (35 nm); Au (50 nm); polystyrene-NH₃ (60 nm) | A549 Luc1 cells (stably transfected with an IL8-luciferate reporter construct) alveolar macrophages Submerged Electron spin resonance (ESR), IL8 | Male Fisher 344 rats Intratracheal instillation (24 h) Neutrophil influx in BALF | In vivo PMN response was strongly correlated with AM ESR (DMPO spin trapping) activity ($r_s = 0.97, p < 0.001$) and A549 Luc-1 cell luciferase response ($r_s = 0.94, p < 0.01$) |
| Wang et al. (2011) | Pristine and functionalized (COOH) MWCNTs (length: 10–30 μm; diameter: 20–30 μm) | BEAS-2B and THP-1a Submerged (24 h) TGF-β1, PDGF-AA, and IL-1β | C57BL/6 mice Oropharyngeal aspiration (24 h and 21 days) TGF-β1, PDGF-AA, and IL-1β in BALF and histology | Better dispersed MWCNT induced more TGF-β1, PDGF-AA, and IL-1β in cells and animals in BALF, which was confirmed by the fibrogenic signs in lung tissue. Results on profibrogenic responses in vitro and in vivo effects showed good agreement |
| Han et al. (2012)  | TiO₂ (8 different samples with different size 3–104 nm)                        | Rat lung epithelial cells line (R3/1) Submerged (24 h) Protein oxidation (carbonyls) and cytotoxicity (LDH) | Male Fisher 344 Intratracheal instillation (24 h) Neutrophil influx in BALF | In vitro-in vivo correlation with steepest slope analysis method, only after doses were normalized by surface area; $r_s = 0.93, p < 0.01$ (protein oxidation versus PMN) and $r_s = 0.88, p < 0.01$ (LDH versus PMN) |
| Zhang et al. (2012)| CuO, Co₃O₄, Co₉O₄, CoO, Cr₂O₃, Fe₂O₃, Fe₃O₄, HfO₂, In₂O₃, Mn₃O₄, NiO, Ni₂O₃, Y₂O₃, ZrO₂ | BEAS-2B and RAW264.7 Submerged (24 h) MTS, ATP, LDH, high-throughput screening for oxidative stress biomarkers | C57BL/6 mice Oropharyngeal aspiration (40 h) | Correlation between cytotoxicity (combined MTS, ATP, LDH) and neutrophil influx ($r_s = 0.76, p < 0.01$) (calculated from mean values reported) |
| Authors          | Nanoparticles                                                                 | Cell models and exposure system endpoints                                                                 | Animal models and exposure endpoints                                                                 | Outcome                                                                                     |
|------------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| Cho et al. (2013) | TiO₂, CeO₂, SiO₂, NiO, Co₃O₄, Cr₂O₃, CuO, ZnO and carbon black                 | • 16-HBE, A549, THP-1 monocytes, THP-1a, primary PBMC, AS49 + THP-1a                                                   | • Neutrophil influx in BALF, MCP-1, IL-6 in BALF                                                      | in article) and oxidative stress biomarker (heme oxygenase 1) in lung tissue (Co₃O₄, Cr₂O₃, Ni₂O₃, Mn₂O₃, and CoO) |
| Li et al. (2013)  | Pristine and f-MWCNT; Carboxylate (COOH), polyethylene glycol (PEG), amine (NH₂), sidewall amine (sw-NH₂), and polyetherimide (PEI)-modified MWCNTs | • BEAS-2B and THP-1a cells • Submerged (24 h) • IL-1β, TGF-β1, and PDGF-AA                                   | • C57Bl/6 male mice • Oropharyngeal aspiration (40 h) • Collagen deposition in lung tissue (Sircol assay, Trichrome staining), TGF-b, IL-1b, TGF-β1, and PDGF-AA in BALF | PEI-MWCNTs induced the highest IL-1β, TGFβ1, and PDGF-AA production in vitro and in vivo, and more pronounced collagen deposition than other MWCNT. MWCNTs with the lowest hazard ranking in vitro (i.e., AF and COOH) did not increase the presence of collagen in lung tissue |
| Pal et al. (2015) | TiO₂ P25, CeO₂, Ni, Ag, CB Printex 90                                         | • THP-1a • Submerged (24 h) • Inflammation (IL-8), cytotoxicity (LDH)                                             | • Male Fischer rats (F344/DuCrI) • Intratracheal instillation (24 h) • PMN cells in the BALF         | Correlation between cell death in THP-1a and PMN in rats (deposited doses: r = 0.98; administered doses: r = 0.80, linear regression). Correlations were not reported for IL-8. However, based on results in the publication, correlation and linear regression coefficient is 0.10, which is not statistically significant |
| Ma et al. (2015)  | Graphene oxides (GO) with differential sizes (small: 50–350 nm; large: 750–1300 nm) | • J774.1A, THP-1a • Submerged (24 h) • IL-6, TNF-α, ROS production (DCFH-DA)                                     | • BALB/c mice • Oropharyngeal aspiration (72 h) • PMN cells in BALF                                  | Large GO had higher proinflammatory response in vitro (TNF-α and IL-6 production in J774.1A cells, TNF-α expression in THP-1a cells) and in vivo (total cells in BALF). Similar ROS production by small and large GOs in J774.1A cells |
| Wiemann et al. (2016) | CeO₂, Fe₂O₃, TiO₂, ZrO₂ (pristine and surface modified), ZnO NMs, amorphous SiO₂ (pristine and surface modified), | • Rat NR8383 alveolar macrophage • Submerged (16 h) • LDH, GLU, TNF-α, and H₂O₂ release                          | • Wistar rat • Inhalation (6 h exposure/day, 5 consecutive days)                                     | The in vitro NR8383 alveolar macrophage assay allows |

Concordance between pulmonary inflammation in lung in vivo and cytotoxicity in PBMC was 89% (11% false negative), followed by primary cultured alveolar macrophages (78% concordance, 11% false negative, and 11% false positive)
| Authors                        | Nanoparticles                                                                 | Cell models and exposure system endpoints                                                                 | Animal models and exposure endpoints                                                                 | Outcome                                                                                                                                                                                                                           |
|-------------------------------|-------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Wiemann et al. (2018)         | SiO₂ (amorphous, 15 or 55 nm)                                                 | Rat NR8383 alveolar macrophage                                                                             | Wistar rats                                                                                             | Higher response in terms of LDH release with 15 nm particles compared with 55 nm particles in both cells and animals (the authors tested also TNF-α, β-glucuronidase, H₂O₂ release)                                          |
| Lorent et al. (2018)          | TiO₂ (NM-100, NM-101, and NM-105)                                            | A549 and THP-1a in co-culture                                                                             | Wistar rats                                                                                             | Same ranking based on cytotoxicity and 20% increase in proinflammatory response in cell cultures and animals, using surface area as dose metric                                            |
| Cui et al. (2019)             | ZnO, SiO₂, TiO₂, and CeO₂                                                     | BEAS-2B                                                                                                     | C57BL/6 mice                                                                                            | Similar response on levels of metabolites after exposure to ENMs in vitro and in vivo (the study also assessed proinflammatory responses in vitro [e.g., IL6 and IL8] and in vivo [e.g., neutrophils in BALF]. The responses do not appear to be congruent, which was not described in the article) |
| Di Ianni, Erdem, et al. (2021)| MWCNTs (length: 443–573 nm; diameter: 11–74 nm)                              | A549, THP-1a, A549:THP-1a:WI-38                                                                            | C57BL/6 mice                                                                                            | Correlation between IL8 secretion in A549 cells (r = 0.92 and 0.95 for A549 (6 and 24 h, respectively) versus BAL cells, and r = 0.96 for THP-1a (both 6 and 24 h) versus BAL cells). Concordant DNA damage between THP-1 cells and BAL cells (83%) |
| Di Ianni, Möller, et al. (2021)| Pristine and surface modified nanoclays; graphene oxide and reduced-graphene oxide | A549, THP-1a                                                                                                | C57BL/6 mice                                                                                            | Correlation between IL8 secretion in A549 cells versus neutrophils; r = 0.96 (p < 0.01) and r = 0.97 (p < 0.01) at 6 and 24 h, respectively; THP-1a cells versus neutrophils: r = 0.93 (p = 0.02) and r = 0.92 (p = 0.02) at 6 and 24 h, respectively |
| Authors          | Nanoparticles                                           | Cell models and exposure system endpoints                                                                 | Animal models and exposure endpoints                                                                 | Outcome                                                                 |
|------------------|---------------------------------------------------------|------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| Karkossa et al.  | SiO$_2$-15 Amino, SiO$_2$-15 unmod, SiO$_2$-7, SiO$_2$-40, TiO$_2$-NM105 | • RLE-6TN (rat alveolar epithelial cells)                                                                 | • Wistar rat                                                                                            | Concordant DNA damage between THP-1a cells and BAL cells (60%) and 3 days (80%) postexposure |
|                  |                                                         | • NR8383 (rat alveolar macrophages)                                                                       | • Short-term inhalation (6 h/day over 5 days, euthanized right after 5th day or after 21 days)          | different changes of markers in all the models; congruent classification of “active” and “passive” materials in vitro and in vivo, based on significantly changed signatures in exposed groups compared with controls |
|                  |                                                         | • Submerged conditions (24 h)                                                                             | • Intratracheal instillation (3 and 6 days)                                                            |                                                                                        |
|                  |                                                         | • Metabolomic and proteomic changes in oxidative stress                                                   | • Metabolomic and proteomic changes in oxidative stress                                                   |                                                                                        |
| Fujita et al.    | Three carbon fibers with different diameters and length: SCF1 diameter: 259 nm length: 12 μm SCF2 diameter: 248 nm length: 7 μm SCF3 diameter: 183 nm length: 14 μm | • NR8383 rat alveolar macrophages                                                                        | • Sprague-Dawley rat                                                                                   |                                                                                        |
|                  |                                                         | • Submerged conditions (24 h)                                                                             | • Intratracheal instillation (1, 3, 7, 30, 90, and 180 days)                                           |                                                                                        |
|                  |                                                         | • Cell viability, MIP-1α, IL-18, TNF-α                                                                      | • BALF analysis and histology                                                                          |                                                                                        |
|                  |                                                         |                                                                                                            |                                                                                        |                                                                                        |

Abbreviation: BALF, broncho-alveolar lavage fluid.
transformed to surface area, using the specific surface area of particles (i.e., nominator) or surface area of the cell culture dish or animal epithelial area (i.e., denominator). In the present article, we refer to the exposure level as dose to keep it simple. Nevertheless, it should be emphasized that the true dose in cell culture systems is uncertain as it depends on the sedimentation rate of particles.

### 2.4 Potency

Certain studies have used potency to rank the hazard of ENMs (Breznan et al., 2016; Loret et al., 2018). Potency is the exposure of a compound in terms of the dose that is needed to produce a predefined effect (e.g., low dose is equivalent to high potency).

### 3 DISCUSSION

#### 3.1 Can cell cultures predict pulmonary toxicity in animals?

A number of studies on ENMs have assessed whether data from in vitro assays predict positive (toxic) or negative (nontoxic) effects in exposed animals. This distinction between toxic and nontoxic materials derives in most studies from a qualitative assessment between in vitro and in vivo responses. Table 3 summarizes the studies that have been used to assess the predictivity of cell culture studies on outcomes in animal models. It is not possible to apply the same analysis of the association or ranking of hazards of ENMs between results from in vitro and in vivo studies because the studies have had different experimental designs. We have segregated the studies into experiments that have resulted in formal analysis of the predictivity or correlation between in vitro and in vivo responses, and those that have used a descriptive approach. Cho et al. (2013) assessed responses of in vitro assays with in vivo acute lung inflammogenicity for a set of metal oxide ENMs with a broad gradient in physicochemical properties. From the in vitro assays, an ENM was classified as “positive” if it caused any statistically significant increase at any dose, otherwise it was classified as “negative” (i.e., no effect) (Cho et al., 2013). For in vivo experiments, statistically significant increases in the number of total granulocytes in bronchoalveolar lavage fluid (BALF) compared with vehicle control were regarded as “positive.”

| Submerged/ ALI | Mono-/ co-cultures | Cell line/ primary cell | Total | Class complexity | References |
|----------------|--------------------|------------------------|-------|------------------|------------|
| 0              | 0                  | 0                      | 0     | A                | Rushton et al. (2010); Wang et al. (2011); Han et al. (2012); Zhang et al. (2012); Cho et al. (2013); Li et al. (2013); Ma et al. (2015); Pal et al. (2015); Wiemann et al. (2016); Loret et al. (2018); Wiemann et al. (2018); Cui et al. (2019); Di Ianni, Erdem, et al. (2021); Di Ianni, Müller, et al. (2021); Karkossa et al. (2021); Fujita et al. (2022) |
| 1              | 0                  | 0                      | 1     | B                | Li et al. (2013)a |
| 0              | 1                  | 0                      | 1     | B                | Sayes et al. (2007); Warheit et al. (2009); Loret et al. (2018) |
| 0              | 0                  | 1                      | 1     | B                | No study |
| 1              | 1                  | 0                      | 2     | C                | Di Ianni, Erdem, et al. (2021); Loret et al. (2018) |
| 1              | 0                  | 1                      | 2     | C                | No study |
| 0              | 1                  | 1                      | 2     | C                | No study |
| 1              | 1                  | 1                      | 3     | D                | No study |

Note: “0” is given to models in “Submerged conditions,” “Monocultures,” or “Cell lines,” whereas “1” is given to models that were in “ALI,” “Co-cultures,” or “Primary cells.” The total (0–3) leads to complexity class (A–D) with class A being the least complex and D the most complex.

aLi et al. (2013) mentions the use of BEAS-2B and THP-1a co-culture, but the results are not specifically described.
The authors made also a comparison of responses at a fixed dose in the in vitro assays (30 cm²/ml, i.e., the only overlapping dose used for all the particles except for CuO, where 3 cm²/ml was the highest dose; Cho et al., 2013). From the comparison of statistically significant effects at any dose, hemolysis assay appeared to be the most predictive (concordance = 100%), followed by cytotoxicity in differentiated peripheral blood mononuclear cells (PBMCs; 78% concordance, 14% false negativity, and 11% false positivity), while other assays showed poorer association (concordance ≤ 67%).

In contrast, with the comparison at equal dose, cytotoxicity in PBMCs was the most predictive, with 89% (8/9) concordance, 11% (1/9) false negativity, and 0% (0/9) false positivity, followed by cytotoxicity using primary cultured alveolar macrophages (78% [7/9] concordance, 11% [1/9] false negativity, and 11% [1/9] false positivity), while all other assays showed a poorer association (concordance ≤ 67%). Similarly, Wiemann et al. assessed the concordance in the effects between in vitro exposure in macrophages (NR8383 rat cells) and rat inhalation exposure for a set of 20 ENMs. The authors defined the no-observed-adverse-effect concentration (NOAEC) of <10 mg/m³ as threshold value indicating ENMs to be positive in vivo (inflammatory reactions defined by hematology, BALF evaluation, and/or lung histopathology), whereas NMs with NOAECs of ≥10 mg/m³ were categorized as being negative. For the in vitro assessment, the authors tested the ENMs in macrophages and quantified lactate dehydrogenase (LDH), β-glucoronidase (GLU), tumor necrosis factor alpha (TNF-α), and ROS production. The authors then assessed the lowest observable effect concentration (LOEC) for each ENM, and normalized it by the ENM Brunauer–Emmett–Teller (BET) surface area. ENMs were classified as positive if the in vitro LOECs were below 6000 mm²/ml for at least two of the four in vitro endpoints (e.g., elevated LDH and TNF-α). With this setup, the in vitro responses predicted the inflammation in vivo with 95% concordance, 91% specificity, and a 100% sensitivity (Wiemann et al., 2016). Karkossa et al. (2021) investigated a group

### TABLE 3  Summary of studies that have assessed the predictivity of cell culture studies on outcomes in animal models

| Author                  | Analysis  | No. of ENMs | Response (biomarker)                                                                 | Key outcome                                                                 |
|-------------------------|-----------|-------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Cho et al. (2013)       | Statistical | 9          | Different in vitro effects (cytotoxicity, proinflammatory cytokines, and hemolysis) versus inflammation (in vivo) | 100% concordance (hemolysis)                                                  |
|                         |           |             |                                                                                    | 59% concordance (cytotoxicity)²                                                 |
|                         |           |             |                                                                                    | 68% concordance (cytokines)²                                                  |
| Wiemann et al. (2016)   | Statistical | 15         | Different in vitro effects (cytotoxicity, proinflammatory cytokines and ROS production) versus inflammation (in vivo) | 95% concordance²                                                               |
| Di Ianni, Erdem, et al. (2021) | Statistical | 8          | Genotoxicity                                                                      | 83% concordance (THP-1a versus BAL cells)                                      |
|                         |           |             |                                                                                    | 32% concordance (THP-1a versus lungs)                                          |
|                         |           |             |                                                                                    | 50% concordance (A549 versus BAL cells)                                        |
|                         |           |             |                                                                                    | 14% concordance (A549 versus lungs)                                            |
| Di Ianni, Möller, et al. (2021) | Statistical | 5          | Genotoxicity                                                                      | 60% concordance (THP-1a versus BAL cells)                                     |
|                         |           |             |                                                                                    | 20% concordance (A549 versus BAL cells)                                        |
| Wiemann et al. (2018)   | Descriptive | 2          | Cytotoxicity (LDH activity)                                                        | Same response in vitro and in vivo (higher LDH activity by 15 nm as compared with 55 nm SiO₂) |
| Ma et al. (2015)        | Descriptive | 2          | Cytotoxicity and inflammation                                                      | Same response in vitro and in vivo (“large” graphene sheets more hazardous than “small” sheets for both cytotoxicity and inflammation biomarkers) |
| Wang et al. (2011)      | Descriptive | 2          | Profibrotic responses                                                              | Same hazard ranking of ENMs in vitro (profibrotic mediators) and in vivo lung (profibrotic mediators and collagen staining) |
| Li et al. (2013)        | Descriptive | 4          | Proinflammatory cytokines (cells) versus fibrosis (animals)                        | Same hazard ranking of ENMs in vitro cytokine production and in vivo lung collagen staining |
| Cui et al. (2019)       | Descriptive | 4          | Metabolites                                                                       | Comparable metabolomics signature in vitro and in vivo                        |

²Mean values of concordances in different cells and proinflammatory cytokines.
³In vitro hazard threshold is based on combined results from cytotoxicity, cytokine release, and ROS production assays.

The authors made also a comparison of responses at a fixed dose in the in vitro assays (30 cm²/ml, i.e., the only overlapping dose used for all the particles except for CuO, where 3 cm²/ml was the highest dose; Cho et al., 2013). From the comparison of statistically significant effects at any dose, hemolysis assay appeared to be the most predictive (concordance = 100%), followed by cytotoxicity in differentiated peripheral blood mononuclear cells (PBMCs; 78% concordance, 11% [1/9] false negativity, and 11% [1/9] false positivity), while other assays showed poorer concordance (≤67%). In contrast, with the comparison at equal dose, cytotoxicity in PBMCs was the most predictive, with 89% (8/9) concordance, 11% (1/9) false negativity, and 0% (0/9) false positivity, followed by cytotoxicity using primary cultured alveolar macrophages (78% [7/9] concordance, 11% [1/9] false negativity, and 11% [1/9] false positivity), while all other assays showed a poorer association (concordance ≤67%). Similarly, Wiemann et al. assessed the concordance in the effects between in vitro exposure in macrophages (NR8383 rat cells) and rat inhalation exposure for a set of 20 ENMs. The authors defined the no-observed-adverse-effect concentration (NOAEC) of <10 mg/m³ as threshold value indicating ENMs to be positive in vivo (inflammatory reactions defined by hematology, BALF evaluation, and/or lung histopathology), whereas NMs with NOAECs of ≥10 mg/m³ were categorized as being negative. For the in vitro assessment, the authors tested the ENMs in macrophages and quantified lactate dehydrogenase (LDH), β-glucoronidase (GLU), tumor necrosis factor alpha (TNF-α), and ROS production. The authors then assessed the lowest observable effect concentration (LOEC) for each ENM, and normalized it by the ENM Brunauer–Emmett–Teller (BET) surface area. ENMs were classified as positive if the in vitro LOECs were below 6000 mm²/ml for at least two of the four in vitro endpoints (e.g., elevated LDH and TNF-α). With this setup, the in vitro responses predicted the inflammation in vivo with 95% concordance, 91% specificity, and a 100% sensitivity (Wiemann et al., 2016). Karkossa et al. (2021) investigated a group
of metal oxide nanomaterials that included four SiO$_2$ NMs and one TiO$_2$. The authors tested with metabolomic and proteomic changes in the pathways involved in the oxidative stress in rat alveolar epithelial cells, macrophages, and rat lung. The authors observed that the changes of markers were different in all the models; however, the classification of “active” or “passive,” based on significantly changed signatures in exposed groups compared with controls, led to the congruent classifications between in vitro and in vivo (Karkossa et al., 2021).

Only one research group has assessed the predictivity of in vitro genotoxicity results on the same endpoint in animal lungs and BALF cells. The categorical classification of ENMs as genotoxic or nongenotoxic agent was based on DNA strand breaks, measured by the standard alkaline comet assay in species exposed to multiwalled carbon nanotubes (MWCNT), graphene oxide (GO), and clay nanomaterials (Di Ianni, Erdem, et al., 2021; Di Ianni, Møller, et al., 2021). A scoring system was developed based on standard deviation fold increase units. The categories were defined within control groups, to which the exposed groups for each ENM were compared. The comparison resulted in good concordance in DNA damage of MWCNT between activated THP-1 macrophages (THP-1a) and BALF cells of mice (83%, 5/6 control groups, to which the exposed groups for each ENM were compared. The comparison resulted in good concordance in DNA damage of MWCNT between activated THP-1 macrophages (THP-1a) and BALF cells of mice (83%, 5/6 correctly assigned), whereas there was less concordance with genotoxic responses in lung tissue (33%, 2/6). Test results for A549 cells were less concordant with the response in BALF cells (concordance = 50%) and lung tissue (concordance = 14%; Di Ianni, Erdem, et al., 2021). Using the same scoring system, there was good concordance in DNA damage caused by flake ENMs (nanoclays, GO and rGO) in THP-1a cells (4/5) and BALF cells of exposed mice groups at day 3 post-exposure (4/5), whereas DNA damage in A549 cells (2/5) was less concordant with the DNA damage in BALF cells (Di Ianni, Møller, et al., 2021).

Some studies have assessed how the variation in one variable/predictor, for example, ENMs size or other physico-chemical property, affects the outcomes quantified in vitro and in vivo. This generates information about the equivalence of the test result relative to the physico-chemical characteristic between the two test systems. Wiemann et al. (2018) showed for 15 and 55 nm SiO$_2$ the levels of cytotoxicity were size dependent and the dependency was concordant in vitro and in vivo (Wiemann et al., 2018). Similarly, Ma et al. (2015) showed that proinflammatory responses of GO increased with increasing lateral size of the particles in both cell cultures and animals. Wang et al. (2011) showed that experiments on well-dispersed MWCNTs led to equivalent conclusions on hazard of fibrosis in vitro and in vivo, whereas experiments using poorly dispersed MWCNTs produced inconsistent responses. When MWCNT were well dispersed in exposure medium, these induced a larger activation of fibrotic biomarkers in epithelial cells and macrophages, which was found to be consistent with the extent of fibrosis detected in lungs in vivo (Wang et al., 2011). Li et al. (2013) assessed in vitro and in vivo the profibrotic changes following exposure to a group of pristine and surface modified MWCNT. The authors identified cationic MWCNT (polyetherimide-modified MWCNT) to cause strongest effects in terms of platelet-derived growth factor AA (PDGF-AA), tissue growth factor β1 and interleukin (IL) 1β in vitro (BEAS-2B and THP-1a cells) and in vivo (same mediators in BALF or collagen staining in lung tissue), compared with other MWCNT, and concluded that the proinflammatory changes were comparable between the two models.

A similar qualitative approach was used to identify the biomarkers that are expressed in both in vitro and in vivo models following ENMs exposure. Cui et al. (2019) found comparable responses in terms of levels of metabolites in bronchial epithelial cells vitro and lung tissue of mice following exposure to a set of metal oxide ENMs, suggesting that the quantification of some metabolites in vitro could be used to predict the release of metabolites in vivo. The same study also assessed proinflammatory responses in cells and animal lungs, but there appeared to be no consistency between the responses seen in vitro and in vivo.

Taken together, binary classification of ENMs toxicity in vitro and in vivo could allow determining whether in vivo testing is necessary for hazard assessment. Results may also be used to group ENMs by biological activity.

### 3.2 Does the magnitude of response in cell cultures correspond to the magnitude of pulmonary toxicity in animals?

The magnitude of ENMs-induced effects in cells and animals have been proven valuable for the in vitro–in vivo correlation assessment. This type of assessment is based on comparisons of dose–response relationships for groups of ENMs, whereby the difference in the slopes is used for hazard rankings (Figure 2). Rushton et al. (2010) compared the steepest slopes of a set of ENMs, because the steepest part of a nonlinear dose–response curve is the response that is induced by intermediate doses and therefore is most sensitive to predict the variation in response. The responses were normalized by the ENM surface area, resulting in a response per unit surface area, as the response-metric. Linear fit models were
used to correlate the surface area-normalized responses in vitro with in vivo. All in vitro responses in terms of ROS production or expression of IL-8, using a luciferase reporter gene, correlated significantly with in vivo results ($r_s = 0.94$) (Rushton et al., 2010). Sayes et al. observed no correlation between in vitro and in vivo responses induced by a set of ENMs, using mass-dose metric (Sayes et al., 2007). Subsequently, Rushton et al. (2010) used the steepest slope-approach with the same dataset and showed good correlation after normalizing the response by use of the ENMs surface area. By using the same approach with slopes, Han et al. (2012) showed that in vitro and in vivo responses following exposure to TiO$_2$ ENM with different size and crystallinity were well correlated ($r_s = 0.93$ and $0.88$ for correlations between protein oxidation or LDH release in cells and influx of polymorphonuclear cells [PMNs] in BALF in vivo, respectively), suggesting that using the approach of steepest slope analysis offers a superior method to correlate in vitro with in vivo results of ENM toxicity and for ranking their toxic potency. Similarly, slopes of IL-8 expression in A549 and THP-1a cells after exposure to multiple MWCNT, GO and nanoclay were significantly correlated with inflammogenicity of the same materials in mice in terms of neutrophil influx ($r_s$-values $\geq 0.92$) after doses were normalized by use of ENMs BET surface area (Di Ianni, Erdem, et al., 2021; Di Ianni, Møller, et al., 2021). In the same dataset, we have observed positive correlation between cell death in A549 cells and THP-1a cells and neutrophil influx in mice (Figure 3). More recently, Fujita et al. tested the pulmonary toxicity of three fibers with different diameters and length (SCF1, SCF2, and SCF3). In vitro, the ranking was observed to be SCF3 > SCF1 > SCF2, based on cell death, MIP-1a, IL-18, and TNF-a. In vivo, the same ranking was observed based on neutrophil influx, eosinophils, total protein, and LDH in BALF. This demonstrated similar potential hazard effects in vitro and in vivo, namely the hazard increases with reduction of diameter and increase of length. It is important to note though that the evaluation included only three materials, which minimizes the chance of rejecting a conclusion of high concordance in hazard ranking in vitro and in vivo (Fujita et al., 2022).

Zhang et al. (2012) tested a set of metal oxide ENMs and demonstrated a concordant ranking of effect in vitro and in vivo by using normalized toxicological size effects. The in vivo effect size was defined by a T-statistic value that describes the difference between the mean neutrophil cell counts of metal oxide-exposed versus nonexposed animals. The in vitro effect size was defined as the predicted probability of cytotoxicity according to a regression model that
assigns a number between 0 and 1 for each ENM, with “1” meaning 100% confidence of a cytotoxic effect (or IL-6 and monocyte chemoattractant protein-1 [MCP-1] upregulation) and “0” meaning no chance. In this comparative analysis, ENMs with an in vitro toxicological probability >50% were statistically more likely to induce a high neutrophil influx in BALF. The in vitro toxicity response (cytotoxicity, MCP-1, and IL-6) correlated well with the T-statistic value of the in vivo inflammation data ($r_p > 0.76$) as well as with oxidative stress biomarkers (i.e., heme oxygenase 1) in lung tissue (Zhang et al., 2012).

Regression analyses have been used for in vitro hazard assessment. Pal et al. investigated the implications of the dosimetry on the hazard ranking, and demonstrated that correlating the slopes of a group of ENMs with “effective doses” instead of administered doses improved the correlation between in vivo lung inflammation and in vitro cell viability, with the regression coefficient ($r_{\text{linear regression}}$) increasing from 0.80 to 0.98. However, a closer inspection of the results shows that the inflammatory potency in cells (based on highest induction of IL-8) and PMNs in BALF did not correlate ($r_s = 0.1, p = 0.87$, results abstracted from figures in the publication). This highlights the uncertainty in

---

**FIGURE 3**  In vitro–in vivo correlations of cell death in A549 and activated THP-1 cells (i.e., macrophage-like) and neutrophil influx following exposure to MWCNT, nanoclays, and graphene oxide. Results on the axis are slopes from linear regression analysis in cells and animals (predictor: Cell death in cells per mass or surface dose; dependent variable: Neutrophils in BALF per mass or surface dose). The correlations were assessed with doses expressed as mass, or normalized against ENM surface area. In vitro correlation with neutrophil influx on mass basis: $r = -0.18$ and $p = 0.57$ with A549, $r = 0.09$ and $p = 0.76$ with THP-1a. In vitro correlation with neutrophil influx on surface area basis: $r = 0.97$ with $p < 0.001$ for A549, $r = 0.67$ with $p = 0.02$ for THP-1a. The correlations have been calculated from previously published data (Di Ianni, Erdem, et al., 2021; Di Ianni, Møller, et al., 2021). Figure created with RStudio using “ggplot2” package (Valero-Mora, 2010)
generalizing associations between specific outcomes to hazards of ENMs. Loret et al. (2018) used benchmark dose–response modeling to determine ENMs (3 TiO₂ and 1 CeO₂) dose intervals related to increase in proinflammatory mediator levels in vitro and in vivo. An effect of 20% increase of proinflammatory mediator level compared with nonexposed controls was used as critical value, and the intervals were defined as Benchmark Dose Lower confidence limit (BMDL) and the Benchmark Dose Upper confidence limit (BMDU) of the interval for a 90% confidence. The authors compared the intervals defined by different dose metrics. Although there were similarities in the ENMs ranking using in vivo and in vitro approaches, the dataset was not sufficiently large to provide firm ranking comparisons (Loret et al., 2018). It might be that including more potent ENMs would have enlarged the gradient in the toxicity and thereby improved the hazard ranking.

In summary, these studies using rankings or correlations on quantitative scale demonstrate that the magnitude of effects induced in vitro can resemble the responses in vivo, however the in vitro–in vivo comparisons need a gradient in the toxicity to reveal strong association of slopes or rankings.

### 3.3 Are advanced cell culture models better predictors of pulmonary toxicity in animals as compared with monocultures?

Inhaled nanoparticles deposit in different regions of the lungs depending on their size. Those in the smaller size range can deposit in the alveoli (Oberdörster et al., 2005). Different cell models can be implemented to mimic the lungs in different regions. Epithelial cells and macrophages are normally used in 2D monoculture and sometimes they are combined in co-culture to increase complexity due to cell–cell communication. To further increase the complexity, fibroblasts have been included in co-culture systems (Table 1). The level of complexity is also defined by the origin of the cells (e.g., primary or cancer cell line), as well as exposure system (submerged or ALI, discussed in the next section). We have arbitrarily defined categories of complexity, which are outlined in Table 2. Increasing the complexity of the cell models should better mimic the cellular microenvironment of the lungs and better predict the occurrence of adverse effects in vivo. As shown in Table 1, different cell culture models have been implemented in studies aiming at in vitro–in vivo comparisons of ENMs-induced oxidative stress, inflammation, and genotoxicity. A large number of studies shows positively correlated oxidative stress and inflammation between bronchial (BEAS-2B or 16-HBE) and alveolar (A549) epithelial cell lines as well as macrophage cell lines (human, THP-1a; mouse, RAW264.7, and J774A.1). Two studies reported also good concordance of genotoxic effects between THP-1a cells and BALF cells from mice lungs, whereas the concordance was poor for genotoxicity in A549 cells and mice lung tissue (Di Ianni, Erdem, et al., 2021; Di Ianni, Moller, et al., 2021). In contrast, there are very few studies showing positive correlation of the responses to ENM exposure in animal lungs and primary cells (human primary blood monocytes derived macrophages, mouse, and rat macrophages) (Cho et al., 2013; Jing et al., 2015), while others were less successful (Sayes et al., 2007; Warheit et al., 2009). The dataset generated in Sayes et al. (2007) was used by Rushton et al. (2010) to assess correlations with the steepest slope method, which resulted in good correlations between macrophage inflammatory protein-2 (MIP-2) in alveolar macrophages (AMs) and neutrophil influx ($r_s = 0.96$, $p = 0.04$), or LDH in L2-AMs co-culture and neutrophil influx ($r_s = 0.96$, $p = 0.002$), when doses were normalized by surface area (Rushton et al., 2010). However, Sayes et al. (2007) reported also other endpoints including TNF-α, IL-6 and cytotoxicity (i.e., MTT assay), but these endpoints were not included in this later analysis. Using cell lines might be beneficial in performing high-throughput screening of toxicity of libraries of ENMs and correlate to in vivo responses.

The combination of different cell types is proposed to be pivotal to increase the physiological relevance of in vitro models. However, of studies reported in Table 1, only Loret et al. (2018) show the same ranking of potency between co-cultures of A549 and THP-1a exposed in the ALI system, yet they did not compare the effects of the co-culture to a monoculture. In line with this, Di Ianni, Erdem, et al. (2021) showed that the co-culture of A549 and THP-1a cells exposed to MWCNT resulted in dose–response relationships that were comparable to those obtained in monocultures. Two other studies did not consist in proinflammation responses between co-cultures an in vivo (Sayes et al., 2007; Warheit et al., 2009). Rushton et al., 2010 used results by Sayes et al. (2007) and showed that normalized concentrations by use of ENM surface area produced equally good correlations between monoculture or co-culture, and neutrophil influx. To date, the limited number of studies showing positive in vitro–in vivo correlations with complex cell models hinders the conclusions regarding whether co-cultures better predict ENMs adverse effects in vivo.

The system in which cultured cells are exposed has been focus of discussions in the nanotoxicology field. Most of the in vitro toxicity tests employ exposure of cells in submerged conditions, whereby ENMs are administered from a...
particle suspension. Proliferating cells like alveolar epithelial cells in submerged condition covered by liquid are far from their real physiological microenvironment, as these cells grow at a gas phase and covered by a thin layer of lung surfactant (Parra & Pérez-gil, 2015). When biological effects of ENMs exposure are tested in submerged conditions, ENMs are dispersed in cell culture medium and administered to the cells, which are adherent to the bottom of the well. After administration, the dispersed ENMs sediment and reach adherent cells over time. The use of inserts (porous membranes) allows for culturing cells in ALI systems and for exposure to aerosolized ENMs. This better mimics the microenvironment of the lung epithelial surface: cells are in contact with medium at the basolateral side of the insert, while exposed to an aerosol at the apical side. Therefore, it has been argued that cells exposed to ENMs in the ALI system might be more predictive of the effects in the lungs. When considering the studies that have to date compared ENM toxicity in vitro with in vivo we have found that the largest majority obtained good correlations with exposure in submerged conditions (Table 1 and Figure 1). One study reported the exposure of cell culture models in the ALI to elicit comparable results to the responses observed in vivo (Loret et al., 2018). Nevertheless, it has been shown that exposure in the ALI of low-density materials like MWCNT has methodological limitations including nebulization of highly concentrated suspensions of materials, and the inflammatory response by cell exposure to MWCNT in the ALI system is comparable to the effective doses (quantified by thermogravimetric analysis) in submerged conditions (Di Ianni, Erdem, et al., 2021). In these two studies (Di Ianni, Erdem, et al., 2021; Loret et al., 2018), the ALI exposure used quantification of deposited doses by quartz-crystal microbalance (QCM), which was shown to be a sensitive and accurate device for real-time dosimetry of the cell-delivered particle dose (Ding et al., 2020). The quantification of “real,” sedimented doses in submerged conditions, determined by In vitro Sedimentation, Diffusion, and Dosimetry (ISDD) model, has been shown to improve hazard assessment of ENMs in vitro and thereby the correlations with hazard rankings in vivo (Pal et al., 2015). Similarly, Loret et al. (2018) found an improvement in the dose-effect estimates when adopting the ISDD model, however non-significant.

4 | PERSPECTIVES

This review aimed at collecting and discussing studies that have compared toxicity responses in cell cultures and animal lungs after exposure to ENMs. Based on the results described, it appears that simple cell cultures might be useful for prediction of adverse effects in vivo, whereas more complex cell culture models do not seem to increase the predictivity remarkably beyond the simple cell culture systems. At the same time, the results show that exposure of cells in submerged condition could be valuable in correlation analysis between in vitro and in vivo responses in animal studies. Studies of toxicological effects in cells exposed to the same insoluble or soluble ENMs in submerged condition and at ALI have shown inconsistent results (Table 4). In addition, these observations are predominantly based on experiments using a limited panel of ENMs in each study. In fact, it seems that only one study has investigated multiple ENMs and that indicated the same ranking of toxicological effect by exposure to TiO2 and CeO2 in submerged condition and ALI system (Loret et al., 2016). The studies comparing effects in cells exposed in submerged conditions with those exposed in the ALI system can be categorized into studies that show a strongest effect in ALI (i.e. higher response at the same dose or lower dose for the same level of response), the opposite or equivocal results. There is similar distribution of studies that have strongest effect in ALI (four studies; Loret et al., 2016; Hilton et al., 2019; Diabaté et al., 2021; Lovén et al., 2021), strongest effect in submerged condition (three studies; Herzog et al., 2014; Panas et al., 2014; Mills-Goodlet et al., 2020) and equivocal outcome of the comparison between ALI and submerged condition (six studies; Lenz et al., 2009; Lenz et al., 2013; Stoehr et al., 2015; Cappellini et al., 2020; Medina-Reyes et al., 2020; Di Ianni, Erdem, et al., 2021). This suggests that it is presently unclear whether the magnitude of effects with the ALI system is higher and whether such exposure systems are superior to submerged conditions. It is our impression that refinements (or improvements) have occurred for both ALI and submerged exposure systems over the period of from 2009 to 2021 (i.e., studies in Table 4). For instance, dispersion protocols are more standardized today than they were 10–15 years ago. The heterogeneity of results in Table 4 suggests that the optimal way forward to assess differences in responses in ALI and exposure condition would be to test the reproducibility of results by using the same ENMs (and same dispersion protocol), cell types, and biomarkers in different laboratories. On the same note, it would be desirable to test the reproducibility of correlations in vitro and in vivo responses in standardized conditions where different laboratories do these tests in parallel.

In vitro models can also include co-culture of multiple cell lines seeded in a 3D-orientation to study the effects of ENMs (Klein et al., 2013), and similar models have been used to model profibrotic potential of MWCNT in vitro
| Author | Exposure model | Engineered nanomaterial | Cell type | Response |
|--------|----------------|-------------------------|-----------|----------|
| Lenz et al. (2009) | ALICE | ZnO | A549 | Equivocal response (strongest response on IL-8 expression in submerged condition; strongest response on HO1 expression by exposure at the ALI) |
| Lenz et al. (2013) | Not applicable | ZnO | A549 | Equivocal response (strongest response on IL-8 expression in submerged condition; strongest response on HO1 expression by exposure at the ALI) |
| Herzog et al. (2014) | ALICE | Ag | Co-culture (A549, dendritic cells and macrophages) | Lower biological response by exposure at ALI than submerged condition |
| Panas et al. (2014) | Invitrocell | SiO₂ (Aerosil200, sample with 50 nm) | A549 | Biological effects observed at higher doses in ALI than in submerged condition |
| Stoehr et al. (2015) | ALICE | ZnO | A549 (with pIL8-Luc vector) | Equivocal response (depending on measurement of IL-8 gene expression or protein levels) |
| Loret et al. (2016) | Invitrocell | TiO₂ (NM-105, NM-101, NM-110) and CeO₂ (NM-212) | Co-culture (A549 and THP-1a) | Biological effects observed at lower doses in ALI than in submerged condition |
| Hilton et al. (2019) | Invitrocell | MWCNT-7 | Co-culture (A549, THP-1a and MRC-5) | Biological effects (proteomics) observed at lower doses in ALI than in submerged condition |
| Cappellini et al. (2020) | XposeALI | CeO₂ (NM-212) | Co-culture (A549 and THP-1a) | Similar level of toxicological effects in ALI and submerged exposure condition |
| Diabaté et al. (2021) | Invitrocell | CeO₂ (three different samples), TiO₂ | A549 | Biological effects (proteomics) observed at lower doses in ALI than in submerged condition |
| Medina-Reyes et al. (2020) | Invitrocell | TiO₂ (nanoparticles and nanofibers) | A549 | Similar level of toxicological effects in ALI and submerged exposure condition |
| Mills-Goodlet et al. (2020) | Invitrocell | SiO₂ (and allergen co-exposure) | Human alveolar epithelial lentivirus immortalized cells (hAELVI) | Biological effects (proinflammatory response) not observed in ALI at dose that caused effect in submerged condition |
| Di Ianni, Erdem, et al. (2021) | Invitrocell | MWCNT (NM-400 and NM-401) | Co-culture (A549, THP-1a, and WI-38) | Equal biological response by exposure at ALI and submerged condition when comparing effective, deposited doses per surface area |
| Lovén et al. (2021) | NACIVT | ZnO | A549 | Biological effects (proteomics) observed at lower doses in ALI than in submerged condition |
following repeated exposures (Barosova et al., 2020). To further increase complexity, microfluidic systems could be used to culture cells in organ-on-chip. In this system, epithelial cells and macrophages are cultured on a membrane exposed to air, and the micro channels include circulating neutrophils. This models could resemble neutrophil influx to the apical side of the chip following ENM deposition (Benam et al., 2016). The microfluidic system has also peristaltic pumps, which compress and stretches the epithelium hence mimicking the movement of the lung during respiration. This system could prove valuable to detect toxicological effects caused by shape of, for example, flake-like ENMs (Di Ianni et al., 2020). In addition, advanced models could include sensitized models that recapitulate the pathophysiology of chronic obstructive pulmonary disease or asthma (Benam et al., 2016; Jimenez-Valdes et al., 2020). Despite the promises of more complex models, there is still a lack of studies showing in vitro–in vivo correlations with more complex in vitro models and therefore more studies are necessary to draw conclusion on their predictivity.

The studies described in this review compare responses in vitro to the responses in vivo following exposure of animals by intratracheal instillation (12 studies), inhalation (3 studies), or oropharyngeal aspiration (5). Intratracheal instillation is broadly used as surrogate for inhalation and has the advantage of delivering the suspended ENMs as a bolus, with known amount of delivered ENM which, on the other hand, would need to be quantified or computationally modeled in the case of exposure via inhalation. Nevertheless, compared with inhalation, instillation is not a physiologically relevant exposure, has higher dose rate, and may cause a different distribution of deposited doses. Further, instillation may cause stronger inflammatory responses compared with inhalation, as it was shown for carbon black at similar deposited doses (Jackson et al., 2012). In oropharyngeal aspiration, the substance is placed in the pharynx, which is more proximal than the trachea in the respiratory tract. The deposition pattern of pharyngeal aspiration is more dispersed than that of intratracheal instillation, making it more similar to inhalation (Kinaret et al., 2017). Using the same surface dose, previous studies showed higher inflammatory and tumor response by instillation than inhalation (Møller et al., 2013; Osier & Oberdörster, 1997). The concern about delivered doses in animal studies leads to a general concern about the assessment of uncertainty in studies on the correlation of responses in vitro and in vivo. For in vitro studies, there might be uncertainty about the delivered dose in submerged exposure condition (i.e., the dose depends on the sedimentation rate), whereas the deposited dose is determined as a part of the experiment in the ALI system. On the contrary, there might be uncertainty about the delivered dose in inhalation studies (i.e., deposition depends on the physico-chemical characteristics of ENMs), whereas the delivered dose is certain in studies on intratracheal instillation. Thus, depending on the comparison, there can be uncertain about the dose of the predictor (i.e., submerged condition) and dependent (i.e., inhalation) variable in statistical analysis. It suggests that the best studies—in terms of alleviating nondifferential exposure misclassification—would be exposure at the ALI versus the same responses in animals after intratracheal instillation. As nondifferential exposure misclassification blunts associations between predictor and outcome in statistical analysis, it can be speculated that studies using submerged condition and inhalation exposure may give rise to lower correlation coefficients or concordances than studies using exposure by ALI and intratracheal instillation.

The determination of effective doses, as well as use of surface area-metric, might improve the assessment of hazard rankings in vitro. Very importantly though, there are too few studies to firmly draw conclusions on whether or not complex cell culture models are superior in predicting in vivo responses. However, based on what is discussed here, there are different approaches that are promising to advance the in vitro–in vivo comparisons of ENMs in the nanotoxicology field and beyond. In general, the ALI systems are still in stages of being validated, although intra- and interlaboratory reproducibility of the CULTEX system has been tested on relatively large panel of compounds, including ENMs (Steinritz et al., 2013; Tsoutsoulopoulos et al., 2019). It should also be noted that the studies so far have assessed biomarkers as proxy-measures of pulmonary diseases. The association between toxicity of ENMs in cell cultures and hard endpoints such as pulmonary fibrosis and neoplastic diseases is unresolved. For instance, evidence suggests that certain ENMs, such as TiO₂ and carbon black, are not carcinogenic in mice following inhalation exposure, whereas rats appear to be a sensitive species. The extrapolation from in vitro and animal models to human hazard assessment may be aided by investigating benchmark particles for which epidemiological and long-term animal studies are available. These include quartz, diesel exhaust particles, and asbestos (IARC, 2012, 2013).

A lot has been discussed about the realistic dose ranges in vitro. For example, 160 μg/ml or 45.7 μg/cm² in Di Ianni, Erdem, et al., 2021 are much larger compared with instilled doses in vivo (e.g., up to 162 μg or 2 μg/cm² of the alveolar surface area [mouse alveolar surface area 82 cm²; Knust & Ochs, 2009]). In a recent report, Ma-Hock et al. (2021) proposed a method to use in vivo organ burden to define the doses to be tested in vitro. Specifically, the lung burden at the LOEC would be used for defining the effective dose to which the cells should be exposed to detect a response. For the case of CeO₂, for example, the in vitro effective concentration range should span around 0.003 μg/cm² based on
extrapolation of the doses used in animals (Ma-Hock et al., 2021). In perspective, Loret et al. (2018) found the BMD of cytokines induction (IL-6, IL-8, IL-1β, and TNF-α) in vitro to be 100–200 times larger than the BMD of the same cytokines in vivo. If the BMDs in vivo shall be applied to in vitro studies, there could be potentially no effect also for particles that are known to be highly potent. Similarly, Teeguarden et al. (2014) showed that higher doses of Fe3O4 were needed to induce inflammatory response in macrophages in vitro (~1.2–4 μg/cm² cell surface area) than when administered in vivo (0.009–0.4 μg/cm² alveolar surface area). Jagiello et al. (2021) determined the BMD to activate inflammatory reactions involved in the development of lung fibrosis (Adverse outcome pathway nr. 173) by exposure to MWCNT, for example, the BMD of a specific type of MWCNT called NM-401 was 0.73 μg in mice, corresponding to 0.0088 μg/cm² of alveolar surface area. This value is still way far from the doses of NM-401 that induced significant effects in submerged conditions (effective doses) and in the ALI (Di Ianni, Erdem, et al., 2021). Accordingly, the dose range tested, minimum 5.25 and 4.45 μg/cm² were necessary to induce a significant effect in A549 and THP-1a cells, respectively, and 9 μg/cm² to the co-culture at the ALI (Di Ianni, Erdem, et al., 2021). It is important to note that the lung complexity in vivo is much higher compared with the mono- and co-culture models creating a large differences in terms of doses required to induce effects. Moreover, the meaningfulness of comparing “realistic exposures” in cell cultures to exposures in animals or humans has been questioned as primary cells undergo a “culture shock” when they are cultured outside the body and those that survive are adapted to ex vivo conditions. In addition, something as simple as the doubling time of the same cell line, obtained in different laboratories, has been shown to differ substantially (Møller et al., 2021). The inclusion of benchmark particles with known toxicity in vivo and in humans may provide scaling between in vitro and in vivo responses.

The mechanism of action and the target cells following ENMs exposure can be pivotal for identifying the right models and select biomarkers of toxicity. Adverse outcome pathways (AOP) are OECD frameworks describing the mechanism of action (oxidative stress, inflammation, genotoxicity, etc.) of a compound into a series of key events (KE) leading to an adverse outcome (AO), via one or multiple KE, interconnected by causative relationships (Halappanavar et al., 2020). AOPs are developed based on analyses of tissue histological changes as well as omics. Further, AOP can also guide the selection of in vitro models to be used in the in vitro assessment. Namely, single cell RNA sequencing (sc-RNA-seq) might be a valuable tool in the development of in vitro models, by identifying the animal lung cell subpopulations that are most reactive, with gene upregulations, following particle deposition in the respiratory regions. These cells and the specific genes could be used in the quantification of ENM-caused hazard in vitro, with in vitro results correlated to the responses in vivo. By knowing the relationship between different KEs (e.g., proinflammatory upregulation and recruitment of neutrophil influx), AOPs can guide the selection of in vitro models and endpoints for the assessment of KE, including biomarkers such as neutrophil chemoattractants (IL-8) in vitro and neutrophil influx in vivo (Di Ianni, Erdem, et al., 2021; Di Ianni, Møller, et al., 2021). In line with this, it has been postulated that battery tests of KE leading to AO could ultimately be included in tiered testing strategies that bypass the use of animals for characterization of ENM hazards (Halappanavar et al., 2020, 2021; Nymark et al., 2021). The mechanism-based hazard assessment aided by AOPs could be pivotal in New Approach Methodologies (NAMs), which represent, according to the European Chemical Agency, “a broad context that includes in silico approaches, in chemico, and in vitro assays, as well as the inclusion of information from the exposure of chemicals in the context of risk assessment.” Divided in different categories, NAMs are meant to reduce or replace the use of animals in high dose toxicity studies (Nymark et al., 2020). These categories are high-throughput screenings, grouping and read across, exposure assessment, and modeling. Quantitative structure–activity relationships, AOPs, dose–response extrapolations, and benchmark dose identification (including EPA method and RIVM PROAST (Slob & Setzer, 2013).

The studies included in this review focused on toxicity of well-characterized ENMs. However, the methods discussed here might be applied to complex mixtures of nanoparticles, such as those in particulate matter from urban environment or occupational settings. Breznan and colleagues explored the relationship between particle effects in vitro and in vivo following exposure to a group of urban particles, minerals and ENMs. By correlating the individual in vitro and in vivo particle potency estimates, as well as the in vitro and in vivo combined particle potency estimates, the authors observed that the correlations of combined inflammatory potential in vitro and vivo, and the integrated potency estimates in cells versus mice approached the threshold for significance $\alpha = 0.05$ ($r_2 = 0.8$, $p = 0.052$) (Breznan et al., 2016).

The present review has described associations between responses in cells and animal lungs, which have been highlighted in the selected publications. In certain instances, the authors have assessed multiple biomarkers in cells, but only focused on the positive associations. Pal et al. (2015) is one such example where strength of the correlation between cytotoxicity (in vitro) and pulmonary inflammation (in vivo) was highlighted, but the correlations between proinflammatory responses in cells and animals were not mentioned. Based on the slopes reported in the article, there was the same ranking of PMN influx in animals and in vitro cell viability (Printex 90 > Ni InCo > nanoAg >
CeO$_2 >$ TiO$_2$), whereas IL-8 release had a different ranking (Printex 90 > nanoAg > Ni InCo > CeO$_2 >$ TiO$_2$). Another example is Cui et al. (2019) that focuses on responses in metabolites in vitro and in vivo; however, the publication also contain data that indicate a lack of consistency in proinflammatory responses in cells and animals. A future approach could be to assess hypothesis-driven associations, where the use of relevant biomarkers is justified before the experimental work is performed. The two examples suggest a risk of reporting bias in the present dataset. In fact, reporting, publication and citation bias need to be carefully assessed in systematic reviews of the association between in vitro and in vivo responses after exposure to ENMs. Null effect or inconsistent findings might be detected in studies with multiple objectives. For instance, reporting bias has been observed in a systematic review of genotoxic effects by TiO$_2$ as a somewhat higher proportion of positive findings in studies on only TiO$_2$ as compared with studies with multiple ENMs (Møller et al., 2017). Publication bias can be difficult to detect, but one way is to look for asymmetry in reported effect sizes versus the study size or precision (e.g., Funnel plots) or assessing the relationship between reports in peer-review journals and other kinds of disseminations (e.g., conference proceedings or project delivery reports). As an example, publication bias has been speculated to occur in studies on genotoxicity of MWCNTs as null effect finding in EU projects do not appear to published at the same rate as effects on genotoxicity (Møller et al., 2021). Reporting bias may occur if authors selectively cite other studies that have reported a positive association between in vitro and in vivo findings. A way to assess reporting bias is to compare citations of articles with opposite findings as for instance has been done in systematic review of DNA damage results by the comet assay (Møller et al., 2010, 2020). Publication and citation biases are flaws that may skew the interpretation of experimental evidence in a certain direction. Thus, it is important to proceed in systematic reviews with predefined hypotheses. Meta-analyses may only be included if there are sufficient number of studies with reasonably similar study design and biomarkers. Publications in the current database are too heterogeneous to be used in meta-analyses and we have used a somewhat explorative search strategy. Especially, it seems prudent to standardize the association between predictors and outcomes. At least three different approaches have been used in the 18 studies that we have systematically evaluated, including binary and continuous outcomes of the biomarkers (e.g., effect size of cytokine release), and doses at fixed effect of biomarkers.

In summary, studies using simple in vitro models with cells exposed to ENMs in submerged condition have been shown to predict the extent of the toxic effects in vivo. There are relatively few studies that have assessed predictivity of the complex models, including ALI, and they have not indicated associations stronger than those observed with simple monoculture systems.

ACKNOWLEDGMENT
This work has received funding from the European Union’s Horizon 2020 research and innovation program [agreement No 760840 (GRACIOUS)] and from the Danish Government FFIKA, Focused Research Effort on Chemicals in the Working Environment. Graphical abstract, Figures 1 and 2 were created with BioRender.com.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
Emilio Di Ianni: Conceptualization (equal); data curation (lead); investigation (lead); methodology (lead); writing – original draft (lead); writing – review and editing (lead). Nicklas Raun Jacobsen: Conceptualization (equal); funding acquisition (equal); methodology (equal); supervision (equal); writing – review and editing (equal). Ulla Vogel: Conceptualization (equal); funding acquisition (equal); supervision (equal); writing – review and editing (equal). Peter Møller: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); supervision (lead); writing – original draft (equal); writing – review and editing (equal).

DATA AVAILABILITY STATEMENT
Data available on request from the authors.

ORCID
Emilio Di Ianni © https://orcid.org/0000-0003-1789-2706
Nicklas Raun Jacobsen © https://orcid.org/0000-0002-2504-2229
Ulla Vogel © https://orcid.org/0000-0001-6807-1524
Peter Møller © https://orcid.org/0000-0002-2021-1249
REFERENCES

Barosova, H., Maione, A. G., Septiadi, D., Sharma, M., Haeni, L., Bulog, S., O’Connell, O., Jackson, G. R., Brown, D., Clippinger, A. J., Hayden, P., Petri-Fink, A., Stone, V., & Rothen-Rutishauser, B. (2020). Use of EpiAlveolar lung model to predict fibrotic potential of multi-walled carbon nanotubes. *ACS Nano*, 14(4), 3941–3956. https://doi.org/10.1021/acsnano.9b06860

Benam, K. H., Villenave, R., Lucchesi, C., Varone, A., Hubeau, C., Lee, H. H., Alves, S. E., Salmon, M., Ferrante, T. C., Weaver, J. C., Bahinski, A., Hamilton, G. A., & Inger, D. E. (2016). Small airway-on-a-chip enables analysis of human lung inflammation and drug responses in vitro. *Nature Methods*, 13(2), 151–157. https://doi.org/10.1038/nmeth.3697

Breznan, D., Karthikeyan, S., Phaneuf, M., Kumarathasan, P., Cakmak, S., Denison, M. S., Brook, J. R., & Vincent, R. (2016). Development of an integrated approach for comparison of in vitro and in vivo responses to particulate matter. *Particle and Fibre Toxicology*, 13(1), 1–24. https://doi.org/10.1186/s12899-016-0152-6

Cappellini, F., Di Buccheri, S., Geraci, R., Latvala, S., Malmlöf, M., Kippner, M., Elihn, K., Hedberg, J., Odnevall Wallinder, I., Gerde, P., & Karlsson, H. L. (2020). Dry generation of CeO2 nanoparticles and deposition onto a co-culture of A549 and THP-1 cells in air-liquid interface—Dosimetry considerations and comparison to submerged exposure. *Nanomaterials*, 10, 618.

Cho, W. S., Duffin, R., Bradley, M., Megson, I. L., MacNee, W., Lee, J. K., Jeong, J., & Donaldson, K. (2013). Predictive value of in vitro assays depends on the mechanism of toxicity of metal oxide nanoparticles. *Particle and Fibre Toxicology*, 10(1), 55. https://doi.org/10.1186/1743-8977-10-55

Cui, L., Wang, X., Sun, B., Xia, T., & Shen, H. (2019). Predictive metabolomic signatures for safety assessment of metal oxide nanoparticles. *ACS Nano*, 13(11), 13065–13082. https://doi.org/10.1021/acsnano.9b05793

Di Ianni, E., Erdem, J. S., Møller, P., Sahlgren, N. M., Poulsen, S. S., Knudsen, K. B., Zienoldldiny, S., Saber, A. T., Wallin, H., Vogel, U., & Jacobsen, N. R. (2021). In vitro-in vivo correlations of pulmonary inflammogenicity and genotoxicity of MWCNT. *Particle and Fibre Toxicology*, 18(1), 1–16. https://doi.org/10.1186/s12899-021-00413-2

Di Ianni, E., Møller, P., Mortensen, A., Szarek, J., Clausen, P. A., Saber, A. T., Vogel, U., & Jacobsen, N. R. (2020). Organomodified nanoclays induce less inflammation, acute phase response, and genotoxicity than pristine nanoclays in mice lungs. *Nanotoxicology*, 14(7), 869–892. https://doi.org/10.1080/17435390.2020.1771786

Di Ianni, E., Møller, P., Vogel, U. B., & Jacobsen, N. R. (2021). Pro-inflammatory response and genotoxicity caused by clay and graphene nanomaterials in A549 and THP-1 cells. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 872, 503405. https://doi.org/10.1016/j.mrgentox.2021.503405

Diabaté, S., Armand, L., Murugadoss, S., Dilger, M., Fritsch-Decker, S., Schlager, C., Béal, D., Arnal, M. E., Biola-Clier, M., Ambrose, S., Müller, S., Paur, H. R., Lynch, I., Valsami-Jones, E., Carriere, M., & Weiss, C. (2021). Air–liquid interface exposure of lung epithelial cells to low doses of nanoparticles to assess pulmonary adverse effects. *Nanomaterials*, 11(1), 1–23. https://doi.org/10.3390/nano11010065

Ding, Y., Weindl, P., Lenz, A. G., Mayer, P., Krebs, T., & Schmid, O. (2020). Quartz crystal microbalances (QCM) are suitable for real-time dosimetry in nanotoxicological studies using VITROCELL® cloud cell exposure systems. *Particle and Fibre Toxicology*, 17(1), 44. https://doi.org/10.1186/s12899-020-00376-w

Evans, S. J., Clift, M. J. D., Singh, N., De Oliveira, J., Mallia, M. B., Wills, J. W., Wilkinson, T. S., Jenkins, G. J. S., & Doak, S. H. (2017). Critical review of the current and future challenges associated with advanced in vitro systems towards the study of nanoparticle (secondary) genotoxicity. *Mutagenesis*, 32(1), 233–241. https://doi.org/10.1093/mutage/gew054

Fujita, K., Obara, S., & Maru, J. (2022). Pulmonary toxicity, cytotoxicity, and genotoxicity of submicron-diameter carbon fibers with different diameters and lengths. *Toxicology*, 466, 153063. https://doi.org/10.1016/j.tox.2021.153063

Gold, L. S., Slone, T. H., Manley, N. B., & Bernstein, L. (1991). Target organs in chronic bioassays of 533 chemical carcinogens. *Environmental Health Perspectives*, 93, 233–246. https://doi.org/10.1289/ehp.919323

Hadrup, N., Zhernovkov, V., Jacobsen, N. R., Voss, C., Strunz, M., Ansari, M., Schiller, H. B., Halappanavar, S., Poulsen, S. S., Khloedenko, B., Stoefer, T., Saber, A. T., & Vogel, U. (2020). Acute phase response as a biological mechanism-of-action of (nano)particle-induced cardiovascular disease. *Small*, 16(21), e1907476.

Halappanavar, S., Nymark, P., Krug, H. F., Clift, M. J. D., Rothen-rutishauser, B., & Vogel, U. (2021). Non-animal strategies for toxicity assessment of nanomaterials: Role of adverse outcome pathways in the selection of endpoints. *Small*, 17, e2007628. https://doi.org/10.1002/smll.202007628

Halappanavar, S., Van Den Brule, S., Nymark, P., Gaté, L., Seidel, C., Valentino, S., Zhernovkov, V., Högh Danielsen, P., de Vizcaya, A., Wolff, H., Stöger, T., Boyadziev, A., Poulsen, S. S., Serli, J. B., & Vogel, U. (2020). Adverse outcome pathways as a tool for the design of testing strategies to support the safety assessment of emerging advanced materials at the nanoscale. *Particle and Fibre Toxicology*, 17(1), 16. https://doi.org/10.1186/s12899-020-00344-4

Han, X., Corson, N., Wade-Mercur, P., Gelein, R., Jiang, J., Sahu, M., Biswas, P., Finkelstein, J. N., Elder, A., & Oberdörster, G. (2012). Assessing the relevance of in vitro studies in nanotoxicology by examining correlations between in vitro and in vivo data. *Toxicology*, 297, 1–9. https://doi.org/10.1016/j.tox.2012.03.006
Herzog, F., Loza, K., Balog, S., Clift, M. J. D., Eppe, M., Gehr, P., Petri-Fink, A., & Rothen-Rutishauser, B. (2014). Mimicking exposures to acute and lifetime concentrations of inhaled silver nanoparticles by two different in vitro approaches. *Beilstein Journal of Nanotechnology, 5*(1), 1357–1370. https://doi.org/10.3762/bjnano.5.149

Hilton, G., Barosova, H., Petri-Fink, A., Rothen-Rutishauser, B., & Bereman, M. (2019). Leveraging proteomics to compare submerged versus air-liquid interface carbon nanotube exposure to a 3D lung cell model. *Toxicology In Vitro, 54*, 58–66. https://doi.org/10.1016/j.tiv.2018.09.010

IARC. (2012). Silica dust, crystalline, in the form of quartz or cristobalite. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans* (Vol. 100C). IARC.

IARC. (2013). Diesel and gasoline engine exhausts and some nitroarenes. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans* (Vol. 105). IARC.

Jackson, P., Hougaard, K. S., Boisen, A. M., Jacobsen, N. R., Jensen, K. A., Møller, P., Brunborg, G., Gutzkow, K. B., Andersen, O., Loft, S., Vogel, U., & Wallin, H. (2012). Pulmonary exposure to carbon black by inhalation or instillation in pregnant mice: Effects on liver DNA strand breaks in dams and offspring. *S(5), 486–500. https://doi.org/10.3109/17453590.2011.587902*

Jagiello, K., Halappanavar, S., Rybinska-Fryca, A., Williams, A., Vogel, U., & Puzyn, T. (2021). Transcriptomics-based and AOP-informed structure-activity relationships to predict pulmonary pathology induced by multiwalled carbon nanotubes. *Small, 17*, e2003465. https://doi.org/10.1002/smll.202003465

Jimenez-Valdes, R. J., Can, U. I., Niemeyer, B. F., & Benam, K. H. (2020). Where we stand: Lung organotypic living systems that emulate human-relevant host-environment/pathogen interactions. *Frontiers in Bioengineering and Biotechnology, 8*, 989. https://doi.org/10.3389/fbioe.2020.00989

Jing, X., Park, J. H., Peters, T. M., & Thorne, P. S. (2015). Toxicity of copper oxide nanoparticles in lung epithelial cells exposed at the air-liquid interface compared with in vivo assessment. *Toxicology In Vitro, 29*(3), 502–511. https://doi.org/10.1016/j.tiv.2014.12.023

Karkossa, I., Bannuscher, A., Hellack, B., Wohlleben, W., Laloy, J., Stan, M. S., Dinischiotti, A., Wiemann, U., Luch, A., Haase, A., von Bergen, M., & Schubert, K. (2021). Nanomaterials induce different levels of oxidative stress, depending on the used model system: Comparison of in vitro and in vivo effects. *Science of the Total Environment, 801*, e149538. https://doi.org/10.1016/j.scitotenv.2021.149538

Kinaret, P., Ilves, M., Fortino, V., Rydman, E., Karisola, P., Lähde, A., Koivistoo, J., Jokiniemi, J., Wolff, H., Savolainen, K., Greco, D., & Alenius, H. (2017). Inhalation and oropharyngeal aspiration exposure to rod-like carbon nanotubes induce similar airway inflammation and biological responses in mouse lungs. *ACS Nano, 11*(1), 291–303. https://doi.org/10.1021/acsnano.6b05652

Klein, S. G., Serchi, T., Hoffmann, L., Blömeke, B., & Gutleb, A. C. (2013). An improved 3D tetraculture system mimicking the cellular orga-

Krems et al., E. I., Delgado-Buenrostro, N. L., Leseman, D. L., Déciga-Alcaraz, A., He, R., Gremmer, E. R., Fokkens, P. H. B., Flores-Flores, J. O., Cassee, F. R., & Chirino, Y. I. (2020). Differences in cytotoxicity of lung epithelial cells exposed to titanium dioxide nanofibers and nanoparticles: Comparison of air-liquid interface and submerged cell cultures. *Toxicology In Vitro, 65*, 104798. https://doi.org/10.1016/j.tiv.2020.104798

Mills-Goodlet, R., Schenck, M., Chary, A., Geppert, M., Serchi, T., Hofer, S., Hofstätter, N., Feinle, A., Hüsing, N., Gutleb, A. C., Himly, M., & Duschl, A. (2020). Biological effects of allergen-nanoparticle conjugates: Uptake and immune effects determined on...
HAELVi cells under submerged: vs. air-liquid Interface conditions. Environmental Science: Nano, 7(7), 2073–2086. https://doi.org/10.1039/cen01353a

Moller, P., Christophersen, D. V., Jensen, D. M., Kermanizadeh, A., Roursgaard, M., Jacobsen, N. R., Hemmingsen, J. G., Danielsen, P. H., Cao, Y., Jantzen, K., Klingberg, H., Hersoug, L. G., & Loft, S. (2014). Role of oxidative stress in carbon nanotube-generated health effects. Archives of Toxicology, 88(11), 1939–1964. https://doi.org/10.1007/s00204-014-1356-x

Moller, P., Danielsen, P. H., Jantzen, K., Roursgaard, M., & Loft, S. (2013). Oxidatively damaged DNA in animals exposed to particles. Critical Reviews in Toxicology, 43(2), 96–118. https://doi.org/10.3109/10408444.2012.756456

Moller, P., Jensen, D. M., Wils, R. S., Andersen, M. H. G., Danielsen, P. H., & Roursgaard, M. (2017). Assessment of evidence for nanosized titanium dioxide-generated DNA strand breaks and oxidatively damaged DNA in cells and animal models. Nanotoxicology, 11, 1237–1256. https://doi.org/10.1080/17435390.2017.1406549

Moller, P., Lehr, M., Folkmann, J. K., Mikkelsen, L., & Loft, S. (2010). Aging and oxidatively damaged nuclear DNA in animal organs. Free Radical Biology and Medicine, 48(10), 1275–1285. https://doi.org/10.1016/j.freeradbiomed.2010.02.003

Moller, P., Stopper, H., & Collins, A. R. (2020). Measurement of DNA damage with the comet assay in high-prevalence diseases: Current status and future directions. Mutagenesis, 35(1), 5–18. https://doi.org/10.1093/mutage/geo018

Moller, P., Wils, R. S., Di Ianni, E., Gutierrez, C. A. T., Roursgaard, M., & Jacobsen, N. R. (2021). Genotoxicity of multi-walled carbon nanotube reference materials in mammalian cells and animals. Mutation Research/Reviews in Mutation Research, 788, 108393. https://doi.org/10.1016/j.mrrev.2021.108393

Nymark, P., Bakker, M., Dekkers, S., Franken, R., Fransman, W., Garcia-Bilbao, A., Greco, D., Gulumian, M., Hadrup, N., Halappanavar, S., Hongisto, V., Hougaard, K. S., Jensen, K. A., Kohonen, P., Koivisto, A. J., Dal Maso, M., Oosterwijk, T., Poikkimiäki, M., Rodríguez-Llopis, I., … Grafström, R. (2020). Toward rigorous materials production: New approach methodologies have extensive potential to improve current safety assessment practices. Small, 16(6), 1–13. https://doi.org/10.1002/smll.201904749

Nymark, P., Karlsson, H. L., Halappanavar, S., & Vogel, U. (2021). Adverse outcome pathway development for assessment of lung carcinogenicity by nanoparticles. Frontiers in Toxicology, 3, 1–11. https://doi.org/10.3389/ftox.2021.653586

Oberdörster, G., Oberdörster, E., & Oberdörster, J. (2005). Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. Environmental Health Perspectives, 13, 1–11. https://doi.org/10.1289/ehp.7339

Osier, M., & Oberdörster, G. (1997). Intratracheal inhalation vs intratracheal instillation: Differences in particle effects. Fundamental and Applied Toxicology, 40(2), 220–227. https://doi.org/10.1006/fati.1997.2390

Pal, A. K., Bello, D., Cohen, J., & Demokritou, P. (2015). Implications of in vitro dosimetry on toxicological ranking of low aspect ratio engineered nanomaterials. Nanotoxicology, 9(7), 871–885. https://doi.org/10.1080/17435390.2014.986670

Panas, A., Comouth, A., Saathoff, H., Leisner, T., Al-Rawi, M., Simon, M., Seemann, G., Dössel, O., Mühlkopf, S., Paur, H. R., Fritsch-Decker, S., Weiss, C., & Diabaté, S. (2014). Silica nanoparticles are less toxic to human lung cells when deposited at the air-liquid Interface compared to conventional submerged exposure. Beilstein Journal of Nanotechnology, 5(1), 1590–1602. https://doi.org/10.3762/bjnano.5.171

Parra, E., & Pérez-gil, J. (2015). Composition, structure and mechanical properties define performance of pulmonary surfactant membranes and films. Chemistry and Physics of Lipids, 185, 153–175. https://doi.org/10.1016/j.chemphyslip.2014.09.002

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 dosometric and chemical/biological response metrics. Journal of Toxicology and Environmental Health - Part A: Current Issues, 73, 445–461. https://doi.org/10.1080/15287390903489422

Saber, A. T., Poulsen, S. S., Hadrup, N., Jacobsen, N. R., & Vogel, U. (2019). Commentary: The chronic inhalation study in rats for assessing lung cancer risk may be better than its reputation. Particle and Fibre Toxicology, 4, 3–6.

Sayes, C. M., Reed, K. L., & Warheit, D. B. (2007). Assessing toxicity of fine and nanoparticles: Comparing in vitro measurements to in vivo pulmonary toxicity profiles. Toxicological Sciences, 97(1), 163–180. https://doi.org/10.1093/toxsci/kfm018

Slob, W., & Setzer, R. W. (2013). Shape and steepness of toxicological dose-response relationships of continuous endpoints. Critical Reviews in Toxicology, 43(7), 270–297. https://doi.org/10.1080/10408444.2013.853726

Steinrütz, D., Möhle, N., Pohl, C., Papritz, M., Stenger, B., Schmidt, A., Kirchhart, C. J., Thiermann, H., Vogel, R., Hoffmann, S., & Auferheide, M. (2013). Use of the Cultex radial flow system as an in vitro exposure method to assess acute pulmonary toxicity of fine dusts and nanoparticles with special focus on the intra- and inter-laboratory reproducibility. Chemos-Biologi-Cal Interactions, 206(3), 479–490. https://doi.org/10.1080/03601277.2013.505001

Stoehr, L. C., Endes, C., Radauer-Preiml, I., Boyles, M. S. P., Casals, E., Balog, S., Pesch, M., Petri-Fink, A., Rothen-Rutishauser, B., Himly, M., Clift, M. J., & Duschl, A. (2015). Assessment of a panel of interleukin-8 reporter lung epithelial cell lines to monitor the pro-inflammatory response following zinc oxide nanoparticle exposure under different cell culture conditions. Particle and Fibre Toxicology, 12(1), 1–12. https://doi.org/10.1186/s12989-015-0104-6

Stone, V., Miller, M. R., Clift, M. J. D., Elder, A., Mills, N. L., Moller, P., Schins, R. P. F., Vogel, U., Kreyling, W. G., Aalstrup Jensen, K., TAJ, K., Schwarze, P. E., Hoet, P., Pietroiusti, A., De Vizcaya-Ruiz, A., Beaune-Squiban, A., Teixeira, J. P., Tran, C. L., & Cassee, F. R. (2017). Nanomaterials versus ambient ultrafine particles: An opportunity to exchange toxicology knowledge. Environmental Health Perspectives, 125(10), 1–17. https://doi.org/10.1289/EHP424

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. Particle and Fibre Toxicology, 11, 46. https://doi.org/10.1186/s12989-014-0046-4
Tsoutsoulopoulos, A., Gohlsch, K., Möhle, N., Breit, A., Hoffmann, S., Krischenowski, O., Mücket, H., Gudermann, T., Thiermann, H., Aufderheide, M., & Steinritz, D. (2019). Validation of the CULTEX® radial flow system for the assessment of the acute inhalation toxicity of airborne particles. Toxicology In Vitro, 58, 245–255. https://doi.org/10.1016/j.tiv.2019.03.020

Upadhyay, S., & Palmberg, L. (2018). Air-liquid interface: Relevant in vitro models for investigating air pollutant-induced pulmonary toxicity. Toxicological Sciences, 164(1), 21–30. https://doi.org/10.1093/toxsci/kfy053

Valero-Mora, P. M. (2010). Ggplot2: Elegant graphics for data analysis. Journal of Statistical Software, 35, 1–3. https://doi.org/10.18637/jss.v035.b01

Wang, X., Xia, T., Ntim, S. A., Ji, Z., Lin, S., Meng, H., Chung, C. H., George, S., Zhang, H., Wang, M., Li, N., Yang, Y., Castranova, V., Mitra, S., Bonner, J. C., & Nel, A. E. (2011). Dispersal state of multiwalled carbon nanotubes elicits profibrogenic cellular responses that correlate with fibrogenesis biomarkers and fibrosis in the murine lung. ACS Nano, 5(12), 9772–9787. https://doi.org/10.1021/nn2033055

Warheit, D. B., Sayes, C. M., & Reed, K. L. (2009). Nanoscale and fine zinc oxide particles: Can in vitro assays accurately forecast lung hazards following inhalation exposures? Environmental Science and Technology, 43(20), 7939–7945. https://doi.org/10.1021/es901453p

Wiemann, M., Sauer, U. G., Vennemann, A., Bäcker, S., Keller, J. G., Ma-Hock, L., Wohlleben, W., & Landsiedel, R. (2018). In vitro and in vivo short-term pulmonary toxicity of differently sized colloidal amorphous SiO2. Nanomaterials, 8(3), 160. https://doi.org/10.3390/nano8030160

Wiemann, M., Vennemann, A., Sauer, U. G., Wiench, K., Ma-Hock, L., & Landsiedel, R. (2016). An in vitro alveolar macrophage assay for predicting the short-term inhalation toxicity of nanomaterials. Journal of Nanobiotechnology, 14(1), 1–27. https://doi.org/10.1186/s12951-016-0164-2

Zeiger, E. (2019). The test that changed the world: The Ames test and the regulation of chemicals. Mutation Research—Genetic Toxicology and Environmental Mutagenesis, 841, 43–48. https://doi.org/10.1016/j.mrgentox.2019.05.007

Zhang, H., Ji, Z., Xia, T., Meng, H., Low-Kam, C., Liu, R., Pokhrel, S., Lin, S., Wang, X., Liao, Y. P., Wang, M., Li, L., Rallo, R., Damoiseaux, R., Telesca, D., Mädler, L., Cohen, Y., Zink, J. I., & Nel, A. E. (2012). Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano, 6(5), 4349–4368. https://doi.org/10.1021/nn3010087

How to cite this article: Di Ianni, E., Jacobsen, N. R., Vogel, U., & Møller, P. (2022). Predicting nanomaterials pulmonary toxicity in animals by cell culture models: Achievements and perspectives. WIREs Nanomedicine and Nanobiotechnology, 14(6), e1794. https://doi.org/10.1002/wnan.1794